Dynamic Modeling and Econometrics in Economics and Finance 21

Carl Chiarella
Xue-Zhong He
Christina Sklibosios Nikitopoulos

# Derivative Security Pricing

## Techniques, Methods and Applications

Springer

# Dynamic Modeling and Econometrics in Economics and Finance

Volume 21

**Editors**

Stefan Mittnik
Ludwig Maximillian University Munich,
Munich, Germany

Seminar für Finanzökonometrie, Institut für Statistik,
Ludwig-Maximilians-Universität München,
München, Germany

Willi Semmler
Bielefeld University
Bielefeld, Germany
and
New School for Social Research
New York, USA

Department of Economics
The New School for Social Research
New York, USA

More information about this series at
http://www.springer.com/series/5859

Carl Chiarella • Xue-Zhong He •
Christina Sklibosios Nikitopoulos

# Derivative Security Pricing

## Techniques, Methods and Applications

Springer

Carl Chiarella
Finance Discipline Group
UTS Business School
University of Technology Sydney
New South Wales
Australia

Xue-Zhong He
Finance Discipline Group
UTS Business School
University of Technology Sydney
New South Wales
Australia

Christina Sklibosios Nikitopoulos
Finance Discipline Group
UTS Business School
University of Technology Sydney
New South Wales
Australia

ISSN 1566-0419 ISSN 2363-8370 (electronic)
Dynamic Modeling and Econometrics in Economics and Finance
ISBN 978-3-662-45905-8 ISBN 978-3-662-45906-5 (eBook)
DOI 10.1007/978-3-662-45906-5

Library of Congress Control Number: 2015934448

Springer Heidelberg New York Dordrecht London


Printed on acid-free paper

Springer-Verlag GmbH Berlin Heidelberg is part of Springer Science+Business Media (www.springer.com)

# Preface

This book is an outgrowth of courses we have offered on stochastic calculus and its applications to derivative securities pricing over the last 15 years. The courses have been offered several times to doctoral students and students in the Master of Quantitative Finance and its forerunner programs in the Finance Discipline Group, UTS Business School at the University of Technology Sydney (UTS), and three times to students in the Financial Engineering program at Nanyang Business School in Singapore. It has also served for shorter courses at the Graduate School of International Corporate Strategy at Hitotsubashi University in Tokyo, the Faculty of Economics at the University of Bielefeld in Germany, the Dipartimento di Matematica per le Decisioni at the University of Florence and the Graduate School of Economics at the University of Kyoto.

The aim of the book is to provide a unifying framework within which many of the key results on derivative security pricing can be placed and related to each other. We have also tried to provide an introductory discussion on stochastic processes sufficient to give a good intuitive feel for Ito's Lemma, martingales and the application of Girsanov's theorem. With the explosion of the literature on option pricing in the last four decades, it would obviously not have been possible to cover in one course even a fraction of the main results. Rather, it was our intention that those completing the course would be able to more confidently approach that literature with a good intuitive understanding of the basic techniques, a good overview of how the different parts of the literature relate to each other, and a knowledge of how to implement the theory for their own particular problems. Judging from the feedback we have received, the book has been successful in these aims, and we have been heartened by the very positive response we have had from people who have read it. This includes not only our immediate circle of research collaborators and doctoral students at UTS but also students, researchers and practitioners both in Australia and overseas. The feedback we have received has left us more convinced than we were 15 years ago that this book fills an important gap in the pedagogical finance literature. There are now many excellent monographs and survey papers that treat the revolution of stochastic methods in finance over the last 40 years. However, many of these treatments require of the reader a high degree of, if not "fluency", then certainly “maturity” with the concepts of measure theory or, are written in the very formal lemma/theorem/proof style of modern mathematics. Readers who are not comfortable with these concepts and formal mathematical approaches are left with a feeling of not understanding the essential foundations of the subject and always lack confidence in applying the techniques of stochastic finance. Our aim in this book is to present to the reader a treatment which emphasises more the financial intuition of the material, and which is at the same mathematical level and uses the same basic hedging arguments of the early papers of Black–Scholes and Merton, which sparked off the revolution to which we have referred above. Uppermost in our mind has been the desire to give the reader an intuitive feel for the many difficult mathematical concepts that will be encountered in working through this book. Whilst the mathematical level is demanding, it should nevertheless be attainable for readers who are comfortable with an intermediate level of calculus and the non-measure theoretic approach to probability theory.

By the foregoing remarks, we do not intend to downplay or denigrate the importance of the modern measure theoretic approach to the theory of diffusion processes and semimartingale integration. We are acutely aware of the fact that many of the subtleties of stochastic finance require these advanced techniques for their proper elucidation. Furthermore, many of the important advances of the last three decades would not have occurred, or would have been much slower in coming, without their use. However, stochastic finance is rapidly evolving from its pure science phase to its applied science and engineering phase. As a result, there is a greater influx into the area of academics and practitioners who are neither “fluent” nor even “comfortable” with measure theoretic arguments and the formal style of modern mathematics. It is to this audience that this book is addressed. Of course the challenge in writing a book at a more intuitive level is to do so in a way that is respectful of the many subtleties that the measure theoretic approach and more formal mathematical approach have been developed to address. We have done our best to meet this challenge; however, we are mindful of many shortcomings that may still exist.

Another feature of the book is the set of problems that has been developed to accompany each chapter. Here, we have tried to include exercises that cover many of the key results and examples that have become significant in applications or in subsequent theoretical developments. As we very firmly believe that a full understanding of stochastic methods in finance can only be attained when one can simulate and compute the quantities that one is discussing, we have also included a number of computational exercises.

The evolution of our thinking about stochastic methods in finance has been greatly assisted by John Van der Hoek of the University of South Australia and our UTS colleague Eckhard Platen. They have been most generous in sharing their knowledge both in private conversations and in courses which they have kindly presented at UTS. We would also like to thank some of our former doctoral students, in particular Nadima El-Hassan, Garry de Jager, Ramaprasad Bhar, Oh Kang Kwon, Adam Kucera, Shenhuai Gao, Thuy Duong Tô and Andrew Ziogas. Numerous discussions and debates with them over recent years have helped, if not them, then certainly ourselves to clarify a number of technical points discussed in this book. Thanks are also due to Andrew Ziogas, Nicole Mingxi Huang and Hing Hung for developing the MATLAB programs used to do the various simulations. We are grateful to Mark Craddok and Boda Kang for checking through a number of mathematical derivations and making some valuable suggestions. We are also indebted to Simon Carlstedt for checking thoroughly through the book and pointed out a number of errors and inconsistencies. Finally, we would like to acknowledge the efforts of Xiaolin Miao, Yuping Wu, Jingfeng He, Xuli Huang, Lifang Zhang, Jenny Yixin Chen, Shing-Yih Chai, Laura Santuz, Gwen Tran, Stephanie Ji-Won Ough and Linh Thuy Tô who have worked diligently and under much pressure to prepare the various drafts and the many graphs. However, all of the aforementioned persons should be totally absolved from any blame for any errors, omissions or confusions that this book may still contain.

University of Technology Sydney
August 14, 2014

Carl Chiarella
Xue-Zhong He
Christina Sklibosios Nikitopoulos

# Contents

## Part II Interest Rate Modelling

# Part I
# The Fundamentals of Derivative Security Pricing

# Chapter 1
# The Stock Option Problem

**Abstract** This chapter outlines the paradigm problem of option pricing and motivates key concepts and techniques that we will develop in Part I when the risk-free rate is deterministic.

## 1.1 Introduction

In this course we shall enter into a lot of the technical details involved in the pricing of derivative securities. We shall first in Part I consider economies in which the risk-free interest rate is deterministic and then in Part II is stochastic.

In this chapter we shall outline the paradigm problem of an option written on a stock in an environment of deterministic interest rates. This paradigm problem will motivate many of the concepts and techniques that we develop in Part I. We also consider other more complicated derivative securities that are of interest in financial markets, and the theory that will allow us to price.

## 1.2 The European Call Option

A European call option on a stock is a contract that gives the purchaser the right (but not the obligation) to purchase the underlying stock at an agreed price (the exercise price) at a fixed date (maturity) in the future. Thus if the stock price at maturity is above the exercise price then the holder would exercise the option, but would not do so if the stock price at maturity is below the exercise price.

The essential characteristics of the European call option situation are illustrated in Fig. 1.1. We use $S$ to denote the underlying stock price, $E$ the exercise price, $C$ the value of the option and $T$ the maturity date. The option contract is assumed to be initiated at time 0, at which point in time our hypothetical investor seeks to value the option. If the stock price realizes path 1 over the life of the option and finishes above $E$ then the payoff to the investor will be $S_T^{(1)} - E$ (since he or she can purchase for $E$ something worth more than $E$). Conversely if the stock price realizes path 2 over the life of the option then the investor will not exercise the option (since he or she



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_1

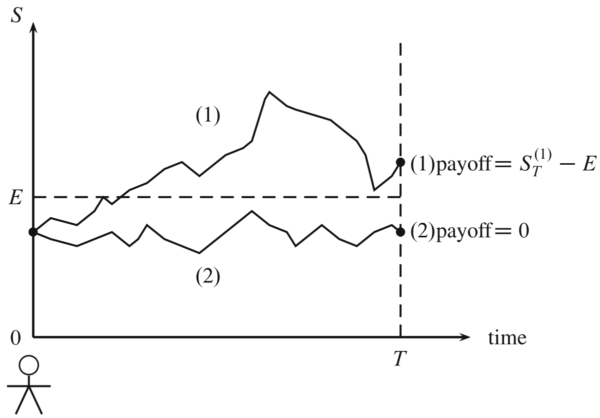


**Fig. 1.1** The characteristics of a European call option

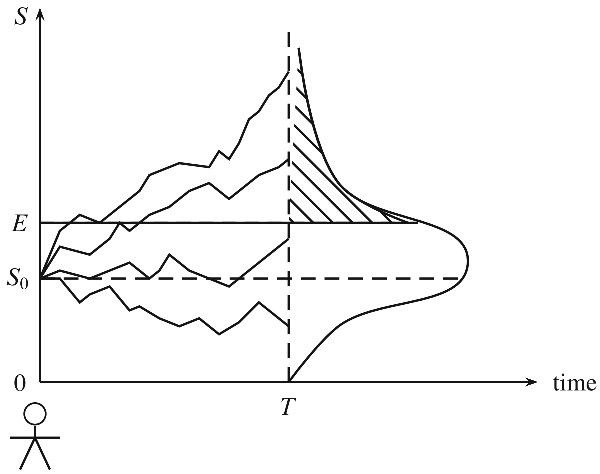


**Fig. 1.2** Visualising the distribution of $S_T$ conditional on $S_0$

will not purchase the stock for $E$ when it is worth less than $E$) and so the payoff to the option will be 0. Using $S_T$ to denote the stock price at maturity and $C_T$ the option payoff at maturity, then the option payoff may be written

$$C_T = \max(S_T - E, 0). \tag{1.1}$$

The problem our investor faces is to determine the value of the option of some time prior to maturity $T$, say time 0 in Fig. 1.2. Let us consider for a moment what is involved in this valuation problem and what will be the mathematical concepts and tools that we need to develop in order to finally carry out the valuation.

The first, and perhaps most important, point to stress is that the option valuation problem is nothing more than a discounted expected cash flow calculation of the type that one encounters in elementary finance courses. The option has a positive payoff if the stock price at maturity is above $E$. In fact for any particular stock price path, the payoff is given by (1.1). Clearly the first thing we need to calculate is the expected value of $\max(S_T - E, 0)$ and then discount this quantity back to time zero. The two main problems (and these are the problems that make option pricing technically complex) are first, to determine the precise distribution to be used in calculating the expectation referred to and then actually perform this calculation, and second, to determine the appropriate discount rate.

As far as the distribution is concerned Fig. 1.2 helps us to visualize what is involved. We show the simulation of a number of stock price paths (think of each of these as a possible realization of the market over the period $(0, T)$). The figure also indicates the distribution of the stock price at final time (calculated using a large number of simulations). Essentially we are interested in the probability that the stock price at $T$ finishes above $E$, and this is essentially the shaded area indicated on the distribution.

It should by now be clear that an essential ingredient of the solution to our option valuation problem is some theory about how stock prices (or asset prices more generally) move stochastically and how to calculate the distributions such as the one shown in Fig. 1.2. We will then need to develop some methods that allow us to calculate expectations of payoffs such at (1.1) with respect to such distributions.

In Chap. 2 we will review the essential aspects of the theory of stochastic processes, from the point of view of obtaining the distribution of the stock price at time $T$ conditional on observing a certain stock price at time 0. In Chap. 3 we shall combine this understanding with the basic notion from finance of discounted expected cash flow to make a first attempt at valuing a European option. The expression we obtain involves (a lengthy) integration, the result of which is the celebrated Black–Scholes option pricing formula. We go on to show how the integral expression can be re-expressed as a partial differential equation. In the course of these derivations we also encounter, albeit in a simple guise, the important concepts of a martingale and the Feynman–Kac formula. Both of these concepts are elaborated upon considerably in later chapters, and indeed play a very important role in the development of the ideas in this book.

However the derivation of the option pricing formula in Chap. 3 leaves a lot of loose ends, in particular it is not at all clear why we need to assume that investors behave as if they were risk neutral. As a prelude to a derivation that tidies up the loose ends we introduce stochastic integrals, stochastic differential equations and Ito's Lemma in Chaps. 4–6. Now the viewpoint on stochastic processes switches from a focus on conditional distributions to a focus on sample paths. It is this focus that in Chap. 7 allows us to derive the option pricing partial differential equation using the continuous hedging argument. Chapter 9 outlines the use of the Fourier transform technique to solve this type of partial differential equations that is frequently encountered in financial economics. In Chap. 8 we develop the martingale interpretation of the continuous hedging argument, which turns out to be a very powerful way of viewing derivative pricing problems as we shall see in later chapters.

Part I deals with embellishments to the basic Black–Scholes–Merton framework (allowing for more general stochastic processes for the underlying asset), with various aspects of solution techniques, or with extensions such as American option pricing. In Part II, we show how to allow for the risk free interest rate to be stochastic in the standard stock option problem. Chapter 20 introduces the very powerful change of numeraire technique. From Chap. 21 onwards we deal with derivatives that gain their value from the stochastic evolution of interest rates. The contents of these chapters will be previewed in Chap. 21 after we introduce the paradigm interest rate option problem.

# Chapter 2
# Stochastic Processes for Asset Price Modelling

**Abstract** This chapter gives an intuitive appreciation and review of many important aspects of the stochastic processes that have been used to model asset price processes. We will be interested in a probabilistic description of the time evolution of asset prices. After imposing some structure on the stochastic process for the return on the asset, this chapter introduces Markov processes, time evolution of conditional probabilities, continuous sample paths, and the Fokker–Planck and Kolmogorov equations.

## 2.1 Introduction

We shall be much concerned with how asset prices evolve over time. It was realised early in the development of the modern theory of finance that since asset prices are evolving randomly over time the best description of price behaviour would be a probabilistic one, which involves using ideas from the theory of stochastic processes. The theory of stochastic processes is not an easy theory to master, since many of its important concepts were developed roughly simultaneously in a variety of disciplines such as electrical engineering, theoretical physics and pure mathematics. The perspective taken in each of these disciplines is slightly different and the same concepts can be presented at vastly different levels of mathematical abstraction.

Our aim in this book is not at all to present a fully rigorous discussion of the theory of stochastic processes. Rather we merely attempt to give an intuitive appreciation and review of those aspects of the theory which have found application in modern finance theory. In putting together the discussion and viewpoint on stochastic processes from this chapter up to Chap. 8 we have drawn heavily on Malliaris and Brock (1982), Astrom (1970), Harrison (1990), Baxter and Rennie (1996) with some ideas from Gardiner (1985) and Horsthemke and Lefever (1984). A more complete mathematical treatment of stochastic processes and stochastic differential equations may be found in the books by Oksendal (2003), Krylov (1995) and Kijima (2002). The book by Oksendal (2003) is particularly recommended.

We will be interested in a probabilistic description of the time evolution of asset prices. Empirical examination of time series of asset prices and asset returns



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_2

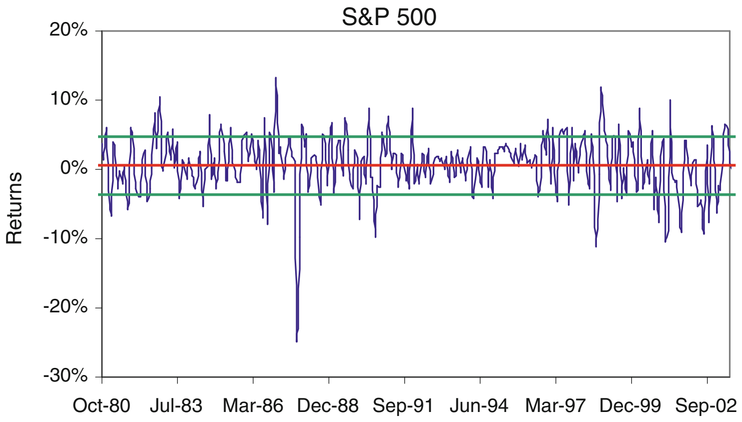


**Fig. 2.1** Intuitive notion of evolution of returns

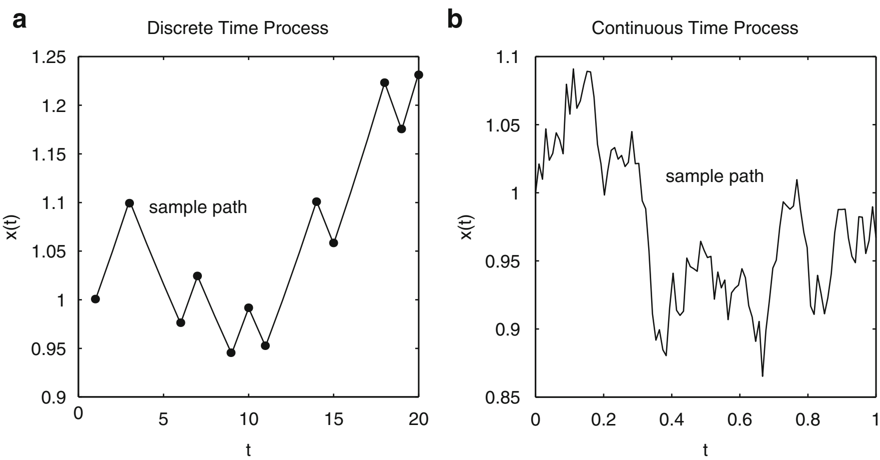


**Fig. 2.2** Sample paths of stochastic processes

suggests that at least intuitively we might initially think of the *return* on the asset as consisting of an average mean component (which we might regard as more or less certain) and some volatile, stochastic component about this mean, as shown in Fig. 2.1 which shows the time series of monthly returns on the $S\&P\,500$ from January 1980 to October 2003 together with the mean and one standard derivation band of the entire series. Already with this intuitive notion we are imposing some structure on the stochastic process for the return on the asset. Much of what we do in Chaps. 2 and 3 will be to give some mathematical description to this intuitive notion.

To make the discussion a little more formal let $x(t)$ denote either the price of or return on the asset at time $t$. In Fig. 2.2 we represent a typical sequence of prices which would be observed over a given time period. Since, depending on the application, we need to concentrate on prices or returns in both *discrete time* and *continuous time*, both are represented in Fig. 2.2.[1]

The stochastic process for the prices (or returns) may be thought of as a family of random variables $\{x(t) \mid t \in \mathscr{T}\}$, where $\mathscr{T}$ denotes the set of values to which the time parameter $t$ belongs, more formally known as the *index set*. For the *discrete time process* in Fig. 2.2a the index set is the set of non-negative integers $\mathscr{T} = \{0, 1, 2, \ldots\}$, whilst for the *continuous time process* in Fig. 2.2b the index set is the set of all $t$ between 0 and infinity, i.e. $\mathscr{T} = \{t \mid 0 \leq t < \infty\}$.

The set of values which $x(t)$ may take is known as the *sample space* and is denoted $\Omega$. For the price paths illustrated in Fig. 2.2 the sample space is all values from 0 to infinity. Generally we will be interested in prices belonging to a subset of $\Omega$, such as the set of prices illustrated in the shaded area in Fig. 1.2. In more formal discussions we may see the stochastic process denoted as $\{x(t, \omega) \mid t \in \mathscr{T}, \omega \in \Omega\}$. A set of particular values arising from the stochastic process is known as a *realisation* of the process, or a *sample path*. In the preceding formal notation, the $\omega$ refers to one sample path out of the set $\Omega$ of all possible sample paths.

A major technical problem in the theory of stochastic processes involves assigning a probability distribution or more formally a probability measure to subsets $\Omega$ of the sample space $\omega$. To do this mathematically correctly requires a great deal of measure theory, however provided the stochastic process assumed for $x(t)$ is not too "wild" then we are able to proceed with a fairly intuitive understanding of probability distributions. However to deal with more sophisticated processes (e.g. Lévy processes) then we do need to resort to a more formal mathematical description.

## 2.2 Markov Processes

In order to put some mathematical flesh on the basic notion of a stochastic process we need to introduce the concept of *joint probability density function*. As we said at the end of the previous section, proving that such a density function can be found is an intricate mathematical problem which we shall not touch on in this book. Malliaris and Brock (1982) outline some of the intricacies involved and give appropriate references (see their Chapter 1 and Section 7).

Given values $x_1, x_2, x_3, \ldots, x_k$ of the asset price $x(t)$ at times $t_1, t_2, t_3, \ldots, t_k$, we assume that we can obtain a joint probability density function

$$p(x_k, t_k; x_{k-1}, t_{k-1}; \ldots; x_1, t_1) \qquad \text{(note that } t_1 \leq t_2 \leq \cdots \leq t_k\text{)}$$

[1]In terms of concepts to be developed later, Fig. 2.2a represents the simulation of a binomial process, whilst Fig. 2.2b is the simulation of a geometric Brownian motion process starting at $x_o = 1$.

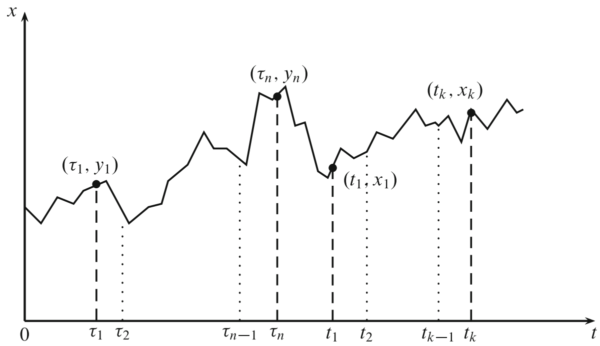


**Fig. 2.3** Conditional probability of a sample path

which measures the joint probability that $x(t_1) = x_1, x(t_2) = x_2, \ldots, x(t_k) = x_k$. Using joint probability density functions, we can also define *conditional probability density functions:*

$$p(x_k, t_k; \ldots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n; \ldots; y_2, \tau_2; y_1, \tau_1) = \frac{p(x_k, t_k; \ldots; x_2, t_2; x_1, t_1; y_n, \tau_n; \ldots; y_2, \tau_2; y_1, \tau_1)}{p(y_n, \tau_n; \ldots; y_2, \tau_2; y_1, \tau_1)}. \tag{2.1}$$

The left-hand side of (2.1) is the probability that the price sequence $(x_1, t_1; x_2, t_2; \ldots; \ x_k, t_k)$ will be observed *given that* the price sequence $(y_1, \tau_1; y_2, \tau_2; \ldots; y_n, \tau_n)$ has just been observed over the previous periods $\{\tau_1, \tau_2, \ldots, \tau_n\}$; see Fig. 2.3.

On the right-hand side of (2.1) the probability on the top line is that of observing the price sequence $(y_1, \tau_1), (y_2, \tau_2), \ldots, (y_n, \tau_n), (x_1, t_1), \ldots, (x_k, t_k)$, whilst the probability on the bottom line is that of observing the price sequence $(y_1, \tau_1), \ldots, (y_n, \tau_n)$. To see the sense of this last formula think of the probabilities as observed frequencies and suppose the $\{y_i\}$ represent a sequence of prices growing at a rate of 3 % and the $\{x_i\}$ represent a sequence of prices growing at a rate of 4 %. Then the formula states that the probability of observing a 4 % rise, given that a 3 % rise has occurred equals the frequency of 3 % rises followed by 4 % rises divided by the frequency of 3 % rises.

A very simple kind of stochastic process that we might deal with is a completely independent one. A stochastic process is said to be completely *independent* if the probability of observing a given price at time $t$ is completely independent of the probability of observing some price at any other time. This allows us to write the joint probability density function as a product of independent probabilities, so that

$$p(x_k, t_k; \ldots; x_2, t_2; x_1, t_1) = p(x_k, t_k) \ldots p(x_2, t_2) p(x_1, t_1). \tag{2.2}$$

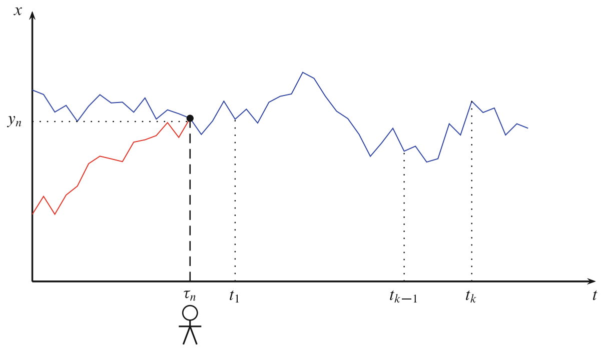


**Fig. 2.4** A Markov process

The next most simple stochastic process we might deal with is a *Markov process* in which knowledge only of the present state of the process is relevant to the future evolution of the process. Referring to the price sequences in Fig. 2.3 this idea may be expressed in terms of conditional probabilities as

$$\begin{aligned} p(x_k, t_k; \dots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n; \dots; y_2, \tau_2; y_1, \tau_1) \\ = p(x_k, t_k; \dots; x_2, t_2; x_1, t_1 \mid y_n, \tau_n). \end{aligned} \tag{2.3}$$

This idea is illustrated in Fig. 2.4 where we see two possible paths arriving at $(y_n, \tau_n)$. Irrespective of which path has been followed to arrive at $(y_n, \tau_n)$, the future evolution from $\tau_n$ is only conditional on $(y_n, \tau_n)$. Markov processes are clearly related to the efficient markets concept.

The Markov assumption (i.e. Eq. (2.3)) is particularly important because it enables us to define all relevant joint probability density functions in terms of simple conditional probabilities, such as $p(x_1, t_1 \mid y_1, \tau_1)$.

To see this consider the following manipulations over the successive times $\tau_n, t_1, t_2$. By the definition of conditional probability density (Eq. (2.1)) and conditioning on $y_n, \tau_n$,

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2; x_1, t_1 \mid y_n, \tau_n) p(y_n, \tau_n). \tag{2.4}$$

But conditioning on $x_1, t_1; y_n, \tau_n$ we have

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2 \mid x_1, t_1; y_n, \tau_n) p(x_1, t_1; y_n, \tau_n). \tag{2.5}$$

Combining these last two equations we obtain

$$\begin{aligned} p(x_2, t_2; x_1, t_1 \mid y_n, \tau_n) &= \frac{p(x_2, t_2 \mid x_1, t_1; y_n, \tau_n) p(x_1, t_1; y_n, \tau_n)}{p(y_n, \tau_n)} \\ &= p(x_2, t_2 \mid x_1, t_1) p(x_1, t_1 \mid y_n, \tau_n). \end{aligned} \tag{2.6}$$

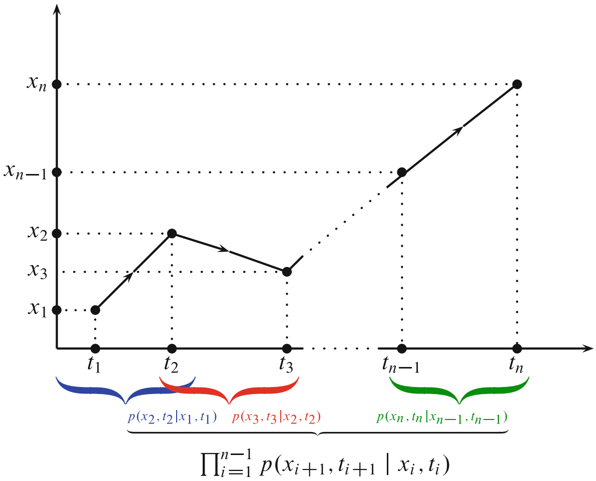


**Fig. 2.5** Illustrating the joint density as a product of conditional density functions

The first equality is due to the definition of conditional probability (2.1), the second follows by the Markovian assumption and uses again of (2.1). Substituting (2.6) to (2.4) we finally have

$$p(x_2, t_2; x_1, t_1; y_n, \tau_n) = p(x_2, t_2 \mid x_1, t_1) p(x_1, t_1 \mid y_n, \tau_n) p(y_n, \tau_n),$$

which simply states that the joint probability density function over the times $\tau_n, t_1, t_2$ is the product of the conditional probability densities over successive time intervals.

The same argument can be extended to any number of realisations of the stochastic process over successive times, to yield[2]

$$p(x_n, t_n; \ldots; x_2, t_2; x_1, t_1) = p(x_1, t_1) \prod_{i=1}^{n-1} p(x_{i+1}, t_{i+1} \mid x_i, t_i). \tag{2.7}$$

The essential feature of (2.7) from the point of view of applications is that the joint density function can be expressed as a product of conditional density functions (over successive time intervals), as illustrated in Fig. 2.5. Thus a statistical description of price and return dynamics is reduced to a description of the conditional density function.

Certainly Markov processes provide the most convenient tool for the modelling of asset prices and returns as we shall see throughout this book. On the one hand, they accord very nicely with the notion of efficiency of financial markets. On the

[2]We recall the product notation $\Pi_{i=1}^{n} X_i = X_1 X_2 \cdots X_n$. Note that in (2.7) we change slightly the notation and take $(x_1, t_1)$ as the initial point.

other hand, they allow us to make use of the highly developed theories of diffusion processes and semi-martingale integration. However it is nevertheless the case that non-Markovian processes do also play a role in the modelling of some aspects of financial markets behaviour. This is the case for instance when considering stochastic volatility models. It is more particularly the case in the modelling of interest rate sensitive derivative securities. It turns out that the most natural (and general) process for modelling the dynamic evolution of the yield curve is a non-Markovian one. We shall see in Part II that rather than working in a non-Markovian framework it turns out to be more convenient to find ways to reduce the non-Markovian process to a Markovian system of higher dimension. In this way the great mathematical convenience of Markovian processes is preserved.

## 2.3 The Time Evolution of Conditional Probabilities

As discussed in the previous sub-section, the Markov assumption implies that in order to obtain a statistical description of prices in a dynamically evolving environment we need to know how the conditional probability density functions evolve over time.

The equation which allows us to do this is the *Chapman–Kolmogorov equation* which is a simple consequence of the Markovian assumption. If $t_1 < t_2 < t_3$ then the Chapman–Kolmogorov equation states that

$$p(x_3, t_3 \mid x_1, t_1) = \int p(x_3, t_3 \mid x_2, t_2) p(x_2, t_2 \mid x_1, t_1) dx_2. \tag{2.8}$$

To see the sense of this equation consider the path $I$ in Fig. 2.6. The probability of going from $x_1$ at $t_1$ to $x_3$ at $t_3$ is the product of the probability of going from $(x_1, t_1)$ to $(x_2, t_2)$, with the probability of going from $(x_2, t_2)$ to $(x_3, t_3)$. The integral in (2.8) sums over all such probabilities by ranging over all possible paths through values $x_2$ at $t_2$. Figure 2.6 shows three such paths.

The Chapman–Kolmogorov equation is in fact a complex, nonlinear functional equation due to the fact that so far the nature of the stochastic process has been left very general. In order to reduce it to a form easier to deal with mathematically we need to put more restrictions on the nature of the stochastic process. In particular the magnitude and type of change that can occur in $x$ from one time period to the next. In particular it can be shown that if the price changes are small over small intervals of time in a way to be made more precise below then the Chapman–Kolmogorov equation reduces to a partial differential equation for the conditional probability which has a remarkable similarity to the partial differential equation governing stock option prices. We will eventually show how these two partial differential equations are related.

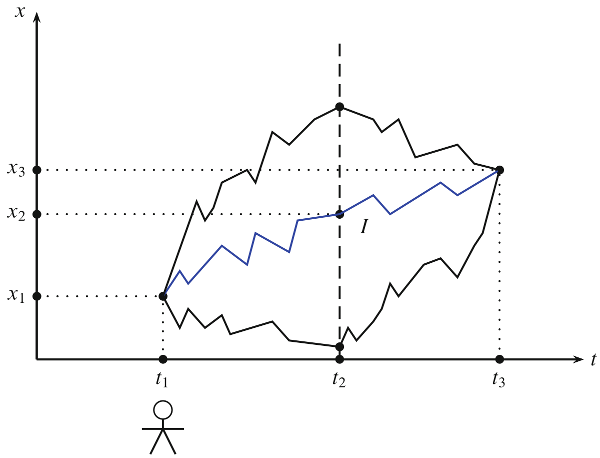


**Fig. 2.6** The Chapman–Kolmogorov equation

## 2.4 Processes with Continuous Sample Paths

As just indicated we shall for the moment focus on price (or return) processes that change by small amounts over small intervals of time. The notion that the price changes by small amounts over a small interval of time is made mathematically more precise by introducing stochastic processes having continuous sample paths.

The mathematical condition that needs to be imposed on the conditional probabilities in order that the sample paths be continuous functions of time is

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} p(x, t + \Delta t \mid z, t) dx = 0, \tag{2.9}$$

for any $\varepsilon > 0$. The sense of this condition is easily understood by referring to Fig. 2.7. Typically $\varepsilon$ would be small and the set $|x - z| > \varepsilon$ (indicated by the hashed region in Fig. 2.7) represents the set of prices $x$ at time $t + \Delta t$ which are further than a distance $\varepsilon$ from the original price of $z$ at time $t$. The probability $p(x, t + \Delta t \mid z, t)$ is the probability of observing the price $x$ at time $t + \Delta t$, given that the price at time $t$ is $z$. Thus the integral in Eq. (2.9) represents the total probability of observing at time $t + \Delta t$ a price which is further than $\varepsilon$ from the current price $z$ at time $t$. The overall condition in Eq. (2.9) states that this probability must decline more rapidly than $\Delta t$ as $\Delta t$ becomes smaller and smaller (e.g. the total probability could be proportional to $(\Delta t)^2$).

The condition in Eq. (2.9) is known as the *Lindeberg condition*, and stochastic processes whose conditional probabilities satisfy it will be expected to experience small changes in $x$ over small intervals of time. Such stochastic processes will also display continuous sample paths as has already been stated.

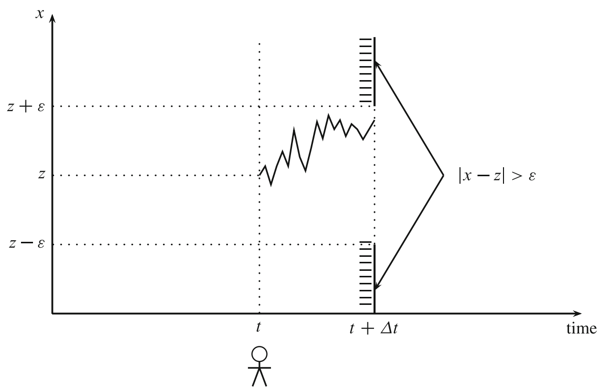


**Fig. 2.7** The Lindeberg condition; the price changes remain within the band $(z+\varepsilon, z-\varepsilon)$

At this point it may be worth considering two particular forms of conditional probability functions, one which satisfies the Lindeberg condition and one which does not.

### *2.4.1 Brownian Motion*

Consider the conditional probability density given by the formula

$$p(x, t+\Delta t \mid z, t) = \frac{1}{\sqrt{2\pi\Delta t}\,\sigma}\exp\left[\frac{-(x-z)^2}{2\sigma^2\Delta t}\right], \tag{2.10}$$

meaning that $x$ is normally distributed, $x \sim N(z, \sigma^2\Delta t)$, centred on the current price $z$ at time $t$ and having variance $\sigma^2\Delta t$. According to this distribution the prices expected at $t+\Delta t$ are distributed normally about the current price $z$ at $t$, as illustrated in Fig. 2.8. As $\Delta t$ becomes smaller, the distribution becomes more peaked, thereby reducing the probability that the price at $t+\Delta t$ will be very far from $z$. It is a simple (albeit tedious) exercise in integration to show that the Lindeberg condition is indeed satisfied for this distribution (the details are given in Appendix 2.2).

The stochastic process whose conditional probabilities are given by Eq. (2.10) is known as *Brownian Motion*, and is widely applied in financial economics, partly because of its mathematical tractability, and partly also because empirical studies indicate that many (though no means all) important asset prices (or returns) are reasonably well modelled by it.

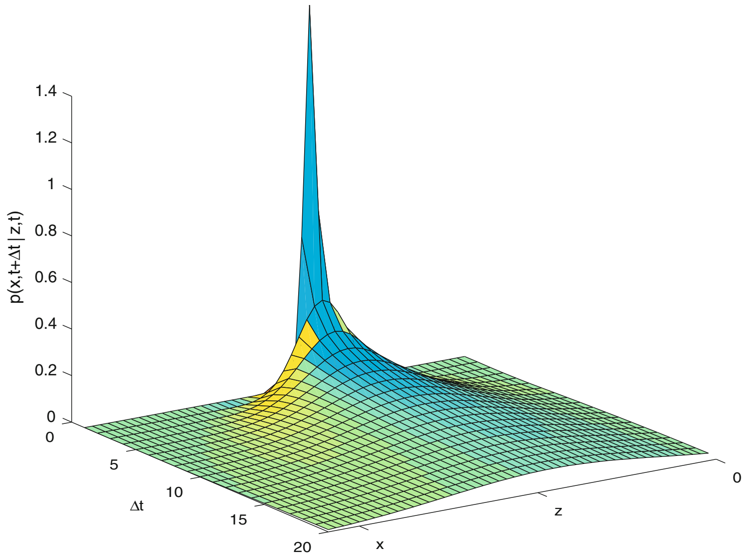


**Fig. 2.8** Conditional probability density of Brownian motion. Note how decreasing $\Delta t$ "squeezes" the distribution around the current price $z$

## 2.4.2 *The Cauchy Process*

In order that the reader not gain the impression that all bell shaped conditional probability distributions satisfy the Lindeberg condition and therefore give rise to stochastic processes having continuous sample paths, consider the distribution with conditional probability density

$$p(x, t + \Delta t \mid z, t) = \frac{\Delta t}{\pi[(x - z)^2 + (\Delta t)^2]}. \tag{2.11}$$

This distribution also has the general bell shape, shown in Fig. 2.8 for the Brownian Motion, which also becomes more peaked as $\Delta t$ becomes smaller. However it can be shown that this distribution, known as the *Cauchy distribution*, does *not* satisfy the Lindeberg condition (see Appendix 2.3) and therefore $x$ is not a continuous process. This means that as $\Delta t$ becomes small, the probability of observing a price well away from current price $z$ does not become small quickly enough, leaving a positive probability that there will be large jumps in the price from time to time.

Indeed simulations of both processes indicate that the Cauchy process exhibits large jumps not infrequently, as illustrated by the simulations in Fig. 2.9.[3] There we

[3]The simulations were calculated as follows. For the Brownian motion process, the path was calculated using $x_{i+1} = x_i + x_i \mu \Delta t + x_i \sigma \sqrt{\Delta t}\ \xi_i$, where $\mu = 0.1$, $\sigma = 0.2$, $\Delta t = 0.002$

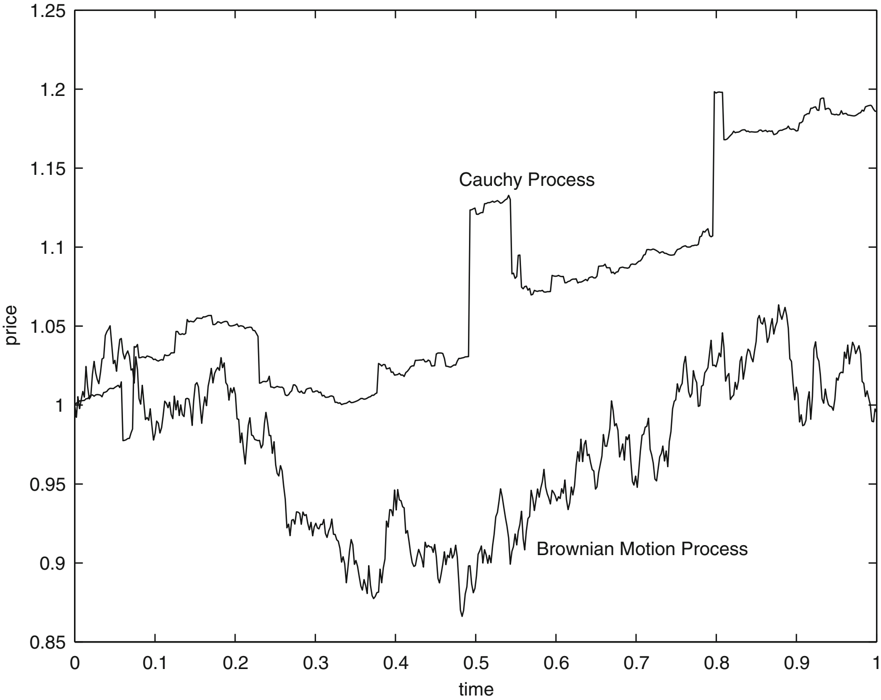


**Fig. 2.9** Typical sample paths for the Brownian motion and Cauchy process

display a typical sample path of both the Brownian motion process and the Cauchy process.

## 2.5 The Dirac Delta Function

The two examples of the proceeding section can also serve as a vehicle for another important concept that we shall need when we come to discuss the solution of option pricing equations, namely the concept of the *Dirac delta function*.

Note first of all that integrating the conditional probability density functions (2.10) and (2.11) with respect to $x$ reveals that the area under the distribution curves is equal to 1, irrespective of the value of $\Delta t$. Of course this is a fundamental requirement of conditional probability density functions. As $\Delta t \rightarrow 0$ the distribution in Fig. 2.8 becomes more and more peaked. Close to the limit $\Delta t = 0$

---

and $\xi_i \sim N(0, 1)$, with $i = 1, 2, \cdots, 500$. For the Cauchy process, the path was calculated using $y_{i+1} = y_i + y_i \mu \Delta t + \mu_i \sigma \Delta t \cot[\pi \xi_i]$.

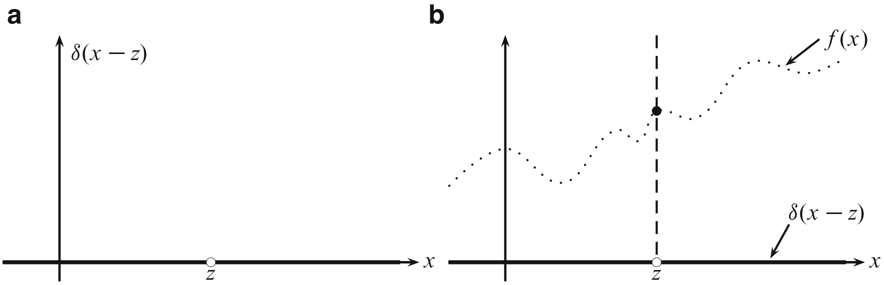


**Fig. 2.10** The sense of the definition of the Dirac delta function. (**a**) $\delta(x-z)=0, \quad x \neq z$. (**b**) $\int_{-\infty}^{\infty} \delta(x-z) f(x) dx = f(z)$

we would see a function which is almost zero everywhere except at $x = z$, and at this point its value becomes very large; all in such a way that the area beneath the distribution curve remains 1 for all $\Delta t$.

In order to formally (as opposed to mathematically rigorously) carry out mathematical operations involving the function obtained from this limiting process, the so-called *Dirac delta function* has been introduced (see e.g. Lighthill 1980). This function is usually denoted by the symbol $\delta$ and is formally defined by:

$$\text{(a)} \;\; \delta(x-z)=0, \quad x \neq z, \qquad \text{(b)} \;\; \int_{-\infty}^{\infty} \delta(x-z) f(x) dx = f(z). \tag{2.12}$$

The sense of both parts of this definition are illustrated in Fig. 2.10.

Both the Brownian motion distribution and the Cauchy distribution in the limit $\Delta t \to 0$ satisfy the formal definition of the Dirac delta function, i.e.

$$\delta(x-z) = \lim_{\Delta t \to 0} p(x, t + \Delta t \mid z, t), \tag{2.13}$$

as illustrated in Fig. 2.8.

To see the economic sense of the second condition in Eq. (2.12), suppose that $f(x)$ is some payoff on the asset at time $t + \Delta t$ if the asset price is $x$ at that point in time. Then

$$\int_{-\infty}^{\infty} p(x, t + \Delta t \mid z, t) f(x)\, dx,$$

is the expected payoff, calculated at time t when the asset price is $z$. The condition (b) of (2.12) states that as $\Delta t \to 0$, the expected payoff becomes the payoff that would be obtained at the current price $z$.

## 2.6 The Fokker–Planck and Kolmogorov Equations

The Chapman–Kolmogorov equation (2.8) tells us how the conditional probabilities are evolving over time. Up to this point in the discussion we have imposed very little structure on the stochastic process apart from the Markov assumption. We would like to reduce the Chapman–Kolmogorov equation to something which is more mathematically tractable and at the same time which involves parameters whose values we could measure from statistical observations on past and current prices.

The above aim could be achieved in a number of ways, but the simplest would be to restrict our attention to those stochastic processes having continuous sample paths, such as those discussed in Sect. 2.4 but with the possibility of sudden large jumps from time to time. In adopting this approach we shall give some mathematical precision to the intuitive notion, which we mentioned in Sect. 2.1, of the asset return consisting of a certain mean component about which there is a stochastic or volatile component.

In particular we shall restrict our attention to stochastic processes whose conditional probability density function satisfies the following three conditions. For all $\varepsilon > 0$,

$$\text{(i)}\quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} (x-z)p(x, t+\Delta t \mid z, t)dx = A(z,t), \tag{2.14}$$

$$\text{(ii)}\quad \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} (x-z)^2 p(x, t+\Delta t \mid z, t)\, dx = B(z,t), \tag{2.15}$$

$$\text{(iii)}\quad \lim_{\Delta t \to 0} \frac{p(x, t+\Delta t \mid z, t)}{\Delta t} = J(x \mid z, t), \tag{2.16}$$

where in (iii) the convergence is uniform in $x$, $z$ and $t$, for $\mid x - z \mid \geq \varepsilon$.

What are these conditions saying? Recall first of all that typically $\varepsilon$ is small, so that $\mid x - z \mid < \varepsilon$ refers to the set of prices $x$ at $t + \Delta t$ which have not moved very far from the current price $z$ at $t$, see Fig. 2.11, while $\mid x - z \mid \geq \varepsilon$ refers to the set of prices $x$ at $t + \Delta t$ which have moved more than a small amount from current price $z$ at $t$.

The integral in condition (i) is the mean of the small (i.e. within $\varepsilon$ of $z$) price changes over the time interval $(t, t + \Delta t)$. The condition thus states that

$$\text{the mean of small price changes over } (t, t+\Delta t) \cong A(z,t)\Delta t.$$

The choice of $A(z,t)$ may be imposed by the financial analyst but would usually be obtained from empirical analysis of past behaviour of asset prices. For example for common stock prices, the form $A(z,t) = \mu z$ where $\mu$ is the mean stock return per unit time fits fairly well with observed price behaviour. On the other hand for short term interest rates empirical evidence suggests the form $A(z,t) = \kappa(\bar{z} - z)$ where $\bar{z}$ is a long run average short term rate and $\kappa$ is a speed of adjustment constant, both of which could be determined from observed interest rate behaviour.

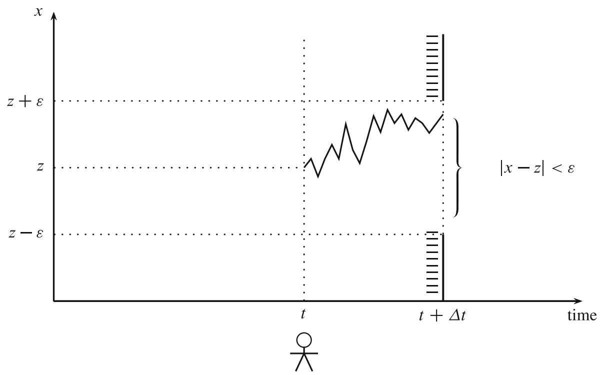


**Fig. 2.11** The price range in conditions (i) and (ii)

The integral in condition (ii) is the second moment of small price changes over $(t, t + \Delta t)$ and thus states that

$$\text{the second moment of small price changes over } (t, t + \Delta t) \cong B(z, t)\Delta t.$$

As with $A(z, t)$ the form of $B(z, t)$ would be determined from the time series behaviour of the asset of interest. Continuing the examples cited above, for common stock prices empirical studies suggest $B(z, t) = \sigma^2 z^2$ ($\sigma$ a constant), whilst for short term interest rates $B(z, t) = \sigma^2 z^{2\gamma}$ ($\sigma, \gamma$ constant, $0 < \gamma < 2$) seems to be the consensus of a range of empirical studies.

The third condition concentrates on price changes which are not small (i.e. are more than a distance $\varepsilon$ from the current price $z$). Condition (iii) states that

$$\text{the probability of large changes over } (t, t + \Delta t) \cong J(x \mid z, t)\Delta t.$$

In essence the quantity $J(x \mid z, t)$ captures the probability that the price will jump from $z$ to $x$ at time $t$. As with the $A$ and $B$ functions, the $J$ function would be obtained from observations on past price movements.

If we assume, or confirm by observation, that $J = 0$, then the asset price will only exhibit small price changes over small time intervals. The stochastic process is then known as a *diffusion process*. We know from the discussion in Sect. 2.4 that the sample paths of such processes are continuous.

If we allow $J \neq 0$ then we are admitting the possibility of sudden large jumps in the asset price. The frequency and magnitude of these jumps determine the functional form of $J$. For instance, in order to model the large jumps in the prices of some of the assets in which we will be interested (e.g. foreign exchange rates), there is empirical evidence that the Poisson process is an appropriate form for $J$.

The $A(z, t)$ term is referred to as *the drift term* of the stochastic process, the $B(z, t)$ term is referred to as *the diffusion term* of the stochastic process and the $J(x \mid z, t)$ term is referred to as *the jump term* of the stochastic process. A stochastic process in which all of these terms are present is known as a *jump-diffusion process*.

The sample paths of such a stochastic process are not continuous in general (unless of course the jump component is zero). Figure 2.9 illustrates the difference between the sample paths of a diffusion process (the Brownian motion process) and a jump-diffusion process (the Cauchy process). For most applications in option pricing we will be concerned with pure diffusion processes. However in some of our applications we will need to allow for jump terms.

Before proceeding to discuss the Fokker–Planck equation we should point out that conditions (i) and (ii) imply that all the higher order moments of the conditional probability density function of the stochastic process vanish, i.e.

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{|x-z|<\varepsilon}(x-z)^k p(x,t+\Delta t \mid z,t)dx = 0$$

*for all* $k \geq 3$. See Appendix 2.4 for an outline of how this result may be proved. *Thus a pure diffusion process is completely specified by the drift and diffusion terms* $A(z,t)$ *and* $B(z,t)$.

Let us assume for the moment that the jump term is zero (i.e. $J = 0$) so that asset prices are following a pure diffusion process. It can be shown that conditions (i) and (ii) in (2.14) and (2.15) reduce the Chapman–Kolmogorov equation for the evolution of conditional probabilities to the partial differential equation (see Appendix 2.5 for details)

$$\frac{\partial}{\partial t}p(z,t \mid y,\tau) = -\frac{\partial}{\partial z}[A(z,t)p(z,t \mid y,\tau)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[B(z,t)p(z,t \mid y,\tau)], \tag{2.17}$$

for the conditional probability $p(z,t \mid y,\tau)$ of observing asset price $z$ at time $t$ given that the current price is $y$ at time $\tau$. This equation is known as the *Fokker–Planck equation* and must be solved for $t \geq \tau$ subject to the initial time condition

$$p(z,\tau \mid y,\tau) = \delta(y-z). \tag{2.18}$$

It is also necessary to impose some *boundary* conditions (e.g. in the case of interest rates, the probability of these becoming negative must be zero) but we will discuss these as they arise in particular applications. It is important to emphasise that the viewpoint adopted with the Fokker–Planck equation is a *forward* one, i.e., we take the current price $y$ and time $\tau$ as fixed and consider the conditional probability as a forward evolving function of the price $z$ at the later time $t$.

However in order to price options we need to adopt the alternative perspective in which the final time is fixed (i.e. the maturity date of the option) and the initial time is varying. In other words we hold $z$ and $t$ fixed and allow $y$ and $\tau$ to vary. In this case it can be shown that under conditions (i) and (ii) the Chapman–Kolmogorov equation becomes (see Appendix 2.5 for details)

$$\frac{\partial}{\partial \tau}p(z,t \mid y,\tau) = -A(y,\tau)\frac{\partial}{\partial y}p(z,t \mid y,\tau) - \frac{1}{2}B(y,\tau)\frac{\partial^2}{\partial y^2}p(z,t \mid y,\tau), \tag{2.19}$$

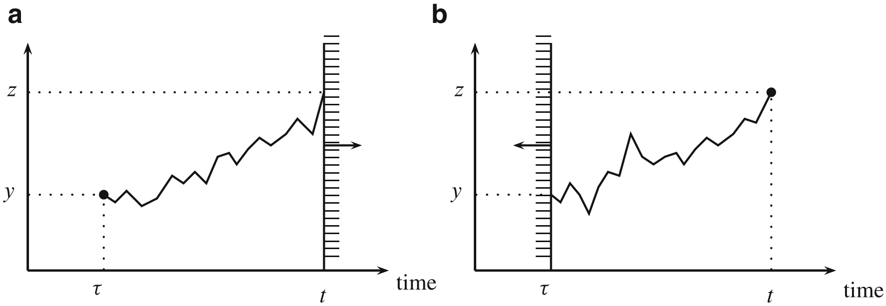


**Fig. 2.12** (**a**) Fokker–Planck equation for forward evolving probability; $(y, \tau)$ fixed. (**b**) Kolmogorov equation for backward evolving probability; $(z, t)$ fixed

which is known as the *Kolmogorov backward equation*, and must also be solved for $\tau \leq t$ subject to the final time condition

$$p(z,t \mid y,t) = \delta(y-z). \tag{2.20}$$

Now $p(z,t \mid y,\tau)$ is a backward evolving function of the price, $y$, at the earlier time $\tau$.

The different viewpoints of the partial differential equations (2.17) and (2.19), i.e. probability evolving forwards or backwards are illustrated in Fig. 2.12. In this figure, the reader should view the hashed wall as moving forward in Fig. 2.12a and backward in Fig. 2.12b.

Since we shall be referring frequently to both the Fokker–Planck equation and the Kolmogorov equation it is useful to introduce some short-hand notation for writing them down. Thus in relation to diffusion processes with drift function $A$ and diffusion function $B$ we introduce the partial differential operators $\mathscr{F}$ and $\mathscr{K}$, defined by

$$\mathscr{F} p = -\frac{\partial}{\partial z}[A(z,t)p(z,t \mid y,\tau)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[B(z,t)p(z,t \mid y,\tau)], \tag{2.21}$$

and

$$\mathscr{K} p = A(y,\tau)\frac{\partial}{\partial y}p(z,t \mid y,\tau) + \frac{1}{2}B(y,\tau)\frac{\partial^2}{\partial y^2}p(z,t \mid y,\tau). \tag{2.22}$$

The Fokker–Planck equation may then be written succinctly as

$$\frac{\partial p}{\partial t} - \mathscr{F}p = 0, \tag{2.23}$$

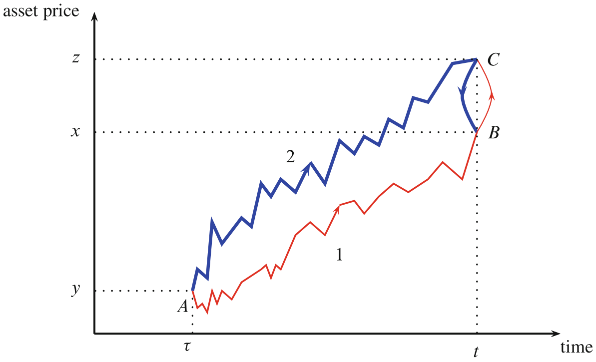


**Fig. 2.13** Explaining the jump term in the Fokker–Planck equation

whilst the Kolmogorov equation as

$$\frac{\partial p}{\partial \tau} + \mathscr{K} p = 0. \tag{2.24}$$

Of course when using this more succinct notation we must be a little cautious in keeping track of whether it is the final point (i.e. $z, t$ in the Fokker–Planck equation) or the initial point (i.e. $y, \tau$ in the Kolmogorov equation) which is varying.

The forward Fokker–Planck equation and the backward Kolmogorov equation are equivalent to each other in the sense that they yield the same conditional probability. They differ only as to whether the initial point or the final point is held fixed. Which form we use depends on the application at hand. In the technical language of the theory of partial differential equations the partial differential operators $\mathscr{F}$ and $\mathscr{K}$ are said to be adjoint operators.

To complete the discussion we show how the forward and backward equations need to be modified in order to allow for a jump component, i.e. when $J \neq 0$. The technical details are also included in Appendix 2.5 where it is shown that the Fokker–Planck equation becomes

$$\frac{\partial p}{\partial t} - \mathscr{F} p = \int_{-\infty}^{\infty} [J(z \mid x, t) p(x, t \mid y, \tau) - J(x \mid z, t) p(z, t \mid y, \tau)] dx. \tag{2.25}$$

Recalling that $J(z \mid x, t)$ essentially measures the probability that the price will jump from $x$ to $z$ at time $t$, the sense of the integral term on the right-hand side of (2.25) can be understood by referring to Fig. 2.13. The jump term, $J$, allows two additional types of events which are not possible under the pure diffusion process described by the operator $\mathscr{F}$. Firstly the price may follow a path to the value $x$ at $t$, where $x$ is not "close" to $z$, and then jump to the value $z$. The product $J(z \mid x, t) p(x, t \mid y, \tau)$ is the probability of going from $A$ to $B$, then $B$ to $C$ along such a path, and the integral of this probability with respect to $x$ measures the probability

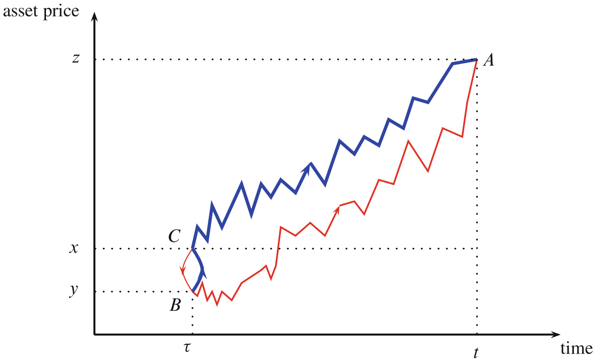


**Fig. 2.14** Explaining the jump term in the Kolmogorov equation

of going from $A$ to $C$ along all such paths. Secondly the price may follow a path to the value $z$ at $t$, and then jump to a value $x$ not "close" to $z$, this would be the path 2 in Fig. 2.13. The product $J(x \mid z,t)p(z,t \mid y,\tau)$ is the probability of going from $A$ to $C$, then $C$ to $B$ along such a path and the integral of this probability with respect to $x$ measures the probability over all such paths of the price reaching $z$ at time $t$ and then immediately jumping away to some other price. The overall term on the right-hand side of (2.25) is the probability of the price being at $z$ at time $t$ once the probabilities of price jumps at $t$ (both to and away from $z$) are fully accounted for.

Inclusion of the jump term in the Kolmogorov equation leads to

$$\frac{\partial p}{\partial \tau} + \mathscr{K}p = \int_{-\infty}^{\infty} J(x \mid y,\tau)[p(z,t \mid y,\tau) - p(z,t \mid x,\tau)]dx, \tag{2.26}$$

as illustrated in Fig. 2.14.

In Chap. 13 we shall consider option pricing when the underlying asset follows a jump-diffusion process. We shall see that the option price is determined by an integro-partial differential equation of the form (2.26). The techniques of the theory of stochastic differential equations that we develop in Chaps. 4, 6 and 8 allow us to conveniently arrive at a specification of the $J(z \mid x,t)$ term.

## 2.7 Appendix

### Appendix 2.1 Probability Density Functions

A function $P(x)$ is a cumulative distribution function if $P$ is a non-decreasing function and satisfies the properties

$$P(x_{\min}) = 0, \quad P(x_{\max}) = 1,$$

where $x_{\min}$ and $x_{\max}$ are respectively the minimum and maximum values attainable by $x$. If $X$ denotes the random variable of interest (e.g. a stock price) then

$$P(x) = \text{Prob.}\{X \leq x\}.$$

Typically we will deal with cumulative distribution functions which are differentiable and we write

$$P'(x) = p(x)$$

so that the cumulative distribution function can be written

$$P(x) = \int_{x_{\min}}^{x} p(\xi)d\xi.$$

With this notation we can readily see that

$$p(x)dx = dP(x) = \text{Prob.}\{x < X \leq x + dx\}.$$

The function $p(x)$ is known as the probability density function and since $P(x_{\max}) = 1$ it satisfies the property

$$\int_{x_{\min}}^{x_{\max}} p(\xi)d\xi = 1. \tag{2.27}$$

Sometimes it is convenient to transform to a new variable $y$ related to $x$ by

$$x = g(y),$$

where $g$ is increasing and differentiable. Making the change of variable $\xi = g(\zeta)$ Eq. (2.27) becomes

$$\int_{y_{\min}}^{y_{\max}} p(g(\zeta))g'(\zeta)d\zeta = 1,$$

where

$$y_{\max} = g^{-1}(x_{\max}), \quad y_{\min} = g^{-1}(x_{\min}).$$

If we set $\pi(y) = p(g(y))$ then the last equation becomes

$$\int_{y_{\min}}^{y_{\max}} \pi(\zeta))g'(\zeta)d\zeta = 1,$$

from which we see that in the new co-ordinates the density function is $\pi(y)g'(y)$. The corresponding c.d.f. is given by

$$\Pi(y) = \int_{y_{\min}}^{y} \pi(\zeta)g'(\zeta)d\zeta.$$

Thus the rule for transforming from $x$ to $y$ is

$$p(x) = \pi(y)g'(y)$$

which in terms of $x$ may be written

$$p(x) = \pi(g^{-1}(x)) \cdot g'(g^{-1}(x)). \tag{2.28}$$

## Appendix 2.2 Brownian Motion is a Continuous Process

We seek to verify that Brownian motion (2.10) satisfies (2.9), that is, Brownian motion is continuous. We need to use the following facts:

(i) In general, for $\alpha > 0$,

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}.$$

(ii) For $\delta$ large enough,

$$\int_{|x|<\delta} e^{-x^2} dx \approx \sqrt{\pi}\sqrt{1 - e^{-\delta^2}}.$$

For the conditional probability density function of the Brownian motion

$$p(x, t + \Delta t|z, t) = \frac{1}{\sqrt{2\pi\Delta t}\sigma} e^{-\frac{(x-z)^2}{2\sigma^2 \Delta t}},$$

we have

$$I := \frac{1}{\Delta t} \int_{|x-z|>\varepsilon} p(x, t + \Delta t|z, t)dx = \frac{1}{\Delta t} \frac{1}{\sqrt{2\pi\Delta t}\sigma} \int_{|x-z|>\varepsilon} e^{-\frac{(x-z)^2}{2\sigma^2 \Delta t}} dx.$$

Let

$$y = \frac{x - z}{\sqrt{2\Delta t}\sigma}$$

then

$$I = \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\int_{|y|>\frac{\varepsilon}{\sqrt{2\Delta t}\sigma}} e^{-y^2}dy = \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\left[\int_{|y|<\infty} e^{-y^2}dy - \int_{|y|<\frac{\varepsilon}{\sqrt{2\Delta t}\sigma}} e^{-y^2}dy\right]$$

$$\approx \frac{1}{\Delta t}\frac{1}{\sqrt{\pi}}\left[\sqrt{\pi} - \sqrt{\pi}\sqrt{1 - e^{-\frac{\varepsilon^2}{2\Delta t\sigma^2}}}\right] = \frac{1}{\Delta t}\left[1 - \sqrt{1 - e^{-\frac{\varepsilon^2}{2\Delta t\sigma^2}}}\right].$$

Let $\alpha = \frac{1}{\Delta t}$ and $A = \frac{\varepsilon^2}{2\sigma^2}$, then

$$\lim_{\Delta t \to 0} I = \lim_{\alpha\to\infty} \alpha\left[1 - \sqrt{1 - e^{-A\alpha}}\right] = \lim_{\alpha\to\infty}\frac{1 - \sqrt{1 - e^{-A\alpha}}}{1/\alpha}$$

$$= \lim_{\alpha\to\infty}\frac{\frac{-Ae^{-A\alpha}}{2\sqrt{1-e^{-A\alpha}}}}{-1/\alpha^2} \quad \text{(applying l'Hôpital's Rule)}$$

$$= \lim_{\alpha\to\infty}\frac{A}{2\sqrt{1 - e^{-A\alpha}}}\frac{\alpha^2}{e^{A\alpha}} = 0.$$

Hence (2.10) satisfies (2.9).

## Appendix 2.3 The Cauchy Process is Not Continuous

We seek to verify that (2.11) does not satisfy (2.9). For the Cauchy process,

$$p(x, t + \Delta t|z, t) = \frac{\Delta t}{\pi((x - z)^2 + (\Delta t)^2)},$$

consider

$$I = \frac{1}{\Delta t}\int_{|x-z|>\varepsilon} p(x, t + \Delta t|z, t)dx = \int_{|x-z|>\varepsilon}\frac{1}{\pi((x - z)^2 + (\Delta t)^2)}dx.$$

Let $y = x - z$, then $dy = dx$ and

$$I = \int_{|y|>\varepsilon}\frac{1}{\pi(y^2 + (\Delta t)^2)}dy.$$

Let $y = \Delta t \tan\theta$ so that $\theta = \tan^{-1}(\frac{y}{\Delta t})$ and $dy = \Delta t \sec^2\theta d\theta$. Hence

$$I = \int_{|y|>\varepsilon}\frac{\Delta t \sec^2\theta}{\pi(\Delta t)^2\sec^2\theta}d\theta$$

$$= \frac{1}{\pi \Delta t}\left[\tan^{-1}\left(\frac{y}{\Delta t}\right)\Big|_{y=-\infty}^{\infty} - \tan^{-1}\left(\frac{y}{\Delta t}\right)\Big|_{y=-\varepsilon}^{\varepsilon}\right]$$
$$= \frac{1}{\pi \Delta t}\left[\pi - 2\tan^{-1}\left(\frac{\varepsilon}{\Delta t}\right)\right].$$

Consider

$$\lim_{\Delta t \to 0} I = \lim_{\Delta t \to 0} \frac{1}{\pi} \frac{\pi - 2\tan^{-1}\left(\frac{\varepsilon}{\Delta t}\right)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{1}{\pi}\left(-2\frac{1}{\sec^2\left(\frac{\varepsilon}{\Delta t}\right)}\right)\left(-\frac{\varepsilon}{(\Delta t)^2}\right) = \infty.$$

Here we use

$$\sec^2\left(\frac{\varepsilon}{\Delta t}\right) = 1 + \tan^2\left(\frac{\varepsilon}{\Delta t}\right) \to 1 + \left(\frac{\pi}{2}\right)^2$$

as $\Delta t \to 0$. Hence (2.11) does not satisfy (2.9).

## Appendix 2.4 The Higher Order Moment Condition

Consider for example the third order moment (i.e. $k = 3$). Note that

$$\left|\int_{|x-z|<\varepsilon} (x-z)^3 p(x, t+\Delta t \mid z, t)\, dx\right| \le \int_{|x-z|<\varepsilon} |x-z|^3 \cdot p(x, t+\Delta t \mid z, t)\, dx$$
$$\le \int_{|x-z|<\varepsilon} |x-z| \cdot |x-z|^2 \cdot p(x, t+\Delta t \mid z, t)\, dx$$
$$\le \varepsilon \int_{|x-z|<\varepsilon} |x-z|^2 \cdot p(x, t+\Delta t \mid z, t)\, dx.$$

Hence

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\left|\int_{|x-z|<\varepsilon} (x-z)^3 p(x, t+\Delta t \mid z, t)\, dx\right|$$
$$\le \varepsilon \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|x-z|<\varepsilon} |x-z|^2 \cdot p(x, t+\Delta t \mid z, t) dx$$
$$\le \varepsilon B(z, t).$$

Since we may choose $\varepsilon$ as small as we like, this last term tends to zero. A similar argument applies for $k > 3$.

## Appendix 2.5 Derivation of Fokker–Plank and Kolmogorov Equations

The derivation is based on Gardiner (1985, Sect. 3.4). Consider the Chapman–Kolmogorov equation in the form (see Fig. 2.15)

$$p(z,t+\Delta t \mid y,\tau) = \int p(z,t+\Delta t \mid x,t)\,p(x,t \mid y,\tau)dx, \tag{2.29}$$

from which we subtract $p(z,t \mid y,\tau)$. Hence

$$\begin{aligned} &p(z,t+\Delta t \mid y,\tau) - p(z,t \mid y,\tau) \\ &\quad = \int p(z,t+\Delta t \mid x,t)\,p(x,t \mid y,\tau)dx - p(z,t \mid y,\tau). \end{aligned} \tag{2.30}$$

Now introduce an arbitrary function $R(x)$, which together with all of its derivatives, vanishes as $x$ approaches the extremities of the range of interest (e.g. 0 and $\infty$ for stock prices). Then multiplying both sides of (2.30) by $R(z)/\Delta t$ and integrating with respect to $z$,

$$\begin{aligned} &\int R(z)\frac{[p(z,t+\Delta t \mid y,\tau) - p(z,t \mid y,\tau)]}{\Delta t}\,dz \\ &\quad = \frac{1}{\Delta t}\int R(z)\left(\int p(z,t+\Delta t \mid x,t)\,p(x,t \mid y,\tau)\,dx\right)dz \\ &\qquad - \frac{1}{\Delta t}\int R(z)\,p(z,t \mid y,\tau)\,dz. \end{aligned} \tag{2.31}$$

The function $R$ is next expanded in a Taylors series about the point $x$, i.e.

$$R(z) = R(x) + \frac{dR(x)}{dx}(z-x) + \frac{1}{2}\frac{d^2R(x)}{dx^2}(z-x)^2 + o(z-x)^3, \tag{2.32}$$

and the integrals are considered over regions $\mid z - x \mid < \varepsilon$ and $\mid z - x \mid \geq \varepsilon$.

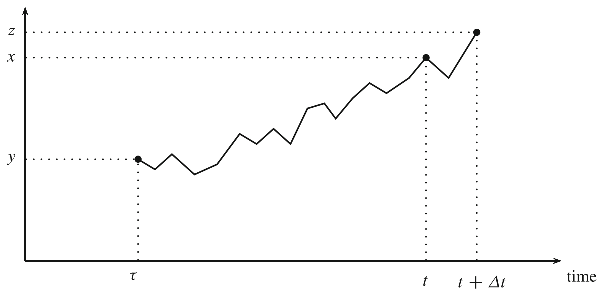


**Fig. 2.15** Evolution from $\tau$ to $t + \Delta t$ via $t$

If we assume a pure diffusion process, i.e. $J = 0$, then the Lindeberg condition allows us to ignore integrals in the region $| z - x | \geq \varepsilon$. The result in Appendix 2.4 allows us to ignore the $o(z-x)^3$ term in integrals over the region $| z - x | < \varepsilon$. We are left with integrals involving terms up to $(z-x)^2$ over the region $| z - x | < \varepsilon$. In the limit $\Delta t \to 0$ the term on the left hand side of (2.31) becomes

$$\int R(z) \frac{\partial p}{\partial t} dz.$$

On the right hand side, after substituting (2.32) and interchanging the order of the integrations, we are left with

$$\begin{aligned} &\frac{1}{\Delta t} \int R(x) p(x,t \mid y,\tau) \left( \int p(z, t+\Delta t \mid x,t) dz \right) dx \\ &+ \sum_{n=1}^{2} \frac{1}{n!} \int \frac{d^n R(x)}{dx^n} p(x,t \mid y,\tau) \left( \frac{1}{\Delta t} \int (z-x)^n p(z, t+\Delta t \mid x,t) dz \right) dx \\ &- \frac{1}{\Delta t} \int R(z) p(z,t \mid y,\tau) dz, \end{aligned} \tag{2.33}$$

where all the integrals are taken over the region $| z - x | < \varepsilon$. Using conditions (i) and (ii) in Eqs. (2.14) and (2.15) and the result in Appendix 2.4 we see that in the limit $\Delta t \to 0$ the terms in the big bracket on the middle line tend to $A(x,t)$ (when $n = 1$), $B(x,t)$ (when $n = 2$) and 0 (when $n \geq 3$). Using the fact that

$$\int p(z, t+\Delta t \mid x,t) dz = 1,$$

Eq. (2.31) now reduces to

$$\int \left[ R(x) \frac{\partial p}{\partial t} - \frac{dR}{dx} A(x,t) p(x,t \mid y,\tau) - \frac{1}{2} \frac{d^2 R}{dx^2} B(x,t) p(x,t \mid y,\tau) \right] dx = 0. \tag{2.34}$$

The final step is to perform an integration by parts on the last two terms so that

$$\int \frac{dR}{dx} A(x,t) p(x,t \mid y,\tau)\, dx = -\int \frac{\partial}{\partial x} (A(x,t) p(x,t \mid y,\tau)) R(x)\, dx$$

and

$$\int \frac{d^2 R}{dx^2} B(x,t) p(x,t \mid y,\tau)\, dx = \int \frac{\partial^2}{\partial x^2} (B(x,t) p(x,t \mid y,\tau)) R(x)\, dx.$$

Note that in performing these integrations by parts we have invoked the properties that R and all of its derivatives vanish at the extremities of the range of interest. So Eq. (2.34) finally becomes

$$\int \left[ \frac{\partial p}{\partial t} + \frac{\partial}{\partial x}(A(x,t)p(x,t \mid y,\tau)) - \frac{1}{2}\frac{\partial^2}{\partial x^2}(B(x,t)p(x,t \mid y,\tau)) \right] R(x)dx = 0.$$

Since this last expression holds for an arbitrary function $R(x)$, the term in the squared bracket must be zero, i.e.

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(A(x,t)p(x,t \mid y,\tau)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(B(x,t)p(x,t \mid y,\tau)),$$

which is the Fokker–Plank forward equation.

If we are dealing with a jump-diffusion process then $J \neq 0$ and we cannot invoke the Lindeberg condition to ignore integrals in the region $\mid x - z \mid \geq \varepsilon$. In this case the expansion at Eq. (2.31) will involve the extra terms

$$\lim_{\varepsilon \to 0} \frac{1}{\Delta t} \left( \int \int_{|x-z| \geq \varepsilon} R(z)p(z,t+\Delta t \mid x,t)p(x,t \mid y,\tau)dzdx \right.$$
$$\left. - \int \int_{|x-z| \geq \varepsilon} R(x)p(z,t+\Delta t \mid x,t)p(x,t \mid y,\tau)dxdz \right).$$

Recalling the definition of $J(x \mid z,t)$ from Eq. (2.16) we can write these terms when $\Delta t \to 0$ as

$$\int \int_{|x-z| \geq \varepsilon} R(z)J(z \mid x,t)p(x,t \mid y,\tau)dzdx$$
$$- \int \int_{|x-z| \geq \varepsilon} R(x)J(z \mid x,t)p(x,t \mid y,\tau)dxdz.$$

Interchanging the roles of $x$ and $z$ in the first integral the two terms can be combined to yield

$$\int \int_{|x-z| \geq \varepsilon} R(x) \left[ J(x \mid z,t)p(z,t \mid y,\tau) - J(z \mid x,t)p(x,t \mid y,\tau) \right] dzdx.$$

Incorporating this term into Eq. (2.34) and proceeding to the limit $\varepsilon \to 0$ we obtain Eq. (2.25) of the main text.

To obtain the Kolmogorov backward equation we consider

$$p(z,t|y,\tau) = \int p(z,t|x,\tau+\Delta t)p(x,\tau+\Delta t|y,\tau)dx, \tag{2.35}$$

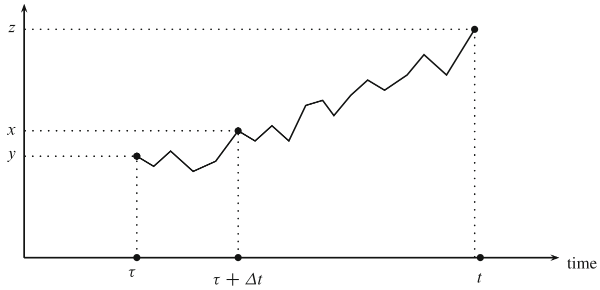


**Fig. 2.16** Evolution from $\tau$ to $t$ via $\tau + \Delta t$

the sense of which is illustrated in Fig. 2.16.

Note that

$$\frac{\partial}{\partial \tau} p(z,t|y,\tau) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}[p(z,t|y,\tau+\Delta t) - p(z,t|y,\tau)]$$

Consider

$$\begin{aligned} I &= \frac{1}{\Delta t}[p(z,t|y,\tau+\Delta t) - p(z,t|y,\tau)] \\ &= \frac{1}{\Delta t}\left[p(z,t|y,\tau+\Delta t) - \int p(z,t|x,\tau+\Delta t)p(x,\tau+\Delta t|y,\tau)dx\right] \\ &= \frac{1}{\Delta t}\int \left[p(z,t|y,\tau+\Delta t) - p(z,t|x,\tau+\Delta t)\right]p(x,\tau+\Delta t|y,\tau)dx \\ &= I_1 + I_2, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[p(z,t|y,\tau+\Delta t) - p(z,t|x,\tau+\Delta t)\right]p(x,\tau+\Delta t|y,\tau)dx, \\ I_2 &= \frac{1}{\Delta t}\int_{|x-y|>\varepsilon} \left[p(z,t|y,\tau+\Delta t) - p(z,t|x,\tau+\Delta t)\right]p(x,\tau+\Delta t|y,\tau)dx. \end{aligned}$$

Assume first that there is no jump, then $I_2 = 0$ as $\Delta t \to 0$ and

$$\begin{aligned} I_1 &= \frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[p(z,t|y,\tau+\Delta t) - p(z,t|x,\tau+\Delta t)\right]p(x,\tau+\Delta t|y,\tau)dx \\ &\approx -\frac{1}{\Delta t}\int_{|x-y|<\varepsilon} \left[\frac{\partial p(z,t|y,\tau+\Delta t)}{\partial y}(x-y) + \frac{1}{2}\frac{\partial^2 p}{\partial y^2}(x-y)^2\right] p(x,\tau+\Delta t|y,\tau)dx \end{aligned}$$

$$= -\frac{\partial p(z,t|y,\tau+\Delta t)}{\partial y}\frac{1}{\Delta t}\int_{|x-y|<\varepsilon}(x-y)p(x,\tau+\Delta t|y,\tau)dx$$
$$-\frac{1}{2}\frac{\partial^2 p(z,t|y,\tau+\Delta t)}{\partial y^2}\frac{1}{\Delta t}\int_{|x-y|<\varepsilon}(x-y)^2 p(x,\tau+\Delta t|y,\tau)dx.$$

Then

$$\lim_{\Delta t\to 0} I_1 = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p}{\partial y^2}.$$

Hence we have established that

$$\frac{\partial p}{\partial \tau}(z,t|y,\tau) = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p(z,t|y,\tau)}{\partial y^2}.$$

In the presence of jump, the only difference is that

$$\lim_{\Delta t\to 0} I_2 = \int_{-\infty}^{\infty} J(x|y,\tau)[p(z,t|y,\tau) - p(z,t|x,\tau)]dx.$$

Hence $p(z,t|y,\tau)$ satisfies

$$\frac{\partial}{\partial \tau}p(z,t|y,\tau) = -A(y,\tau)\frac{\partial p(z,t|y,\tau)}{\partial y} - \frac{1}{2}B(y,\tau)\frac{\partial^2 p(z,t|y,\tau)}{\partial y^2}$$
$$+\int_{-\infty}^{\infty} J(x|y,\tau)[p(z,t|y,\tau) - p(z,t|x,\tau)]dx,$$

which is the result given in (2.26).

## Appendix 2.6 The Mean Value Theorem

Suppose that $f(x)$ is a continuous and non-negative function on the interval $[a,b]$. There exists a value $\xi$ satisfying $a \le \xi \le b$ such that

$$\int_a^b f(x)dx = (b-a)f(\xi).$$

Basically this result says that we can always find $\xi$ such that the area under the rectangle shown in the figure equals the area under the curve (Fig. 2.17).

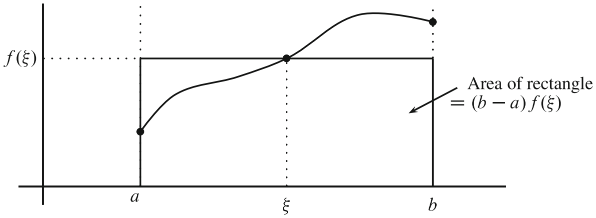


**Fig. 2.17** Illustrating the mean value theorem for integrals

## 2.8 Problems

### Problem 2.1

(a) Consider the Brownian notion whose transition probability density is given by

$$p(x,t|y,\tau) = \frac{1}{\sqrt{2\pi(t-\tau)}\sigma} \exp\left[\frac{-(x-y)^2}{2\sigma^2(t-\tau)}\right].$$

By direct differentiation show that $p$ satisfies the partial differential equations

$$\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial y^2} + \frac{\partial p}{\partial \tau} = 0,$$

and

$$-\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial t} = 0.$$

What is the boundary condition for both of these partial differential equations?

(b) Consider the Cauchy distribution whose transitional partial differential function is given by

$$p(x,t|y,\tau) = \frac{(t-\tau)}{\pi\left[(x-y)^2 + (t-\tau)^2\right]}.$$

Show that the first and second moments of this distribution are given by

$$\int_{-\infty}^{\infty} (x-y)p(x,t|y,\tau)dx = 0,$$

$$\int_{-\infty}^{\infty} (x-y)^2 p(x,t|y,\tau)dx = \infty.$$

Explain why these results indicate that it would not be possible to obtain Kolmogorov or Fokker–Planck equations as in (a).

**Problem 2.2** Consider the function $\delta_\varepsilon(x)$ defined by

$$\delta_\varepsilon(x) = \begin{cases} 0, |x| > \varepsilon/2, \\ 1/\varepsilon, |x| \le \varepsilon/2. \end{cases}$$

(a) Sketch this function.
(b) Show that

$$\int_{-\infty}^{\infty} \delta_\varepsilon(x)dx = 1$$

for all $\varepsilon$.

Consider a function $f$, defined on $\mathbb{R}$, which is continuous.

(c) Sketch the function $f(x)\delta_\varepsilon(x)$.
(d) Use the mean value theorem of integral calculus to show that

$$\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(\theta),$$

where $-\varepsilon/2 \le \theta \le \varepsilon/2$.
(e) Explain the intuition of this result in the sketch you have just drawn. Hence show that

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(0).$$

(f) Explain how to use the foregoing arguments to establish the result

$$\lim_{\varepsilon \to 0} \delta_\varepsilon(x) = \delta(x),$$

where $\delta(x)$ is the Dirac delta function.

**Problem 2.3** Consider the transitional probability density function

$$p(x, t|y, \tau) = \frac{1}{\sqrt{2\pi(t-\tau)}\sigma} \exp\left[-\frac{(x - y - a(t-\tau))^2}{2\sigma^2(t-\tau)}\right].$$

Show by direct differentiation that $p$ satisfies the partial differential equations

$$\frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial y^2} + a\frac{\partial p}{\partial y} + \frac{\partial p}{\partial \tau} = 0$$

and

$$-\frac{1}{2}\sigma^2\frac{\partial^2 p}{\partial x^2} + a\frac{\partial p}{\partial x} + \frac{\partial p}{\partial t} = 0.$$

**Problem 2.4** Consider the function $\delta_\varepsilon(x)$ defined as follows:

$$\delta_\varepsilon(x) = \begin{cases} 0, & x < -\varepsilon, \\ \frac{3}{4\varepsilon^3}(\varepsilon^2 - x^2), & -\varepsilon \le x \le \varepsilon, \\ 0, & x > \varepsilon. \end{cases}$$

(a) Sketch this function and show that

$$\int_{-\infty}^{\infty} \delta_\varepsilon(x)dx = 1$$

for all $\varepsilon$.

(b) Consider a function $f$, defined on $\mathbb{R}$, which is continuous. Use the Mean Value Theorem of integral calculus to show that

$$\int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(\theta),$$

where $-\varepsilon \le \theta \le \varepsilon$. Hence show that

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = f(0).$$

(c) Explain how to use the foregoing result to establish the result

$$\lim_{\varepsilon \to 0} \delta_\varepsilon(x) = \delta(x),$$

where $\delta(x)$ is the Dirac delta function.

# Chapter 3
# An Initial Attempt at Pricing an Option

**Abstract** This chapter uses the concepts developed in Chap. 2 to illustrate the problem of option pricing as a discounted expected option payoff. By assuming that investors are risk neutral and using the Kolmogorov equation for the conditional probability, we demonstrate how the Black–Scholes option formula can be arrived. We also illustrate how the option price can be viewed in a quite natural way as a martingale and the Feynman–Kac formula, two very important concepts of continuous time finance.

## 3.1 Option Pricing as a Discounted Cash Flow Calculation

To someone trained in basic finance, in particular discounted cash flow concept, the problem of pricing an option would seem (at least conceptually) to be a fairly straight forward one. After all it should just be the expected future payoff of the option, discounted back to current time with an appropriate discount rate.

We shall now use the concepts developed in Chap. 2 to attempt to apply this viewpoint to the problem of pricing a European option. Consider the simple example of a European call option on a stock. Suppose maturity is at time $T$ and exercise price is $E$. If the stock price at maturity is $S_T$, then as we have seen in Chap. 2, the payoff on the option is $h(S_T) = \max(S_T - E, 0) \equiv (S_T - E)^+$.

In order to calculate the expected payoff at $T$, the investor requires the conditional probability density function, $p(S_T, T|S, t)$, for the stock price $S_T$ at $T$ given the price $S$ at current time $t$. Assuming processes with continuous sample paths this density function is obtained from the solution of the Kolmogorov backward equation (2.19). We have already pointed out in Chap. 2 that for common stock a great deal of empirical evidence suggests that the $A$ and $B$ functions appearing in the Kolmogorov equation are given by

$$A(z,t) = \mu z, \qquad B(z,t) = \sigma^2 z^2, \tag{3.1}$$

where $\mu$ is the expected stock return per unit time and $\sigma^2$ is the instantaneous variance of stock returns per unit time. The Kolmogorov backward equation for the

 
C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_3

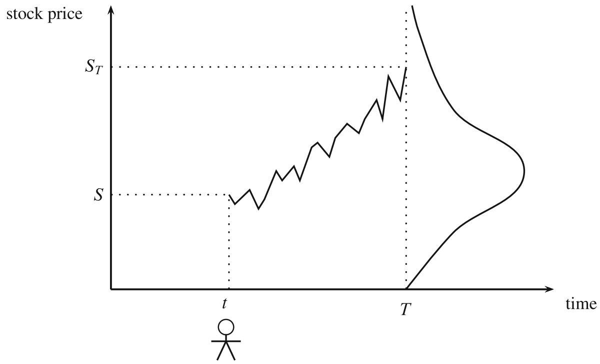


**Fig. 3.1** The investor obtains $p(S_T, T|S, t)$ from the Kolmogorov equation

conditional probability distribution $p(S_T, T \mid S, t)$ in our current notation assumes the form (put $z = S_T, t = T$ and $y = S, t' = t$ in Eq. (2.19)),

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 p}{\partial S^2} + \mu S \frac{\partial p}{\partial S} + \frac{\partial p}{\partial t} = 0. \tag{3.2}$$

Consider an investor who at time $t$ wishes to value the option which matures at $T$ (see Fig. 3.1). If the stock price at maturity is $S_T$ then the expected payoff in the stock price interval $(S_T, S_T + dS_T)$ is $h(S_T)p(S_T, T \mid S, t)dS_T$. Integrating over all possible values of $S_T$ then

$$\left.\begin{array}{r}\textit{expected payoff from}\\ \textit{the option at maturity}\end{array}\right\} = \int_0^\infty h(S_T)p(S_T, T \mid S, t)dS_T. \tag{3.3}$$

This expected payoff is measured in dollars at time $T$. In order to convert it into dollars at time $t$ the investor must discount this amount at an appropriate rate.

The choice of this discount rate is a crucial one in the approach to option pricing that we are adopting in this section. The theory of finance tells us that utility maximising investors discount risky payoffs at a rate $\rho$, which is the sum of the prevailing risk free rate, $r$, and a risk premium term depending on the investor's attitude to risk, i.e.

$$\rho = r + \pi, \tag{3.4}$$

where $\pi$ represents the risk premium term. If investors are risk averse $\pi > 0$, if they are risk lovers $\pi < 0$, and if they are risk neutral $\pi = 0$. Let us assume (without any reason being offered at this stage)[1] investors value options as if they

[1]Actually one could give the justification for the investor doing the option valuation problem this way as follows. The investor realises that the option valuation problem is a *relative* pricing problem

were risk neutral then $\pi = 0$ and $\rho = r$. The discount term applied to the options expected payoff would then be $e^{-r(T-t)}$. Hence the investor would value the option as

$$\left.\begin{array}{c}\textit{value of option}\\ \textit{at time } t \textit{ when}\\ \textit{stock price is } S\end{array}\right\} = C(S,t) = e^{-r(T-t)}\int_0^\infty h(S_T)p(S_T,T \mid S,t)dS_T. \tag{3.5}$$

If we solve the Kolmogorov equation (3.2) for the conditional probability $p(S_T,T \mid S,t)$ then the problem of valuing the option is reduced to a problem in integration. We will demonstrate later how the Black–Scholes European stock call option formula can be arrived at via this route. However for the moment we will pursue another approach which allows us to see the connection with the continuous hedging argument approach to be developed in Chap. 7.

If we differentiate $C$ as defined in (3.5) with respect to $S$ and $t$ we will see how the option price changes as the current stock price changes and as we move closer to maturity. In particular

$$\begin{aligned}\frac{\partial C}{\partial S} &= e^{-r(T-t)}\int_0^\infty h(S_T)\frac{\partial p}{\partial S}(S_T,T \mid S,t)dS_T,\\ \frac{\partial^2 C}{\partial S^2} &= e^{-r(T-t)}\int_0^\infty h(S_T)\frac{\partial^2 p}{\partial S^2}(S_T,T \mid S,t)dS_T,\\ \frac{\partial C}{\partial t} &= e^{-r(T-t)}\int_0^\infty h(S_T)\frac{\partial p}{\partial t}(S_T,T \mid S,t)dS_T + rC(S,t).\end{aligned}$$

With an eye to the structure of the Kolmogorov equation (3.2) for $p(S_T,T \mid S,t)$ note that

$$\begin{aligned}&\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} + \mu S\frac{\partial C}{\partial S} + \frac{\partial C}{\partial t}\\ &\quad = e^{-r(T-t)}\int_0^\infty h(S_T)\left(\frac{1}{2}\sigma^2 S^2\frac{\partial^2 p}{\partial S^2} + \mu S\frac{\partial p}{\partial S} + \frac{\partial p}{\partial t}\right)dS_T + rC.\end{aligned}$$

By virtue of Eq. (3.2) the square bracket under the integral sign is zero, and hence the option price follows the partial differential equation

$$\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} + \mu S\frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} = rC. \tag{3.6}$$

---

i.e. he/she seeks to value the option given that the market's current assessment of the value of the underlying stock is $S$ (which will impound in it the market's assessment of the risk-premium). The relation between the stock and the option brings no additional source of risk. So the pricing relation between them can be viewed as a risk neutral one.

Finally we note that a risk neutral investor would expect the stock price to grow at the risk free rate $r$, i.e. $\mu = r$. Hence a risk neutral investor would expect that as the stock price $S$ and current time $t$ change, the option price would change according to

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} = rC, \tag{3.7}$$

which must be solved subject to the final time condition

$$C(S_T, T) = \max(S_T - E, 0). \tag{3.8}$$

The partial differential equation (3.7) is the celebrated Black–Scholes equation whose solution (solution techniques will be discussed in Chap. 9) is

$$C(S,t) = S\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2), \tag{3.9}$$

where $\mathcal{N}(d)$ is the cumulative normal function, defined by

$$\mathcal{N}(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{x^2}{2}} dx,$$

and

$$d_1 = \frac{\ln(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \tag{3.10}$$

and

$$d_2 = d_1 - \sigma\sqrt{T-t}. \tag{3.11}$$

The quantity $\mathcal{N}(d)$ is the probability that the standard normal variate lies in the interval $(-\infty, d)$. As is clear from its definition it is in fact the area under the normal density function curve from $-\infty$ to $d$ and is illustrated in Fig. 3.2. This figure displays $\mathcal{N}(-d)$ on the left and on the right makes clear the relation

$$\mathcal{N}(-d) = 1 - \mathcal{N}(d), \tag{3.12}$$

which is often used in later manipulations. The result (3.12) is easily proved by formal mathematical manipulations.

The approach to option pricing we have just discussed, which is the so-called risk-neutral valuation approach, seems to have been first proposed by Cox and Ross (1976b), who extended it to value options under a variety of assumptions about the stochastic process followed by the underlying asset price.

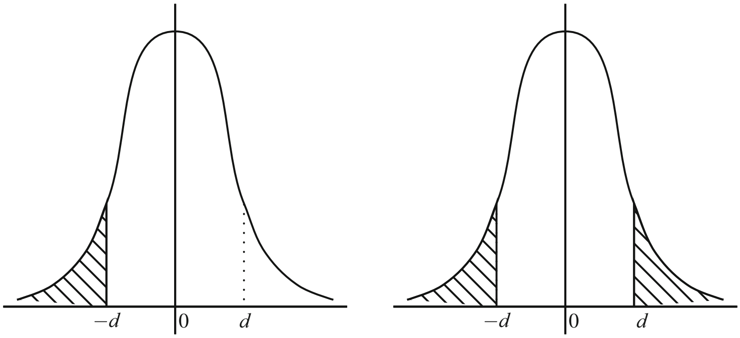


**Fig. 3.2** Symmetry of $\mathcal{N}(d)$

We could also have proceeded to obtain the option value directly from Eq. (3.5). If we solve the Kolmogorov backward equation (3.2) for the conditional probability function we would obtain the solution[2]

$$p(S_T, T \mid S, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma S_T} \exp\left[-\frac{\{\ln(S_T/S) - (\mu - \frac{1}{2}\sigma^2)(T-t)\}^2}{2\sigma^2(T-t)}\right]. \tag{3.13}$$

The problem of pricing the option then reduces to performing the integration (3.5) using Eq. (3.13), with $\mu$ replaced by $r$ and $p$ by $\tilde{p}$, by using

$$\tilde{p}(S_T, T|S, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma S_T} \exp\left[-\frac{\{\ln(S_T/S) - (r - \frac{1}{2}\sigma^2)(T-t)\}^2}{2\sigma^2(T-t)}\right], \tag{3.14}$$

where we use $\tilde{p}$ to denote the transition density function obtained when $\mu$ is replaced by $r$. This density shall be referred to as the risk-neutral density. The option price is thus given by

$$C(S,t) = e^{-r(T-t)} \int_0^\infty h(S_T)\tilde{p}(S_T, T|S, t)dS_T. \tag{3.15}$$

[2] Whilst we leave a formal discussion of the solution techniques for the partial differential equations of financial economics to Chap. 9, the reader can easily verify that (3.13) is indeed the solution to (3.2) as follows. First transform the partial differential equation (3.2) by the transformation $y_T = \ln S_T$, $y = \ln S$. The resulting partial differential equation is then of the form encountered in Problem 2.3 whose solution was verified there. See Problem 3.3.

Once Eq. (3.14) is substituted into (3.15) we are left to perform an exercise in integration and the details are given in Appendix 3.1. The Black–Scholes formula (3.9) is again the result. For more complicated stochastic processes for the asset price, if we can find the conditional distribution $p$ we can value the option by integration and by making the assumption that investors behave as if they are risk neutral. This is the so-called principle of risk-neutral valuation.

A loose end with this approach to option pricing is that it doesn't seem an obvious step to treat each investor as if he or she were risk neutral. However when considering how the investor might react to small changes in the stock and option prices over a small interval of time, the risk neutral argument occurs quite naturally. In order to carry out such an approach we need to concentrate our attention on the sample path rather than the conditional probability density function. We are then led to a study of the stochastic differential equation, to which we turn in the next chapter.

We shall devote the remainder of this chapter to demonstrating how the expressions for the option price obtained above can be viewed in various alternative ways that allow us to encounter in a quite natural way two very important concepts of continuous time finance, namely the martingale concept and the Feynman–Kac formula.

## 3.2 Our First Glimpse of a Martingale

We have argued in the previous subsection that if we assume investors are risk neutral then the option is priced according to

$$C(S,t) = e^{-r(T-t)} \int_0^\infty h(S_T)\tilde{p}(S_T, T|S,t)dS_T, \tag{3.16}$$

where the risk-neutral density $\tilde{p}$ satisfies

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \tilde{p}}{\partial S^2} + rS\frac{\partial \tilde{p}}{\partial S} + \frac{\partial \tilde{p}}{\partial t} = 0,$$

subject to

$$\tilde{p}(S_T, T \mid S, T) = \delta(S_T - S).$$

We could also use the notation of expectation operators and write

$$C(S,t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[h(S_T)], \tag{3.17}$$

where the expectation operator $\tilde{\mathbb{E}}_t$ is defined by

$$\tilde{\mathbb{E}}_t[h(S_T)] = \begin{array}{l}\text{expectation of } h(S_T) \text{ calculated at time } t\\ \text{using the probability density function } \tilde{p}(S_T, T \mid S, t).\end{array}$$

Recall that

$$h(S_T) \equiv C(S_T, T),$$

so we can also write

$$C(S,t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[C(S_T, T)],$$

or alternatively

$$e^{-rt}C(S,t) = \tilde{\mathbb{E}}_t[e^{-rT}C(S_T, T)]. \tag{3.18}$$

Note that

$$e^{rt} = \begin{cases}\text{Value at } t \text{ of \$1 invested at 0 and}\\ \text{continuously compounded at } r.\end{cases}$$

This quantity is known as the money market account. Thus

$$V(S,t) = \frac{C(S,t)}{e^{rt}} = \begin{cases}\text{Value of the option at time t measured in}\\ \text{units of the money market account.}\end{cases} \tag{3.19}$$

Then (3.18) (after interchanging LHS and RHS) can be re-expressed as

$$\tilde{\mathbb{E}}_t[V(S_T, T)] = V(S,t) \tag{3.20}$$

i.e.

$$\begin{cases}\text{the expected value of } V \text{ for time } T,\\ \text{calculated at current time } t\end{cases} = \begin{cases}\text{the value of } V\\ \text{at current time } t.\end{cases}$$

A stochastic process satisfying such a property (i.e. that its expected value at a future time is just its current value) is known as a *martingale*.

Essentially such processes have the property that at each point in time, the conditional distribution for future points in time is "centred at" (i.e. has as mean value) the current point at which the process has arrived. Figure 3.3 illustrates this concept. In Fig. 3.3a the conditional distribution of $V(S_T, T)$ at time $t$ is "centred" at the current value $V(S_t, t)$. On the following trading day, $t + \Delta t$, the conditional distribution for $T$ is now "centred at" $V(S_{t+\Delta t}, t + \Delta t)$, the value that $V$ has attained at this point in time, as illustrated in Fig. 3.3b. In a sense that could be

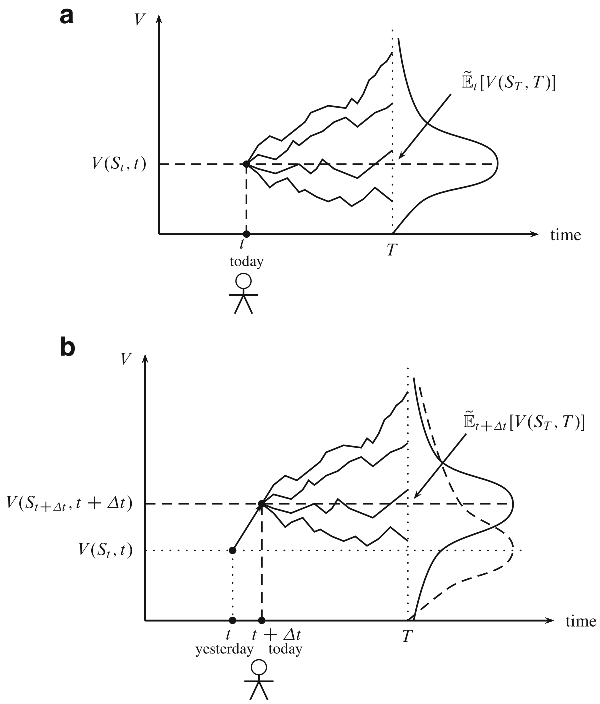


**Fig. 3.3** Illustrating the martingale concept over two successive "days". (**a**) The martingale relation at day $t$. (**b**) The martingale relation at day $(t + \Delta t)$

made mathematically more precise, this restraint on the stochastic process stops its sample paths from becoming "too wild".[3]

Martingales will be formally introduced in Chap. 8 where we will also demonstrate how to arrive at Eq. (3.20) via the continuous hedging argument.

## 3.3 Our First Glimpse of the Feynman–Kac Formula

We have argued in Sect. 3.1 that $C(S, t)$ satisfies the partial differential equation (3.7),

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} = rC$$

[3]This is the sense of the term martingale. One of the meanings of this word in French is the harness that one places around the head of a horse. This allows the jockey to control the horse and prevent its movements from becoming too wild.

with

$$C(S_T, T) = \max(S_T - E, 0) \equiv h(S_T).$$

It is of interest to enquire as to what partial differential equation the adjusted option price $V(S,t)$, defined in (3.19), satisfies. To this end we note that

$$\frac{\partial}{\partial S}V(S,t) = \frac{\partial}{\partial S}(e^{-rt}C(S,t)) = e^{-rt}\frac{\partial C}{\partial S},$$
$$\frac{\partial^2}{\partial S^2}V(S,t) = \frac{\partial^2}{\partial S^2}(e^{-rt}C(S,t)) = e^{-rt}\frac{\partial^2 C}{\partial S^2},$$
$$\frac{\partial}{\partial t}V(S,t) = \frac{\partial}{\partial t}(e^{-rt}C(S,t)) = -re^{-rt}C(S,t) + e^{-rt}\frac{\partial C}{\partial t}.$$

Hence after some algebra and making use of (3.7), we find that

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} = 0, \tag{3.21}$$

or, using the notation for the Kolmogorov partial differential operator $\mathscr{K}$,

$$\mathscr{K}V + \frac{\partial V}{\partial t} = 0.$$

Using the terminal conditions for $C(S_T, T)$ we find that

$$V(S_T, T) = e^{-rT}\max(S_T - E, 0) \equiv g(S_T, T). \tag{3.22}$$

On the other hand we have shown in Sect. 3.2 that $V$ can be expressed as a conditional expectation,

$$V(S,t) = \tilde{\mathbb{E}}_t[V(S_T, T)] = \tilde{\mathbb{E}}_t[g(S_T, T)], \tag{3.23}$$

where $\tilde{\mathbb{E}}_t$ is calculated using the risk-neutral probability density function $\tilde{p}(S_T, T \mid S, t)$ which is the solution of (see Eqs. (3.5), (3.13) and (3.14))

$$\mathscr{K}\tilde{p} + \frac{\partial \tilde{p}}{\partial t} = 0, \tag{3.24}$$

with terminal condition

$$\tilde{p}(S_T, T \mid S, T) = \delta(S_T - S). \tag{3.25}$$

What we have effectively shown with these various manipulations is that the solution to the partial differential equation (3.21) generated by the operator $\mathscr{K}$ subject to the terminal (payoff) condition (3.22) may be expressed as an expectation of the payoff as in Eq. (3.23). The Kolmogorov partial differential equation yielding the probability density function under which the expectation is calculated is also generated by the same operator $\mathscr{K}$ under the conditional expectation defined by Eq. (3.20).

We will show in Chap. 8 that this is in fact a specific instance of a quite general result known as the Feynman–Kac formula. This will allow us to have two alternative representations of the option price, either via the solution of a partial differential equation subject to a terminal (payoff) condition or as an expectation of the payoff, with the calculation of the expectation requiring a density function which is the solution of the Kolmogorov partial differential equation. Depending on the application under consideration one or other of these representations of the option price may be more convenient. The equivalence between these two viewpoints is shown by two of the boxes in Fig. 3.4. This figure shows a link to a third box, this last box involves the viewpoint from using stochastic differential equations that basically gives us another way to calculate the expectation $\tilde{\mathbb{E}}_t$ in Eq. (3.23) (namely simulation) and also will allow us to develop the continuous hedging argument. This viewpoint and the necessary technical apparatus shall be developed over the next three chapters.

In this chapter we have already obtained what we will appreciate later the two of the main representations of the option price, the partial differential equation representation and the martingale representation. We have also seen how going back and forth between these two representations throws up the results that we shall formalise in Chap. 8 as the Feynman–Kac formula.

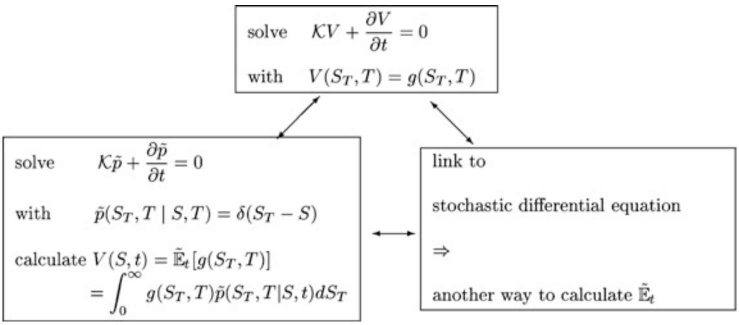


**Fig. 3.4** The holy trinity of option pricing: different viewpoints on the option price representation

We have been able to arrive at these representations, by combining the asset price modelling concepts we developed in Chap. 2 with the very basic notion of discounted expected cash flow. The only assumption we have had to make is that investors act as if they are risk-neutral when pricing options. This assumption is of course completely ad hoc (though to some extent justified by the comment in footnote 1 that option pricing is a relative pricing problem) at this stage and must be regarded as a loose end. However the rationale for the risk-neutral pricing principle will become clear by the time we reach Chap. 8 .

## 3.4 Appendix

### Appendix 3.1 Derivation of the Black–Scholes Formula

Letting $\tau = T - t$, the risk-neutral density (3.14) becomes

$$\tilde{p}(S_T, T \mid S, t) = \frac{1}{\sqrt{2\pi\tau}\sigma S_T} \exp\left[-\frac{\{\ln(S_T/S) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right]. \tag{3.26}$$

The option value (3.15) is then given by[4]

$$C(S,t) = e^{-r\tau} \int_E^\infty (S_T - E) \frac{1}{\sqrt{2\pi\tau}\sigma S_T} \exp\left[-\frac{\{\ln(S_T/S) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right] dS_T. \tag{3.27}$$

Equation (3.27) is the difference of two integrals, which we evaluate separately. Thus we define

$$A_1 = \frac{1}{\sqrt{2\pi\tau}\sigma} \int_E^\infty \exp\left[-\frac{\{\ln(S_T/S) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right] dS_T, \tag{3.28}$$

and

$$A_2 = \frac{E}{\sqrt{2\pi\tau}\sigma} \int_E^\infty \exp\left[-\frac{\{\ln(S_T/S) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right] \frac{dS_T}{S_T}. \tag{3.29}$$

Then we can write

$$C(S,t) = e^{-r\tau}[A_1 - A_2], \tag{3.30}$$

and consider in turn the evaluation of $A_1$ and $A_2$.

[4] Note that we still denote the dependence of the option price on $t$ rather than $\tau$. Some readers may be tempted to write $C(S, \tau)$, but the dependence of $C$ on $\tau$ is properly denoted as $C(S, T - t) = C(S, \tau)$.

**The evaluation of $A_1$.** Make the change of variable

$$S_T = e^x \qquad (\text{ i.e.} \qquad x = \ln S_T,)$$

so that

$$\begin{aligned} A_1 &= \frac{1}{\sqrt{2\pi\tau}\sigma}\int_{\ln E}^{\infty} \exp\left[-\frac{\{x-\ln S-(r-\sigma^2/2)\tau\}^2}{2\sigma^2\tau}\right] e^x dx \\ &= \frac{1}{\sqrt{2\pi\tau}\sigma}\int_{\ln E}^{\infty} \exp\left[\frac{-(x-\alpha)^2+2\sigma^2\tau x}{2\sigma^2\tau}\right] dx, \end{aligned} \tag{3.31}$$

where we let

$$\alpha \equiv \ln S + (r-\sigma^2/2)\tau,$$

and have rearranged slightly. Overall, under the integral sign $e$ is being raised to a quadratic power, which may be simplified by using the procedure of completing the square.[5] Thus identifying the $\beta$ in footnote 5 with $-2\sigma^2\tau$ we have

$$\begin{aligned} (x-\alpha)^2 - 2\sigma^2\tau x &= (x-(\alpha+\sigma^2\tau))^2 - 2\sigma^2\tau\alpha - (\sigma^2\tau)^2 \\ &= (x-(\alpha+\sigma^2\tau))^2 - 2\sigma^2\tau(\ln S + r\tau) + (\sigma^2\tau)^2 - (\sigma^2\tau)^2 \\ &= (x-(\alpha+\sigma^2\tau))^2 - 2\sigma^2\tau(\ln S + r\tau). \end{aligned} \tag{3.32}$$

So substituting (3.32) into (3.31) we now write

$$A_1 = \frac{1}{\sqrt{2\pi\tau}\sigma}\int_{\ln E}^{\infty} \exp\left[-\frac{\{x-[\ln S+(r+\sigma^2/2)\tau]\}^2}{2\sigma^2\tau}\right] e^{\ln S + r\tau} dx,$$

i.e.

$$A_1 = \frac{Se^{r\tau}}{\sqrt{2\pi\tau}\sigma}\int_{\ln E}^{\infty} \exp\left[-\frac{\{x-[\ln S+(r+\sigma^2/2)\tau]\}^2}{2\sigma^2\tau}\right] dx.$$

---

[5]The procedure of completing the square simply makes use of the algebraic identity

$$x^2 + ax = (x+a/2)^2 - a^2/4$$

for any $a$. This procedure may be used to perform the following type of re-arrangement:

$$(x-\alpha)^2 + \beta x = x^2 - 2\alpha x + \alpha^2 + \beta x = x^2 + (\beta - 2\alpha)x + \alpha^2 = (x + (\frac{\beta}{2} - \alpha))^2 + \alpha\beta - \frac{\beta^2}{4},$$

which is what is being used here to rearrange the exponent.

Now let

$$y = \frac{x - [\ln S + (r + \sigma^2/2)\tau]}{\sigma\sqrt{\tau}},$$

so that

$$A_1 = \frac{Se^{r\tau}}{\sqrt{2\pi\tau}\sigma} \int_{-d_1}^{\infty} e^{-y^2/2} \sigma\sqrt{\tau} dy,$$

where we have put

$$d_1 = \frac{\ln S + (r + \sigma^2/2)\tau - \ln E}{\sigma\sqrt{\tau}}. \tag{3.33}$$

That is

$$A_1 = Se^{r\tau} \frac{1}{\sqrt{2\pi}} \int_{-d_1}^{\infty} e^{-y^2/2} dy.$$

Now it follows from the properties of the normal distribution (see Fig. 3.2) that

$$\frac{1}{\sqrt{2\pi}} \int_{-d_1}^{\infty} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-y^2/2} dy \equiv \mathcal{N}(d_1).$$

Hence finally

$$A_1 = Se^{r\tau} \mathcal{N}(d_1). \tag{3.34}$$

**The evaluation of $A_2$.** Recall that

$$A_2 = \frac{E}{\sqrt{2\pi\tau}\sigma} \int_{E}^{\infty} \exp\left[ -\frac{\{\ln(S_T/S) - (r - \sigma^2/2)\tau\}^2}{2\sigma^2\tau} \right] \frac{dS_T}{S_T}.$$

Make the change of variable

$$x = \ln(S_T/S) - (r - \sigma^2/2)\tau.$$

Then

$$A_2 = \frac{E}{\sqrt{2\pi\tau}\sigma} \int_{\ln(E/S)-(r-\sigma^2/2)\tau}^{\infty} \exp\left( -\frac{x^2}{2\sigma^2\tau} \right) dx.$$

Make the further change of variable

$$y = \frac{x}{\sigma\sqrt{\tau}},$$

then

$$A_2 = E\frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{-y^2/2}dy,$$

where

$$-d_2 = \frac{\ln(E/S) - (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}},$$

i.e.

$$d_2 = \frac{\ln(S/E) + (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}. \tag{3.35}$$

Using the same manipulations with the normal distribution as in the calculation of $A_1$ we find that

$$A_2 = E\mathcal{N}(d_2). \tag{3.36}$$

Substituting (3.34) and (3.36) into (3.30) we finally obtain

$$C(S,t) = S\mathcal{N}(d_1) - Ee^{-r\tau}\mathcal{N}(d_2).$$

Observe from (3.33) and (3.35) that

$$d_2 = d_1 - \sigma\sqrt{\tau}.$$

## 3.5 Problems

**Problem 3.1** (a) Consider the integral

$$I(b,x) = \int_x^{\infty} e^{-(v^2+bv)}dv.$$

By using the "trick" of completing the square show that

$$I(b,x) = \sqrt{\pi}e^{b^2/4}\mathcal{N}(-\sqrt{2}x - \frac{b}{\sqrt{2}}),$$

where the function $\mathcal{N}$ is defined as

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy.$$

(b) Consider the integral

$$J(a, b, x) = \int_{x}^{\infty} e^{-(v+a)^2 - bv} dv.$$

By making a suitable change of variable show that

$$J(a, b, x) = e^{ba} I(b, x + a).$$

Hence write out the expression for $J$ in terms of the function $\mathcal{N}$.

(c) Consider the function

$$K(b, x) = \int_{x}^{\infty} v e^{-(v^2+bv)} dv.$$

Use the technique of integration by parts to show that

$$K(b, x) = \frac{1}{2} e^{-(x^2+bx)} - \frac{b}{2} I(b, x).$$

Hence express $K(b, x)$ in terms of the function $\mathcal{N}$.

**Problem 3.2** It has been observed empirically that the price, $x$, of an asset has the following properties:

(i) The mean of small price changes over time interval $(t, t + \Delta t) \simeq a(x, t)\Delta t$;
(ii) The second moment of small price changes over $(t, t + \Delta t) \simeq b(x, t)\Delta t$.

(a) Write down the Kolmogorov backward equation for the conditional transition probability density function $p(x_T, T|x, t)$ for $t < T$.

(b) A derivative instrument written on the asset and which matures at time $T$ has payoff $g(x_T, T)$.
If investors are assumed to be risk neutral, write out the integral expression for the value, $f(x, t)$, of the derivative security at time $t(< T)$ when the price of the asset is $x$.

(c) By making use of the Kolmogorov backward equation, show that the function $f(x, t)$ satisfies the partial differential equation

$$\frac{1}{2} b(x, t) \frac{\partial^2 f}{\partial x^2} + a(x, t) \frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} = rf,$$

subject to the boundary condition

$$f(x, T) = g(x, T).$$

**Problem 3.3** Consider the partial differential equation (3.2). Use the transformation $y = \ln S$ (and $y_T = \ln S_T$) and the transformed transitional density function

$$p(S_T, T|S, t) = p(e^{y_T}, T|e^y, t) \equiv q(y_T, T|y, t)$$

to express it as

$$\frac{1}{2}\sigma^2 \frac{\partial^2 q}{\partial y^2} + \left(\mu - \frac{1}{2}\sigma^2\right) \frac{\partial q}{\partial y} + \frac{\partial q}{\partial t} = 0 \tag{3.37}$$

which must be solved subject to the initial condition

$$q(y_T, T|y, T) = \delta(y_T - y).$$

Use the result of Problem 2.3 to show that the solution to (3.37) is

$$q(y_T, T|y, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma} \exp\left[-\frac{(y_T - y - (\mu - \frac{1}{2}\sigma^2)(T-t))^2}{2\sigma^2(T-t)}\right].$$

Hence obtain the solution (3.13).

*Hint*: Recall the result in Appendix 2.1 about how the density function transforms under a change of variable.

**Problem 3.4** Consider an option with the following payoff structure:-

$$h(S_T) = \begin{cases} 0, & S_T \le E_1, \\ H\left(\frac{S_T - E_1}{E_2 - E_1}\right)^2, & E_1 < S_T \le E_2, \\ H, & E_2 < S_T \end{cases}$$

where the (positive) quantities $E_1$, $E_2$ and $H$ would be specified in the contract. Use the techniques of Appendix 3.1 and integration by parts to show that the value of this option contract is

$$C(S, \tau) = \frac{H}{(E_2 - E_1)^2}\Big[S^2 e^{(r+\sigma^2)\tau}\{\mathcal{N}[d_3(E_1)] - \mathcal{N}[d_3(E_2)]\}$$
$$- 2SE_1\{\mathcal{N}[d_1(E_1)] - \mathcal{N}[d_1(E_2)]\} + E_1^2 e^{-r\tau}\{\mathcal{N}[d_2(E_1)] - \mathcal{N}[d_2(E_2)]\}\Big]$$
$$+ He^{-r\tau}\mathcal{N}[d_2(E_2)],$$

where

$$d_1(K) = \frac{\ln \frac{S}{K} + \left(r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}},$$

$$d_2(K) = d_1(K) - \sigma\sqrt{\tau},$$

and

$$d_3(K) = \frac{\ln \frac{S}{K} + \left(r + \frac{3\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}.$$

# Chapter 4
# The Stochastic Differential Equation

**Abstract** To develop the hedging argument of Black and Scholes, this chapter introduces stochastic differential equations to model the evolution of the price path itself and the statistical properties of small price changes over small changes in time. We then consider the stochastic differential equations for the Wiener process, Ornstein–Uhlenbeck process and Poisson process and examine the autocovariance behaviour of the Wiener process. Furthermore we introduce stochastic integrals to define the stochastic differential equations.

## 4.1 Introduction

In the previous chapter our focus was on the conditional probability density function and how it evolves through time. Once we know how this time evolution occurs we are then able to obtain the expected value of some payoff contingent on future values of the asset (e.g. an option). With the approach developed there we can consider a wide range of contingent claim valuation problems that arise in practice, and indeed we shall consider a number of these in later chapters. In this chapter we develop an alternative approach to the valuation of contingent claims, which arises out of a consideration of the statistical properties of the sample paths of the asset price. The valuation approach to which we refer is the hedging argument developed by Black and Scholes. In order to develop it we need to concentrate on the evolution of the price path itself and on the statistical properties of small price changes over small changes in time. We also need to develop a technique for determining how the value of the contingent claim (e.g. an option) changes when the asset price undergoes a small change. Consideration of these problems will lead us to the development of stochastic integrals, stochastic differential equations and Ito's lemma.

In order to develop techniques and concepts which shall be useful in this and later chapters we shall consider some particular Markov processes such as the Wiener process, the Poisson process and the Ornstein–Uhlenbeck process. In order to clarify the significance of the white noise assumption and its relation to the Wiener process, and the Ornstein–Uhlenbeck process we introduce a section on autocorrelation functions, spectra and white noise. We then tackle the task of defining the stochastic integral and the stochastic differential equation. In Chap. 6



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_4

we undertake the derivation of Ito's lemma. Once all of this machinery is assembled we are then able to value a European call option by use of the continuous hedging argument.

## 4.2 A First Encounter with the Stochastic Differential Equation

If we consider a pure diffusion process then the jump term $J$ (see Sect. 2.6) is absent and the Lindeberg condition (see Sect. 2.4) is satisfied and so the sample paths are continuous. We know from our discussion in Chap. 2 that the conditional probability density function $p(z, t \mid y, \tau)$ satisfies the Fokker–Planck forward equation (2.17). Since our focus in this chapter is on how the statistical properties of the sample path evolve over a small interval of time $(t, t + \Delta t)$, we are particularly interested in the conditional probability density function $p(z, t + \Delta t \mid y, t)$. We may calculate this latter quantity by using the Fokker–Planck equation for which the appropriate initial condition is

$$p(z, t \mid y, t) = \delta(z - y). \tag{4.1}$$

Figure 4.1 illustrates the initial $\delta$ function distribution and the corresponding distribution $\Delta t$ time units later. For $\Delta t$ small the effect of the initial $\delta$ function will still be quite pronounced and so the solution of the Fokker–Planck will be very

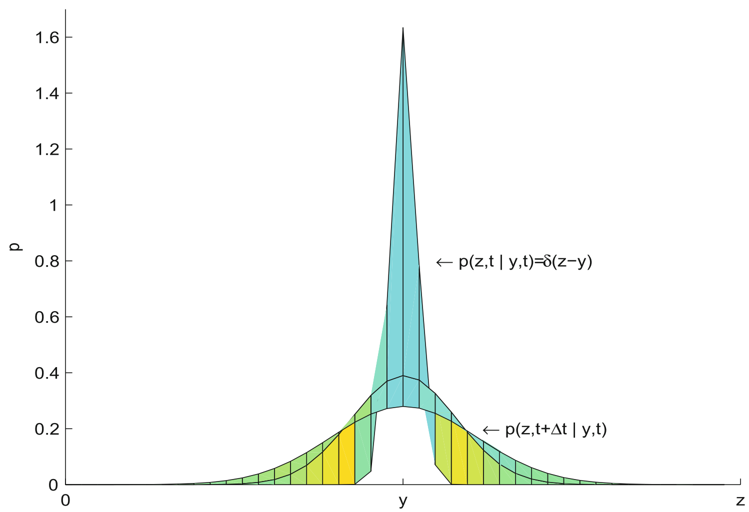


**Fig. 4.1** Evolution of conditional probability over time interval $\Delta t$

peaked. This means that the derivatives of the drift and diffusion terms $A$ and $B$ will be quite small in magnitude compared to those of $p$.

A reasonable approximation of the Fokker–Planck equation for the situation we are considering may be obtained by replacing $A(z,t)$ and $B(z,t)$ by $A(y,t)$ and $B(y,t)$ respectively. Thus the Fokker–Planck equation becomes

$$\frac{\partial}{\partial t}p(z,t \mid y,\tau) = -A(y,t)\frac{\partial}{\partial z}p(z,t \mid y,\tau) + \frac{1}{2}B(y,t)\frac{\partial^2}{\partial z^2}p(z,t \mid y,\tau). \qquad (4.2)$$

The solution to Eq. (4.2) subject to the $\delta$ function initial condition turns out to be (after identifying $\tau$ with $t$ and $t$ with $t+\Delta t$)[1]

$$p(z,t+\Delta t \mid y,t) = \frac{1}{\sqrt{2\pi\Delta t B(y,t)}}\exp\left[-\frac{(z-y-A(y,t)\Delta t)^2}{2B(y,t)\Delta t}\right], \qquad (4.3)$$

since we are assuming that $A$ and $B$ undergo negligible changes over the small time interval $(t, t+\Delta t)$.

Equation (4.3) is a Gaussian distribution with variance $B(y,t)\Delta t$ and mean $y + A(y,t)\Delta t$. The picture of the asset price movement that is implied by (4.3) is one in which there is an average price change of $A(y,t)\Delta t$ over the time interval $(t, t+\Delta t)$. This average price change is superimposed a Gaussian fluctuation with variance $B(y,t)\Delta t$ as illustrated in Fig. 4.2.

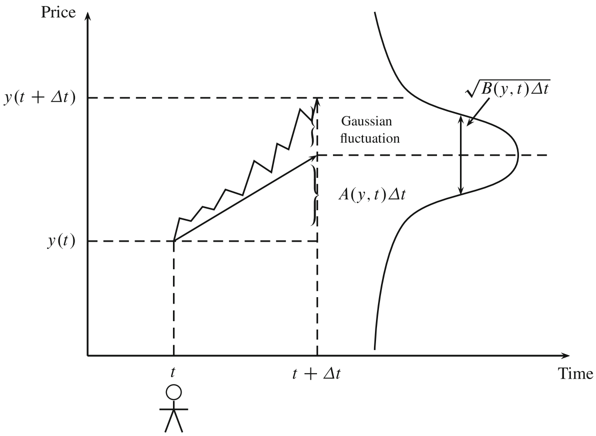


**Fig. 4.2** Price change broken down into a mean and a Gaussian fluctuation

[1] Here we are applying the result obtained in Problem 2.3.

This type of price movement over the time interval $(t, t + \Delta t)$ can be written in the form

$$y(t + \Delta t) = y(t) + A(y, t)\Delta t + \sqrt{B(y, t)}\sqrt{\Delta t}\xi(t), \tag{4.4}$$

where the random variable $\xi(t)$ is normally distributed and has the properties

$$\mathbb{E}[\xi(t)] = 0 \qquad \text{and} \qquad \text{var}[\xi(t)] = 1.$$

Here the expectation and variance are calculated at time $t$. In other words the price change over the interval $(t, t + \Delta t)$ can be written

$$\Delta y(t) = A(y, t)\Delta t + \sqrt{B(y, t)}\sqrt{\Delta t}\xi(t). \tag{4.5}$$

Equation (4.5) describes the stochastic evolution of the price over the small time interval $(t, t + \Delta t)$. Notice how our focus has now switched from following the evolution of the conditional probability $p(z, t \mid y, \tau)$ to following the stochastic evolution of the price path (i.e. the sample path of the stochastic process). Equation (4.5) is a discretised version of the stochastic differential equation for $y(t)$, which we shall describe more formally later on.

To see why we cannot switch to the standard differential equation notation of ordinary deterministic calculus note first of all that since

$$\Delta y(t) \to 0 \quad \text{as} \quad \Delta t \to 0,$$

the sample paths are everywhere continuous (this we already know from Lindeberg's condition). However if we attempt to form the derivative of $y(t)$ by considering the quotient $\Delta y(t)/\Delta t$ and let $\Delta t \to 0$, we encounter some difficulty. Note that

$$\frac{\Delta y(t)}{\Delta t} = A(y, t) + \frac{\sqrt{B(y, t)}\xi(t)}{\sqrt{\Delta t}}. \tag{4.6}$$

As we let $\Delta t \to 0$, the first term on the right-hand side of (4.6) is a well defined limit. However the second term presents some complications. We note that

$$\mathbb{E}\left[\frac{\xi(t)}{\sqrt{\Delta t}}\right] = 0 \quad \text{and} \quad \text{var}\left[\frac{\xi(t)}{\sqrt{\Delta t}}\right] = \frac{1}{\Delta t}.$$

Thus $\xi(t)/\sqrt{\Delta t}$ is a normal random variable with mean zero and variance $1/\Delta t$ as is illustrated in Fig. 4.3. As $\Delta t \to 0$ the variance tends to $\infty$ so that there is an equally likely probability of this stochastic variable drawing a value anywhere between $-\infty$ and $+\infty$. In other words this limit does not exist. What we are seeing here in an informal way is that the sample paths are not differentiable. We shall formally demonstrate this in Sect. 4.3.1. This non-differentiability is the reason why

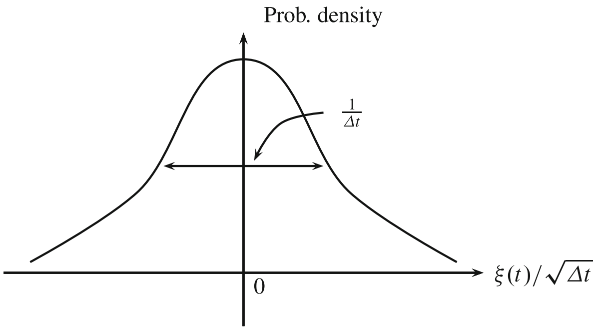


**Fig. 4.3** The distribution of $\xi(t)/\sqrt{\Delta t}$

we need to describe the stochastic evolution of the asset price with an equation such as (4.5) which concentrates on the price change rather than its derivative. We proceed to the limit $\Delta t \to 0$ by writing Eq. (4.5) in the form

$$dy(t) = A(y,t)dt + \sqrt{B(y,t)}dz(t), \tag{4.7}$$

where

$$\mathbb{E}[dz(t)] = 0, \quad \text{var}[dz(t)] = dt. \tag{4.8}$$

This notation implies

$$\mathbb{E}[dy(t)] = A(y,t)dt \quad \text{and} \quad \text{var}[dy(t)] = B(y,t)dt.$$

However at this stage this notation is purely formal and can only be given a proper mathematical interpretation once we define the stochastic integral.

## 4.3 Three Examples of Markov Processes

In this section we consider some fundamental solutions of the Fokker–Planck and Kolmogorov equations for certain special cases which have found wide application in finance theory.

### *4.3.1 The Wiener Process*

The Wiener process is a pure diffusion process whose drift coefficient is zero and diffusion coefficient is unity, i.e.

$$A = 0 \quad \text{and} \quad B = 1.$$

The Fokker–Planck equation in this situation reduces to

$$\frac{\partial}{\partial t} p(z, t \mid y, \tau) = \frac{1}{2} \frac{\partial^2}{\partial z^2} p(z, t \mid y, \tau), \tag{4.9}$$

subject to the initial condition

$$p(z, \tau \mid y, \tau) = \delta(z - y). \tag{4.10}$$

The solution turns out to be[2]

$$p(z, t \mid y, \tau) = \frac{1}{\sqrt{2\pi(t - \tau)}} \exp\left[-\frac{(z - y)^2}{2(t - \tau)}\right], \tag{4.11}$$

which is a Gaussian distribution with the properties

$$\mathbb{E}[z(t)] = y, \tag{4.12}$$

$$\text{var}[z(t)] = \mathbb{E}[(z(t) - y)^2] = (t - \tau). \tag{4.13}$$

Thus the initial sharp of $\delta$ function distribution spreads out in time as shown in Fig. 4.4. The Wiener process is also referred to as Brownian motion in the stochastic processes literature.

Some notable features of the Wiener process are firstly, the *highly irregular nature of its sample paths*. This arises from the fact that the variance becomes infinite as $(t - \tau) \to \infty$. Figure 4.4 shows some different sample paths arising out of the same initial price.[3] Figure 4.4a shows just 4 paths, whilst Fig. 4.4b shows a bundle of 1,000 paths. In the latter diagram we note the "spreading out" of the range over which the paths vary. This is just another way of seeing the increasing variance of the transitional probability density function as the time horizon increases.

Secondly the sample paths of the Wiener process are *non-differentiable*. We know from the way we introduced diffusion processes that the sample paths are continuous (because of the Lindeberg condition) and we saw at an intuitive level in Sect. 4.2 why they are not differentiable. To see more formally that they are not differentiable consider (see Fig. 4.5)

$$I(h) = \text{Prob}\left(\left|\frac{z(t + h) - z(t)}{h}\right| > k\right). \tag{4.14}$$

[2] Again we apply the result obtained in Problem 2.3.

[3] See Problem 4.6 for a discussion on how such paths are simulated.

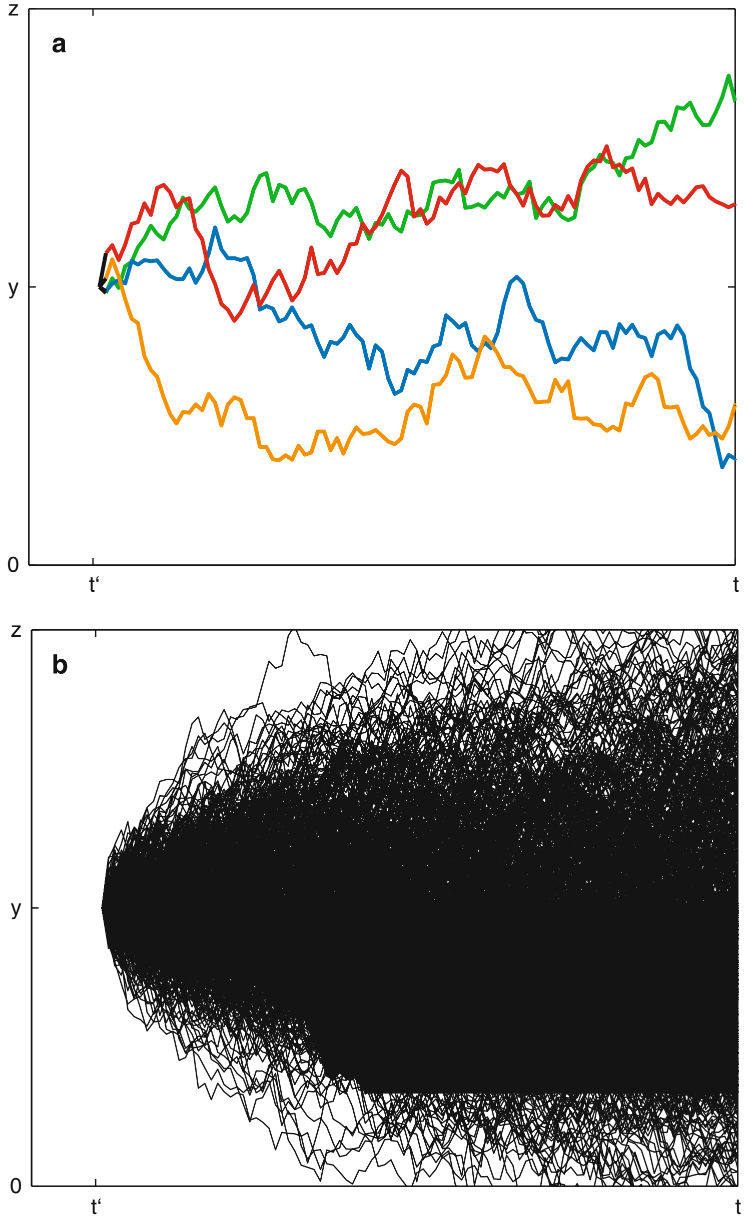


**Fig. 4.4** Irregular nature of simulated sample paths of the Wiener process; (**a**) shows 4 typical paths, (**b**) shows a bundle of 1,000 paths. Here $t - t' = 1$

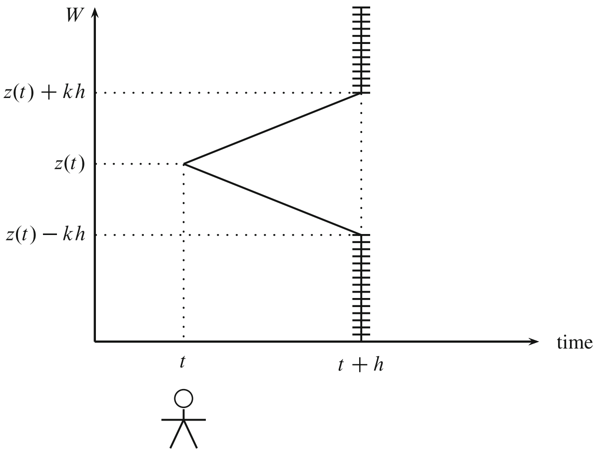


**Fig. 4.5** The set $\left|\frac{z(t+h)-z(t)}{h}\right| > k$

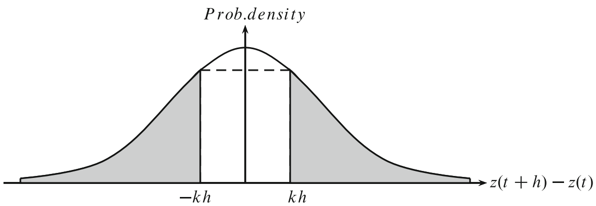


**Fig. 4.6** Shaded area = Prob($[z(t+h) - z(t)] > kh$)

Using the solution (4.11) for the conditional probability, this probability turns out to be (see Fig. 4.6)

$$I(h) = 2\int_{kh}^{\infty} \frac{1}{\sqrt{2\pi h}} e^{-z^2/2h} dz,$$

which tends to 1 as $h \to 0$. In other words no matter what value of $k$ we choose, $\left|\frac{z(t+h)-z(t)}{h}\right|$ is almost certain to be greater than it, which means that the derivative at any point in time is almost certainly infinite. This argument renders more mathematically precise the discussion in Sect. 2.4 showing intuitively that the sample paths of a diffusion process are nowhere differentiable (put $A = 0$ and $B = 1$ in Sect. 4.2 to make the two situations equivalent).

A third important property of the Wiener process is the *independence of the increments* of $z(t)$. By the Markov property of the Wiener process the joint probability of a price sequence $(z_0, t_0), (z_1, t_1), \ldots, (z_n, t_n)$ can be written

$$p(z_n, t_n; z_{n-1}, t_{n-1}; \ldots; z_0, t_0) = \prod_{i=0}^{n-1} p(z_{i+1}, t_{i+1} \mid z_i, t_i) p(z_0, t_0). \tag{4.15}$$

Using the solution (4.11) and setting

$$\Delta z_i = z(t_i) - z(t_{i-1}), \qquad \Delta t_i = t_i - t_{i-1},$$

we see that

$$p(z_{i+1}, t_{i+1} \mid z_i, t_i) = \frac{1}{\sqrt{2\pi \Delta t_{i+1}}} \exp\left(-\frac{\Delta z_{i+1}^2}{2\Delta t_{i+1}}\right). \tag{4.16}$$

Hence the joint probability density for the price changes, $\{\Delta z_i\}_{i=1}^n$, is

$$p(\Delta z_n; \Delta z_{n-1}; \Delta z_{n-2}; \ldots; \Delta z_1; z_0) = \left[\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \Delta t_i}} \exp\left(-\frac{\Delta z_i^2}{2\Delta t_i}\right)\right] p(z_0, t_0), \tag{4.17}$$

which shows that the price changes $\Delta z_i$ are independent of each other and of $z(t_0)$. The property of the independence of the increments $\Delta z_i$ turns out to be very important in the definition of the stochastic integral that we undertake below.

### *4.3.2 The Ornstein–Uhlenbeck Process*

The Wiener process does not have a stationary distribution, since as $t \to \infty$ the distribution at any point $y$ tends to zero, which means that ultimately the price moves between $-\infty$ and $+\infty$ with probability one. Hence the Wiener process is not stationary.

A process which does tend to a stationary distribution and is related to the Wiener process in a way to be explained later is the Ornstein–Uhlenbeck (O–U) process. This process has a linear drift term, $-kz$ (for $k$ a positive constant) and a diffusion coefficient $D$. The Fokker–Planck equation for this process is then

$$\frac{\partial}{\partial t} p(z, t \mid y, \tau) = \frac{\partial}{\partial z}[kz p(z, t \mid y, \tau)] + \frac{1}{2} D \frac{\partial^2}{\partial z^2} p(z, t \mid y, \tau), \tag{4.18}$$

with the same $\delta$ function initial condition. It can be shown (see Gardiner for technical details) that the solution is a Gaussian distribution with mean

$$\mathbb{E}[z(t) \mid y, \tau] = y e^{-k(t-\tau)}, \tag{4.19}$$

and variance

$$\operatorname{var}[z(t) \mid y, \tau] = \frac{D}{2k}[1 - e^{-2k(t-\tau)}]. \tag{4.20}$$

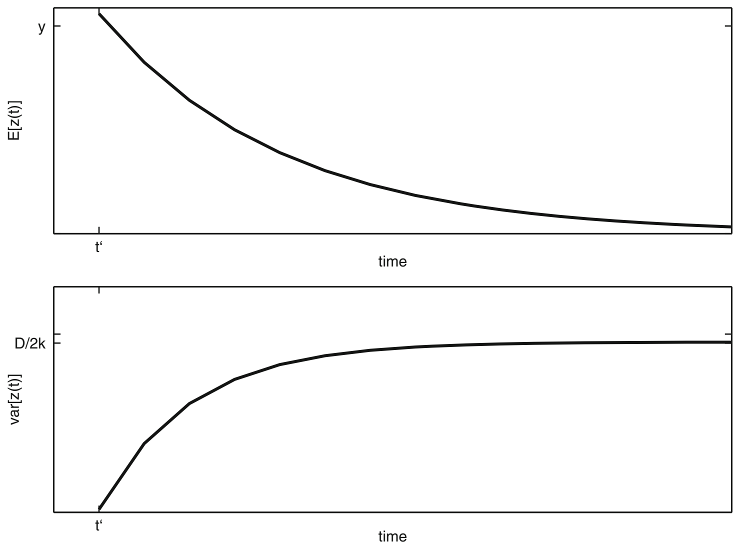


**Fig. 4.7** Time behaviour of mean and variance of the Ornstein–Uhlenbeck process

As $t \to \infty$ the mean tends to 0 and the variance tends to the constant value of $D/2k$, as illustrated in Fig. 4.7.

In a later chapter we shall discuss some empirical evidence which suggests that, at least in some bond markets, the variance behaviour shown in Fig. 4.7 is more likely to be observed than that implied by the Wiener process.

Finally note that in the limit $t \to \infty$, the Ornstein–Uhlenbeck process has a stationary distribution with mean 0 and variance $D/2k$. Using $p_s$ to denote the stationary distribution we can write

$$p_s(z) = \frac{\sqrt{k}}{\sqrt{\pi}\sqrt{D}} \exp(-\frac{kz^2}{D}). \tag{4.21}$$

In Fig. 4.8 we compare the density functions for the Wiener process and the Ornstein–Uhlenbeck process at various points in time. We have taken $\tau = 0$, $y(0) = 1$, $D = 1$ (same diffusion coefficient as the Wiener process) and $k = 1$ (for this value the stationary distribution is approached fairly rapidly). We see clearly how the distribution for the O–U process settles down fairly quickly onto the stationary distribution, whilst the distribution for the Wiener process continues to spread out.

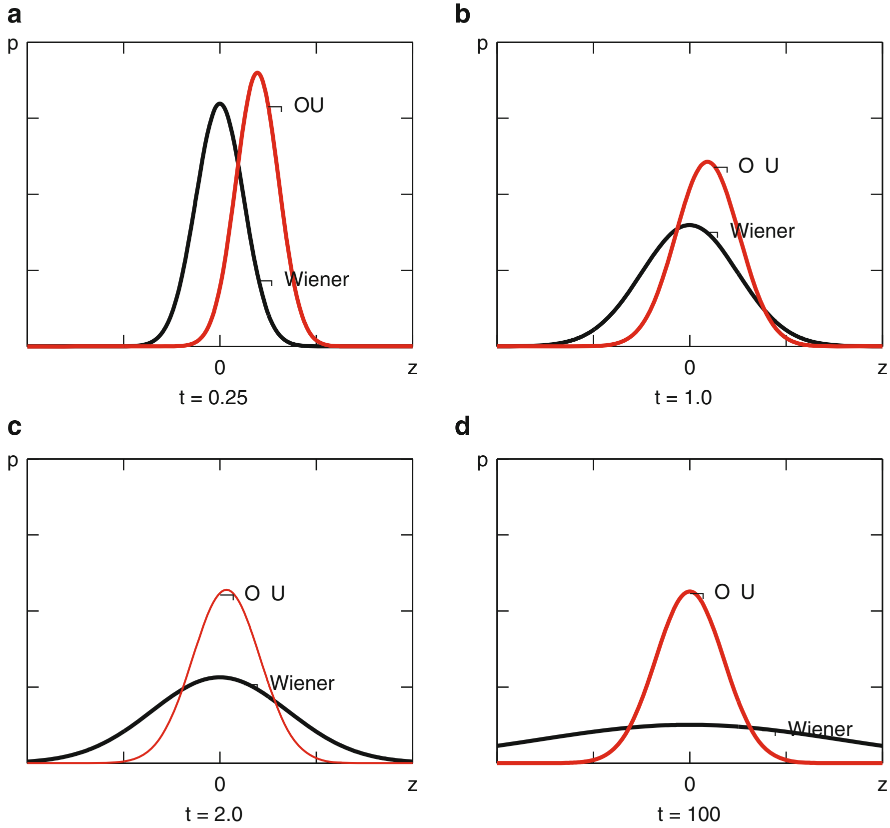


**Fig. 4.8** Comparing the distributions of the Wiener and O–U processes at different times

### *4.3.3 The Poisson Process*

We shall discuss the Poisson process within the framework of the Fokker–Planck and Kolmogorov equations of Sect. 2.6, in particular, the Fokker–Planck equation which allows for a jump component (i.e. Eq. (2.25)).

Let us concentrate on a *pure jump process* for the price movement, i.e. the drift $A = 0$ and the diffusion $B = 0$. The Fokker–Planck equation (2.25) reduces to

$$\frac{\partial}{\partial t} p(z,t \mid y,\tau) = \int_{-\infty}^{\infty} [J(z \mid x,t) p(x,t \mid y,\tau) - J(x \mid z,t) p(z,t \mid y,\tau)] dx. \tag{4.22}$$

Suppose, on the basis of past observations, we know that a price jump occurs with a frequency of $\lambda$ per time unit. Furthermore in order to concentrate on the essential feature of the process, assume that when the price jumps it jumps by a positive amount of one unit. In Chap. 12, when we price options whose underlying asset follows a jump process, we allow the magnitude of the jump to be also drawn from a probability distribution. Thus after the occurrence of $n$ jumps the price would be at $n$, and the conditional probability in which we are interested in this situation is[4]

$$\mathbb{P}(n,t \mid n',\tau) = \textit{probability that } n \textit{ price jumps have occurred by time } t,$$
$$\textit{given that } n' \textit{ have occurred up to the earlier time } \tau.$$

Since the integral is now evaluated at one point the integral equation in (4.22) collapses to

$$\frac{\partial}{\partial t}\mathbb{P}(n,t \mid n',\tau) = \lambda[\mathbb{P}(n-1,t \mid n',\tau) - \mathbb{P}(n,t \mid n',\tau)]. \tag{4.23}$$

Assuming there is no initial jump at time zero, the solution for this probability is (again see Gardiner for details)

$$\mathbb{P}(n,t \mid 0,0) = \frac{(\lambda t)^n}{n!}e^{-\lambda t}, \tag{4.24}$$

which gives us the probability that $n$ price jumps have occurred in the time interval $(0,t)$.

## 4.4 Autocovariance Behaviour and White Noise

Market analysts are naturally interested in how the price at one point in time may be related to the price at some other point in time. Such an effect is measured by the autocovariance function $\mathbb{E}[z(t)z(s) \mid z_0, t_0]$ which in a sense measures the amount of memory that the dynamics driving the price display. We give a brief review of this concept in Appendix 4.1.

The autocovariance function of the Wiener process turns out to be

$$\mathbb{E}[z(t)z(s) \mid z_0, 0] = \min(t,s) + z_0^2. \tag{4.25}$$

The result (4.25) is easily shown by rewriting the left-hand side as (for convenience assume $s < t$)[5]

$$\mathbb{E}[z(t)z(s)|z_0,0] = \mathbb{E}\big[z(s)\mathbb{E}_s(z(t))|z_0,0\big] = \mathbb{E}[z^2(s)|z_0,0] = s + z_0^2. \tag{4.26}$$

[4]Note that we are using $\mathbb{P}$ to denote probability of the number of jumps, whereas $p$ refers to probability of the value of the price. For this particular price jump process the two are equal of course.

[5]Note the law of iterated expectation $\mathbb{E}(X) = \mathbb{E}[\mathbb{E}(X|Y)]$.

Similarly if $t < s$, we find that

$$\mathbb{E}[z(t)z(s)|z_0, 0] = t + z_0^2. \tag{4.27}$$

Hence we have established the result (4.25).

Many of the calculations of stochastic calculus reduce to calculating the autocovariance of the increment of the Wiener process. Arguing at an informal level this amounts to considering $\mathbb{E}[\Delta z(t)\Delta z(s)]$ and then passing to the stochastic limit as $\Delta t \to 0$. To be done mathematically precisely we would need to consider stochastic integrals.

Adopting the informal approach we first assume $s \neq t$ and $s < t$ and consider

$$\mathbb{E}[\Delta z(t)\Delta z(s)] = \mathbb{E}\big[\Delta z(s)\mathbb{E}_s(\Delta z(t))\big] = 0.$$

The case $s \neq t$ and $s > t$ will yield the same result. Next consider $s = t$, then we have

$$\mathbb{E}[(\Delta z(t))^2] = \mathbb{E}[(z(t + \Delta t) - z(t))^2] = \Delta t. \tag{4.28}$$

Hence we have shown that

$$\mathbb{E}[\Delta z(t)\Delta z(s)] = \begin{cases} 0 & \text{if} \quad s \neq t, \\ \Delta t & \text{if} \quad s = t, \end{cases} \tag{4.29}$$

and so passing to the stochastic limit

$$\mathbb{E}[dz(t)dz(s)] = \begin{cases} 0 & \text{if} \quad s \neq t, \\ dt & \text{if} \quad s = t. \end{cases} \tag{4.30}$$

We note that (4.30) is another way (heuristically) of seeing that the Wiener increments are independent. The result (4.30) may also be written

$$\mathbb{E}[dz(t)dz(s) \mid z_0, 0] = \delta(t - s)\sqrt{dtds}, \tag{4.31}$$

showing that the autocovariance of the increment of the Wiener process is a Dirac delta function.

For the Ornstein–Uhlenbeck process, the stationary autocovariance function turns out to be

$$\mathbb{E}[x(t)x(s)] = \frac{D}{2k}e^{-k|t-s|} \tag{4.32}$$

and is illustrated in Fig. 4.9.

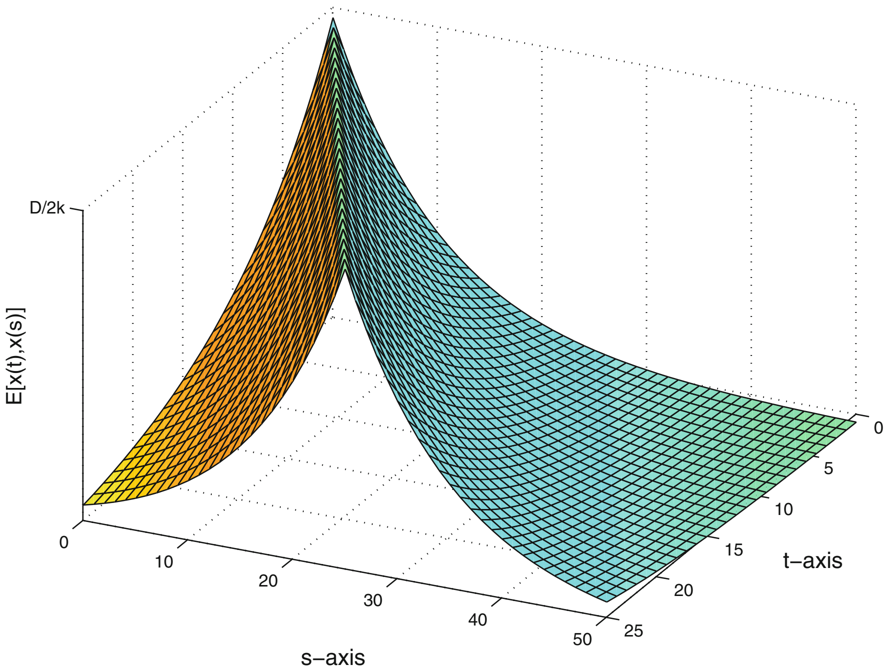


**Fig. 4.9** Autocovariance function of the Ornstein–Uhlenbeck process; here $D = 1$ and $k = 0.5$

It is more convenient to express the autocovariance function as a measure of the distance between $t$ and $s$, i.e. (see Appendix 4.1)

$$G(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt = \frac{D}{2k} e^{-k|\tau|}. \tag{4.33}$$

Recall that $G(\tau)$ measures the amount of memory in the stochastic process $x(t)$. Figure 4.9 indicates that this memory drops off exponentially, the more rapidly the larger is $k$. In this connection it is useful to have a measure of the average amount of memory time of the stochastic process, which is defined as

$$\tau_{cor} = \frac{1}{G(0)} \int_0^\infty G(\tau)d\tau = \frac{1}{k}. \tag{4.34}$$

Clearly as $k \to \infty$, the average memory time $\tau_{cor} \to 0$. This is another way of seeing that the Ornstein–Uhlenbeck process tends to white noise as $k \to \infty$. However this limiting process needs to be treated with some caution. To appreciate this comment consider the power spectrum (that broadly speaking, measures the strength of the noise) of the Ornstein–Uhlenbeck process which from Appendix 4.1 is given by

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\tau)e^{-i\omega\tau}d\tau = \frac{D}{2\pi}\,\frac{1}{\omega^2 + k^2}. \tag{4.35}$$

If we proceed naively to the limit $k \to \infty$ in (4.35) we will see that $S(\omega) \to 0$. However a zero power spectrum implies no noise at all! So the naive passage to the limit $k \to \infty$ ($\tau_{cor} \to 0$) poses the dilemma that the limit represents a no noise situation. The way out of this dilemma is to make the diffusion coefficient $D$ (which measures the intensity of the noise) also depend on $k$. With an eye to a further discussion of this issue in Sect. 6.3.5, we set

$$D = \sigma^2 k^2, \tag{4.36}$$

where $\sigma$ is a constant. Now we find that

$$\lim_{k\to\infty} S(\omega) = \frac{\sigma^2}{2\pi}. \tag{4.37}$$

Thus with the choice of diffusion coefficient (4.36), the limit $k \to \infty$ ($\tau_{cor} \to 0$) yields a noise process which has a flat power spectrum. Since a flat power spectrum is a characteristic of white light (and is a consequence of the fact that in white light all frequencies are present with equal power) this limiting noise process is known as *white noise*.

With $D$ defined as in (4.36) we see that the maximum value of the autocovariance function of the Ornstein–Uhlenbeck process in Fig. 4.9 is $\sigma^2 k/2$. So as $k \to \infty$ the autocovariance $G(\tau)$ becomes more peaked around the origin in such a way that

$$\lim_{k\to\infty} G(\tau) = \sigma^2\delta(\tau). \tag{4.38}$$

Thus white noise can be viewed as an Ornstein–Uhlenbeck process with zero memory time and having infinite intensity all in such a way that the limits (4.37) and (4.38) are satisfied. Of course white noise is a highly idealised process and probably does not exist in the real world. Real market noise is probably best modelled as an Ornstein–Uhlenbeck process with a high value of $k$ (such a noise process is often referred to as coloured noise). However it is mathematically more convenient to use white noise. Whilst modern financial markets are highly efficient they are nevertheless likely to display some small but nonzero memory time dependent on some inevitable market frictions. The important issue is whether the idealised white noise with $\tau_{cor} = 0$ is a robust approximation to real noise with $\tau_{cor} \simeq 0$. It turns out that the answer to this question is yes. It is this fact that allows us to proceed to build a theory of uncertainty in financial markets based on the theory of Markov diffusion processes and semimartingale integration.

## 4.5 Modelling Uncertain Price Dynamics

In a world of certainty we would use a differential equation such as

$$\frac{dx}{dt} = \mu(x, t), \tag{4.39}$$

to model the price movements in continuous time, where the form of the function $\mu$ would depend on the characteristics of the particular market under consideration. For example for stock prices we could write

$$\frac{dx}{dt} = \mu x, \tag{4.40}$$

where $\mu$ (constant) is the average growth rate per unit time in stock prices over an appropriate period. For $\mu > 0$ this would result in the upward trending price path shown in Fig. 4.10a. On the other hand for short term interest rates we might use

$$\frac{dx}{dt} = \kappa(x_0 - x), \tag{4.41}$$

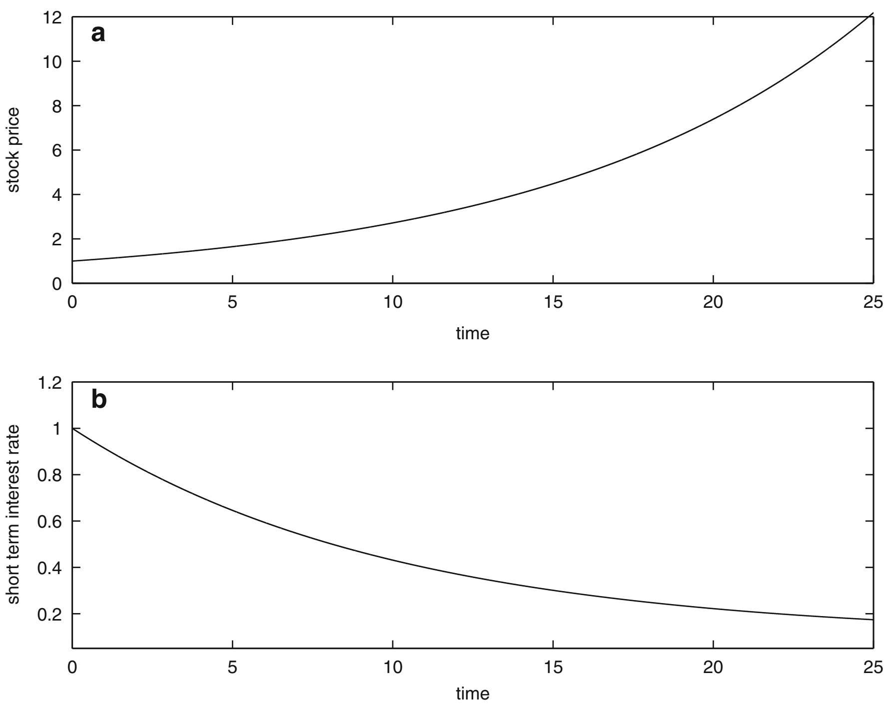


**Fig. 4.10** Deterministic long run price paths; (**a**) stock price path with $\mu = 0.1$, (**b**) short term interest rate path with $x_0 = 0.1$ and $k = 0.1$

where $x_0$ is the equilibrium short term rate and $\kappa > 0$ is the speed at which this rate is approached. The path of the short run rate is as shown in Fig. 4.10b.

Deterministic models of the type in Eq. (4.39) are adequate for an analysis of long-run trends when the day to day fluctuations even out. However in option pricing situations, the day-to-day fluctuations are of prime importance. Hence we need to add to the right hand side of (4.39) a term which captures the short-term (i.e. day-to-day) fluctuations caused by the market noise discussed in the previous section, and whose parameters could eventually be estimated by statistical techniques.

Our initial temptation would be to add another term to the right-hand side of (4.39) so that

$$\frac{dx}{dt} = \mu(x,t) + \sigma(x,t)\zeta(t), \tag{4.42}$$

where $\zeta(t)$ is a rapidly fluctuating random term and $\sigma(x,t)$ measures the magnitude of the fluctuations and its dependence on price and time. If we want to model the noise as a white noise process we would demand of the random terms $\zeta(t)$ that they be statistically independent at different points in time, i.e.

$$\mathbb{E}[\zeta(t)\zeta(s)] = \delta(t-s). \tag{4.43}$$

The difficulty with trying to proceed with this approach is that if we wanted the price path to be a Markov diffusion process (we would want to do this so as to be able to make use of much of the established theory of stochastic processes discussed earlier) then we are not really able to write down the derivative $dx/dt$, since, the sample paths of the now random processes are not differentiable, as we saw in Sect. 4.2 when we tried to form the derivative. Indeed we see from the analysis in that section (in particular equation (4.6)) the case in which the randomly fluctuating term is modelled by

$$\zeta(t) = \lim_{\Delta t \to 0} \frac{\xi(t)}{\sqrt{\Delta t}}, \tag{4.44}$$

where the $\xi(t)$ is a standard normal variable (i.e. $\mathbb{E}[\xi(t)] = 0$, $\text{var}[\xi(t)] = 1$), which is statistically independent across time (i.e. $\mathbb{E}[\xi(t)\xi(s)] = 0$, $t \neq s$). When the rapidly fluctuating random $\zeta(t)$ term is viewed in this way, the origin of the $\delta$ function term in (4.43) becomes clear since $\text{var}(\zeta(t)) = 1/\Delta t$. In fact $\zeta(t)$ has the same probability density function as the one illustrated in Fig. 4.3.

Since the sample paths of diffusion processes are not differentiable we cannot satisfactorily proceed with the differential equation (4.42). In order to cater for the fact that sample paths are not differentiable what we need to do is go back to the step before taking the limit $\Delta t \to 0$ and consider the price change over a small time interval $(t, t + \Delta t)$. We would then have

$$x(t+\Delta t) - x(t) = \mu(x,t)\Delta t + \sigma(x,t)\Psi(t), \tag{4.45}$$

where $\Psi(t)$ is an appropriate stochastic process. The first term on the right-hand side represents the average price change whilst the second term represents the random fluctuations around this average. The main point at issue is which stochastic process is the most appropriate to use for the $\Psi(t)$ term. To properly answer this question we really need to look at the empirical behaviour of asset price time series. The vast literature on market efficiency, initiated by Fama (1970), suggests that the price changes at different points in time are statistically independent. This evidence suggests that an appropriate stochastic process for $\Psi(t)$ would be the increments of the Wiener process, which we saw in Sect. 4.3.1 enjoy the property of statistical independence. We shall see in later chapters that a richer class of stochastic processes can be used in addition to the Wiener process. For instance we could include in $\Psi(t)$ a Poisson jump process, which also enjoys the independence of increments properly. We could also allow the standard deviation of $\Psi(t)$ to be driven by another diffusion process and so obtain stochastic volatility models. At a most general level we could use Lévy processes (see Eberlein 2001) to model $\Psi(t)$. All of those embellishments contribute to making asset returns exhibit fat tails, skewness and peakedness that is characteristic of financial market data. Thus the stochastic process term we shall use in (4.45) is

$$\Psi(t) = \Delta z(t), \tag{4.46}$$

where $\Delta z(t)$ is the increment of the Wiener process $z(t)$ over the interval $(t, t+\Delta t)$. From our discussion in Sect. 4.3.1 (see in particular equations (4.10) or (4.16))

$$\mathbb{E}[\Delta z(t)] = 0, \tag{4.47}$$

$$\text{var}[\Delta z(t)] = \Delta t. \tag{4.48}$$

Hence, the equation for the price change (4.45) becomes

$$x(t + \Delta t) - x(t) = \mu(x,t)\Delta t + \sigma(x,t)\Delta z(t), \tag{4.49}$$

which of course is (4.5) in a different notation.

## 4.6 Proceeding to the Continuous Time Limit

In order to proceed to the continuous time limit in (4.49), we let $\Delta t \to 0$ as we do in ordinary deterministic calculus (except that now we do *not* divide throughout by $\Delta t$ before taking the limit). At this point we encounter some definitional problems. Imitating the differential notation of ordinary calculus, it is easy enough to write, in the limit,

$$dx = \mu(x,t)dt + \sigma(x,t)dz, \tag{4.50}$$

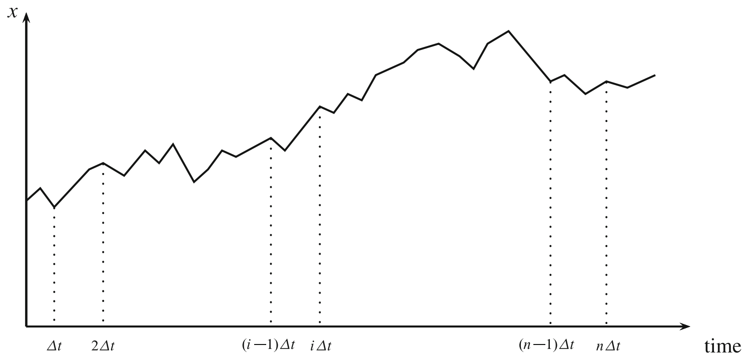


**Fig. 4.11** Time subdivisions for the price changes

however what meaning should we attach to these symbols? For instance, how do we interpret $\lim_{\Delta t \to 0} \Delta z(t)$ which we have formally written as $dz(t)$?

The answer turns out to be that we should reinterpret the equation for price change, (4.49), as an integral. To see how an integral may arise consider the interval $(0, t)$ divided into $n$ sub-periods each of length $\Delta t$, as shown in Fig. 4.11.

Summing the price changes over successive intervals we see that

$$x(t) - x(0) = \sum_{i=0}^{n-1} \mu(x_i, i\Delta t)\Delta t + \sum_{i=0}^{n-1} \sigma(x_i, i\Delta t)\Delta z_i. \tag{4.51}$$

In the limit as $\Delta t \to 0$ the sums on the right-hand side will clearly become integrals, so that we could write

$$x(t) - x(0) = \int_0^t \mu(x, s)ds + \int_0^t \sigma(x, s)dz(s). \tag{4.52}$$

Since $x(s)$ is continuous and the function $\mu$ is assumed to be reasonably well behaved,[6] the first integral on the right-hand side is readily seen to be the standard Riemann integral of ordinary calculus. This integral is defined by considering the maximum $(\bar{\mu}_i)$ and minimum $(\underline{\mu}_i)$ values of $\mu$ over each subinterval $[(i-1)\Delta t, i\Delta t]$ as shown in Fig. 4.12. The area under $\mu$ over this interval, denoted $A_i$, satisfies

$$\underline{\mu}_i \Delta t \leq A_i \leq \bar{\mu}_i \Delta t.$$

By summing over all $n$ subintervals and taking the limit $\Delta t \to 0$, $n \to \infty$ in an appropriate manner we are able to properly define $\int_0^t \mu(x, s)ds$.

[6]That is, $\mu$ maps continuous functions into functions with at most a countable number of discontinuities.

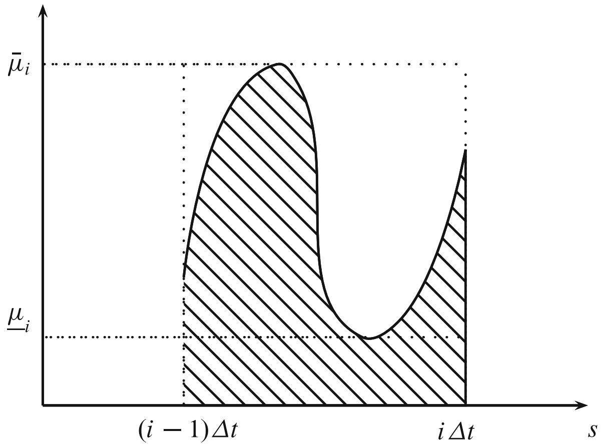


**Fig. 4.12** The typical subinterval for a Riemann integral; note the upper and lower bounds

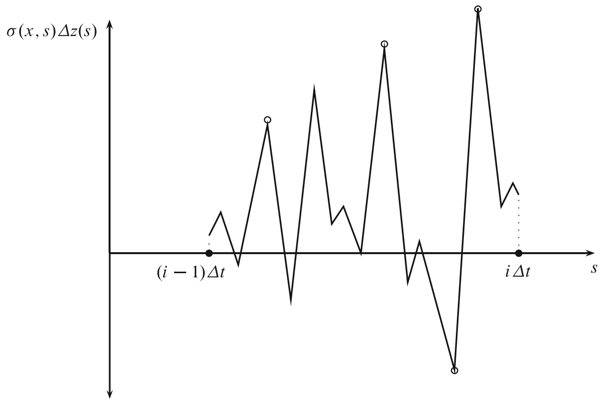


**Fig. 4.13** The typical subinterval for the stochastic integral; the *open circle* indicate points of unboundedness

The problem that needs to be addressed is how to define and interpret the second integral on the right-hand side. To see the nature of the problem here consider the function $\sigma(x, s)\Delta z(s)$ over the subinterval $[(i-1)\Delta t, i\Delta t]$. The term $\Delta z(s)$ has mean zero but can with positive probability take values (either up or down) larger than any bound over this subinterval. Hence it is not possible to find the equivalent of $\bar{\mu}_i$, $\underline{\mu}_i$ for this term. Also, because of the random term $\Delta z(s)$ the function is highly irregular over the subinterval. Figure 4.13 attempts to illustrate this situation by plotting the integrand $\sigma(x, s)\Delta z(s)$ as a function of $s$, with the $\circ$ indicating points where it becomes unbounded.

To solve this problem properly we need the technical apparatus of measure theory by means of which a new kind of integral, known as the stochastic integral, has been developed. In the next section we merely develop the intuition of this concept using ideas of classical analysis.

Before setting off on the task of defining the stochastic integral and after that deriving Ito's lemma, a word of caution should be sounded about an assumption we have made on how to model the fluctuations of the price process. In deciding to model the stochastic term, $\Psi(t)$, in the price change equation (4.45) by the increments of a Wiener process, $\Delta z(t)$, we refer to the literature on efficient markets to justify this assumption. In spite of the wide spread acceptance within the academic finance literature of the efficient markets hypothesis (and the attendant hypothesis of random walk behaviour of asset price movements), there has nevertheless continued to exist what could be described as a "counter-culture" literature which casts doubt on the efficient markets hypothesis, at least in its purest form. This literature contains a number of strands. One strand is the so-called anomalies literature which claims to provide evidence that excess stock returns can be related to a variety of factors such as firm size, market sentiment, day of the week, month of the year, dividend yield, to name but a few. A good survey of this literature can be found in Keim (1986). Another strand of literature studies variance properties of stock returns and finds that these are not consistent with a random walk model; see Lo and Mackinlay (1988). In yet another strand there is presented evidence that stock returns contain a sizable predictable component e.g. French et al. (1987) and Fama and French (1988). There is of large literature on the excess volatility of asset prices, e.g. West (1988), that puts into question the view that asset prices are determined solely from rationally expected fundamental values, which according to Samuelson (1965) should follow a random walk. Then there are continual claims that it is possible to "beat the market" with technical analysis. The article by Leroy (1989) surveys many of those issues, and concludes that there does seem to be some dependency in returns.

It is well to bear in mind that in using the increments of the Wiener process (i.e. white noise) to model the stochastic fluctuations we are excluding the types of behaviour referred to in the above empirical studies. This is another reflection of the fact that pure white noise is really a mathematical idealisation which does not exist in the real world. Extending the $\Psi(t)$ term of Eq. (4.45) to incorporate Poisson jump processes, stochastic volatility or more generally Lévy processes can go some way to explaining many (though not all) of the apparent anomalies.

## 4.7 The Stochastic Integral

Let us recall first of all the definition of the Riemann integral of ordinary deterministic calculus. In this case the integral corresponds to the area under the curve described by the function. Consider the function $\mu$ on the interval $[0, T]$. Divide

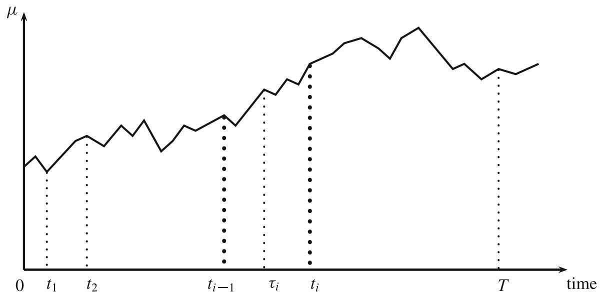


**Fig. 4.14** Partitioning the time interval

this interval into $n$ sub-intervals $[t_0, t_1], [t_1, t_2], \ldots, [t_{i-1}, t_i], \ldots, [t_{n-1}, t_n]$. Note that the sub-intervals are not necessarily of equal length (Fig. 4.14).

Concentrating on the $i$th sub-interval $[t_{i-1}, t_i]$, we take a point $\tau_i$ in this sub-interval such that

$$t_{i-1} \leq \tau_i \leq t_i,$$

and form

$$\mu(\tau_i)(t_i - t_{i-1}).$$

This quantity represents an approximation to the area under the curve between $t_{i-1}$ and $t_i$. Now form the sum of all such terms, i.e.

$$S_n = \sum_{i=1}^{n} \mu(\tau_i)(t_i - t_{i-1}).$$

The quantity $S_n$ is known as the *partial sum*. If, as we let the number of partitions tend to $\infty$, the partial sum $S_n$ tends to a limit, then we say the function $\mu$ is Riemann integrable over $[0, T]$ and the limit to which the partial sums converge is known as the integral of $f$ over $[0, T]$. Mathematically this is written

$$\lim_{n\to\infty} S_n = S,$$

or

$$\lim_{n\to\infty} \sum_{i=1}^{n} \mu(\tau_i)(t_i - t_{i-1}) = \int_0^T \mu(t)dt. \tag{4.53}$$

Figure 4.15 illustrates the convergence of the partial sums. It can be shown that if the function $\mu$ is continuous, or is piece wise continuous with a finite number of

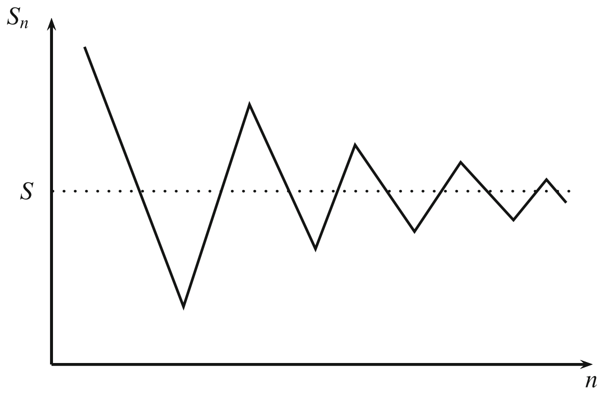


**Fig. 4.15** Convergence of the partial sums converging

discontinuous points, then $\mu$ is integrable in the above sense. It can also be shown that the same result is obtained, no matter what the choice of $\tau_i$ in the interval $[t_{i-1}, t_i]$. As we shall see below, the choice of $\tau_i$ does make a difference when we come to define the stochastic integral.

Now let us set out to define a stochastic integral in the same way. Let $\sigma(x(t), t)$ be defined on $[0, T]$ and let $z(t)$ be a Wiener process. Using the same partition as in Fig. 4.14 form, over the interval $[t_{i-1}, t_i]$, the quantity

$$\sigma(x(\tau_i), \tau_i)[z(t_i) - z(t_{i-1})] = \sigma(x(\tau_i), \tau_i)\Delta z(t_i). \tag{4.54}$$

Given the properties of the increments of the Wiener process, $\Delta z(t_i)$, displayed in Eqs. (4.12) and (4.13), the quantity in Eq. (4.54) is a random variable with mean 0 and variance $\sigma(x(\tau_i), \tau_i)^2(t_i - t_{i-1})$. The distribution of this term is illustrated in Fig. 4.16.

Now form the sum of all such terms over $[0, T]$ to obtain the partial sum

$$S_n = \sum_{i=1}^{n} \sigma(x(\tau_i), \tau_i)[z(t_i) - z(t_{i-1})] = \sum_{i=1}^{n} \sigma(x(\tau_i), \tau_i)\Delta z(t_i). \tag{4.55}$$

We will say that the stochastic integral of $\sigma$ exists if the partial sums $S_n$ converge as $n \to \infty$. Now however the $S_n$ are random variables and we have to define what we mean by the limit of a sequence of random variables.

There is no unique way of defining the limit of a sequence of random variables, however the way that turns out to be most useful in practical applications is the *mean square limit*. The sequence of random variables $S_n$ is said to converge to the random variable $S$ in *mean square* if

$$\lim_{n \to \infty} \mathbb{E}[(S_n - S)^2] = 0. \tag{4.56}$$

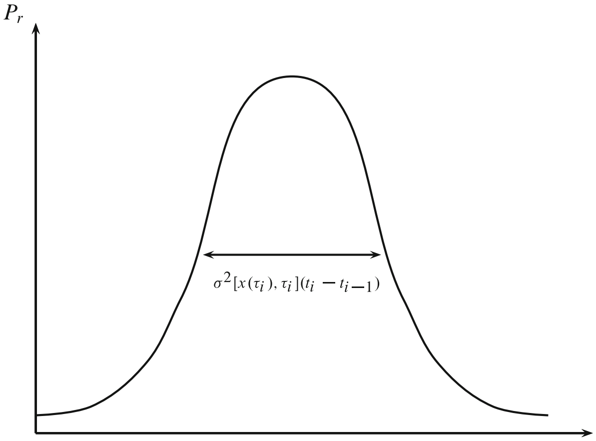


**Fig. 4.16** The distribution of a typical term in the partial sum of the stochastic integral

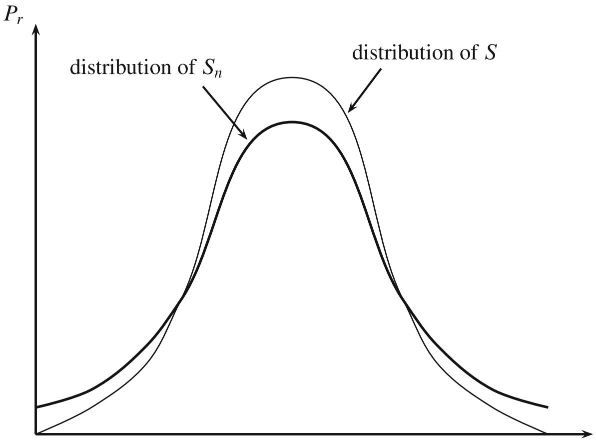


**Fig. 4.17** The distributions of $S_n$ and $S$

Figure 4.17 attempts to illustrate this concept, showing that in the limit there will be some limiting distribution. This limit is usually indicated by writing

$$\underset{n\to\infty}{\text{ms-lim}}\, S_n = S. \tag{4.57}$$

If the partial sums in (4.55) converge in the mean square sense to a limit $S$ in (4.56) then we call this limit the *stochastic integral* and write

$$\underset{n\to\infty}{\text{ms-lim}}\, S_n = \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \sigma(x(\tau_i), \tau_i)[z(t_i) - z(t_{i-1})] = \int_o^T \sigma(x(t), t)dz(t). \tag{4.58}$$

However, unlike the Riemann integral of deterministic calculus, the limit obtained in (4.58) above *does* depend on the choice of the intermediate point $\tau_i$. This point is most simply illustrated by an example.

## 4.8 An Example of Stochastic Integral

Let $\sigma$ itself be a Wiener process $z(t)$, so that the partial sum is

$$S_n = \sum_{i=1}^{n} z(\tau_i)[z(t_i) - z(t_{i-1})].$$

We can obtain different values for the mean of the integral to which $S_n$ converges, depending on the choice of $\tau_i$. Taking expectations of the last equation we have

$$\begin{aligned}\mathbb{E}[S_n] &= \sum_{i=1}^{n} \{\mathbb{E}[z(\tau_i)z(t_i)] - \mathbb{E}[z(\tau_i)z(t_{i-1})]\} \\ &= \sum_{i=1}^{n} [\min(\tau_i, t_i) - \min(\tau_i, t_{i-1})] \quad \text{(using Eq. (4.25))} \\ &= \sum_{i=1}^{n} (\tau_i - t_{i-1}). \end{aligned} \tag{4.59}$$

If we choose

$$\tau_i = \alpha t_i + (1-\alpha)t_{i-1}, \qquad (0 \le \alpha \le 1) \tag{4.60}$$

so that by letting $\alpha$ vary from 0 to 1, $\tau_i$ varies from $t_{i-1}$ to $t_i$, then

$$\mathbb{E}[S_n] = \sum_{i=1}^{n} \alpha(t_i - t_{i-1}) = \alpha T \tag{4.61}$$

i.e.

$$\lim_{n\to\infty} \mathbb{E}[S_n] = \alpha T. \tag{4.62}$$

Thus the mean value of the integral can vary between 0 and $T$, depending on the choice of $\alpha$.

We are thus faced with a wide variety of possible definitions of the stochastic integral. One definition which has turned out to be very useful in practical applications is obtained by setting $\alpha = 0$ so that

$$\tau_i = t_{i-1}. \tag{4.63}$$

The stochastic integral defined with this choice of $\alpha$ is known as the *Ito stochastic integral* of the function $\sigma$. Formally the Ito stochastic integral is defined as

$$\int_0^T \sigma(x(t), t)dz(t) = \underset{n\to\infty}{\text{ms-lim}} \left\{ \sum_{i=1}^{n} \sigma(x(t_{i-1}), t_{i-1})[z(t_i) - z(t_{i-1})] \right\}. \tag{4.64}$$

As a simple example of an Ito stochastic integral, we consider again the example when the function $\sigma$ is the Wiener process $z(t)$ to illustrate some of the computational techniques involved. For the Ito stochastic integral the relevant partial sum is

$$S_n = \sum_{i=1}^{n} z_{i-1}[z_i - z_{i-1}] = \sum_{i=1}^{n} z_{i-1}\Delta z_i. \tag{4.65}$$

This expression may be manipulated as follows

$$\begin{aligned} S_n &= \frac{1}{2}\sum_{i=1}^{n}[(z_{i-1} + \Delta z_i)^2 - z_{i-1}^2 - (\Delta z_i)^2] \\ &= \frac{1}{2}\sum_{i=1}^{n}[z_i^2 - z_{i-1}^2 - (\Delta z_i)^2] \\ &= \frac{1}{2}[z(T)^2 - z(0)^2] - \frac{1}{2}\sum_{i=1}^{n}(\Delta z_i)^2. \end{aligned}$$

We show in Appendix 4.2 that the mean square limit of the last term on the right-hand side is $T$, i.e.

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\sum_{i=1}^{n}(\Delta z_i)^2 - T\right)^2\right] = 0. \tag{4.66}$$

Hence

$$\underset{n\to\infty}{\text{ms-lim}}\, S_n = \frac{1}{2}[z(T)^2 - z(0)^2] - \frac{1}{2}T, \tag{4.67}$$

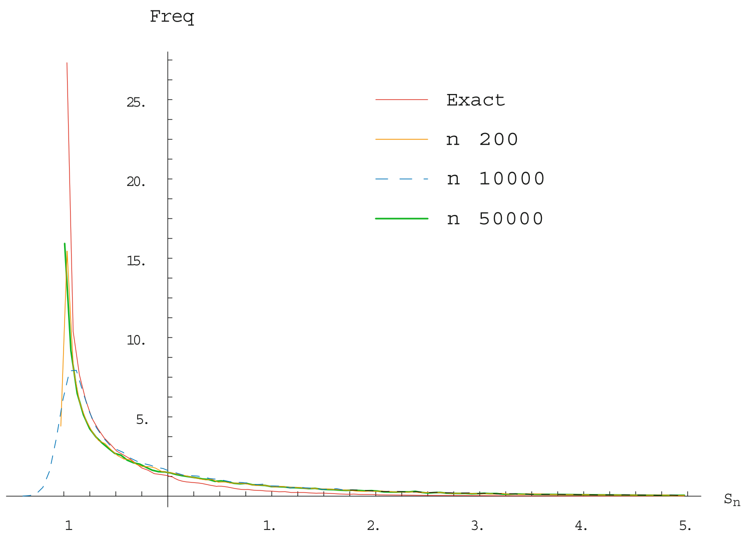


**Fig. 4.18** Approximating $\int_0^T z(t)dz(t)$ with the partial sums $S_n$

that is

$$\int_0^T z(t)dz(t) = \frac{1}{2}[z(T)^2 - z(0)^2] - \frac{1}{2}T. \tag{4.68}$$

In order to give some idea of how quickly there is convergence, in Fig. 4.18 we display the distribution for $S_n$ (defined in (4.65)) with $T = 2$ for various value of $n$ where for each value of $n$ we have simulated 100,000 sample paths for the Wiener process. In this example we are able to compare with the exact distribution calculated from (4.68). We see that for this example by 50,000 subdivisions convergence to the exact distribution is reasonable. In computing the exact distribution we note that $z(T) \sim N(0, T)$ and that the square of a normal random variate has a chi-squared distribution.

## 4.9 The Proper Definition of the Stochastic Differential Equation

With the Ito stochastic integral properly defined we are now able to give a meaning to the limit of the price change equation from Eq. (4.49) to (4.50). We shall say that the stochastic variable $x(t)$ satisfies the *Ito stochastic differential equation*

$$dx(t) = \mu(x(t), t)dt + \sigma(x(t), t)dz(t), \tag{4.69}$$

if for all $t$ from initial time 0 to current time $t$ it evolves according to the stochastic integral equation

$$x(t) = x(0) + \int_0^t \mu(x(s), s)ds + \int_0^t \sigma(x(s), s)dz(s), \tag{4.70}$$

where the second integral is interpreted as an Ito stochastic integral. The first integral term represents the accumulation of the mean changes over $(0, t)$, whilst the second (stochastic) integral term represents the accumulation of the random shock terms.

It can be shown that provided the drift coefficient $\mu(x, t)$ and the diffusion (or volatility) coefficient $\sigma(x, t)$ are reasonably smooth and do not grow too quickly, in ways that can be made mathematically precise, then the Ito stochastic differential equation (4.69) generates a stochastic process $x(t)$ that is a Markov process.

Comparing with the analysis in Sect. 4.2 we see that the stochastic differential equation (4.69) is the proper limit of the stochastic difference equation (4.5). We recall that this latter equation is associated with the Fokker–Planck equation (4.2). Thus the stochastic process described by the stochastic differential equation (4.69) can also be described by the conditional probability density function $p(x, t \mid x_0, t_0)$ of the price reaching $x$ at time $t$ given that it was at $x_0$ at time $t_0$. This density function satisfies the forward Fokker–Planck equation

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}[\mu(x, t)p(x, t \mid x_0, t_0)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma^2(x, t)p(x, t \mid x_0, t_0)], \tag{4.71}$$

or the Kolmogorov backward equation

$$\frac{\partial p}{\partial t_0} = -\mu(x_0, t_0)\frac{\partial}{\partial x_0}p(x, t \mid x_0, t_0) - \frac{1}{2}\sigma^2(x_0, t_0)\frac{\partial^2}{\partial x_0^2}p(x, t \mid x_0, t_0). \tag{4.72}$$

We emphasize that by analysing the process for $x(t)$ via the stochastic differential equation (4.69) we are focusing on the sample path evolution, whilst with the Fokker–Planck equation (4.71) and (4.72) we are focusing on the evolution of the conditional probability density function.

## 4.10 The Stratonovich Stochastic Integral

As we have already pointed out there is no unique way of defining the stochastic integral. Another definition of the stochastic integral which has also proved useful in applications is that due to Stratonovich (1963). In this definition the function

$x(t)$ is averaged over the interval $[t_{i-1}, t_i]$. The formal definition of the Stratonovich stochastic integral is

$$S\int_0^T \sigma(x(s), s)dz(s) = \underset{n\to\infty}{\text{ms-lim}}\left[\sum_{i=1}^{n}\sigma\left(\frac{x(t_{i-1}) + x(t_i)}{2}, t_{i-1}\right)[z(t_i) - z(t_{i-1})]\right]. \tag{4.73}$$

Note the $S$ in front of the integral sign to denote the Stratonovich integral. Referring again to the example when the function $\sigma$ is a Wiener process , we find by applying the definition that

$$S\int_0^T z(t)dz(t) = \underset{n\to\infty}{\text{ms-lim}}\sum_{i=1}^{n}\frac{z(t_{i-1}) + z(t_i)}{2}[z(t_i) - z(t_{i-1})] = \frac{1}{2}[z(T)^2 - z(0)^2], \tag{4.74}$$

which is different from the result for the Ito stochastic integral (see Eq. (4.68)). In fact (4.74) is the result that would be obtained by using the integration rules of ordinary calculus. This is a general feature of the Stratonovich stochastic integral.

We can use the Stratonovich definition of the stochastic integral to define the Stratonovich stochastic differential equation. Thus the *Stratonovich stochastic differential equation*

$$dx(t) = \mu(x(t), t)dt + \sigma(x(t), t) \circ dz(t),$$

(note use of the symbol $\circ$ to denote the Stratonovich stochastic differential equation) is to be interpreted as

$$x(t) = x(0) + \int_0^t \mu(x(s), s)ds + S\int_0^t \sigma(x(s), s)dz(s).$$

It can be shown by applying to its definition that (see Appendix 4.3) the *Stratonovich stochastic differential equation*

$$dx = \mu(x(t), t)dt + \sigma(x(t), t) \circ dz(t),$$

is equivalent to the *Ito stochastic differential equation*

$$dx = \left[\mu(x(t), t) + \frac{1}{2}\sigma(x(t), t)\frac{\partial\sigma}{\partial x}(x(t), t)\right]dt + \sigma(x(t), t)dz(t).$$

So situations requiring use of the Stratonovich stochastic differential equation can still be reduced to a consideration of an appropriate Ito stochastic differential equation.

## 4.11 Appendix

### Appendix 4.1 Autocovariance Functions, Spectra and White Noise

The measurements that we can perform on the output $x(t)$ of some market are rather limited. We have discussed the distributions that $x$ might follow and how these distributions will evolve. In order to make use of these distributions normally all that we have available for the time series of $x(t)$ are the mean, $\mathbb{E}[x(t)]$ and the variance, $\text{var}[x(t)]$.

The mean and the variance only give us limited information about the underlying dynamics. What we would also like to have is some quantity which is a measure of the influence of $x$ at time $t$ on the value of $x$ at time $t + \tau$. Loosely speaking such a function would measure the amount of memory that the dynamics driving the market display.

The mathematical concept which captures this memory effect is the *autocovariance function* which is defined as

$$G(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)x(t+\tau)dt.$$

An important related concept is that of the *spectrum* of the variable $x(t)$. This concept arises if we consider the fluctuations of $x(t)$ to be represented as a linear combination of sines and cosines. The spectrum picks out the frequencies of these underlying sines and cosines which are dominant.

First we take the Fourier transform of $x(t)$

$$y(\omega) = \int_0^T e^{-i\omega t} x(t)dt.$$

So if the fluctuations of $x(t)$ could be decomposed into a discrete number of frequencies, i.e.

$$x(t) = \sum_{n=1}^{N} a_n e^{i\omega_n t},$$

then since

$$\int_0^T e^{-i(\omega-\omega_n)t}\,dt = \delta(\omega-\omega_n),$$

we would have

$$y(\omega) = \sum_{n=1}^{N} a_n \delta(\omega-\omega_n),$$

which would look like Fig. 4.19a.

It is more likely that $x(t)$ would require a continuous spectrum of frequencies to adequately represents its fluctuations. In which case the graph of $y(\omega)$ might typically look like Fig. 4.19b.

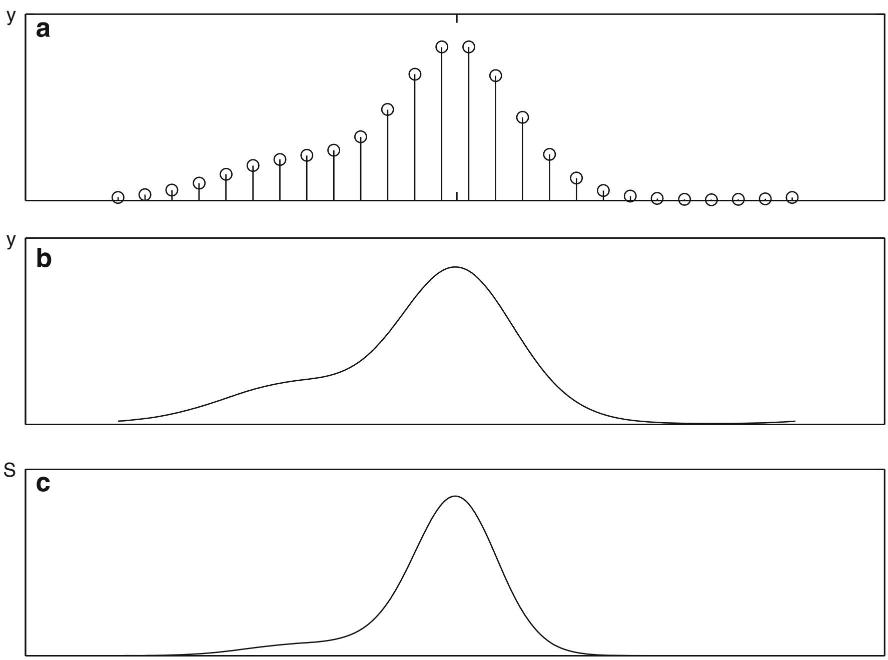


**Fig. 4.19** The Fourier transform, $y$, of the process $x$ (**a**, **b**); and the spectrum (**c**). (**a**) The Fourier transform at discrete frequencies, (**b**) the Fourier transform at continuous frequencies, (**c**) the spectrum

The *spectrum* is then defined in terms of the Fourier transform

$$S(\omega) = \lim_{T \to \infty} \frac{1}{2\pi T} \mid y(\omega) \mid^2 .$$

One reason for not staying with the function $y$ is that we want a measure of the spectrum which is independent of $T$, a parameter whose value may not be obvious. As with $y$, the spectrum $S(\omega)$ will peak around the dominant frequencies as shown in Fig. 4.19c.

Using the mathematical techniques of Fourier transforms it can be shown that the spectrum $S$ and autocovariance functions $G$ are related via

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\tau) e^{-i\omega\tau} d\tau,$$

$$G(\tau) = \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} d\omega.$$

A fluctuating function $x(t)$ which has a completely flat spectrum is known as *white noise*. The name arises from the theory of light. White light has all colours of the spectrum (i.e. all frequencies) present in equal proportions and hence will display a flat spectrum. By using the above formulae relating $G$ and $S$ we find that the autocovariance function, $G$, corresponding to a constant $S$ is a Dirac delta function. The spectrum and autocovariance function of white noise are illustrated in Fig. 4.20.

Of course, the autocovariance function for white noise (delta function) corresponds precisely with the efficient markets notion, that the $x$ are completely independent at different times.

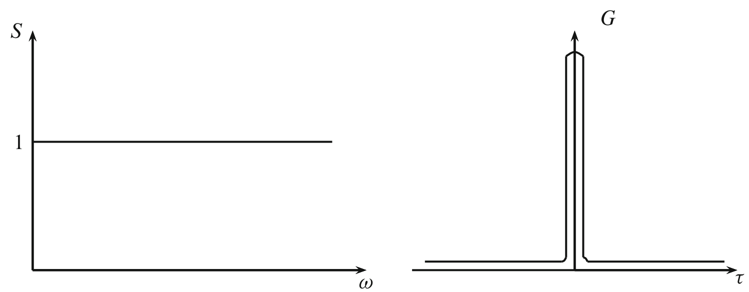


**Fig. 4.20** The spectrum ($S$) and autocovariance ($G$) of white noise

## Appendix 4.2 Evaluating the Ito Stochastic Integral $\int_0^T z(t)dz(t)$

The relevant partial sum is

$$S_n = \sum_{i=1}^{n} z_{i-1}(z_i - z_{i-1}) = \sum_{i=1}^{n} z_{i-1}\Delta z_i = \frac{1}{2}\sum_{i=1}^{n}[(z_{i-1} + \Delta z_i)^2 - z_{i-1}^2 - (\Delta z_i)^2]$$

$$= \frac{1}{2}\sum_{i=1}^{n}[z_i^2 - z_{i-1}^2 - (\Delta z_i)^2] = \frac{1}{2}[z(T)^2 - z(0)^2] - \frac{1}{2}\sum_{i=1}^{n}(\Delta z_i)^2.$$

We need to calculate the mean square limit of the last term on the right-hand side. First of all

$$\mathbb{E}\left[\sum_{i=1}^{n}(\Delta z_i)^2\right] = \sum_{i=1}^{n}\mathbb{E}[(\Delta z_i)^2] = \sum_{i=1}^{n}(t_i - t_{i-1}) = T$$

and then

$$\mathbb{E}\left[\left(\sum_{i=1}^{n}(\Delta z_i)^2 - T\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n}(\Delta z_i)^2\right)^2 - 2T\sum_{i=1}^{n}(\Delta z_i)^2 + T^2\right]$$

$$= \mathbb{E}\left[\sum_{i=1}^{n}(\Delta z_i)^4 + 2\sum_{j=1}^{n}\sum_{i>j}(\Delta z_i)^2(\Delta z_j)^2 - 2T\sum_{i=1}^{n}(\Delta z_i)^2 + T^2\right]. \tag{4.75}$$

We have already noted the property that the increments of the Wiener process are independent. Hence for $i \neq j$

$$\mathbb{E}[(\Delta z_i)^2(\Delta z_j)^2] = \mathbb{E}[(\Delta z_i)^2]\cdot\mathbb{E}[(\Delta z_j)^2] = (t_i - t_{i-1})(t_j - t_{j-1}).$$

Using the fact that the Wiener process is Gaussian and a property of the fourth moment of the Gaussian distribution we have that[7]

$$\mathbb{E}\{(\Delta z_i)^4\} = 3\{\mathbb{E}[(\Delta z_i)^2]\}^2 = 3(t_i - t_{i-1})^2.$$

[7] Note that, for $x \sim N(\mu, \sigma^2)$, $\mathbb{E}[x^n] = \mu\mathbb{E}[x^{n-1}] + (n-1)\sigma^2\mathbb{E}[x^{n-2}]$.

Assembling all of the above results we find that the right hand side of (4.75) becomes

$$3\sum_{i=1}^{n}(t_i - t_{i-1})^2 + \sum_{i,j=1,i\neq j}^{n}(t_i - t_{i-1})(t_j - t_{j-1}) - 2T^2 + T^2$$

$$= 2\sum_{i=1}^{n}(t_i - t_{i-1})^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}(t_i - t_{i-1})(t_j - t_{j-1}) - T^2$$

$$= 2\sum_{i=1}^{n}(t_i - t_{i-1})^2.$$

Hence we have shown that

$$\mathbb{E}\left[\left(\sum_{i=1}^{n}(\Delta z_i)^2 - T\right)^2\right] = 2\sum_{i=1}^{n}(t_i - t_{i-1})^2,$$

This last term on the right-hand side tends to 0 as $n \to \infty$. That is

$$\underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n}(\Delta z_i)^2 = T,$$

and so

$$\int_0^T z(t)dz(t) = \frac{1}{2}[z(T)^2 - z(0)^2 - T].$$

## Appendix 4.3 Link Between Ito and Stratonovich

Assume $x(t)$ follows the Ito stochastic differential equation

$$dx = \mu(x,t)dt + \sigma(x,t)dz(t). \tag{4.76}$$

For the Stratonovich integral,

$$\int_{t_0}^{t} \sigma(x,s)odz(s) \simeq \sum_{i=1}^{n}\sigma\left(\frac{x(t_i) + z(t_{i-1})}{2}, t_{i-1}\right)(z(t_i) - z(t_{i-1})). \tag{4.77}$$

In (4.77) set $x(t_i) = x(t_{i-1}) + dx(t_{i-1})$ so that the Ito stochastic differential equation can be written as

$$dx(t_i) = \mu(x(t_{i-1}), t_{i-1})(t_i - t_{i-1}) + \sigma(x(t_{i-1}), t_{i-1})(z(t_i) - z(t_{i-1})).$$

We can apply Taylor's expansion to

$$\begin{aligned}\sigma\left(\frac{x(t_i) + x(t_{i-1})}{2}, t_{i-1}\right) &= \sigma\left(x(t_{i-1}) + \frac{1}{2}dx(t_{i-1}), t_{i-1}\right) \\ &= \sigma(t_{i-1}) + \frac{1}{2}\frac{\partial \sigma}{\partial x}(t_{i-1})dx(t_{i-1}) \\ &= \sigma(t_{i-1}) + \frac{1}{2}\frac{\partial \sigma}{\partial x}(t_{i-1})\Big[\mu(t_{i-1})dt_{i-1} + \sigma(t_{i-1})dz(t_{i-1})\Big],\end{aligned} \tag{4.78}$$

where to alleviate the notation we write $dx(t_{i-1})$, $\sigma(t_{i-1})$ and $\mu(t_{i-1})$ to denote $x(t_i) - x(t_{i-1})$, $\sigma(x(t_{i-1}), t_{i-1})$ and $\mu(x(t_{i-1}), t_{i-1})$ respectively. Substituting (4.78) into (4.77), setting $dz^2 = dt$ and ignoring the $(dt)^2$ and $dtdz$ terms we obtain

$$S\int \sigma(x,s)dz(s) \simeq \sum_i^1 \sigma(t_{i-1})(z(t_i) - z(t_{i-1})) + \frac{1}{2}\sum_i^1 \sigma(t_{i-1})\frac{\partial \sigma}{\partial x}(t_{i-1})(t_i - t_{i-1})$$

which yields

$$S\int_{t_0}^t \sigma(x,s)dz(s) \simeq \frac{1}{2}\int_0^t \sigma(x,s)\frac{\partial \sigma}{\partial x}(x,s)ds + \int_{t_0}^t \sigma(x,s)dz(s).$$

Thus we obtain the result that the Stratonovich stochastic differential equation

$$dx = \mu(x,t)dt + \sigma(x,t) \circ dz$$

is equivalent to the Ito stochastic differential equation

$$dx = \left(\mu(x,t) + \frac{1}{2}\sigma(x,t)\frac{\partial \sigma}{\partial x}(x,t)\right)dt + \sigma(x,t)dz(t).$$

## 4.12 Problems

**Problem 4.1** Consider the Ornstein–Uhlenbeck process described in Sect. 4.3.2. Obtain the solution (4.21).

**Problem 4.2** Derive the autocorrelations functions (4.25) and (4.32) for the Wiener process and Ornstein–Uhlenbeck process respectively.

**Problem 4.3** For the Wiener process $z(t)$ show that

$$corr(z(t), z(t+s)) = \sqrt{\frac{t}{t+s}}, \quad t \geq 0, s \geq 0.$$

**Problem 4.4** Adapt the approach of Appendix 4.2 to show that

$$\int_0^T z(t)^2 dz(t) = \frac{1}{3}[z(t)^3 - z(0)^3] - \int_0^T z(s)ds.$$

**Problem 4.5** Using the techniques of Appendix 4.2 and Problem 4.4 show that

$$\int_0^T \left(\int_0^{s_2} dz(s_1)\right) dz(s_2) = \frac{1}{2}\left((z(T) - z(0))^2 - T\right).$$

(See Oksendal (2003, Problem 3.7) for a more general version of this problem.)

**Problem 4.6 Computational Problem**—Consider the process $x(t)$, driven in discrete time by

$$\Delta x(t) = x(t + \Delta t) - x(t) = \mu(x,t)\Delta t + \sigma(x,t)\Delta z(t)$$

where $z(t)$ is a Wiener process. This is Eq. (4.49) before passing to the limit. Here $0 \leq t \leq T$ and $\Delta t = T/n$ where $n$ is user specified.

Write a program to simulate this process on $[0, T]$. The program structure should allow the user to specify $T$, $n$, the functions $\mu(x,t)$ and $\sigma(x,t)$. As well the number of simulations $M$ should be an input variable.

(a) Take $x(0) = 1$, $\mu(x,t) = 0$, $\sigma(x,t) = 1$ and $T = 1$ so that $x(t)$ is a pure Wiener process on $[0, 1]$. Initially take $n = 100$ and simulate $M = 1{,}000$ paths.

   (i) Use the output to compare graphically the distribution of $x(t)$ at $t = 0.5$ and $t = 1$.
   (ii) Experiment with $n$ and $M$ to try to obtain better approximations to the known distributions.
   (iii) Constructing a table to compare these with the known theoretical distributions for $x(t)$ and comment on the effect of $n$ and $M$.

(b) Take $x(0) = 1$, $\mu(x,t) = \mu x$, $\sigma(x,t) = \sigma x$ and $T = 2$ so that $x(t)$ is the geometric Brownian motion for the stock price.

   (i) Use $\mu = 0.15$ and $\sigma = 0.20$ and take $n = 100$ and simulate $M = 1{,}000$ paths to obtain better approximations.

(ii) Use the output to graph the distributions of $x(t)$ and $ln(x(t)/x(0))$ at $t = 2$.

(iii) Constructing a table to compare these with the known true distributions for $x(T)$ and $ln(x(T)/x(0))$. As in the previous question play with the values of $n$ and $M$ and comment on the effect of $n$ and $M$.

(c) Repeat the exercise by taking $x(0) = 0.06$,

$$\mu(x,t) = k(\overline{x} - x)$$

with $k = 0.5$, $\overline{x} = 0.065$ and $\sigma(x,t) = \sigma$ with $\sigma = 0.02$, $n = 100$ and $M = 1{,}000$. Calculate and graph the distribution of $x$ at $T = 6$ months and $T = 12$ months, and compare these with the theoretical distributions. Comment on the effect of $n$ and $M$.

# Chapter 5
# Manipulating Stochastic Differential Equations and Stochastic Integrals

**Abstract** Many of the calculations of derivative security pricing involve formal manipulations of stochastic differential equations and stochastic integrals. This chapter derives those that are most frequently used. We also consider transformation of correlated Wiener processes to uncorrelated Wiener processes for higher dimensional stochastic differential equations.

## 5.1 The Basic Rules of Stochastic Calculus

In our application of stochastic differential equations to the continuous hedging argument in later chapters we shall frequently encounter the increment of the Wiener process, $dz(t)$, raised to various powers, i.e. terms of the type $(dz(t))^2$, $(dz(t))^3$ etc. as well as terms of the type $(dt)dz(t)$, $(dt)(dz(t))^2$ etc. In general we need to interpret $(dz(t))^{2+\alpha}$ for $\alpha \geq 0$ and $(dt)^{\beta} dz(t)^{2+\alpha}$ for $\alpha \geq 0, \beta \geq 1$ and we take $\alpha$ and $\beta$ as integers. Interpreting and computing these quantities is the key to the formal mechanical manipulations of the stochastic calculus.

It is useful to think of these quantities on two levels. One is a mathematically informal level which is useful to have in mind when one is doing the manipulations of stochastic calculus. The other is the more mathematically formal level where all manipulations are justified in terms of the convergence of appropriate partial sums to stochastic integrals. We present the informal view first and follow with a discussion of the more mathematically correct approach. We note that in all of our manipulations in stochastic calculus we ignore terms of order higher than $dt$.

Consider first the random quantity $(dz)^2$, which may also be written

$$(dz)^2 = (dt)\xi^2, \tag{5.1}$$

where $\xi \sim N(0, 1)$. We note first that

$$\mathbb{E}[(dz)^2] = (dt)\mathbb{E}[\xi^2] = dt, \tag{5.2}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_5

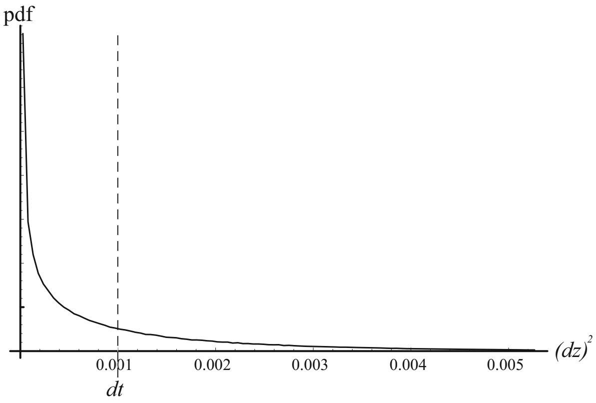


**Fig. 5.1** The probability mass of $(dz)^2$ concentrated at $dt$

and

$$\begin{aligned}\text{var}[(dz)^2] &= \mathbb{E}[(dt\xi^2 - dt)^2] = (dt)^2\mathbb{E}[(\xi^2 - 1)^2] \\ &= (dt)^2[\mathbb{E}(\xi^4) - 2\mathbb{E}(\xi^2) + 1] = 2(dt)^2.\end{aligned} \tag{5.3}$$

Thus we can view $(dz)^2$ as a random variable distributed with mean $dt$ and variance $2(dt)^2$; see Fig. 5.1.[1] However as we are ignoring terms of order higher than $dt$, we can regard this variance as zero. So the distribution of $(dz)^2$ can be considered as a $\delta$ function with all probability mass concentrated at $dt$. Thus to order $dt$ we can regard $(dz)^2$ as a deterministic variable equal to $dt$, and we write

$$(dz(t))^2 = dt. \tag{5.4}$$

Consider next $(dt)dz(t)$ which can be written

$$(dt)dz(t) = (dt)^{3/2}\xi, \tag{5.5}$$

where $\xi \sim N(0, 1)$. We calculate

$$\mathbb{E}[(dt)dz(t)] = (dt)^{3/2}\mathbb{E}[\xi] = 0, \tag{5.6}$$

$$\text{var}[(dt)dz(t)] = \mathbb{E}[(dt)^3\xi^2] = (dt)^3\mathbb{E}[\xi^2] = (dt)^3. \tag{5.7}$$

[1] In fact since the square of a normal random variable is chi-squared distributed, $(dz)^2 \sim (dt)^2\chi_1^2$ where $\chi_1^2$ represents the chi-squared distribution with one degree of freedom.

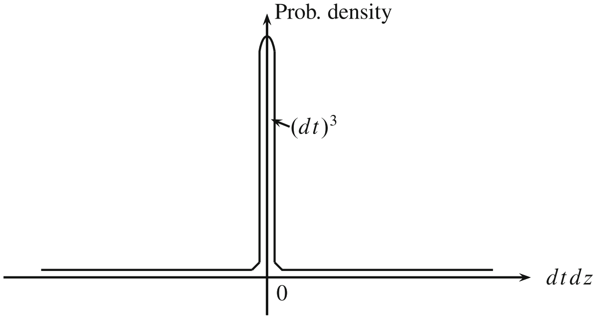


**Fig. 5.2** The probability mass of *dtdz* concentrated at 0

Thus $(dt)dz(t)$ can be regarded as a stochastic variable distributed with mean 0 and variance $(dt)^3$; see Fig. 5.2. Again since we ignore terms of order higher than $dt$ we can regard this distribution as a $\delta$ function with all probability mass concentrated at 0. Thus to order $dt$ we can regard $dtdz$ as a deterministic variable equal to 0. We leave as an exercise for the reader to analyse in a similar fashion terms such as $(dz)^3$, $(dt)(dz)^2$ etc.

We saw in the last chapter that stochastic differential equations are interpreted in terms of stochastic integrals. Thus expressions of the type $(dz(t))^{2+\alpha}$, where $\alpha$ is a positive integer, would occur in connection with the stochastic integral

$$\int_o^t \sigma(x(s),s)(dz(s))^{2+\alpha} \equiv \underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \sigma(x(t_{i-1}),t_{i-1})(\Delta z_i)^{2+\alpha}. \tag{5.8}$$

In Appendix 5.1 we indicate how it may be proved that

$$\int_0^t \sigma(x(s),s)(dz(s))^{2+\alpha} = \begin{cases} \int_0^t \sigma(x(s),s)ds, & \text{for } \alpha = 0 \\ 0, & \text{for } \alpha > 0. \end{cases} \tag{5.9}$$

It may similarly be proved that, for $\beta$ also a positive integer,

$$\int_0^t \sigma(x(s),s)(ds)^{\beta}(dz(s))^{\alpha} = 0 \qquad \text{for } \alpha > 0, \beta > 0. \tag{5.10}$$

The foregoing results are usually expressed in the shorthand notation

$$(dz(t))^2 = dt, \tag{5.11}$$

$$(dz(t))^{2+\alpha} = 0, \qquad \text{for} \quad \alpha > 0, \tag{5.12}$$

$$(dt)^{\beta}(dz(t))^{\alpha} = 0, \qquad \text{for} \quad \alpha > 0, \beta > 0. \tag{5.13}$$

The results (5.11)–(5.13) are the basic rules of stochastic calculus. We stress again that it needs to be appreciated that the quantities on the left-hand side are random variables and that these formulae apply when they appear under integral signs or when we take their expectations (which are also integral operations).

## 5.2 Some Basic Stochastic Integrals

Consider the stochastic integral term in Eq. (4.70), namely

$$Y(t) \equiv \int_0^t \sigma(x(s), s, t)dz(s), \tag{5.14}$$

which we have generalised slightly to allow a dependence of $\sigma$ on $t$ as well. Recall that by definition

$$Y(t) = \text{ms-}\lim_{n\to\infty} \sum_{i=1}^{n} \sigma(x_{i-1}, s_{i-1}, t)\Delta z_i, \tag{5.15}$$

where we use $x_i$ to denote $x(s_i)$ and $\Delta z_i$ to denote $z(s_i) - z(s_{i-1})$. It is often the case that we need to calculate the mean and the variance of stochastic integrals of the type (5.14), i.e. the quantities

$$M(t) = \mathbb{E}_0[Y(t)] = \mathbb{E}_0\left[\int_0^t \sigma(x(s), s, t)dz(s)\right], \tag{5.16}$$

and

$$V^2(t) = \text{var}_0[Y(t)] = \mathbb{E}_0\left[\left(\int_0^t \sigma(x(s), s, t)dz(s) - M(t)\right)^2\right]. \tag{5.17}$$

Consider first that calculation of $M(t)$. Assuming we can interchange the operations ms-$\lim_{n\to\infty}$ and $\mathbb{E}_0$ we have[2]

$$\begin{aligned} M(t) &= \lim_{n\to\infty} \mathbb{E}_0 \sum_{i=1}^{n} \sigma(x_{i-1}, s_{i-1}, t)\Delta z_i \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}_0(\mathbb{E}_{i-1}\sigma(x_{i-1}, s_{i-1}, t)(\Delta z_i)), \end{aligned} \tag{5.18}$$

[2]The operation ms-$\lim_{n\to\infty}$ becomes simply $\lim_{n\to\infty}$ since the quantities on the right hand side in (5.18) will be deterministic after application of the expectation operation.

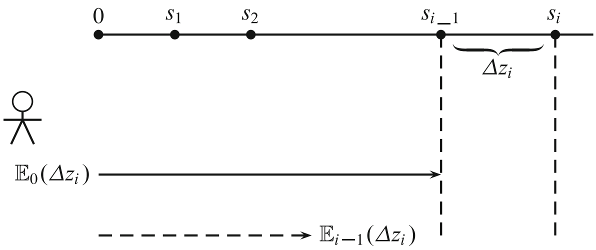


**Fig. 5.3** The investor's perception of $\mathbb{E}_0(\Delta z_i)$ and the law of iterated expectations

where we have used the law of iterated expectations. See Fig. 5.3 for an illustration of this concept. Since at time $s_{i-1}$ the only stochastic quantity relevant to the calculation of $\mathbb{E}_{i-1}$ in (5.18) is $\Delta z_i$ and $\mathbb{E}_{i-1}(\Delta z_i) = 0$, we conclude that

$$M(t) = \mathbb{E}_0\left[\int_0^t \sigma(x(s), s, t)dz(s)\right] = 0. \tag{5.19}$$

Thus from (5.17)

$$\begin{aligned} V^2(t) &= \mathbb{E}_0\left[\left(\int_0^t \sigma(x(s), s, t)dz(s)\right)^2\right] \\ &= \mathbb{E}_0\left[\left(\text{ms-}\lim_{n\to\infty}\sum_{i=1}^n \sigma(x_{i-1}, s_{i-1}, t)\Delta z_i\right)^2\right]. \end{aligned} \tag{5.20}$$

Using some informal mathematical reasoning, which should be seen as an intuitive approach to understanding the calculation of $V(t)$ we write (5.20) as

$$\begin{aligned} V^2(t) &= \lim_{n\to\infty} \mathbb{E}_0\left(\sum_{i=1}^n \sigma(x_{i-1}, s_{i-1}, t)\Delta z_i\right)^2 \\ &= \lim_{n\to\infty} \mathbb{E}_0\left(\sum_{i=1}^n\sum_{j=1}^n \sigma(x_{i-1}, s_{i-1}, t)\sigma(x_{j-1}, s_{j-1}, t)\Delta z_i \Delta z_j\right) \\ &= \lim_{n\to\infty}\left(\sum_{i=1}^n \mathbb{E}_0[\sigma(x_{i-1}, s_{i-1}, t)\Delta z_i]^2\right) \\ &\quad +2\lim_{n\to\infty}\left(\sum_{i,j=1, i<j}^n \mathbb{E}_0[\sigma(x_{i-1}, s_{i-1}, t)\sigma(x_{j-1}, s_{j-1}, t)\Delta z_i \Delta z_j]\right). \end{aligned} \tag{5.21}$$

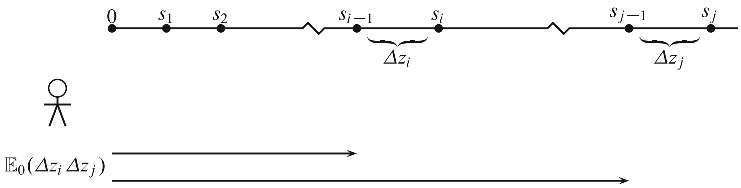


**Fig. 5.4** The time-line for the calculation of $\mathbb{E}_0(\Delta z_i \Delta z_j)$; note that $\Delta z_i$, $\Delta z_j$ are independent for $i \neq j$

Figure 5.4 illustrates the time-line involved in the calculation of the expectations in (5.21). By the independence of Wiener increments we have that

$$\mathbb{E}_0(\Delta z_i \Delta z_j) = \begin{cases} 0, & i \neq j, \\ \Delta s_i = s_i - s_{i-1}, & i = j. \end{cases}$$

Hence, for $i < j$,

$$\begin{aligned} &\mathbb{E}_0[\sigma(x_{i-1}, s_{i-1}, t)\sigma(x_{j-1}, s_{j-1}, t)\Delta z_i \Delta z_j] \\ &\quad = \mathbb{E}_0\Big[\sigma(x_{i-1}, s_{i-1}, t)(\Delta z_i)\sigma(x_{j-1}, s_{j-1}, t)\mathbb{E}_{j-1}(\Delta z_j)\Big] = 0. \end{aligned}$$

Also

$$\begin{aligned} \mathbb{E}_0[\sigma(x_{i-1}, s_{i-1}, t)\Delta z_i]^2 &= \mathbb{E}_0\Big[\sigma^2(x_{i-1}, s_{i-1}, t)\mathbb{E}_{i-1}(\Delta z_i)^2\Big] \\ &= \mathbb{E}_0[\sigma^2(x_{i-1}, s_{i-1}, t)]\Delta s_i\,. \end{aligned}$$

Therefore

$$V^2(t) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}_0[\sigma^2(x_{i-1}, s_{i-1}, t)]\Delta s_i\,. \tag{5.22}$$

The limit in (5.22) is just the Riemann integral so we finally have

$$V^2(t) = \int_0^t \mathbb{E}_0[\sigma^2(x(s), s, t)]ds.$$

We have thus demonstrated somewhat informally the important results

$$\mathbb{E}_0\left(\int_0^t \sigma(x(s), s, t)dz(s)\right) = 0, \tag{5.23}$$

and

$$\mathbb{E}_0 \left( \int_0^t \sigma(x(s), s, t) dz(s) \right)^2 = \int_0^t \mathbb{E}_0[\sigma^2(x(s), s, t)] ds. \tag{5.24}$$

The interchange of limit operations in going from (5.20) to (5.21) is more subtle than we have indicated here. However in Chap. 8 we shall give a more formal derivation of the result (5.24) using martingale ideas.

Finally we point out that in the special case when $\sigma$ is a function of $s$ and $t$ only, that is a deterministic function, so that

$$Y(t) = \int_0^t \sigma(s, t) dz(s), \tag{5.25}$$

then the partial sums in (5.15) form a weighted sum of quantities which are normally distributed. Standard results from statistics tell us that in this case the partial sums must also be normally distributed. Thus we argue that in the mean-square limit the quantity $Y(t)$ is normally distributed (though this involves some subtle limit arguments) with mean 0 and variance given by

$$\int_0^t \sigma^2(s, t) ds. \tag{5.26}$$

## 5.3 Higher Dimensional Stochastic Differential Equations

In later chapters we shall encounter situations in which both the stochastic variable $x(t)$ is a vector (i.e. think of several price processes) and the Wiener process $z(t)$ also becomes a vector (e.g. think of several sources of risk such as stochastic volatility, inflation, etc.). In the most general situation we could have $n$ stochastic variables, i.e. $\mathbf{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^\top$ and $m$ Wiener processes, i.e. $\mathbf{z}(t) = (z_1(t), z_2(t), \ldots, z_m(t))^\top$. The Wiener processes may be correlated, thus for $i, j = 1, \cdots, m$,

$$\mathbb{E}[dz_i(t)] = 0, \quad \mathbb{E}[(dz_i(t))^2] = dt, \quad \mathbb{E}[dz_i(t) dz_j(t)] = \rho_{ij} dt \quad (i \neq j). \tag{5.27}$$

We consider the system of $n$ stochastic differential equations

$$dx_i(t) = \mu_i(\mathbf{x}(t), t) dt + \sum_{j=1}^{m} \sigma_{ij}(\mathbf{x}(t), t) dz_j(t), \qquad \text{for} \qquad i = 1, 2, \ldots, n. \tag{5.28}$$

Each stochastic differential equation needs to be interpreted mathematically as the stochastic integral equation

$$x_i(t) = x_i(0) + \int_0^t \mu_i(\mathbf{x}(s), s)ds + \sum_{j=1}^{m} \int_0^t \sigma_{ij}(\mathbf{x}(s), s)dz_j(s) \tag{5.29}$$

for $i = 1, 2, \ldots, n$, where the integrals under the $\sum$ sign are interpreted as Ito stochastic integrals. The definition of these integrals in terms of mean-square limits as in Sect. 4.7 needs to take account of the correlation between the Wiener processes.

In some applications it turns out to be more convenient to convert the vector of correlated Wiener processes to a vector of uncorrelated Wiener processes. Let $(w_1(t), w_2(t), \ldots, w_m(t))^\top$ denote a vector of uncorrelated Wiener processes. We seek a transformation

$$\begin{bmatrix} dz_1(t) \\ dz_2(t) \\ \vdots \\ dz_m(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{bmatrix} \begin{bmatrix} dw_1(t) \\ dw_2(t) \\ \vdots \\ dw_m(t) \end{bmatrix}, \tag{5.30}$$

where the $(a_{ij})_{m\times m}$ are to be chosen so as to preserve the correlation structure of the $z_i(t)$ as given by Eq. (5.27).

Consider each of the conditions of Eq. (5.27) in turn. Since the $w_i(t)$ for $i = 1, 2, \cdots, m$ are Wiener processes the condition

$$\mathbb{E}[dz_i(t)] = 0$$

is satisfied for any choice of the $a_{ij}$. Next we note that the conditions

$$\mathbb{E}[(dz_i)^2] = dt$$

imposes the $m$ conditions

$$\sum_{j=1}^{m} a_{ij}^2 = 1, \qquad \text{for } i = 1, 2, \ldots, m. \tag{5.31}$$

Finally the condition

$$\mathbb{E}[dz_i(t)dz_j(t)] = \rho_{ij}dt \qquad i \neq j, i, j = 1, \cdots, m$$

imposes the $\frac{(m-1)m}{2}$ conditions

$$\sum_{k=1}^{m} a_{ik}a_{jk} = \rho_{ij}, \qquad \text{for } i = 1, 2, \ldots, m-1, \quad j = 1, 2, \ldots, m. \tag{5.32}$$

In all, we have a set of $m + \frac{m(m-1)}{2}$ conditions from which to determine the $m^2$ quantities $a_{ij}$. Clearly these cannot be chosen uniquely. Normalising conditions will help us to determine an appropriate set of $a_{ij}$. We give the explicit results for the $m = 2$ and $m = 3$ cases which will cover most of the applications we shall encounter in later chapters.

Note that if we use $A$ to denote the matrix in (5.30) whose $ij$th element is $a_{ij}$, then the conditions (5.31) and (5.32) can be succinctly written as

$$AA^T = \boldsymbol{\rho} = \begin{bmatrix} 1 & \rho_{12} & \ldots & \rho_{1n} \\ \rho_{12} & 1 & \ldots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{1n} & \rho_{2n} & \ldots & 1 \end{bmatrix}. \tag{5.33}$$

### *5.3.1 The Two-Noise Case*

For $m = 2$, the conditions (5.31) and (5.32) reduce to

$$a_{11}^2 + a_{12}^2 = 1, \tag{5.34}$$

$$a_{21}^2 + a_{22}^2 = 1, \tag{5.35}$$

$$a_{11}a_{21} + a_{22}a_{12} = \rho_{12}. \tag{5.36}$$

One possible characterisation of the $a_{ij}$ is to set

$$a_{11} = a_{12} = \frac{1}{\sqrt{2}}. \tag{5.37}$$

From Eqs. (5.35) and (5.36) we would then obtain

$$a_{21} = \frac{\rho_{12} - \sqrt{1-\rho_{12}^2}}{\sqrt{2}}, \qquad a_{22} = \frac{\rho_{12} + \sqrt{1-\rho_{12}^2}}{\sqrt{2}}. \tag{5.38}$$

Thus the transformation becomes

$$\begin{aligned} dz_1(t) &= \frac{1}{\sqrt{2}}\, dw_1(t) + \frac{1}{\sqrt{2}}\, dw_2(t), \\ dz_2(t) &= \frac{\rho_{12} - \sqrt{1-\rho_{12}^2}}{\sqrt{2}}\, dw_1(t) + \frac{\rho_{12} + \sqrt{1-\rho_{12}^2}}{\sqrt{2}}\, dw_2(t). \end{aligned} \tag{5.39}$$

Alternatively, we could set

$$a_{11} = 1, \qquad a_{12} = 0, \tag{5.40}$$

in which case we would obtain

$$a_{21} = \rho_{12}, \qquad a_{22} = \sqrt{1-\rho_{12}^2}. \tag{5.41}$$

With this latter specification we would retain the first of the original noise processes. Thus

$$\begin{aligned} dz_1(t) &= dw_1(t), \\ dz_2(t) &= \rho_{12}\, dw_1(t) + \sqrt{1-\rho_{12}^2}\, dw_2(t). \end{aligned} \tag{5.42}$$

To see how the transformations (5.42) applies to specific models we consider a one asset model and a two asset model both with two noise terms. We leave as exercise for the reader the task of applying the transformation (5.39).

#### 5.3.1.1 A One Asset-Two Noise Term Model

Consider the stochastic differential equation

$$dx = \mu(x,t)dt + \sigma_1(x,t)dz_1(t) + \sigma_2(x,t)dz_2(t) \tag{5.43}$$

where, as above, $\mathbb{E}[dz_1(t)dz_2(t)] = \rho_{12}dt$. After using the transformation (5.42) to express the stochastic dynamics for $x(t)$ in terms of the independent Wiener processes $w_1(t)$, $w_2(t)$ we obtain the stochastic differential equation

$$dx = \mu(x,t)dt + s_1(x,t)dw_1(t) + s_2(x,t)dw_2(t), \tag{5.44}$$

where

$$s_1(x,t) = \sigma_1(x,t) + \rho_{12}\sigma_2(x,t), \tag{5.45}$$

$$s_2(x,t) = \sqrt{1-\rho_{12}^2}\,\sigma_2(x,t). \tag{5.46}$$

#### 5.3.1.2 A Two Asset-Two Noise Term Model

Consider the two asset-two noise term stochastic differential equation system which we express in matrix notation

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(\mathbf{x},t) \\ \mu_2(\mathbf{x},t) \end{bmatrix} dt + \begin{bmatrix} \sigma_{11}(\mathbf{x},t) & \sigma_{12}(\mathbf{x},t) \\ \sigma_{21}(\mathbf{x},t) & \sigma_{22}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} dz_1(t) \\ dz_2(t) \end{bmatrix}. \tag{5.47}$$

Consider the second term on the right-hand side and apply (in matrix notation) the transformation (5.42). Thus

$$\begin{bmatrix} \sigma_{11}(\mathbf{x},t) & \sigma_{12}(\mathbf{x},t) \\ \sigma_{21}(\mathbf{x},t) & \sigma_{22}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} dz_1(t) \\ dz_2(t) \end{bmatrix} = \begin{bmatrix} \sigma_{11}(\mathbf{x},t) & \sigma_{12}(\mathbf{x},t) \\ \sigma_{21}(\mathbf{x},t) & \sigma_{22}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} \end{bmatrix} \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}.$$

Thus in terms of the independent Wiener processes $w_1(t)$, $w_2(t)$ the stochastic differential equation system (5.47) becomes

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(\mathbf{x},t) \\ \mu_2(\mathbf{x},t) \end{bmatrix} dt + \begin{bmatrix} s_{11}(\mathbf{x},t) & s_{12}(\mathbf{x},t) \\ s_{21}(\mathbf{x},t) & s_{22}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix}, \tag{5.48}$$

where

$$s_{11}(\mathbf{x},t) = \sigma_{11}(\mathbf{x},t) + \rho_{12}\sigma_{12}(\mathbf{x},t), \tag{5.49}$$

$$s_{12}(\mathbf{x},t) = \sqrt{1-\rho_{12}^2}\,\sigma_{12}(\mathbf{x},t), \tag{5.50}$$

$$s_{21}(\mathbf{x},t) = \sigma_{21}(\mathbf{x},t) + \rho_{12}\sigma_{22}(\mathbf{x},t), \tag{5.51}$$

$$s_{22}(\mathbf{x},t) = \sqrt{1-\rho_{12}^2}\,\sigma_{22}(\mathbf{x},t). \tag{5.52}$$

### 5.3.2 The Three-Noise Case

For $m = 3$ the conditions (5.31) and (5.32) reduce to

$$\begin{aligned} a_{11}^2 + a_{12}^2 + a_{13}^2 &= 1 \\ a_{21}^2 + a_{22}^2 + a_{23}^2 &= 1 \\ a_{31}^2 + a_{32}^2 + a_{33}^2 &= 1 \\ a_{11}a_{21} + a_{12}a_{22} + a_{13}a_{23} &= \rho_{12} \\ a_{11}a_{31} + a_{12}a_{32} + a_{13}a_{33} &= \rho_{13} \\ a_{21}a_{31} + a_{22}a_{32} + a_{23}a_{33} &= \rho_{23}. \end{aligned} \tag{5.53}$$

We have six conditions for nine unknown coefficients. A convenient choice for the remaining three conditions is

$$a_{12} = a_{13} = a_{23} = 0. \tag{5.54}$$

With this choice, system (5.53) reduces to

$$\begin{aligned} a_{11}^2 &= 1 \\ a_{21}^2 + a_{22}^2 &= 1 \\ a_{31}^2 + a_{32}^2 + a_{33}^2 &= 1 \\ a_{11}a_{21} &= \rho_{12} \\ a_{11}a_{31} &= \rho_{13} \\ a_{21}a_{31} + a_{22}a_{32} &= \rho_{23} \end{aligned} \tag{5.55}$$

which is readily solved to yield the transformation

$$\begin{bmatrix} dz_1(t) \\ dz_2(t) \\ dz_3(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ \rho_{13} & \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} & \sqrt{\frac{1-\rho_{13}^2-\rho_{12}^2-\rho_{23}^2+2\rho_{12}\rho_{23}\rho_{13}}{1-\rho_{12}^2}} \end{bmatrix} \begin{bmatrix} dw_1(t) \\ dw_2(t) \\ dw_3(t) \end{bmatrix}. \tag{5.56}$$

Consider for example a two underlying asset and three noise term model

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(\mathbf{x},t) \\ \mu_2(\mathbf{x},t) \end{bmatrix} dt + \begin{bmatrix} \sigma_{11}(\mathbf{x},t) & \sigma_{12}(\mathbf{x},t) & \sigma_{13}(\mathbf{x},t) \\ \sigma_{21}(\mathbf{x},t) & \sigma_{22}(\mathbf{x},t) & \sigma_{23}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} dz_1(t) \\ dz_2(t) \\ dz_3(t) \end{bmatrix}. \tag{5.57}$$

Using (5.56) this transforms to

$$\begin{bmatrix} dx_1(t) \\ dx_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1(\mathbf{x},t) \\ \mu_2(\mathbf{x},t) \end{bmatrix} dt + \begin{bmatrix} s_{11}(\mathbf{x},t) & s_{12}(\mathbf{x},t) & s_{13}(\mathbf{x},t) \\ s_{21}(\mathbf{x},t) & s_{22}(\mathbf{x},t) & s_{23}(\mathbf{x},t) \end{bmatrix} \begin{bmatrix} dw_1(t) \\ dw_2(t) \\ dw_3(t) \end{bmatrix}, \tag{5.58}$$

where the elements $s_{ij}$ may be calculated from the matrix relationship

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2} & 0 \\ \rho_{13} & \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{1-\rho_{12}^2}} & \sqrt{\frac{1-\rho_{13}^2-\rho_{12}^2-\rho_{23}^2+2\rho_{12}\rho_{23}\rho_{13}}{1-\rho_{12}^2}} \end{bmatrix}. \tag{5.59}$$

## 5.4 The Kolmogorov Equation for an n-Dimensional Diffusion System

In later chapters we shall need to deal with multi-dimensional diffusion processes. In particular, we will require the Kolmogorov backward equation for such processes. In this section we merely summarise the relevant results from Oksendal (2003) (see Chaps. 7 and 8).

Consider the n-dimensional Ito stochastic differential system

$$dx_i = \mu_i dt + \sum_{j=1}^{n} \sigma_{ij} dw_j(t) \quad (i = 1, 2, \ldots, n),$$

where $w_1(t)$, $w_2(t)$,..., $w_n(t)$ are independent Wiener processes. Let $\sigma$ denote the matrix whose elements are the $\sigma_{ij}$ and define the matrix $S$ as

$$S = (s_{ij})_{n\times n} = \sigma\sigma^\top. \tag{5.60}$$

The infinitesimal generator $\mathscr{K}$ for the process $\mathbf{x}$ is given by

$$\mathscr{K} = \sum_{i=1}^{n} \mu_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \tag{5.61}$$

so that the Kolmogorov backward equation for the transition probability density function $p(\mathbf{x}(T), T \mid \mathbf{x}(t), t)$ is given by

$$\frac{\partial p}{\partial t} + \mathscr{K} p = 0. \tag{5.62}$$

Note that if we are dealing with correlated Wiener processes then a transformation discussed in Sect. 5.3 allows us to transform them to a set of independent Wiener processes and the correlation coefficients will appear in the $\sigma_{ij}$ coefficients.

## 5.5 The Differential of a Stochastic Integral

The result of this section turns out to be very useful in manipulations required when we come to study term structure of interest rate models.

Consider the stochastic quantity

$$x(t) = \int_0^t g(s,t)dz(s), \tag{5.63}$$

where $g(s,t)$ is a function that is at least once differentiable in both arguments. What is the stochastic differential equation satisfied by $x(t)$? To answer this question we proceed directly and form

$$\begin{aligned} dx(t) &= \int_0^{t+dt} g(s,t+dt)dz(s) - \int_0^t g(s,t)dz(s) \\ &= \int_0^t [g(s,t+dt) - g(s,t)]dz(s) + \int_t^{t+dt} g(s,t+dt)dz(s) \\ &= \int_0^t \left[\frac{\partial g(s,t)}{\partial t}dt\right]dz(s) + \int_t^{t+dt}\left[g(s,t) + \frac{\partial g(s,t)}{\partial t}dt\right]dz(s) + o(dt), \end{aligned} \tag{5.64}$$

where we have applied Taylor's expansion and used the basic rules of stochastic calculus in obtaining (5.64).

By applying to the definition of the Ito stochastic integral (in particular evaluation of the integral at the left hand limit of each subinterval) we see that

$$\int_t^{t+dt} g(s,t)dz(s) \simeq g(t,t)dz(t), \tag{5.65}$$

and

$$\int_t^{t+dt} \frac{\partial g(s,t)}{\partial t}dtdz(s) \simeq \frac{\partial g(t,t)}{\partial t}dz(t)dt = o(dt), \tag{5.66}$$

where the last equality follows by the basic rules of stochastic calculus. Putting together the result (5.64)–(5.66) we finally obtain (to $o(dt)$) the result

$$dx(t) = \left[\int_0^t \frac{\partial g(s,t)}{\partial t}dz(s)\right]dt + g(t,t)dz(t). \tag{5.67}$$

## 5.6 Appendix

### Appendix 5.1 Proof of the Fundamental Rules of Stochastic Calculus

Putting $\sigma_{i-1} = \sigma(x(t_{i-1}), t_{i-1})$, we are considering

$$\underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \sigma_{i-1}(\Delta z_i)^{2+\alpha} \text{ for } \alpha \geq 0.$$

Consider first $\alpha = 0$. Anticipating the result, define

$$\begin{aligned} I &= \lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{n} \sigma_{i-1}[(\Delta z_i)^2 - \Delta t_i]\right]^2 \qquad (\Delta t_i \equiv t_i - t_{i-1}) \\ &= \lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{n} \sigma_{i-1}^2[(\Delta z_i)^2 - \Delta t_i]^2 + \sum_{j=1}^{n}\sum_{i>j}^{n} 2\sigma_{i-1}\sigma_{j-1} \right.\\ &\quad \left. \times [(\Delta z_j)^2 - \Delta t_j][(\Delta z_i)^2 - \Delta t_i]\right]. \end{aligned}$$

To obtain the noted statistical independencies we are assuming that the $\sigma_i$ are independent of all $\Delta z_j$ for $j > i$. (i.e. the function $\sigma$ is non-anticipating). Making use of the results

(a) $\mathbb{E}[(\Delta z_i)^2] = \Delta t_i$
(b) $\mathbb{E}[(\Delta z_i)^2 - \Delta t_i]^2 = 2(\Delta t_i)^2$

we find that

$$I = 2 \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}(\sigma_{i-1}^2)(\Delta t_i)^2 = 0 \qquad \text{(provided the function } \sigma \text{ is bounded).}$$

Since

$$\underset{n\to\infty}{\text{ms-lim}} \sum_{i=1}^{n} \sigma_{i-1}\Delta t_i = \int_0^t \sigma(x(s), s)ds,$$

we have in effect shown that

$$\int_0^t \sigma(x(s), s)(dz(s))^2 = \int_0^t \sigma(x(s), s)ds.$$

The proof that

$$\int_0^t \sigma(x(s), s)(dz(s))^{2+\alpha} = 0, \text{ for } \alpha > 0$$

follows similar lines and makes use of standard results for the higher moments of the Gaussian distribution.

## 5.7 Problems

**Problem 5.1** Consider the stochastic differential equation system

$$dx_1 = \mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2,$$
$$dx_2 = \mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2,$$

where $w_1$ and $w_2$ are independent Wiener processes. For this system determine the matrix $S$ referred to in Eq. (5.60). Hence write out the Kolmogorov backward equation for the transition probability density function for the joint process $x_1$ and $x_2$.

**Problem 5.2** Consider the stochastic differential equation system

$$dx_1 = \mu_1 dt + s_{11} dz_1 + s_{12} dz_2,$$
$$dx_2 = \mu_2 dt + s_{21} dz_1 + s_{22} dz_2,$$

where $z_1$ and $z_2$ are correlated Wiener processes satisfying

$$\mathbb{E}[dz_1 dz_2] = \rho dt.$$

Transform this system to one involving uncorrelated Wiener processes. Hence obtain the Kolmogorov backward equation for the transition probability density function for the joint process $x_1$ and $x_2$.

**Problem 5.3** Consider the quantity

$$\zeta(t) = dw_1(t) dw_2(t)$$

where $w_1(t)$ and $w_2(t)$ are two independent Wiener processes. Analyse this quantity in the same way that $(dz(t))^2$ is analysed in Sect. 5.1 and show why it can be ignored to $o(dt)$.

**Problem 5.4** Consider the quantity $(dz)^4$. Calculate its mean and variance. Why can it be ignored to $o(dt)$?

**Problem 5.5** Consider the two asset and two-noise model

$$dx_1 = \mu_1 x_1 dt + \sigma_{11} x_1 dz_1 + \sigma_{12} \sqrt{x_2 x_1} dz_2$$
$$dx_2 = \mu_2 x_2 dt + \sigma_{21} x_2 dz_1 + \sigma_{22} x_2 dz_2,$$

where

$$\mathbb{E}[dz_1 dz_2] = \rho dt.$$

Write this system in terms of two independent Wiener processes $w_1$, $w_2$. Use the transformation that expresses $z_1$ in terms of just $w_1$ and, $z_2$ in terms of both $w_1$ and $w_2$. Write out the Kolmogorov backward equation for this system.

**Problem 5.6** Consider the stochastic integral

$$x(t) = \int_0^t \sigma(s,t) dW(s).$$

Calculate $\mathbb{E}_0[x(t_1)x(t_2)]$, where $t_1 < t_2$.

**Problem 5.7** Consider the stochastic integral

$$x(t) = \int_0^t \sigma_1(s,t) dW_1(s) + \int_0^t \sigma_2(s,t) dW_2(s)$$

where $W_1$, $W_2$ are independent Wiener processes. Calculate $\mathbb{E}_0[x(t_1)x(t_2)]$, where $t_1 < t_2$.

**Problem 5.8** Repeat the calculation of Sect. 5.1 for $(dz)^3$ and $dt(dz)^2$. Draw graphs to illustrate the distribution of these quantities and show that, to order $dt$, they can be regarded as deterministic quantities equal to 0.

**Problem 5.9** Consider the one asset-two noise term model

$$dx = \mu dt + \sigma_1 dz_1 + \sigma_2 dz_2,$$

where $\mathbb{E}[dz_1 dz_2] = \rho dt$. Write out the Kolmogorov backward equation for the transition probability density function $p(x_T, T|x, t)$. Be sure in particular to give the form of the infinitesimal generator $\mathscr{K}$ in terms of the coefficients $\mu, \sigma_1, \sigma_2$ and $\rho$.

**Problem 5.10** Consider the two asset-two noise model

$$dx_1 = \mu_1 x_1 dt + \sigma_{11} x_1 x_2 dz_1,$$
$$dx_2 = \mu_2 x_2 dt + \sigma_{21} x_2 dz_1 + \sigma_{22} x_2 dz_2,$$

where $\mathbb{E}[dz_1, dz_2] = \rho dt$. Write this system in terms of two independent Wiener processes $w_1$ and $w_2$ such that the stochastic differential equation for $x_1$ contains only *one* of these independent Wiener processes. Write out the Kolmogorov backward equation for this system as well.

**Problem 5.11** Re-work the proof of Appendix 5.1 for the case when $\alpha = 1$.

**Problem 5.12** **Computational Problem**—Consider the stochastic integral

$$Y(t) = \int_0^t e^{-k(t-s)} dz(s),$$

where $z(s)$ is a Wiener process.

(i) Write a program that will approximate $Y(t)$ by

$$Y_n(t) = \sum_{i=0}^{n-1} e^{-k(n\Delta t - i\Delta t)} \Delta z_i$$

where $\Delta t = t/n$ and $\Delta z_i = z((i+1)\Delta t) - z(i\Delta t)$. The values of $k$, $t$, $n$ and the number $M$ of simulated paths should be user defined inputs. Initially take $k = 0.5$, $t = 1$, $n = 100$ and $M = 1{,}000$.

(ii) Compare the simulated distribution of $Y_n(1)$ with the true distribution of $Y(1)$ [see Sect. 5.2 if you have forgotten how to calculate this]. Experiment with the values of $n$ and $M$.

# Chapter 6
# Ito's Lemma and Its Applications

**Abstract** This chapter introduces Ito's lemma, which is one of the most important tools of stochastic analysis in finance. It relates the change in the price of the derivative security to the change in the price of the underlying asset. Applications of Ito's lemma to geometric Brownian motion asset price process, the Ornstein–Uhlenbeck process, and Brownian bridge process are discussed in detail. Extension and applications of Ito's lemma in several variables are also included.

## 6.1 Introduction

We saw in Chap. 4 that the *stochastic differential equation*

$$dx(t) = \mu(x(t), t)dt + \sigma(x(t), t)dz(t), \tag{6.1}$$

describing the price change $dx(t)$ over the time interval $dt$ in the limit as $dt \to 0$ is properly interpreted mathematically as the *stochastic integral equation*

$$x(t) = x(0) + \int_0^t \mu(x(s), s)ds + \int_0^t \sigma(x(s), s)dz(s). \tag{6.2}$$

The first integral is the standard Riemann integral. The second integral is the stochastic integral which may be interpreted in a number of ways. For most practical modelling situations the choice reduces to the definition of Ito and the definition of Stratonovich.

The Stratonovich definition is probably more satisfactory from the model building perspective as the random shock term associated with it corresponds to modelling market noise with a small correlation time between successive price shocks and then allowing the correlation to shrink to zero. Such a modelling procedure would be suggested by a lot of the empirical literature that we briefly surveyed in Chap. 4.

The Ito definition, on the other hand, does make the proofs of the mathematical theorems easier (relatively!). This definition does however rely on the concept of non-anticipating functions, which rules out any of the price dependency apparently found in some of the empirical literature.



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_6

We pointed out in Chap. 4 that every Stratonovich stochastic differential equation has a corresponding Ito one. The difference being the effect of the volatility on the drift term. It is therefore possible to develop the theory in terms of the Ito definition and decide at the point of application which definition is most appropriate.

In Chap. 5 we reviewed some of the pertinent properties of Ito stochastic differential equations. In this chapter we will derive Ito's lemma, which is one of the most important tools of stochastic analysis in finance. Its importance from the view point of our applications lies in the fact that it will enable us to relate the change in the price of the derivative security to the change in the price of the underlying asset. We are then naturally led into Chap. 7 that deals with the continuous hedging argument and the derivation of the partial differential equation determining the price of the derivative security.

## 6.2 Ito's Lemma

### *6.2.1 Introduction*

In keeping with our assumption that the noise in financial markets can be robustly approximated by the mathematically idealized white noise process, we model asset price movements as Ito stochastic differential equations. We shall henceforth refer to the Ito stochastic differential equation simply as the stochastic differential equation.

In order to price a derivative security on the asset we shall make the assumption that the derivative security price is a function of the price of the underlying asset. If we wish to set up a hedged portfolio of the derivative security and the underlying asset (as we will do in Chap. 7) then we would need to answer the question: given the stochastic differential equation driving the asset price, what is the stochastic differential equation satisfied by the price of the derivative security?

Mathematically, this problem may be stated as follows: Let $x(t)$ denote the asset price at time $t$, and let $y = y(x, t)$ denote the price of the derivative security. Given that $x(t)$ satisfies the stochastic differential equation

$$dx = \mu(x, t)dt + \sigma(x, t)dz,$$

what is the stochastic differential equation followed by $y$, the price of the derivative security?

### 6.2.2 Statement and Proof of Ito's Lemma

The answer to the question at the end of the previous subsection is provided by

**Ito's Lemma** *Let the stochastic process $x$ satisfy the stochastic differential equation*

$$dx = \mu(x,t)dt + \sigma(x,t)dz, \tag{6.3}$$

*and let $y(x,t)$ be a function of $x$ and $t$ (which is continuously differentiable in $t$ and twice continuously differentiable in $x$), then $y$ satisfies the stochastic differential equation*

$$dy = \left[\frac{\partial y}{\partial t} + \mu(x,t)\frac{\partial y}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 y}{\partial x^2}\right]dt + \sigma(x,t)\frac{\partial y}{\partial x}dz. \tag{6.4}$$

The proof of Ito's lemma involves use of Taylor's expansion and the rules for manipulating the increments of the Wiener process which we discussed in Sect. 5.1. The change in $y$ is given by

$$dy = y(x(t) + dx(t), t + dt) - y(x(t), t). \tag{6.5}$$

Expanding the first term on the right-hand side by Taylor's theorem

$$dy = \frac{\partial y}{\partial x}dx + \frac{1}{2}\frac{\partial^2 y}{\partial x^2}(dx)^2 + \frac{\partial y}{\partial t}dt + o(dt), \tag{6.6}$$

where $o(dt)$ denotes higher order terms in $dt$. By squaring (6.3) we observe that

$$(dx)^2 = \mu^2(x,t)(dt)^2 + 2\mu(x,t)\sigma(x,t)dtdz + \sigma^2(x,t)(dz)^2. \tag{6.7}$$

Applying the basic rules of stochastic calculus from Sect. 5.1 that

$$dtdz = 0, \quad (dz)^2 = dt,$$

and ignoring terms of higher order than $dt$, we see that (6.7) simply becomes

$$(dx)^2 = \sigma^2(x,t)dt.$$

Substituting this last expression into (6.6) and ignoring terms of $o(dt)$ the expression for $dy$ becomes

$$dy = \frac{\partial y}{\partial x}dx + \left[\frac{\partial y}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 y}{\partial x^2}\right]dt. \tag{6.8}$$

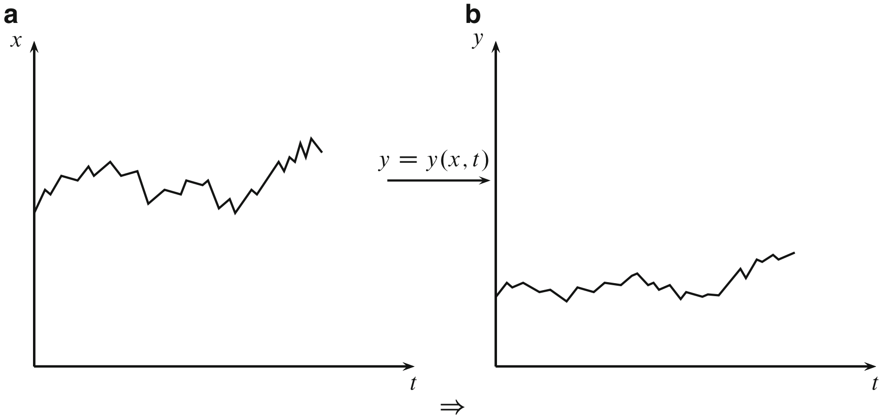


**Fig. 6.1** Ito's lemma; the tool that determines the stochastic differential equation for $y$ given the stochastic differential equation for $x$. (**a**) $dx = \mu dt + \sigma dz$. (**b**) $dy = \mu_y dt + \sigma_y dz$

Finally substitute the expression (6.3) for $dx$ into (6.8) to obtain

$$dy = \mu_y dt + \sigma_y dz, \tag{6.9}$$

where

$$\mu_y = \mu(y,t) = \frac{\partial y}{\partial t} + \mu(x,t)\frac{\partial y}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 y}{\partial x^2},$$
$$\sigma_y = \sigma(y,t) = \sigma(x,t)\frac{\partial y}{\partial x}.$$

This leads to Ito's lemma.

Figure 6.1 illustrates the basic content of Ito's lemma. In the expansion (6.6) we have not gone beyond the $(dx)^2$ term. Inclusion of higher order terms, such as $(dx)^3$, leads to terms of the type $(dz)^{2+\alpha}$ and/or $(dt)^\beta (dz)^\alpha$ where $\alpha$ and $\beta$ are positive integers. As we have shown in Sect. 5.1 such terms are zero to order $dt$.

## 6.3 Applications of Ito's Lemma

In this section we outline a number of applications of Ito's lemma which are frequently used in stochastic finance applications.

### 6.3.1 Function of a Geometric Stock Price Process

A very relevant application for our purposes is to consider the price process

$$dx = \mu x dt + \sigma x dz, \tag{6.10}$$

(i.e. $\mu(x,t) = \mu x$, $\sigma(x,t) = \sigma x$). This process is known as **geometric Brownian motion (GBM)** or **log-normal** since as we shall see below that $\ln x$ is distributed according to the distribution for Brownian motion that we encountered in Sect. 2.4.1. As we pointed out in Sect. 2.6 there is a deal of empirical evidence to suggest that this is the process followed by common stock prices, where $\mu$ is the expected stock return per unit time and $\sigma^2$ is the instantaneous variance of stock returns per unit time.

If $y(x,t)$ is the price of an option written on the common stock then by direct application of Ito's lemma in (6.9)

$$dy = \left[\frac{\partial y}{\partial t} + \mu x \frac{\partial y}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 y}{\partial x^2}\right] dt + \sigma x \frac{\partial y}{\partial x} dz, \tag{6.11}$$

a result we shall have occasion to use shortly.

### 6.3.2 The Lognormal Asset Price Process

Let

$$y = \ln x, \quad (\text{so that} \quad x = e^y), \tag{6.12}$$

then a straight forward mechanical application of Ito's lemma yields that $y$ satisfies the stochastic differential equation

$$dy = [\mu(x,t)e^{-y} - \frac{1}{2}\sigma^2(x,t)e^{-2y}]dt + \sigma(x,t)e^{-y} dz. \tag{6.13}$$

If in particular $x$ follows the stock price process (6.10) so that

$$\mu(x,t) = \mu x = \mu e^y, \qquad \sigma(x,t) = \sigma x = \sigma e^y,$$

then (6.13) simplifies to

$$dy = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dz. \tag{6.14}$$

The process for $y(= \ln x)$ is of interest when $x$ is a stock price process since changes in log-price is one way to measure the stock return.

Since $\mu$ and $\sigma$ are here assumed to be constant we are able to integrate (6.14) from 0 to $t$ to obtain

$$y(t) = y(0) + \int_0^t (\mu - \frac{1}{2}\sigma^2)ds + \int_0^t \sigma dz(s),$$

i.e.

$$y(t) = y(0) + (\mu - \frac{1}{2}\sigma^2)t + \sigma(z(t) - z(0)). \tag{6.15}$$

Recalling that $x = e^y$ we may also express (6.15) in terms of the stock price $x$ to obtain

$$x(t) = x(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(z(t) - z(0))}. \tag{6.16}$$

This last expression (6.16) is the solution[1] to the lognormal stock price stochastic differential equation (6.10). This is one of the rare occasions in which we can obtain an analytical solution to a stochastic differential equation. We note that Eq. (6.16) allows us to simulate the process for $x(t)$ up to time $t$ without resorting to discretisation (see Problem 6.16). Equation (6.16) shows explicitly how the sample paths for the stock price can be viewed as random excursions around a growing trend.

Equation (6.16) may be written

$$\ln\left(\frac{x(t)}{x(0)}\right) = (\mu - \frac{1}{2}\sigma^2)t + \sigma(z(t) - z(0)). \tag{6.17}$$

Since $z(t)$ is normally distributed it follows from (6.17) that $\ln(x(t)/x(0))$ is also normally distributed with mean and variance readily calculated as

$$\mathbb{E}_0\left[\ln\left(\frac{x(t)}{x(0)}\right)\right] = (\mu - \frac{1}{2}\sigma^2)t, \tag{6.18}$$

$$\text{var}_0\left[\ln\left(\frac{x(t)}{x(0)}\right)\right] = \sigma^2 t. \tag{6.19}$$

If we use $\Phi(\mu, \sigma^2)$ to denote the density function of the normal distribution with mean $\mu$ and variance $\sigma^2$ then we have just shown that

$$\ln\left(\frac{x(t)}{x(0)}\right) \sim \Phi\left((\mu - \frac{1}{2}\sigma^2)t, \sigma^2 t\right). \tag{6.20}$$

[1] We do not give a formal mathematical definition of what is meant by the term "a solution of a stochastic differential equation". But essentially it means what we see in Eq. (6.16). The value of the process for $x$ up to time $t$ is a function of the initial condition, time $t$ and the underlying driving stochastic process (here $z(t) - z(0)$).

If in going from (6.14) to (6.15) we integrate from $t$ to $T$ we would obtain the result

$$\ln\left(\frac{x(T)}{x(t)}\right) \sim \Phi\left((\mu - \frac{1}{2}\sigma^2)(T-t), \sigma^2(T-t)\right). \tag{6.21}$$

We note (6.21) is in effect the transition density function for the process $y(t) = \ln x(t)$, which we denote $q(y_T, T|y, t)$. Using the expression for the density function for the normal distribution, the function $q$ can be expressed as

$$q(y_T, T|y, t) = \frac{1}{\sqrt{2\pi}(T-t)\sigma} \exp\left[-\frac{\{y(T) - y(t) - \left(\mu - \frac{1}{2}\sigma^2\right)(T-t)\}^2}{2\sigma^2(T-t)}\right]. \tag{6.22}$$

In most applications we require the transition density in terms of $x$, namely $p(x_T, T|x, t)$. This is obtained by using the rule for transforming p.d.f.s, that is (see Appendix 2.1)

$$p(x_T, T|x, t) = q(y_T, T|y, t)\frac{dy_T}{dx_T}$$

from which

$$p(x_T, T|x, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma x_T} \exp\left[-\frac{\{\ln(x_T/x) - (\mu - \frac{1}{2}\sigma^2)(T-t)\}^2}{2\sigma^2(T-t)}\right], \tag{6.23}$$

the lognormal density function. This is precisely the transition probability density function that we used at Eq. (3.13), obtained by solving the Kolmogorov equation, see Problem 2.3. Here we have obtained it by solving the stochastic differential equation for $x(t)$. In Chap. 9 we shall see how to obtain the same result by solving the Kolmogorov backward equation.

It is a simple matter to extend the result (6.23) to the case when $\mu$ and $\sigma$ are both functions of time. It turns out that we need replace the $\mu$ and $\sigma$ in Eqs. (6.20) and (6.21) by the time averaged drift and diffusion, namely

$$\overline{\mu}(t, T) = \frac{1}{(T-t)} \int_t^T \mu(s)ds, \tag{6.24}$$

$$\overline{\sigma}^2(t, T) = \frac{1}{(T-t)} \int_t^T \sigma^2(s)ds. \tag{6.25}$$

### 6.3.3 Exponential Functions

Suppose $x$ follows the process (6.3). Let

$$y = e^x, \quad (\text{so that} \quad x = \ln y).$$

Now an application of Ito's lemma reveals that

$$dy = y\left[\mu(x,t) + \frac{1}{2}\sigma^2(x,t)\right]dt + y\sigma(x,t)dz. \tag{6.26}$$

Assume $\mu$ and $\sigma$ in (6.3) are constant and consider the more general exponential function

$$m = e^{x-(\mu+\frac{1}{2}\sigma^2)t}. \tag{6.27}$$

Then after application of Ito's lemma the stochastic differential equation for $m$ turns out to be

$$dm = \sigma m dz. \tag{6.28}$$

Thus the quantity $m$ satisfies a stochastic differential equation with zero drift term. As we shall see in the next chapter, $m$ is an example of a martingale.

### 6.3.4 Calculating $\mathbb{E}[e^{x(t)}]$

Quite often in stochastic finance we need to calculate the expectation of some function of a Wiener process, for example $\mathbb{E}_0[e^{z(t)}]$ where $z(t)$ is a Wiener process and $\mathbb{E}_0$ is the expectation operator conditional on information at time 0.

To proceed with this calculation we see that $z(t)$ may be viewed as the solution of the stochastic differential equation

$$dx(t) = dz(t), \quad x(0) = 0 \tag{6.29}$$

i.e. we have a stochastic differential equation with $\mu = 0$ and $\sigma = 1$. Applying Ito's lemma we find that $y = e^x$ satisfies the stochastic differential equation

$$dy = \frac{1}{2}ydt + ydz(t). \tag{6.30}$$

Bearing in mind that $\mathbb{E}_0[dz(t)] = 0$ we see that

$$m(t) = \mathbb{E}_0[y(t)] = \mathbb{E}_0[e^{x(t)}] = \mathbb{E}_0[e^{z(t)}],$$

satisfies the ordinary deterministic differential equation

$$dm = \frac{1}{2} m dt, \quad m(0) = y(0) = e^{x(0)} = 1, \tag{6.31}$$

whose solution is

$$m(t) = e^{\frac{1}{2}t}.$$

Hence we have established the result

$$\mathbb{E}_0[e^{z(t)}] = e^{\frac{1}{2}t}. \tag{6.32}$$

Note how we may generalize the above result. Suppose we wish to calculate $\mathbb{E}_0[e^{x(t)}]$ where $x(t)$ is a diffusion process with time varying drift and diffusion coefficients, so that

$$dx(t) = \mu(t)dt + \sigma(t)dz(t). \tag{6.33}$$

Let $y = e^x$ then by Ito's lemma

$$dy = [\mu(t) + \frac{1}{2}\sigma^2(t)]y dt + \sigma(t) y dz.$$

Integrating between 0 and $t$ the last equation implies

$$y(t) = y(0) + \int_0^t (\mu(s) + \frac{1}{2}\sigma^2(s)) y(s) ds + \int_0^t \sigma(s) y(s) dz(s).$$

Let $m(t) = \mathbb{E}_0[y(t)]$ and using the result (5.19) then[2]

$$m(t) = y(0) + \int_0^t (\mu(s) + \frac{1}{2}\sigma^2(s)) m(s) ds.$$

Thus, differentiating[3] with respect to $t$

$$\frac{dm}{dt} = (\mu(t) + \frac{1}{2}\sigma^2(t)) m,$$

which has solution

$$m(t) = m(0) \exp\left[\int_0^t \left(\mu(s) + \frac{1}{2}\sigma^2(s)\right) ds\right] = y(0) \exp\left[\int_0^t \left(\mu(s) + \frac{1}{2}\sigma^2(s)\right) ds\right].$$

[2]Note that the operations $\mathbb{E}_0$ and $\int_0^t$ commute (i.e. their order may be interchanged) as may be easily shown by appealing to the definition of the integral.

[3]Note that $m(t)$ is a deterministic quantity so we can use ordinary calculus here.

That is, we have shown that

$$\mathbb{E}[e^{x(t)}] = \exp\left[x(0) + \int_0^t \left(\mu(s) + \frac{1}{2}\sigma^2(s)\right)ds\right]. \tag{6.34}$$

To further interpret (6.34) we recall that (6.33) can be expressed as

$$x(t) = x(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dz(s). \tag{6.35}$$

For reasons we discussed in Sect. 5.2 the stochastic integral $\int_0^t \sigma(s)dz(s)$ is normally distributed, hence $x(t)$ is normally distributed. We readily calculate that the mean of $x(t)$ is given by

$$\mathbb{E}_0[x(t)] = x(0) + \int_0^t \mu(s)ds \equiv M(t),$$

and the variance of $x(t)$ is given by[4]

$$\mathbb{E}_0[(x(t) - M(t))^2] = \mathbb{E}_0\left[\left(\int_0^t \sigma(s)dz(s)\right)^2\right] = \int_0^t \sigma^2(s)ds \equiv V^2(t).$$

Thus we can assert that

$$x(t) \sim N(M(t), V^2(t)), \tag{6.36}$$

and the result (6.34) may be re-stated as

$$\mathbb{E}\left[e^{x(t)}\right] = e^{M(t)+\frac{1}{2}V^2(t)}. \tag{6.37}$$

This is a result that will be used not infrequently in later chapters.

We can use the foregoing result to calculate $\mathbb{E}_t[x(T)]$ when $x$ follows the GBM (6.10). We note first from (6.12) that $x = e^y$, so that

$$\mathbb{E}[x(T)] = \mathbb{E}_t[e^{y(T)}]. \tag{6.38}$$

By integrating (6.14) from $t$ to $T$ we obtain

$$y(T) = y(t) + (\mu - \frac{1}{2}\sigma^2)(T - t) + \sigma(z(T) - z(t)). \tag{6.39}$$

[4]The result at the second equality uses results demonstrated in Sect. 5.2.

Hence we calculate the mean and variance of $y(T)$ as

$$\mathbb{E}_t[y(T)] \equiv M_y(T) = y(t) + \left(\mu - \frac{1}{2}\sigma^2\right)(T-t),$$

and

$$\text{var}_t[y(T)] \equiv V_y^2(T) = \sigma^2(T-t).$$

Thus by a direct application of (6.37) we obtain the result

$$\mathbb{E}_t[e^{y(T)}] = e^{y(t)+\mu(T-t)}. \tag{6.40}$$

Recalling again the relation (6.12) between $x$ and $y$ we write (6.40) as

$$\mathbb{E}_t[x(T)] = x(t)e^{\mu(T-t)}. \tag{6.41}$$

Some simulated paths for $x(t)$ as well as the trend are shown in Fig. 6.2. These paths have been simulated using Eq. (6.16).

If $x$ is interpreted as a stock price then $dy$ is the instantaneous return (i.e. over the time interval $(t, t+dt)$) and $y(t) - y(0)$ from Eq. (6.15) is the accumulated return

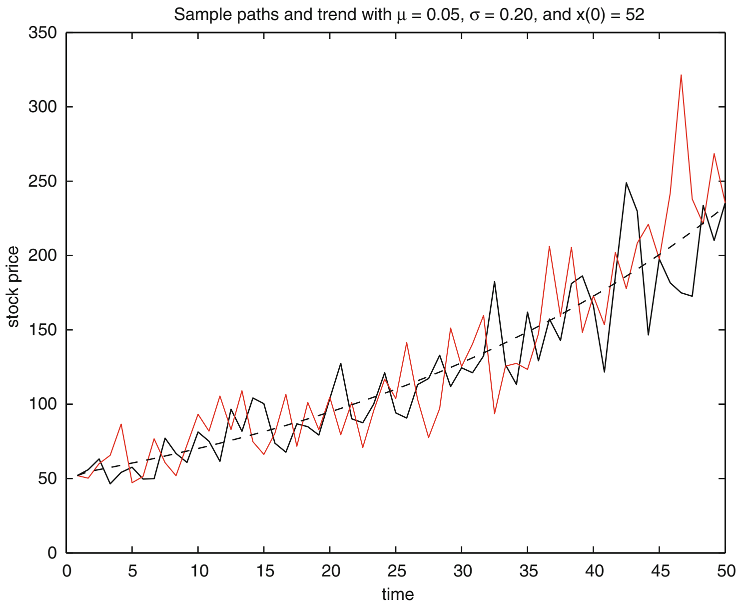


**Fig. 6.2** Trend and simulated paths of the asset price process. Here $\mu = 0.05$, $\sigma = 0.20$ and $x(0) = 52$

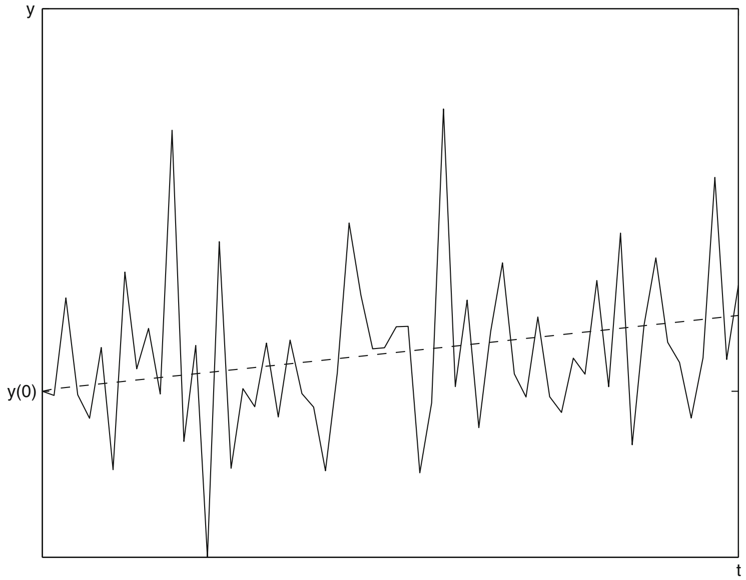


**Fig. 6.3** Simulated paths of the accumulated return $y(t)$ over $(0, t)$

over $(0, t)$, which is plotted in Fig. 6.3. These simulations correspond to the paths shown in Fig. 6.2.

### *6.3.5 The Ornstein–Uhlenbeck Process*

We have discussed in Sect. 4.3.2 the Ornstein–Uhlenbeck process. It is characterized by a linear drift coefficient and a constant diffusion coefficient. Thus the stochastic differential equation for an Ornstein–Uhlenbeck process is[5]

$$dx(t) = -kxdt + \sigma dz(t). \tag{6.42}$$

With a view as to how we would solve (6.42) if it were an ordinary deterministic differential equation with the second term on the right-hand side considered as a forcing term we define the quantity

$$y = xe^{kt}.$$

[5]Note that the $\sigma$ of the discussion of this section is equivalent to $\sqrt{D}$ of the discussion in Sect. 4.3.2.

By Ito's lemma this quantity satisfies the stochastic differential equation

$$dy(t) = \sigma e^{kt} dz(t). \tag{6.43}$$

Integrating (6.43) from 0 to $t$ we obtain

$$y(t) - y(0) = \sigma \int_o^t e^{ks} dz(s).$$

In terms of the original variable $x$ this result can be expressed

$$x(t) = x(0)e^{-kt} + \sigma \int_0^t e^{-k(t-s)} dz(s). \tag{6.44}$$

We shall frequently have occasion to use this form of the solution to the stochastic differential equation (6.42).

It is also of interest to calculate the mean and the variance of the Ornstein–Uhlenbeck process. Taking expectations across (6.44) we calculate

$$\mathbb{E}_0[x(t)] = x(0)e^{-kt}, \tag{6.45}$$

and

$$\begin{aligned} \text{var}_0[x(t)] &= \sigma^2 \mathbb{E}_0 \left[ \left( \int_0^t e^{-k(t-s)} dz(s) \right)^2 \right] \\ &= \sigma^2 \int_0^t e^{-2k(t-s)} ds = \frac{\sigma^2}{2k}(1 - e^{-2kt}). \end{aligned} \tag{6.46}$$

Equations (6.45) and (6.46) correspond to the results (4.19) and (4.20) when we make the identification $\tau \to 0$, $y \to x(0)$ and $D \to \sigma^2$.

Figure 6.4 displays some simulated paths of the Ornstein–Uhlenbeck process as well as the mean trend ($x(0)e^{-kt}$; see Eq. (6.45)) and the standard deviation bands ($x(0)e^{-kt} \pm 2\sigma(1 - e^{-2kt})^{1/2}/\sqrt{2k}$; see Eq. (6.46)). The parameter values are the same as those used to generate the distributions in Fig. 4.8.[6]

### 6.3.6 *Brownian Bridge Processes*

In later chapters we will want to consider stochastic differential equations which model bond prices. One of the key characteristics of bond prices is that at maturity

[6] In the current notation the values are $\sigma = 1$, $k = 1$, $x(0) = 1$.

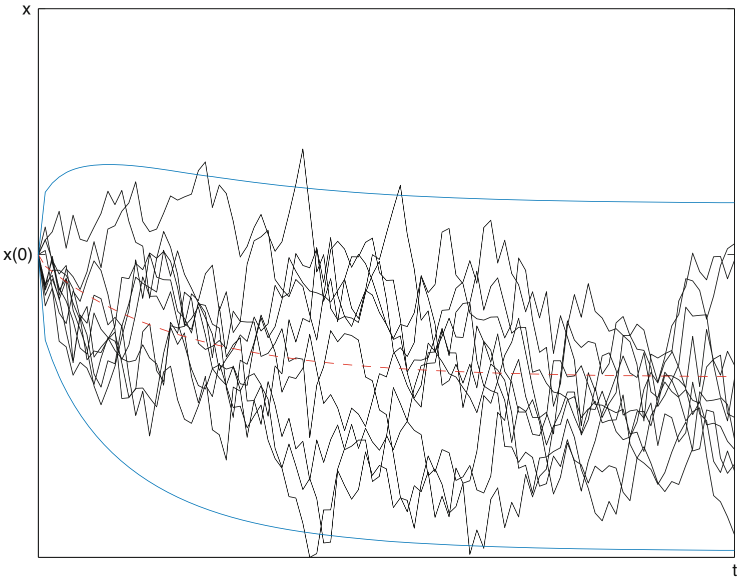


**Fig. 6.4** Simulated paths of the Ornstein–Uhlenbeck process. Also displayed are the mean and the two standard deviation bands calculated from Eqs. (6.45) and (6.46)

the price must equal the par-value. This behaviour of bond prices is known as the pull-to-par-effect.

One approach to modelling this behaviour uses stochastic processes that are "tied-down" at both ends i.e. both the value at $t = 0$ and the value at $t = T$ are specified. The processes we have used to date to model stock prices do not enjoy this property. Hence we need to introduce a new type of process.

The Brownian bridge is one such process. For convenience units are standardized so that it is defined on the unit interval $0 \leq t < 1$. The Brownian bridge process is the process $y(t)$ satisfying the stochastic differential equation

$$dy(t) = \frac{b - y(t)}{1 - t}dt + \sigma dz(t), \quad 0 \leq t < 1, \quad y(0) = a. \tag{6.47}$$

We now want to show that[7]

$$\lim_{t \to 1} y(t) = b. \tag{6.48}$$

[7]In this subsection the $\lim_{t \to 1} y(t)$ should be interpreted as in the mean square sense.

To do this first consider the subsidiary variable

$$u = \frac{y}{1-t}. \tag{6.49}$$

An application of Ito's lemma easily shows that $u$ satisfies the stochastic differential equation

$$du(t) = \frac{b}{(1-t)^2}dt + \frac{\sigma}{1-t}dz(t). \tag{6.50}$$

Integrating (6.50) on $(0, t)$ we find that

$$\begin{aligned} u(t) &= a + b\int_0^t \frac{ds}{(1-s)^2} + \int_0^t \frac{\sigma}{1-s}dz(s) \\ &= a + b\left(\frac{1}{1-t} - 1\right) + \int_0^t \frac{\sigma}{1-s}dz(s). \end{aligned}$$

Re-expressing this last equation in terms of the original variable $y$ we obtain

$$y(t) = a(1-t) + bt + \int_0^t \sigma\left(\frac{1-t}{1-s}\right)dz(s). \tag{6.51}$$

Whilst some subtle technical arguments are required to prove that $\lim_{t\to 1} y(t) = b$, the intuition is fairly clear from (6.51). Clearly the deterministic term $[a(1-t)+bt] \to b$ as $t \to 1$.

As for the stochastic integral, it represents the sum of shock terms over $(0, t)$. But each shock term is multiplied by the coefficient $(1-t)/(1-s)$ which $\to 0$ as $t \to 1$. By considering appropriate limiting processes it can indeed be shown that

$$\lim_{t\to 1}\int_0^t \left(\frac{1-t}{1-s}\right)dz(s) = 0. \tag{6.52}$$

Figure 6.5 illustrates 100 simulations of (6.47) for $b = 1$ (i.e. for par value = \$1) and $\sigma = 0.05$ (i.e. bond price volatility = 5 %).

We observe that there is a large probability of bond prices exceeding \$1 which is not satisfactory for zero-coupon bond prices that pay \$1 for sure at maturity. Note further that although negative bond prices are possible with this process, they are not very likely with realistic values of $\sigma$. We shall see in a later chapter how the modelling procedure of Heath et al. (1992a) gives us processes for bond prices that are more likely to keep zero-coupon bond prices in the range $(0, 1)$. The square root process for the instantaneous spot interest rate introduced by Cox et al. (1985b) certainly guarantees that interest rates remain non-negative and hence bond prices remain below 1. All of these processes will be discussed in Chap. 22.

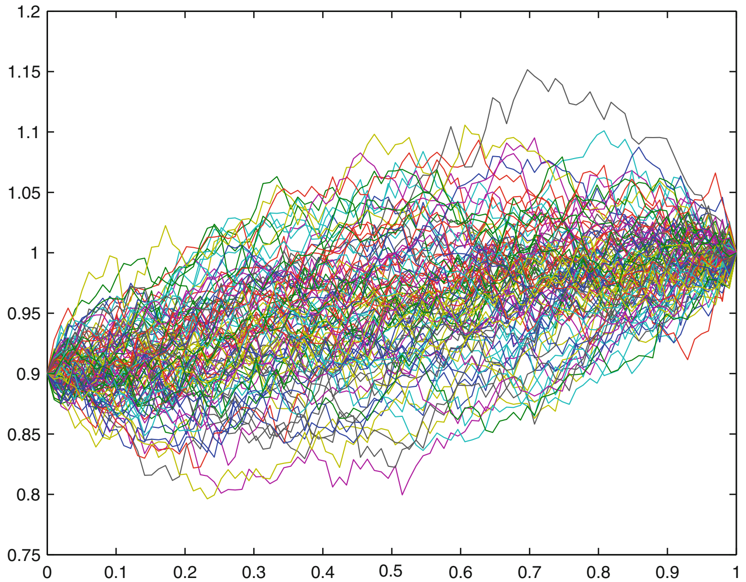


**Fig. 6.5** Simulation of the Brownian bridge process

### 6.3.7 White Noise and Colored Noise Processes

In Sect. 4.4 we discussed the autocorrelation behaviour of noise processes. There we saw that the Ornstein–Uhlenbeck process having drift $-kx$ and diffusion coefficient $\sigma^2 k^2$ has a correlation time of $1/k$. Such a noise process is known as coloured noise. We also saw that as $k \to \infty$, the correlation time $\tau_{\text{corr}} \to 0$ and the coloured noise process tends to white noise with an "intensity" of $\sigma^2$. In this subsection we want to make explicit the relationship between white noise and coloured noise via the language of stochastic differential equations.

Consider the following system of stochastic differential equations

$$dx = \mu(x,t)dt + \frac{\rho}{\varepsilon}dt, \tag{6.53}$$

$$d\rho = -\frac{\rho}{\varepsilon^2}dt + \frac{\sigma}{\varepsilon}dz, \tag{6.54}$$

with $\rho(0) = 0$. The process (6.54) for $\rho$ is the stochastic differential equation representation of the Ornstein–Uhlenbeck process having correlation time $\tau_{corr} = \varepsilon^2$.

The output of the Ornstein–Uhlenbeck process $\rho$ generates the noise term $\rho dt/\varepsilon$ driving the process (6.53) for $x$. In this case $x$ is said to be driven by the process $\rho$. Note that by Ito's lemma we can re-express (6.54) as

$$d\left(\rho e^{t/\varepsilon^2}\right) = \frac{\sigma}{\varepsilon}e^{t/\varepsilon^2}dz(t), \tag{6.55}$$

so that

$$\rho(t) = \frac{\sigma}{\varepsilon} \int_0^t e^{-(t-s)/\varepsilon^2} dz(s). \tag{6.56}$$

It is thus possible to re-express the pair of stochastic differential equations (6.53) and (6.54) as the single equation

$$dx = \mu(x,t)dt + \frac{\sigma}{\varepsilon^2} \int_0^t e^{-(t-s)/\varepsilon^2} dz(s)dt. \tag{6.57}$$

We note that in (6.57) the noise term depends on the path followed up to $t$. Hence the future evolution of the process *does* depend on the path history. Thus Eq. (6.57) is our first example of a non-Markovian process.

Since the non-Markovian stochastic differential equation (6.57) is equivalent to the pair of Markovian stochastic differential equations (6.53) and (6.54), we see that at least in some instances it may be possible to express a non-Markovian system as a higher order Markovian system. This is in fact a technique which we shall often exploit when we come to consider term structure of interest rate models. There we shall quite naturally encounter non-Markovian systems when we try to model the stochastic evolution of the yield curve.

One final point about (6.57). We know that to give it proper mathematical meaning it should be expressed as a stochastic integral equation. Integrating we obtain

$$x(t) = x(0) + \int_0^t \mu(x,\tau)d\tau + \frac{\sigma}{\varepsilon^2} \int_0^t \left( \int_0^\tau e^{-(\tau-s)/\varepsilon^2} dz(s) \right) d\tau. \tag{6.58}$$

In order that this equation be in the form of a standard stochastic integral equation we need the *dz* term of the last (double) integral to appear in the outer integral. In order to do this we need to perform on this stochastic double integral the operation of change of order of integration which is routine for Riemann integrals. A theorem known as Fubini's theorem[8] allows us (under suitable conditions) to manipulate such stochastic double integrals in the same way that we manipulate double Riemann integrals. Thus

$$\begin{aligned}\int_0^t \left( \int_0^\tau e^{-(\tau-s)/\varepsilon^2} dz(s) \right) d\tau &= \int_0^t \left( \int_s^t e^{-(\tau-s)/\varepsilon^2} d\tau \right) dz(s) \\ &= \varepsilon^2 \int_0^t [1 - e^{-(t-s)/\varepsilon^2}] dz(s).\end{aligned}$$

[8]Fubini's theorem is discussed in Sect. 22.4.

So the stochastic integral equation equivalent to (6.58) is

$$x(t) = x(0) + \int_0^t \mu(x,\tau)d\tau + \sigma \int_0^t [1 - e^{-(t-s)/\varepsilon^2}]dz(s). \tag{6.59}$$

The Fubini theorem technique used to go from (6.58) to (6.59) is frequently used in term structure of interest rate modelling as we shall see in Chaps. 22–26.

Finally, let us return to the white noise, coloured noise issue. From our discussion in Sect. 4.4 we know that as $\varepsilon \to 0$, the stochastic differential equation (6.53) should tend to one driven by the white noise process $dz(t)$. We see from Eq. (6.59) that this is indeed the case since $\lim_{\varepsilon\to 0} \exp[-(t-s)/\varepsilon^2] = 0$ (for $s < t$) and so the stochastic integral in (6.59) becomes $\int_0^t dz(s)$. Thus Eq. (6.59) becomes

$$x(t) = x(0) + \int_0^t \mu(x,\tau)d\tau + \sigma \int_0^t dz(s), \tag{6.60}$$

which is equivalent to the stochastic differential equation

$$dx(t) = \mu(x,t)dt + \sigma dz(t). \tag{6.61}$$

In order to show the difference on sample paths of coloured noise and pure white noise we display in Fig. 6.6 the simulation of

$$\frac{dx}{x} = \mu dt + \frac{\rho}{\varepsilon}dt$$

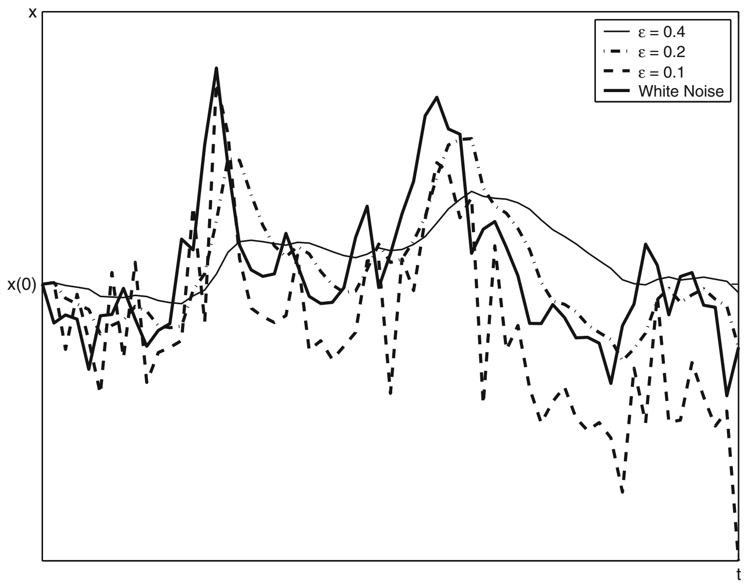


**Fig. 6.6** Comparing white noise and coloured noise for different values of $\varepsilon$

with $\rho$ given by (6.54). We use the same parameters as for Fig. 6.4 and using the same random sequence for $dz$ display paths for $\varepsilon = 0.1, 0.2, 0.4$ and the pure white noise case (i.e. the limit $\varepsilon \to 0$), we see how the sample paths are fairly smooth for large $\varepsilon$ but become more like the non-differentiable sample paths of white noise as $\varepsilon$ decreases towards zero.

Horsthemke and Lefever (see Sect. 8.4) consider the more general case in which (6.53) is replaced by

$$dx = \mu(x,t)dt + \frac{\rho}{\varepsilon}g(x)dt$$

and $\rho$ still follows (6.54). They show that in the limit as $\varepsilon \to 0$ the stochastic differential equation followed by $x$ is the Stratonovich stochastic differential equation

$$dx = \mu(x,t)dt + \sigma g(x) \circ dz$$

which is equivalent to the Ito stochastic differential equation

$$dx = (\mu(x,t) + \frac{1}{2}\sigma^2 g(x)g'(x))dt + \sigma g(x)dz. \tag{6.62}$$

For example if $x$ is our standard stock price process then (6.62) would become

$$dS = \left(\mu + \frac{1}{2}\sigma^2\right) S dt + \sigma S dz. \tag{6.63}$$

Of course for option pricing the slightly different drift term $(\mu + \frac{1}{2}\sigma^2)$ is not of importance since it disappears in the continuous hedging argument. However the issue of whether noise in financial markets is best modelled as pure white noise or as coloured noise with a very short correlation time is perhaps not yet clearly resolved in the finance literature.

## 6.4 A More Formal Statement of Ito's Lemma

As we have stressed in Chap. 4 the notation for the stochastic differential equation

$$dx(t) = \mu(x(t),t)dt + \sigma(x(t),t)dz(t),$$

is merely a convenient shorthand notation for the more precise expression in terms of the stochastic integral equation, namely,

$$x(t) = x(0) + \int_0^t \mu(x(s),s)ds + \int_0^t \sigma(x(s),s)dz(s).$$

Thus our statement of Ito's lemma in Sect. 6.2.2 has been in terms of the shorthand notation of the stochastic differential equation. The reason for this is that it is in this form, i.e. Eq. (6.4), that Ito's lemma finds its most common applications in finance.

However the more precise statement of Ito's lemma is in terms of stochastic integrals and should be written

$$y(x(t),t) = y(x(0),0) + \int_0^t \left[\frac{\partial y}{\partial s} + \mu(x(s),s)\frac{\partial y}{\partial x} + \frac{1}{2}\sigma^2(x(s),s)\frac{\partial^2 y}{\partial x^2}\right] ds + \int_0^t \sigma(x(s),s)\frac{\partial y}{\partial x} dz(s). \tag{6.64}$$

To formally prove this result we partition the interval $(0, t)$ into subintervals as in Sect. 4.7. Then

$$y(x(t),t) = y(x(0),0) + \sum_{i=1}^{n} \Delta y(x(i\Delta t), i\Delta t), \tag{6.65}$$

where $\Delta y[x(i\Delta t), i\Delta t]$ represents the change in $y$ over the subinterval $[(i-1)\Delta t, i\Delta t]$. This change may be approximated by the Taylor expansion

$$\Delta y(x(i\Delta t), i\Delta t) \simeq \frac{\partial y}{\partial t}\Delta t + \frac{\partial y}{\partial x}\Delta x_i + \frac{1}{2}\frac{\partial^2 y}{\partial x^2}(\Delta x_i)^2, \tag{6.66}$$

where the partial derivatives are evaluated at $(i-1)\Delta t$ (i.e. at the beginning of the subinterval as is the case with Ito stochastic integrals). Note that we have omitted the terms

$$\frac{\partial^2 y}{\partial x \partial t}(\Delta t)(\Delta x_i) \qquad \text{and} \qquad \frac{\partial^2 y}{\partial t^2}(\Delta t)^2,$$

since by the results of Appendix 5.1 integrals over these are zero, at least to $o(\Delta t)$. Substituting (6.66) into (6.65) we obtain terms of the type

$$\sum_{i=1}^{n} \frac{\partial y}{\partial x}\Delta x_i \qquad \text{and} \qquad \sum_{i=1}^{n} \frac{\partial^2 y}{\partial x^2}(\Delta x_i)^2.$$

By taking the limit $n \to \infty$ in the way described in Sect. 4.7 the stochastic integral equation (6.64) will emerge. For full details we refer the reader to Theorem 4.5 in Oksendal (2003).

The more formal statement of Ito's lemma in integral form is required when we come to consider the setting up of self-financing strategies in Chap. 7.

## 6.5 Ito's Lemma in Several Variables

In later chapters we shall have occasion to consider derivative securities dependent on more than one asset price or factor, e.g. options written on two stocks (so as to benefit from their correlation structure), or options on debt instruments that may depend on long term and short term interest rates. In this case the derivative security price, $y$, may depend on several stochastic price equations and so we shall need a multi-dimensional version of Ito's lemma. This is also derived by use of Taylor's expansion and the rules of Sect. 5.1 for manipulating the $dz$ terms. Here we only sketch the main ideas, for full details we refer the reader to Oksendal (2003).

We break our discussion up into two subsections. In Sect. 6.5.1 we consider the case in which each factor is driven by a separate Wiener process, but the Wiener processes may all be correlated. In Sect. 6.5.2 each factor is driven by a number of uncorrelated Wiener processes. One case can always be reduced to the other by changes of variable as discussed in Sect. 5.3 but it is useful for later applications to state both cases separately.

### *6.5.1 Correlated Wiener Processes*

In order to clarify the notation we shall first write out fairly explicitly the result when there are two factors and then state the result in the case of $n$ factors.

First consider the case of *two asset prices*. Let the stochastic processes $x_1, x_2$ satisfy the stochastic differential equations

$$dx_1 = \mu_1(x_1, x_2, t)dt + \sigma_1(x_1, x_2, t)dz_1, \tag{6.67}$$

$$dx_2 = \mu_2(x_1, x_2, t)dt + \sigma_2(x_1, x_2, t)dz_2, \tag{6.68}$$

where $dz_1$ and $dz_2$ are the increments of Wiener processes which may be correlated i.e.

$$\begin{aligned} \mathbb{E}(dz_1) = \mathbb{E}(dz_2) &= 0, \\ \mathbb{E}(dz_1^2) = \mathbb{E}(dz_2^2) &= dt, \\ \mathbb{E}(dz_1 dz_2) &= \rho dt. \end{aligned} \tag{6.69}$$

Here $\rho$ measures the degree of correlation between the two Wiener processes.

Let $y(x_1, x_2, t)$ be a function of $x_1, x_2$ and $t$. Using an entirely analogous procedure to that used in Sect. 6.2.2 we calculate $dy$ according to

$$dy = y(x_1 + dx_1, x_2 + dx_2, t + dt) - y(x_1, x_2, t). \tag{6.70}$$

Expanding (6.70) by Taylor's theorem, applying the rules of stochastic calculus (keeping in mind the relationships (6.69)) we find that $y$ satisfies the stochastic differential equation

$$dy = \left[\frac{\partial y}{\partial t} + \mu_1 \frac{\partial y}{\partial x_1} + \mu_2 \frac{\partial y}{\partial x_2} + \frac{1}{2}\left(\sigma_1^2 \frac{\partial^2 y}{\partial x_1^2} + 2\rho\sigma_1\sigma_2 \frac{\partial^2 y}{\partial x_1 \partial x_2} + \sigma_2^2 \frac{\partial^2 y}{\partial x_2^2}\right)\right] dt$$
$$+ \sigma_1 \frac{\partial y}{\partial x_1} dz_1 + \sigma_2 \frac{\partial y}{\partial x_2} dz_2. \tag{6.71}$$

To handle the general case of *n asset prices* we need to introduce vector notation. Let

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$$

denote the vector of asset prices. The stochastic differential equation for the price of the $i$ th asset is

$$dx_i = \mu_i(\mathbf{x}, t)dt + \sigma_i(\mathbf{x}, t)dz_i, \tag{6.72}$$

and the increments of the Wiener process, $dz_i$, satisfy

$$\mathbb{E}(dz_i) = 0, \quad \text{var}(dz_i) = dt, \quad \mathbb{E}(dz_i dz_j) = \rho_{ij} dt.$$

Now we need to specify $(\rho_{ij})_{n\times n}$ and the variance-covariance matrix of the increments of the various Wiener processes. Let

$$y = y(\mathbf{x}, t),$$

be a function of $\mathbf{x}$ and $t$. We form $dy = y(\mathbf{x} + d\mathbf{x}, t + dt) - y(\mathbf{x}, t)$, expand by Taylor's theorem and applying the rules of stochastic calculus find that $y$ satisfies the stochastic differential equation

$$dy = \left[\frac{\partial y}{\partial t} + \sum_{i=1}^{n} \mu_i \frac{\partial y}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 y}{\partial x_i \partial x_j}\right] dt + \sum_{i=1}^{n} \sigma_i \frac{\partial y}{\partial x_i} dz_i. \tag{6.73}$$

This is the $n$-factor version of Ito's lemma in the case of correlated Wiener processes.

### 6.5.2 Independent Wiener Processes

In this subsection we consider the case in which each diffusion process is driven by several Wiener processes. Since we have shown in Sect. 5.3 how to transform such stochastic differential systems to ones in which the Wiener processes are independent we only consider this latter situation here.

We first consider the two asset price case,

$$dx_1 = \mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2, \tag{6.74}$$

$$dx_2 = \mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2, \tag{6.75}$$

where we have suppressed the dependence of the $\mu_i$ and $\sigma_{ij}$ on $x_1, x_2$ and $t$. Now the $w_1(t)$ and $w_2(t)$ are independent Wiener processes so that $\mathbb{E}[dw_1 dw_2] = 0$. Again we let $y(x_1, x_2, t)$ be a function of $x_1, x_2$ and $t$. As in the previous subsection we calculate $dy$ by use of a Taylor's expansion and applying the rules for manipulating stochastic differentials, thus

$$\begin{aligned} dy =& \frac{\partial y}{\partial t} dt + \frac{\partial y}{\partial x_1}(\mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2) + \frac{\partial y}{\partial x_2}(\mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2) \\ &+ \frac{1}{2}\frac{\partial^2 y}{\partial x_1^2}(\sigma_{11}^2 dt + \sigma_{12}^2 dt) + \frac{1}{2}\frac{\partial^2 y}{\partial x_2^2}(\sigma_{21}^2 dt + \sigma_{22}^2 dt) \\ &+ \frac{\partial^2 y}{\partial x_1 \partial x_2}(\mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2)(\mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2). \end{aligned} \tag{6.76}$$

Note that by using the basic rules of stochastic calculus for manipulating increment of Wiener processes

$$\begin{aligned} (\mu_1 dt+&\sigma_{11} dw_1 + \sigma_{12} dw_2)(\mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2) \\ &= \sigma_{11}\sigma_{21}(dw_1)^2 + (\sigma_{12}\sigma_{21} + \sigma_{11}\sigma_{22}) dw_1 dw_2 + \sigma_{12}\sigma_{22}(dw_2)^2 \\ &= (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}) dt. \end{aligned} \tag{6.77}$$

Using this last result and gathering together the $dw_1$ and $dw_2$ Eq. (6.76) simplifies to

$$\begin{aligned} dy = \Bigg[& \frac{\partial y}{\partial t} + \mu_1 \frac{\partial y}{\partial x_1} + \mu_2 \frac{\partial y}{\partial x_2} + \frac{1}{2}(\sigma_{11}^2 + \sigma_{12}^2)\frac{\partial^2 y}{\partial x_1^2} + \frac{1}{2}(\sigma_{21}^2 + \sigma_{22}^2)\frac{\partial^2 y}{\partial x_2^2} \\ &+(\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})\frac{\partial^2 y}{\partial x_1 \partial x_2}\Bigg] dt \\ &+(\sigma_{11}\frac{\partial y}{\partial x_1} + \sigma_{21}\frac{\partial y}{\partial x_2}) dw_1 + (\sigma_{12}\frac{\partial y}{\partial x_1} + \sigma_{22}\frac{\partial y}{\partial x_2}) dw_2. \end{aligned} \tag{6.78}$$

In the case of $n$ asset prices we are dealing with the stochastic differential system

$$dx_i = \mu_i(\mathbf{x}, t)dt + \sum_{k=1}^{m} \sigma_{ik} dw_k(t) \tag{6.79}$$

for $i = 1, 2, \cdots, n$. Note that we allow $m \neq n$ so that the number of independent Wiener processes may not be the same as the number of price processes. Now expanding by use of Taylor's expansion we obtain

$$\begin{aligned} dy =& \frac{\partial y}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \left( \mu_i dt + \sum_{k=1}^{m} \sigma_{ik} dw_k \right) \\ &+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 y}{\partial x_i \partial x_j} \left( \mu_i dt + \sum_{k=1}^{m} \sigma_{ik} dw_k \right) \left( \mu_j dt + \sum_{k=1}^{m} \sigma_{jk} dw_k \right). \end{aligned} \tag{6.80}$$

Consider more closely the product in the last term

$$\begin{aligned} &\left( \mu_i dt + \sum_{k=1}^{m} \sigma_{ik} dw_k \right) \left( \mu_j dt + \sum_{k=1}^{m} \sigma_{jk} dw_k \right) \\ &= \left( \sum_{k=1}^{m} \sigma_{ik} dw_k \right) \left( \sum_{k=1}^{m} \sigma_{jk} dw_k \right) = \left( \sum_{k=1}^{m} \sigma_{ik} \sigma_{jk} \right) dt, \end{aligned} \tag{6.81}$$

where we have used the independence of the Wiener processes to obtain the last equality. Ignoring terms of $o(dt)$ we finally obtain from (6.80) that

$$dy = \left[ \frac{\partial y}{\partial t} + \sum_{i=1}^{n} \mu_i \frac{\partial y}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 y}{\partial x_i \partial x_j} \sum_{k=1}^{m} \sigma_{ik} \sigma_{jk} \right] dt + \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \sum_{k=1}^{m} \sigma_{ik} dw_k. \tag{6.82}$$

Note that if we define the matrix $\sigma = (\sigma)_{n \times m}$ and set

$$S = (s_{ij})_{n \times n} = \sigma \sigma^\top,$$

then

$$s_{ij} = \sum_{k=1}^{m} \sigma_{ik} \sigma_{jk}.$$

Referring to the definition of $\mathscr{K}$ in (5.61) we can write (6.82) more compactly as

$$dy = \left(\frac{\partial y}{\partial t} + \mathscr{K} y\right)dt + \sum_{k=1}^{m}\left(\sum_{i=1}^{n}\sigma_{ik}\frac{\partial y}{\partial x_i}\right)dw_k. \tag{6.83}$$

Equation (6.83) is the most general, multivariable version of Ito's lemma that we shall require. In the more formal stochastic integral equation form (see Sect. 6.4) it would be written

$$y(\mathbf{x}(t), t) = y(\mathbf{x}(0), 0) + \int_0^t \left(\frac{\partial y}{\partial s} + \mathscr{K} y(s)\right) ds + \sum_{k=1}^{m}\sum_{i=1}^{n}\int_0^t \sigma_{ik}(\mathbf{x}(s), s)\frac{\partial y}{\partial x_i}dw_k(s).$$

## 6.6 The Stochastic Differential Equation Followed by the Quotient of Two Diffusions

As an example of an application of Ito's Lemma for several variables, consider the two diffusion processes

$$dx_1 = \mu_1 x_1 dt + \sigma_1 x_1 dw, \tag{6.84}$$

$$dx_2 = \mu_2 x_2 dt + \sigma_2 x_2 dw. \tag{6.85}$$

We want to determine the stochastic differential equation followed by the quotient of these two processes, namely

$$y = \frac{x_1}{x_2}. \tag{6.86}$$

To apply Ito's lemma, we need to calculate

$$\frac{\partial y}{\partial x_1} = \frac{y}{x_1}, \quad \frac{\partial y}{\partial x_2} = -\frac{y}{x_2},$$

$$\frac{\partial^2 y}{\partial x_1^2} = 0, \quad \frac{\partial^2 y}{\partial x_1 \partial x_2} = -\frac{1}{x_2^2} \text{ and } \frac{\partial^2 y}{\partial x_2^2} = \frac{2x_1}{x_2^3}.$$

Thus (see Eq. (6.73))

$$dy = \left[\mu_1 x_1 \frac{1}{x_2} + \mu_2 x_2\left(-\frac{y}{x_2}\right) + \frac{1}{2}\left\{0 + 2\sigma_1 x_1 \sigma_2 x_2\left(\frac{-1}{x_2^2}\right) + \sigma_2^2 x_2^2 \frac{2x_1}{x_2^3}\right\}\right]dt$$
$$+ \left[\sigma_1 x_1 \frac{y}{x_1} + \sigma_2 x_2\left(\frac{-y}{x_2}\right)\right]dw, \tag{6.87}$$

i.e.

$$dy = [(\mu_1 - \mu_2)y + \{-\sigma_1\sigma_2 y + \sigma_2^2 y\}]dt + (\sigma_1 - \sigma_2)ydw.$$

Simplifying further we finally have

$$dy = [(\mu_1 - \mu_2) + \sigma_2(\sigma_2 - \sigma_1)]ydt + (\sigma_1 - \sigma_2)ydw.$$

Another useful way to obtain the same result is by formal manipulations of stochastic differentials using the basic rules of stochastic calculus. Thus calculate $dY$ directly as

$$dy = \frac{x_1 + dx_1}{x_2 + dx_2} - \frac{x_1}{x_2} = \frac{x_1}{x_2}\frac{\left(1 + \frac{dx_1}{x_1}\right)}{\left(1 + \frac{dx_2}{x_2}\right)} - \frac{x_1}{x_2}. \tag{6.88}$$

Recalling the definition of $y$ (Eq. (6.86)) this last equation can be written

$$\begin{aligned}
\frac{dy}{y} &= \frac{1 + \frac{dx_1}{x_1}}{1 + \frac{dx_2}{x_2}} - 1 \\
&= \left(1 + \frac{dx_1}{x_1}\right)\left(1 - \frac{dx_2}{x_2} + \left(\frac{dx_2}{x_2}\right)^2\right) - 1 + o(dt) \\
&= \frac{dx_1}{x_1} - \frac{dx_2}{x_2} - \frac{dx_1}{x_1}\frac{dx_2}{x_2} + \left(\frac{dx_2}{x_2}\right)^2 + o(dt).
\end{aligned} \tag{6.89}$$

By the rules of stochastic calculus

$$\begin{aligned}
\frac{dx_1}{x_1}\frac{dx_2}{x_2} &= (\mu_1 dt + \sigma_1 dw)(\mu_2 dt + \sigma_2 dw) = \sigma_1\sigma_2(dw)^2 = \sigma_1\sigma_2 dt, \\
\left(\frac{dx_2}{x_2}\right)^2 &= (\mu_1 dt + \sigma_2 dw)^2 = \sigma_2^2(dw)^2 = \sigma_2^2 dt.
\end{aligned} \tag{6.90}$$

Thus

$$\frac{dy}{y} = [(\mu_1 - \mu_2) + \sigma_2(\sigma_2 - \sigma_1)]dt + (\sigma_1 - \sigma_2)dw \tag{6.91}$$

as obtained previously.

## 6.7 Problems

**Problem 6.1** In relation to the log-normal asset price process discussed is Sect. 6.3.4, show that

$$\mathrm{var}_t[x(T)] = x(t)^2 e^{2\mu(T-t)}[e^{\sigma^2(T-t)} - 1].$$

Prove that more generally

$$\mathbb{E}_t[x(T)^n] = x(t)^n \exp\left[n\mu(T-t) + \frac{\sigma^2 n(n-1)}{2}(T-t)\right].$$

**Problem 6.2** Verify the results (6.24) and (6.25).

**Problem 6.3** Consider again the Ornstein–Uhlenbeck Process (6.42) but now with time varying coefficients, viz.

$$dx(t) = -k(t)xdt + \sigma(t)dz(t).$$

Show that

$$x(t) = x(0)e^{-\int_0^t k(s)ds} + \int_0^t e^{-\int_u^t k(s)ds}\sigma(u)dz(u),$$

and hence that

$$\mathrm{var}_0[x(t)] = \int_0^t e^{-2\int_u^t k(s)ds}\sigma^2(u)du.$$

**Problem 6.4** Consider the process $y$ defined by

$$y(t) = \sigma \int_0^t (\alpha + (1-\alpha)e^{-\kappa(t-s)})dw(s) + y_0 - (1-\alpha)(1-e^{-\kappa t})x_0, \tag{6.92}$$

where $\sigma, \alpha, \kappa, y_0, x_0$ are constants and $w$ is a Wiener process.

(i) Explain why this process in non-Markovian;
(ii) Show that by defining a second process $x$ as

$$x(t) = x_0 e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-s)}dw(s), \tag{6.93}$$

the system for $x, y$ can be written as the jointly Markovian system

$$dx = -\kappa x(t)dt + \sigma dw(t), \tag{6.94}$$

$$dy = -\kappa(1-\alpha)x(t)dt + \sigma dw(t). \tag{6.95}$$

**Problem 6.5** Consider the stochastic differential equation

$$dx = \mu(x)dt + \sigma(x)dz,$$

and define the transformation

$$y = g(x) \equiv \int_{xo}^{x} \frac{du}{\sigma(u)},$$

for some arbitrary $x_0$. Apply Ito's lemma to show that $y$ satisfies the stochastic differential equation

$$dy = m(y)dt + dz,$$

where

$$m(y) \equiv \left( \frac{\mu(g^{-1}(y))}{\sigma(g^{-1}(y))} - \frac{1}{2}\sigma'(g^{-1}(y)) \right).$$

**Problem 6.6** Consider the so-called CEV (constant elasticity of variance) process for the process $Y$ is of the form

$$dy = \mu_1 y dt + \sigma_1 y^{\beta} dw,$$

where $w$ is a Wiener process under the measure $\mathbb{P}$. Define a new process

$$x = y^{2-2\beta}$$

and show that $x$ satisfies

$$dx = (\mu x + \kappa)dt + \sigma\sqrt{x}dw,$$

where

$$\mu = (2-2\beta)\mu_1, \quad \sigma = (2-2\beta)\sigma_1 \quad \text{and} \quad \kappa = (1-\beta)(1-2\beta)(\sigma_1)^2.$$

**Problem 6.7** Consider the two diffusion processes

$$\frac{dx_1}{x_1} = \mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2,$$
$$\frac{dx_2}{x_2} = \mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2,$$

where the $\mu_i$ and $\sigma_{ij}$ are constant and $w_1$ and $w_2$ are independent Wiener processes. Show that $y_1 = \ln x_1$, and $y_2 = \ln x_2$ satisfy

$$dy_1 = (\mu_1 - \frac{1}{2}(\sigma_{11}^2 + \sigma_{12}^2))dt + \sigma_{11}dw_1 + \sigma_{12}dw_2,$$

$$dy_2 = (\mu_2 - \frac{1}{2}(\sigma_{21}^2 + \sigma_{22}^2))dt + \sigma_{21}dw_1 + \sigma_{22}dw_2.$$

By integrating from $t$ to $T$ show that the vector of random variables

$$(y_1(T) - y_1(t), \quad y_2(T) - y_2(t))$$

are jointly (bi-variate) normally distributed with mean

$$(\mu_1 - \frac{1}{2}(\sigma_{11}^2 + \sigma_{12}^2))(T - t), \quad (\mu_2 - \frac{1}{2}(\sigma_{21}^2 + \sigma_{22}^2))(T - t)$$

and variance-covariance structure

$$var_t[y_1(T)] = (\sigma_{11}^2 + \sigma_{12}^2)(T - t),$$

$$var_t[y_2(T)] = (\sigma_{21}^2 + \sigma_{22}^2)(T - t),$$

$$cov_t[y_1(T), y_2(T)] = (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})(T - t).$$

It is known (see Abramowitz and Stegun 1970) that if $u_1, u_2$ are two normal variables with

$$\mathbb{E}[u_1] = \mathbb{E}[u_2] = 0,$$

$$var[u_1] = s_1, \quad var[u_2] = s_2,$$

$$corr[u_1, u_2] = \frac{cov[u_1, u_2]}{s_1 s_2} = \rho,$$

then they have joint (or bivariate) probability density function

$$\pi(u_1, u_2) = \frac{1}{2\pi s_1 s_2 \sqrt{1 - \rho^2}} \exp\left[\frac{-(u_1^2 - 2\rho u_1 u_2 + u_2^2)}{2 s_1^2 s_2^2 (1 - \rho^2)}\right].$$

Use this result to show that the joint (or bivariate) probability density function for $y_1, y_2$ is

$$p(y_1(T), y_2(T), T | y_1(t), y_2(t), t)$$
$$= \frac{1}{2\pi s_1 s_2 \sqrt{1 - \rho^2 (T - t)}} \exp\left[\frac{-(u_1^2 - 2\rho u_1 u_2 + u_2^2)}{2 s_1^2 s_2^2 (1 - \rho^2)(T - t)}\right],$$

where

$$u_i = y_i(T) - y_i(t) - (\mu_i - \frac{1}{2}s_i^2), (i = 1, 2),$$
$$s_1^2 = \sigma_{11}^2 + \sigma_{12}^2, \;\; s_2^2 = \sigma_{21}^2 + \sigma_{22}^2, \;\; \rho = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}.$$

**Problem 6.8** Consider the two diffusion processes

$$dx_1 = x_1(a_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2),$$
$$dx_2 = x_2(a_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2),$$

where $w_1$ and $w_2$ are independent Wiener processes. Find the diffusion process followed by

$$y = x_1 x_2.$$

First apply Ito's lemma for several variables. Then use the definition of $dy$, namely

$$dy = (x_1 + dx_1)(x_2 + dx_2) - x_1 x_2$$

and the rules of stochastic calculus, to obtain the same result.

**Problem 6.9** Consider the quantities

$$y_i(t) = \int_0^t \sigma_i(u) dw_i(u), \qquad y_j(t) = \int_0^t \sigma_j(u) dw_j(u),$$

where $w_i(t), w_j(t)$ are independent Wiener processes when $i \neq j$. Use Ito's lemma to find the stochastic differential equation followed by

$$z(t) = y_i(t) y_j(t).$$

Hence show that

$$\mathbb{E}_0\left[\left(\int_0^t \sigma_i(u) dw_i(u)\right)\left(\int_0^t \sigma_j(u) dw_j(u)\right)\right] = \begin{cases} 0, i \neq j \\ \int_0^t \mathbb{E}_0[\sigma_i{}^2(u)] du, i = j \end{cases}.$$

**Problem 6.10** Consider the two diffusion processes

$$dx_1(t) = m_1(t) dt + \sigma_{11}(t) dw_1(t) + \sigma_{12}(t) dw_2(t),$$
$$dx_2(t) = m_2(t) dt + \sigma_{21}(t) dw_1(t) + \sigma_{22}(t) dw_2(t),$$

where $w_1(t)$ and $w_2(t)$ are independent Wiener processes. Use the techniques of Problem 6.9 to show that

$$\text{cov}_0\left[x_1(t), x_2(t)\right] = \int_0^t \sigma_{11}(u)\sigma_{21}(u)du + \int_0^t \sigma_{12}(u)\sigma_{22}(u)du.$$

**Problem 6.11** Suppose $x$ is driven by the Ito stochastic differential equation

$$dx(t) = \mu(x,t)dt + \sigma(x,t)dw(t).$$

The process $y(t)$ is defined by

$$y(t) = \int_0^t g(x(s))ds$$

where $g$ is suitably well-defined function. Calculate

$$\mathbb{E}_0[y(t)] \quad \text{and} \quad \text{var}_0[y(t)].$$

**Problem 6.12** Consider the stochastic differential equation

$$dx(t) = (\alpha(t) + \beta(t)x)dt + \sigma(t)dw,$$

where $\alpha$, $\beta$ and $\sigma$ are time deterministic functions. Show that

$$x(t) = e^{\int_0^t \beta(s)ds}\left[x(0) + \int_0^t \alpha(s)e^{-\int_0^s \beta(u)du}ds + \int_0^t \sigma(s)e^{-\int_0^s \beta(u)du}dw(s)\right].$$

Explain why $x(t)$ is normally distributed and show that

$$\mathbb{E}_0[x(t)] = e^{\int_0^t \beta(s)ds}\left[x(0) + \int_0^t \alpha(s)e^{-\int_0^s \beta(u)du}ds\right],$$

and

$$\text{var}_0[x(t)] = \int_0^t \sigma^2(s)e^{-2\int_t^s \beta(u)du}ds.$$

**Problem 6.13** Consider the linear stochastic differential equation

$$dx = (\alpha(t) + \beta(t)x)dt + (\gamma(t) + \delta(t)x)dw,$$

for some time deterministic functions $\alpha$, $\beta$, $\gamma$ and $\delta$. Consider also the related stochastic differential equations

$$du = \beta(t)udt + \delta(t)udw, \quad u(0) = 1,$$

and

$$dv = a(t)dt + b(t)dw, \quad v(0) = x(0),$$

where $a(t)$ and $b(t)$ are also time deterministic functions. Use the approach to Sect. 6.3.2 to solve for $u(t)$. Then show by appropriate choice of the $a(t)$ and $b(t)$ that it is possible to express $x(t)$ as

$$x(t) = u(t)v(t).$$

Solve for $v(t)$ and hence show that $x(t)$ can be expressed as

$$x(t) = u(t)\left(x(0) + \int_0^t \frac{(\alpha(s) - \delta(s)\gamma(s))}{u(s)}ds + \int_0^t \frac{\gamma(s)}{u(s)}dw(s)\right).$$

**Problem 6.14** Consider the set of diffusion processes

$$\frac{dS_i}{S_i} = \mu_i(S,t)dt + \sum_{j=0}^{m} s_{ij}(S,t)dw_j(t),$$

for $i = 0, 1, \ldots, n$, where $S$ denote the vector $(S_0, S_1, \ldots, S_n)$ and the $w_j$ are independent Wiener processes. Define the set of processes

$$y_i(t) = S_i(t)/S_0(t), \ (i = 1, 2, \ldots, n).$$

Show that these processes satisfy

$$\frac{dy_i}{y_i} = \left[(\mu_i - \mu_0) - \sum_{j=0}^{m} s_{oj}(s_{ij} - s_{oj})\right]dt + \sum_{j=0}^{m}(s_{ij} - s_{oj})dw_j.$$

**Problem 6.15** Consider the two diffusion processes

$$\frac{dx_1}{x_1} = \mu_1 dt + \sigma_{11}dw_1 + \sigma_{12}dw_2, \tag{6.96}$$

$$\frac{dx_2}{x_2} = \mu_2 dt + \sigma_{21}dw_1 + \sigma_{22}dw_2. \tag{6.97}$$

Show that the process $y = x_1/x_2$ satisfies the stochastic differential equation

$$\begin{aligned}\frac{dy}{y} =& [\mu_1 - \mu_2 + \left(\sigma_{21}^2 + \sigma_{22}^2\right) - (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})]dt \\ &+ (\sigma_{11} - \sigma_{21})dw_1 + (\sigma_{12} - \sigma_{22})dw_2.\end{aligned}$$

**Problem 6.16 Computational Problem**—Use the solutions (6.16) and (6.17) to simulate the stock price process $x(t)$ and the return process $\ln(x(t)/x(0))$ in the interval $(0, T)$ and in particular obtain the simulated distribution for these quantities. Use the same values for $x(0)$, $\mu$, $\sigma$ and $T$ as used in Problem 4.6(b).

Since now it is possible to draw the $(z(T) - z(0))$ directly from a normal distribution, discretisation error is avoided. Gauge the impact of the discretisation error by comparing the distributions obtained here with the ones obtained in Problem 4.6(b) and the true distribution.

# Chapter 7
# The Continuous Hedging Argument

**Abstract** This chapter develops a continuous hedging argument for derivative security pricing. Following fairly closely the original Black and Scholes (1973) article, we make use of Ito's lemma to derive the expression for the option value and exploit the idea of creating a hedged position by going long in one security, say the stock, and short in the other security, the option. Alternative hedging portfolios based on Merton's approach and self financing strategy approach are also introduced.

## 7.1 The Continuous Hedging Argument: The Black–Scholes Approach

Black and Scholes in surveying some of the earlier attempts to price options, discuss how the idea that they develop so effectively had already arisen in some of the earlier literature. It is worth quoting here one of the key paragraphs in their now celebrated paper where they reveal their key insight that the expected return on a hedged position must equal the return on the risk free asset in equilibrium:

> *One of the concepts that we use in developing our model is expressed by* Thorp and Kassouf (1967). *They obtain an empirical valuation formula for warrants by fitting a curve to actual warrant prices. Then they use this formula to calculate the ratio of shares of stock to options needed to create a hedged position by going long in one security and short in the other. What they fail to pursue is the fact that in equilibrium, the expected return on such a hedged position must be equal to the return on a riskless asset. What we show below is that this equilibrium condition can be used to derive a theoretical valuation formula.*

Let $S$ denote the stock price and $C$ the value of a European call option written on the stock. The following "ideal market conditions" are assumed:

(a) the short term risk free interest rate is known and is constant through time;
(b) the stock price follows a stochastic process described by the stochastic differential equation

$$\frac{dS}{S} = \mu dt + \sigma dz; \tag{7.1}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_7

(c) the stock pays no dividend;
(d) there are no transaction costs in buying or selling the stock or the option;
(e) it is possible to borrow, at the short-term risk-free interest rate, any fraction of the price of a security to buy it or hold it;
(f) there are no penalties to short selling.

Our starting assumption is that the value of the option will depend on the stock price $S$ i.e.

$$C = C(S, t).$$

Denote

$$\theta = \frac{\partial C}{\partial t}, \quad \Delta = \frac{\partial C}{\partial S}, \quad \Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 C}{\partial S^2}. \tag{7.2}$$

Given that the stock price follows the stochastic differential equation (7.1), an application of Ito's lemma in the previous chapter leads to that the option price satisfies the stochastic differential equation

$$\frac{dC}{C} = \mu_c dt + \sigma_c dz, \tag{7.3}$$

where

$$\mu_c = \frac{\left(\theta + \mu S\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma\right)}{C}, \tag{7.4}$$

can be interpreted as the option return per unit time, and

$$\sigma_c = \frac{\sigma S \Delta}{C}, \tag{7.5}$$

as the option volatility per unit time.

Now set up a hedging portfolio containing quantities $Q_s$ stock and $Q_c$ options (we shall adopt the convention that $Q > 0$ indicates a long position whilst $Q < 0$ indicates a short position). If we use $V_H$ to denote the value of this portfolio then

$$V_H = SQ_s + CQ_c. \tag{7.6}$$

Over the time interval $(t, t + dt)$ the change in the value of the hedging portfolio is

$$dV_H = Q_s dS + Q_c dC,$$

which after use of the expressions (7.1) and (7.3) for $dS$ and $dC$ becomes

$$dV_H = [\mu S Q_s + \mu_c C Q_c]\, dt + [\sigma S Q_s + \sigma_c C Q_c] dz. \tag{7.7}$$

As with the stochastic differential equations for the stock price and option price, the one for the value of the hedging portfolio has a drift term and a diffusion (or stochastic) term. The crucial observation is that by varying $Q_s$ and $Q_c$ we can vary the coefficient of the diffusion term. In fact, by an appropriate choice of $Q_s/Q_c$ we can make the (stochastic) diffusion term vanish entirely. This result is achieved by choosing $Q_s$ and $Q_c$ so that

$$\sigma S Q_s + \sigma_c C Q_c = 0,$$

i.e.

$$\frac{Q_s}{Q_c} = -\Delta, \tag{7.8}$$

where we have used the definition of $\sigma_c$ from (7.5). This result tells us that for every long (short) option position in the hedging portfolio, we must take $\Delta$ short (long) positions in the underlying stock. If this ratio of stock to options is continually maintained then the hedging portfolio yields a certain return over time interval $(t, t + dt)$ given by

$$dV_H = [-\mu S \Delta Q_c + \mu_c C Q_c]\, dt = Q_c(-\mu S \Delta + \mu_c C)dt, \quad \text{(using (7.8))}. \tag{7.9}$$

Since this return is riskless, it must be the case that in an efficient capital market the original hedging portfolio (with proportion of stock to option satisfying $Q_s = -\Delta Q_c$) must earn the short-term risk-free rate, i.e.

$$dV_H = r\, V_H dt, \tag{7.10}$$

where we use $r$ to denote the short term risk-free rate of interest. If this were not the case then it would be possible to set up a riskless arbitrage strategy to profit from the strategy (either investment in the hedging portfolio or in the riskless instrument) giving the higher sure return.

With $Q_s = -\Delta Q_c$ the expression for $V_H$ at Eq. (7.6) becomes

$$V_H = Q_c(-S\Delta + C). \tag{7.11}$$

Use of (7.9) and (7.11) reduces the no riskless arbitrage condition (7.10) to

$$-\mu S \Delta + \mu_c C = r(-S\Delta + C), \tag{7.12}$$

i.e.

$$\mu_c - r = (\mu - r)\frac{S\Delta}{C}, \tag{7.13}$$

which upon use of (7.5) may be written

$$\frac{\mu_c - r}{\sigma_c} = \frac{\mu - r}{\sigma}. \tag{7.14}$$

Recalling the expressions (7.4) and (7.5) for $\mu_c$ and $\sigma_c$, (7.14) becomes

$$\theta + rS\Delta + \frac{1}{2}\sigma^2 S^2 \Gamma = rC. \tag{7.15}$$

Referring back to the expressions for $\theta$, $\Delta$ and $\Gamma$ this last condition appears as

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC, \tag{7.16}$$

which is the partial differential equation for the option price, which we obtained in Chap. 3 (Eq. (3.7)) by discounting, at the risk free rate, the expected payoff of the option.

In Chap. 9 we outline a systematic approach to the solution of the partial differential equations of financial economics, of which (7.16) with boundary conditions for a European call option is a particular example. The solution is of course the same as obtained in Chap. 3 (Eq. (3.9)) by integration. In Appendix 9.1 we also show the original solution technique based on transforming (7.16) into the heat equation, whose solution may then be applied to obtain the Black–Scholes formula.

Finally we make the observation that (7.9) may be written

$$dV_H = Q_c\left(\theta + \frac{1}{2}\sigma^2 S^2 \Gamma\right) dt \tag{7.17}$$

upon use of the expression for $\mu_c$. Equation (7.17) is important for practical considerations of hedging strategies. It indicates how the value of the portfolio that is now hedged instantaneously changes as the value of the underlying stock changes. This change is in particular sensitive to the option's $\theta$ (i.e. its time decay) and its $\Gamma$ (i.e. its convexity). We refer the reader to Hull (2000) for further discussion of these issues.

## 7.2 Interpreting the No-Arbitrage Condition

Equation (7.14) can be given a simple economic interpretation. The right-hand side is the expected excess return on the stock, risk-adjusted by its standard deviation of return, whilst the left-hand side is the corresponding quantity for the option. The no-arbitrage condition (7.14) merely states that in an efficient capital market the option will be priced such that the expected excess return on the option risk-adjusted equals the expected excess return on the underlying stock risk adjusted.

We shall denote this common risk-adjusted expected excess return by $\lambda$ , so that from (7.14) we may write

$$\mu = r + \lambda\sigma, \tag{7.18}$$

and

$$\mu_c = r + \lambda\sigma_c. \tag{7.19}$$

The terms $\lambda\sigma$ and $\lambda\sigma_c$ can be interpreted as risk premia above the risk free rate required by investors to hold the stock and option respectively. We may in fact interpret $\lambda$ as the **market price of risk** or **Sharpe ratio** for the uncertainty associated with the $dz$ term in the stochastic differential equations for $S$ and $C$. The sense of this interpretation can be appreciated by noting that $\lambda$ is the extra premium required for a unit increase in $\sigma$ in the stock market and a unit increase in $\sigma_c$ in the option market.

It is also possible to relate the condition (7.14) to the capital asset pricing model (CAPM) of modern portfolio theory. We show in Appendix 7.1 that if $\beta$ and $\beta_c$ represent the CAPM beta of the stock and of the option respectively, then

$$\frac{\beta_c}{\beta} = \frac{\sigma_c}{\sigma}. \tag{7.20}$$

Thus the condition (7.14) may be rewritten

$$\frac{\mu_c - r}{\beta_c} = \frac{\mu - r}{\beta}. \tag{7.21}$$

This last condition has a simple interpretation in terms of the CAPM. The top line on each side is the excess return of each security above the risk free rate, the bottom line is the security's beta which is a measure of its riskiness. The condition (7.21) merely states that the excess return on each security, risk adjusted by its beta, is the same for both the option and the stock. This result shows that the hedging argument technique is compatible with modern portfolio theory. According to this theory the above risk adjusted excess returns should equal the excess return on the market portfolio, i.e.

$$\frac{\mu_c - r}{\beta_c} = \frac{\mu - r}{\beta} = \mu_m - r, \tag{7.22}$$

where $\mu_m$ is the expected return on the market portfolio. By comparing (7.18) and (7.19) with (7.22) and making use of (7.20) we see that the market price of risk can also be related to the expected excess return on the market portfolio and the stock beta, $\beta$, or the option beta, $\beta_c$, according to

$$\lambda = (\mu_m - r)\frac{\beta}{\sigma}, \tag{7.23}$$

or

$$\lambda = (\mu_m - r)\frac{\beta_c}{\sigma_c}. \tag{7.24}$$

Black and Scholes, in their original article give an alternative derivation that starts from the CAPM relationship for the stock and the option, i.e.

$$\mu = r + (\mu_m - r)\beta, \tag{7.25}$$

$$\mu_c = r + (\mu_m - r)\beta_c, \tag{7.26}$$

where $\mu_m$ is the expected return on the market portfolio. Eliminating $(\mu_m - r)$ yields (7.21), the relation (7.20) connecting $\beta_c/\beta$ to $\sigma_c/\sigma$ may be used to obtain (7.14), which upon use of the expression for $\mu_c$ and $\sigma_c$ from Ito's lemma reduces to (7.15) which is the Black–Scholes option pricing partial differential equation once again.

The approach of starting from (7.14), the condition that risk adjusted excess return should be the same across all securities, turns out to be a very helpful idea and will lead us into the general approach for pricing derivative securities, which we will discuss in Chap. 10.

## 7.3 Alternative Hedging Portfolios: The Merton's Approach

It is possible to arrive at the no-riskless arbitrage condition (7.14) by setting up a number of alternative hedging portfolios. The hedging portfolio that we set up in Sect. 7.1 consisted of positions in the stock and option and focused on the quantities of these in the portfolio. This section considers a slightly different way of setting up the hedging portfolio, which was used by Merton in an appendix to Samuelson (1973). This approach sets up a hedging portfolio consisting of positions in the stock, option and risk-free instrument. It also makes explicit how the hedging portfolio may be set up with zero net investment and focuses on the dollars invested in each asset rather than the quantity of assets purchased.

We set up a portfolio $V$ of stock, options and riskless bonds with

$Q_1$ = dollar amount invested in the stock,
$Q_2$ = dollar amount invested in the option,
$Q_3$ = dollar amount invested in the bond.

By short selling or borrowing we can constrain the portfolio to require net zero investment, which implies that

$$V = Q_1 + Q_2 + Q_3 = 0. \tag{7.27}$$

The instantaneous proportional change in the portfolio value is given by

$$\begin{aligned}\frac{dV}{V} &= Q_1\frac{dS}{S} + Q_2\frac{dC}{C} + Q_3 r\,dt \\ &= Q_1\left(\frac{dS}{S} - r\,dt\right) + Q_2\left(\frac{dC}{C} - r\,dt\right) \quad [\text{ using (7.27)}] \\ &= [Q_1(\mu - r) + Q_2(\mu_c - r)]dt + (Q_1\sigma + Q_2\sigma_c)dz \end{aligned} \tag{7.28}$$

using the stochastic differential equations for $S$ and $C$. Choose $Q_1/Q_2$ so that the stochastic term in (7.28) vanishes, yielding the condition

$$\frac{Q_1}{Q_2} = -\frac{\sigma_c}{\sigma}, \tag{7.29}$$

then the change in portfolio value is

$$\frac{dV}{V} = Q_2\left[-\frac{\sigma_c}{\sigma}(\mu - r) + (\mu_c - r)\right]dt. \tag{7.30}$$

Since no net investment was required to set up the portfolio $V$, by no arbitrage condition, it must be zero

$$-\frac{\sigma_c}{\sigma}(\mu - r) + (\mu_c - r) = 0,$$

which reduces to

$$\frac{\mu - r}{\sigma} = \frac{\mu_c - r}{\sigma_c}.$$

The last equation is (7.14) again, and is equivalent to the Black–Scholes option pricing equation as we have already derived.

## 7.4 Self Financing Strategy: The Modern Approach

In this section we derive the option pricing equation by using so-called self-financing strategies. The approach we develop here will also make use of the stochastic integral form of Ito's lemma (see Sect. 6.4). Since this is the mathematically correct way to view Ito's lemma, we avoid the "sloppy" mathematical reasoning that accompanied the use of the stochastic differentials in the arguments of the previous sections. The approach we develop in this section is the basis of what is now termed the modern approach, that would be regarded as mathematically more complete.

Suppose we invest in two financial assets whose values at time $t$ are $x_1(t)$ and $x_2(t)$ that are driven by the Ito stochastic differential equations

$$dx_1 = \mu_1(x,t)dt + \sigma_1(x,t)dz(t), \tag{7.31}$$

and

$$dx_2 = \mu_2(x,t)dt + \sigma_2(x,t)dz(t), \tag{7.32}$$

where $x$ denotes the vector $(x_1, x_2)$ and we assume the same Wiener process $z(t)$ drives both prices. Denote by $Q_1(t)$ and $Q_2(t)$ the amount invested in assets 1 and 2 respectively at time $t$, and by $V(t)$ the value of the portfolio of the two assets. We consider the investment strategy over a time interval $(0, \tau)$, when $\tau$ will later be the option maturity. We will initially develop the trading strategy in terms of the discrete partitioning of the time interval used in Sect. 4.7, only here we assume a fixed time step of length $\Delta t$.

Consider the evolution of the portfolio value over the time interval $(i\Delta t, (i+1)\Delta t)$. At time $i\Delta t$ we decide to invest the amounts $Q_1(i\Delta t)$ and $Q_2(i\Delta t)$ in $x_1$ and $x_2$ respectively and hold these amounts until time $(i+1)\Delta t$, when we will revise our portfolio balance. The change in the portfolio value over the time interval will be given by

$$\begin{aligned} V((i+1)\Delta t) - V(i\Delta t) = {} & Q_1(i\Delta t)\left[x_1((i+1)\Delta t) - x_1(i\Delta t)\right] \\ & + Q_2(i\Delta t)\left[x_2((i+1)\Delta t) - x_2(i\Delta t)\right]. \end{aligned} \tag{7.33}$$

If we let time $\tau = n\Delta t$, then summing over $n$ subintervals we obtain the value of the portfolio at time $\tau$ as

$$\begin{aligned} V(\tau) - V(0) = {} & \sum_{i=0}^{n-1} Q_1(i\Delta t)\left[x_1((i+1)\Delta t) - x_1(i\Delta t)\right] \\ & + \sum_{i=0}^{n-1} Q_2(i\Delta t)\left[x_2((i+1)\Delta t) - x_2(i\Delta t)\right]. \end{aligned} \tag{7.34}$$

Using the definition of the stochastic differential prior to going to the limit (see Eq. (4.49) we can also express (7.34) as

$$\begin{aligned} V(\tau) - V(0) = {} & \sum_{i=0}^{n-1} \left[Q_1(i\Delta t)\mu_1(x(i\Delta t), i\Delta t) + Q_2(i\Delta t)\mu_2(x(i\Delta t), i\Delta t)\right]\Delta t \\ & + \sum_{i=0}^{n-1} \left[Q_1(i\Delta t)\sigma_1(x(i\Delta t), i\Delta t) + Q_2(i\Delta t)\sigma_2(x(i\Delta t), i\Delta t)\right]\Delta z(i\Delta t). \end{aligned} \tag{7.35}$$

By applying the idea of the mean square limit (see Sect. 4.7) Eq. (7.35) in the limit can be written in integral form as

$$
\begin{aligned}
V(\tau) - V(0) = \int_0^\tau & [Q_1(s)\mu_1(x(s), s) + Q_2(s)\mu_2(x(s), s)]\, ds \\
& + \int_0^\tau [Q_1(s)\sigma_1(x(s), s) + Q_2(s)\sigma_2(x(s), s)]\, dz(s).
\end{aligned}
\tag{7.36}
$$

Equation (7.36) merely tells us how our portfolio value evolves if we follow the trading strategy $(Q_1(t), Q_2(t))$ at each time $t$. A trading strategy is said to be **self-financing** if the increase in value of the portfolio $V(t)$ arises only from changes in the prices of the assets $x_1$ and $x_2$. In other words we do not require the inflow of some external source of cash to finance the strategy.[1] This condition imposes on $Q_1(t)$ and $Q_2(t)$ that

$$
V(t) = Q_1(t)x_1(t) + Q_2(t)x_2(t), \tag{7.37}
$$

so that we cannot spread across the two assets any more wealth than what we have in our portfolio.

We have deliberately gone back to the definition of the stochastic integral in deriving Eq. (7.36) for the evolution of portfolio value. In subsequent discussion of the setting up of self-financing portfolio strategies we proceed by going directly to the limit in Eq. (7.34) to write

$$
V(\tau) - V(0) = \int_0^\tau Q_1(s)dx_1(s) + \int_o^\tau Q_2(s)dx_2(s), \tag{7.38}
$$

which upon use of the stochastic differential equations (7.31) and (7.32) and some re-arrangement reduces to (7.36).

We now apply the foregoing concept of a self-financing trading strategy to the option pricing problem. We shall consider two portfolios, one consisting of a position in the option, the other consisting of a self-financing strategy portfolio in the underlying stock and a bond that pays the risk-free rate of interest.

Consider first the portfolio $V_1$ consisting solely of a position in the option over the interval $(t, T)$. By definition we have

$$
V_1(T) - V_1(t) = C(S_T, T) - C(S, t), \tag{7.39}
$$

[1] Nor is the strategy required to generate some cash outflow, say in the form of dividends.

which by application of Ito's lemma in integral form (see (6.64)) may be written

$$V_1(T) - V_1(t) = \int_t^T \left[ \frac{\partial C(S,u)}{\partial u} + \mu S \frac{\partial C(S,u)}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C(S,u)}{\partial S^2} \right] du + \int_t^T \sigma S \frac{\partial C(S,u)}{\partial S} dz(u). \tag{7.40}$$

Now consider a portfolio $V_2$ consisting of a self-financing strategy in the underlying stock and the bond. If the value of the bond at time $t$ is denoted $B(t)$ then its value evolves according to

$$dB = rBdt, \tag{7.41}$$

which does not involve any stochastic term. Letting $Q_S$ and $Q_B$ denote the positions in the stock and bond respectively, so the evolution of the self-financing portfolio $V_2$ over the interval $t$ to $T$ can be obtained from (7.36) (after appropriate re-interpretation of symbols[2]) as

$$V_2(T) - V_2(t) = \int_t^T [Q_S(u)\mu S(u) + Q_B(u) r B(u)]\, du + \int_t^T Q_S(u)\sigma S(u) dz(u),$$

or, after application of the self-financing condition (7.37)[3]

$$V_2(T) - V_2(t) = \int_t^T [Q_S(u)(\mu - r)S(u) + rV(u)]\, du + \int_t^T Q_S(u)\sigma S(u) dz(u). \tag{7.42}$$

If the evolution of the self-financing strategy portfolio $V_2$ is to be the same as the portfolio $V_1$ consisting of the position in the option (i.e. (7.40) and (7.42) yield the same evolution for $V_1(t) = V_2(t) = V(t)$) then the corresponding integrands in (7.40) and (7.42) must be equal. Thus, from comparison of the stochastic integrals (involving the $dz(u)$ term) we see that the stock position at time $t$ must be chosen to so that

$$Q_S(t) = \frac{\partial C(S,t)}{\partial S}. \tag{7.43}$$

This of course is the hedge ratio that we obtained in Eq. (7.8), the difference in sign merely being a reflection of the fact that here we are taking a long position in the option. Using (7.43) and the fact that $V(t) = C(S,t)$, comparison of the first

---

[2] In (7.36) we set $\tau \to T$, $0 \to t$, $Q_1 \to Q_S$, $Q_2 \to Q_B$, $\mu_1 \to \mu S$, $\mu_2 \to rB$, $\sigma_1 \to \sigma S$, $\sigma_2 \to 0$.

[3] Which here becomes $V(u) = Q_S(u)S(u) + Q_B(u)B(u)$.

integrals in (7.40) and (7.42) yield the condition

$$\frac{\partial C}{\partial S}(\mu - r)S + rC = \frac{\partial C}{\partial t} + \mu S\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2},$$

which after a little manipulation reduces to

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} = rC, \tag{7.44}$$

the familiar Black–Scholes partial differential equation.

## 7.5 Appendix

### Appendix 7.1 Relation Between Stock and Option Betas

Recall the stochastic differential equations for the stock price and the option price which may be written in return form as

$$\frac{dS}{S} = \mu dt + \sigma dz, \tag{7.45}$$

$$\frac{dC}{C} = \mu_c dt + \sigma_c dz. \tag{7.46}$$

We let $M$ denote the value of the market portfolio and assume that it follows the same type of stochastic process as the stock, i.e.

$$\frac{dM}{M} = \mu_m dt + \sigma_m dz. \tag{7.47}$$

From portfolio theory the definition of the beta factors is

$$\beta = \frac{\text{cov}\left(\frac{dS}{S}, \frac{dM}{M}\right)}{\text{var}\left(\frac{dM}{M}\right)}, \qquad \beta_c = \frac{\text{cov}\left(\frac{dC}{C}, \frac{dM}{M}\right)}{\text{var}\left(\frac{dM}{M}\right)}. \tag{7.48}$$

Since

$$\mathbb{E}\left(\frac{dS}{S}\right) = \mu dt \quad \text{and} \quad \mathbb{E}\left(\frac{dM}{M}\right) = \mu_m dt,$$

it follows that

$$\begin{aligned}\text{cov}\left(\frac{dS}{S}, \frac{dM}{M}\right) &= \mathbb{E}\left[\left(\frac{dS}{S} - \mu dt\right)\left(\frac{dM}{M} - \mu_m\, dt\right)\right] \\ &= \mathbb{E}[(\sigma dz)(\sigma_m dz)] = \sigma\sigma_m \mathbb{E}(dz^2) = \sigma\sigma_m\, dt.\end{aligned}$$

Also var $\left(\frac{dM}{M}\right) = \sigma_m^2 dt$, hence,

$$\beta = \frac{\sigma\sigma_m dt}{\sigma_m^2 dt} = \frac{\sigma}{\sigma_m}. \tag{7.49}$$

Similarly

$$\beta_c = \frac{\sigma_c}{\sigma_m}. \tag{7.50}$$

Eliminating $\sigma_m$ between (7.49) and (7.50) yields

$$\frac{\beta}{\beta_c} = \frac{\sigma}{\sigma_c}, \tag{7.51}$$

which is the result used in the main text.

## 7.6 Problems

**Problem 7.1** Rework the hedging argument of Sect. 7.1 but now the hedging portfolio consists of physical positions $Q_S$ in the stock, $Q_C$ in the option and $Q_B$ in risk free bonds. The risk free bond is an instrument whose market value is $B$ and whose return process is given by

$$\frac{dB}{B} = rdt.$$

**Problem 7.2** (a) Rework the continuous hedging argument of Sect. 7.1 but now allow the underlying asset to pay a continuously compounded dividend at the rate $q$.

Show that in this case the condition of no-riskless arbitrage (7.14) becomes

$$\frac{\mu_c - r}{\sigma_c} = \frac{\mu - (r - q)}{\sigma}.$$

Thus show that Eq. (7.16) becomes

$$\frac{\partial C}{\partial t} + (r - q)S\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC.$$

(b) Adjust the self financing strategy argument of Sect. 7.4 in the case that the underlying asset pays a continuously compounded dividend at the rate $q$.

# Chapter 8
# The Martingale Approach

**Abstract** The martingale approach is widely used in the literature on contingent claim analysis. Following the definition of a martingale process, we give some examples, including the Wiener process, stochastic integral, and exponential martingale. We then present the Girsanov's theorem on a change of measure. As an application, we derive the Black–Scholes formula under risk neutral measure. A brief discussion on the pricing kernel representation and the Feynman–Kac formula is also included.

## 8.1 Martingales

### *8.1.1 Introduction*

The seminal papers of Harrison and Kreps (1979) and Harrison and Pliska (1981) ushered in a new approach to contingent claim pricing. This approach, the martingale approach, which is now widely used in the literature on contingent claim analysis is expressed in the language of the modern theory of stochastic integration, which relies on the abstract theory of semi-martingale integration. To properly come to grips with these concepts would require a lengthy and high level mathematical course involving measure theory. All we shall do here is to define, and try to explain, at the level of mathematical discussion used in previous chapters the main themes of this approach which occur in the modern contingent claims literature. Similar approaches may be found in Sundaran (1997), Baxter and Rennie (1996) and Neftci (2000). An excellent and more rigorous account of the modern martingale approach to derivative security pricing can be found in Musiela and Rutkowski (1997).

The Wiener processes, which drive the stochastic differential equations that are used to model asset prices have two main mathematical characteristics, namely (i) the independence of increments, and (ii) the continuity of the sample paths. If $M(t)$ represents such a stochastic process then the first property essentially says that the probability distribution of $M(t) - M(s)(s < t)$ is not affected by any information about what the process does up to time $s$. This property is also shared by the Poisson process.



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_8

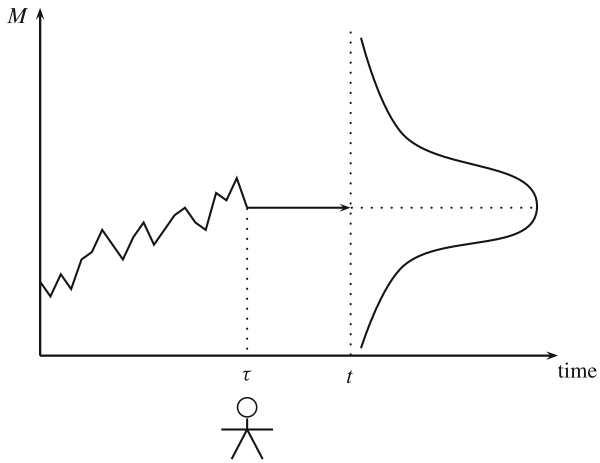


**Fig. 8.1** The distribution of $M(t)$ given information at time $\tau$

The process $M$ is said to be a **martingale process** if

$$\mathbb{E}_\tau(M(t)) = M(\tau), \tag{8.1}$$

for all $t > \tau$. This notion is illustrated in Fig. 8.1 where, given the information at time $\tau$, we regard $M(t)$ as a distribution whose mean is $M(\tau)$. We also refer the reader back to Fig. 3.3 and the related discussion for more intuition on martingale processes. Often prices, or price related functions, turn out to be martingales. In fact the major task in applying martingale methods is to find the price related function that is a martingale.

## *8.1.2 Examples of Martingales*

**(i) The Wiener Process**—Since the increments of the Wiener process have the property

$$\mathbb{E}[dz(t)] = 0, \quad \text{for all } t,$$

it follows that the Wiener process itself is a martingale i.e.

$$\mathbb{E}_\tau[z(t) - z(\tau)] = 0, \quad t > \tau,$$

or

$$\mathbb{E}_\tau[z(t)] = z(\tau), \quad t > \tau.$$

**(ii) The Stochastic Integral**—Consider the stochastic integral

$$x(t) = \int_\tau^t \sigma(x(s), s)dz(s). \tag{8.2}$$

It can alternatively be written as the stochastic differential equation

$$dx(t) = \sigma(x(t), t)dz(t). \tag{8.3}$$

From this last equation we see that

$$\mathbb{E}[dx(t)] = \mathbb{E}[\sigma(x(t), t)dz(t)] = 0.$$

As in the previous example it follows that

$$\mathbb{E}_\tau[x(t) - x(\tau)] = 0,$$

i.e.

$$\mathbb{E}_\tau[x(t)] = x(\tau). \tag{8.4}$$

Hence the stochastic integral (8.2) is a martingale. By further noting that $x(\tau) = 0$ we can also express this result as

$$\mathbb{E}_\tau\left[\int_\tau^t \sigma(x(s), s)dz(s)\right] = 0. \tag{8.5}$$

### *8.1.3 The Exponential Martingale*

Consider the process $u(t)$ defined by

$$du(t) = \mu(u, t)dt + \sigma(u, t)dz(t), \qquad u(0) = 0. \tag{8.6}$$

Recall that (8.6) may also be expressed as the stochastic integral equation

$$u(t) = \int_0^t \mu(u(s), s)ds + \int_0^t \sigma(u(s), s)dz(s). \tag{8.7}$$

Now define the quantity

$$
\begin{aligned}
M(t) &= \exp\left(u(t) - \int_0^t [\mu(u(s), s) + \frac{1}{2}\sigma^2(u(s), s)]ds\right) \\
&= \exp\left(-\frac{1}{2}\int_0^t \sigma^2(u(s), s)ds + \int_0^t \sigma(u(s), s)dz(s)\right).
\end{aligned} \tag{8.8}
$$

For notational convenience define the subsidiary process

$$
v(t) = -\frac{1}{2}\int_0^t \sigma^2(u(s), s)ds + \int_0^t \sigma(u(s), s)dz(s), \tag{8.9}
$$

which can be written as the stochastic differential equation

$$
dv(t) = -\frac{1}{2}\sigma^2(u, t)dt + \sigma(u, t)dz(t).
$$

Then $M(t)$ in Eq. (8.8) can be simply written

$$
M(t) = e^{v(t)}. \tag{8.10}
$$

It is a straight forward application of Ito's lemma (see Eq. (6.26) in Sect. 6.3) to see that $M(t)$ satisfies the stochastic differential equation

$$
dM(t) = \sigma(u, t)M(t)dz(t). \tag{8.11}
$$

It follows that

$$
\mathbb{E}[dM(t)] = 0, \tag{8.12}
$$

and hence $M(t)$ is a martingale. Equation (8.12) may also be expressed in the form

$$
\mathbb{E}_0[M(t)] = M(0) = 1,
$$

which using the definition (8.8) states that

$$
\mathbb{E}_0\left[\exp\left(-\frac{1}{2}\int_0^t \sigma^2(u(s), s)ds + \int_0^t \sigma(u(s), s)dz(s)\right)\right] = 1. \tag{8.13}
$$

In many of the applications of the exponential martingale that we encounter in later chapters the drift coefficient and diffusion coefficient $\sigma$ are either constants or functions of time only. In fact we have already encountered the constant case in example (iii) of Sect. 6.3 and the time-varying case in example (iv) of Sect. 6.3.

### 8.1.4 Quadratic Variation Processes

Consider again the stochastic integral

$$x(t) = \int_0^t \sigma(x(s), s)dz(s),$$

or

$$dx(t) = \sigma(x(t), t)dz(t). \tag{8.14}$$

As we have seen in example (ii) above $x(t)$ is a martingale. For reasons which we describe later we consider the process

$$y(t) = x^2(t). \tag{8.15}$$

An application of Ito's lemma shows that $y$ satisfies the stochastic differential equation

$$dy = \sigma^2 dt + 2\sqrt{y}\sigma dz(t). \tag{8.16}$$

Clearly $y(t)$ cannot be a martingale because of the nonzero drift term in Eq. (8.16). However the process

$$m(t) = x^2(t) - \int_0^t \sigma^2(x(s), s)ds \tag{8.17}$$

is a martingale. This result follows since

$$dm(t) = dy(t) - \sigma^2(x(t), t)dt,$$

which upon use of (8.16) reduces to

$$dm(t) = 2x(t)\sigma(x(t), t)dz(t), \tag{8.18}$$

from which it follows that $m(t)$ is a martingale. Hence

$$\mathbb{E}_0(m(t)) = m(0) = x^2(0) = 0.$$

Using the definition of $m(t)$ in (8.17) the last equation implies

$$\mathbb{E}_0(x^2(t)) = \int_0^t \mathbb{E}_0[\sigma^2(x(s), s)]ds,$$

or, in terms of the definition of $x(t)$,

$$\mathbb{E}_0\left[\left(\int_0^t \sigma(x(s),s)dz(s)\right)^2\right] = \int_0^t \mathbb{E}_0\left[\sigma^2(x(s),s)\right]ds. \tag{8.19}$$

This result has already been derived using informal mathematical arguments in Sect. 5.2.

The quantity $\int_0^t \sigma^2[x(s),s]ds$ which appears in (8.17) is known as the *quadratic variation process* of the martingale $x(t)$. It is often denoted $\langle x,x\rangle_t$ i.e.

$$\langle x,x\rangle_t \equiv \int_0^t \sigma^2(x(s),s)ds, \tag{8.20}$$

though strictly speaking we only require this notation when considering a vector of Wiener processes. In this notation the martingale $m(t)$ of Eq. (8.17) may be written

$$m(t) = x^2(t) - \langle x,x\rangle_t. \tag{8.21}$$

The result we have established in this example is often stated that the *square of a martingale minus its quadratic variation process is a martingale*. Note that since $m(t)$ is a martingale $\mathbb{E}_0[m(t)] = 0$ and hence we have established the result that

$$\mathbb{E}[x^2(t)] = \mathbb{E}_0[< x,x >_t].$$

The concept of quadratic variation arises when we seek to determine the length of a sample path of a Wiener process.

First let us review the notion of length of a path in ordinary calculus. Consider the function $f$ on the interval $[0,t]$ and take the subdivision $[t_{i-1},t_i]$ $(i = 1,\ldots,n)$ $(t_n \equiv t)$. The length of the graph of $f$ over $(0,t)$ is approximated by (see Fig. 8.2)

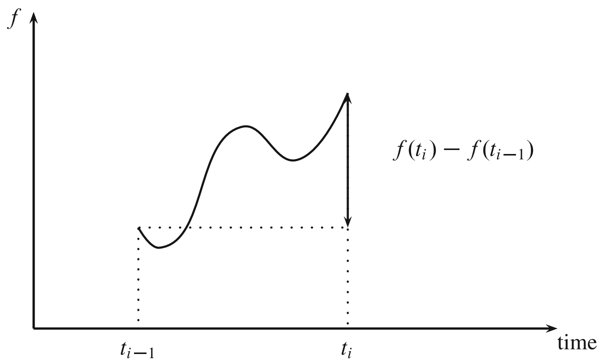


**Fig. 8.2** Approximating the path length over $[t_{i-1},t_i]$

$$A_n = \sum_{i=1}^{n} \mid f(t_i) - f(t_{i-1}) \mid . \tag{8.22}$$

If $A_n$ tends to a finite limit as we refine the partition (i.e. let $n \to \infty$ and the length of the subintervals tend to zero) then the graph of $f$ has finite length over $[0, t]$. In such cases the function $f$ is said to be *rectifiable*. A function which is rectifiable is said to have *bounded variation*. As the name implies such functions cannot have a behaviour which is too "wild" over the interval $[0, t]$. It can be proved using results of standard analysis that a function which has bounded variation is almost everywhere differentiable.

As far as the sample paths of Wiener processes are concerned we have shown in Sect. 4.3.1 that they are non-differentiable. It follows that no sample path of a Wiener process has bounded variation. This means that if the function $f$ in (8.22) is a Wiener process then the quantity $A_n$ tends to infinity as the partition is refined.

We have seen in Chap. 4 that we can define stochastic integrals by considering the mean-square limit. This leads us to consider the idea of considering the square of the function differences in (8.22). If this limit exists then the sample path has "length" in a mean-square sense.

If $x$ is a Wiener process we define the function $Q_n$ ($Q \equiv$ quadratic)

$$Q_n = \sum_{i=1}^{n} \mid x(t_i) - x(t_{i-1}) \mid^2 . \tag{8.23}$$

It can be shown that we can find $Q$ such that

$$\operatorname*{ms-lim}_{n \to \infty} Q_n = Q. \tag{8.24}$$

The quantity $Q$ is known as the *quadratic variation* of $x$ on $(0, t)$. The quadratic variation of $x$ on $(0, t)$ is usually denoted $\langle x, x \rangle_t$.

Suppose that $x$ is the sample path of the stochastic differential equation

$$dx = \mu(x, t)dt + \sigma(x, t)dz, \tag{8.25}$$

on $(0, t)$. Then by the rules of stochastic calculus (see Sect. 5.1)

$$(dx)^2 = \sigma^2(x, t)dt. \tag{8.26}$$

The calculation of the quadratic variation essentially consists of summing all the $(dx)^2$ over $(0, t)$, i.e. integrating the right-hand side of (8.26). Thus

$$\langle x, x \rangle_t = \int_0^t \sigma^2(x, s) ds. \tag{8.27}$$

For example if $\sigma(x, t)$ is a constant $\sigma$ then

$$\langle x, x \rangle_t = \sigma^2 t. \tag{8.28}$$

Of course to properly define the quadratic variation in (8.27) we should adopt proper limiting arguments and this is done in Chung and Williams (1990). However the beauty of the rules of stochastic calculus that we have adopted is that we can perform the formal manipulations (8.26) to (8.27) confidently that they are valid.

### *8.1.5 Semimartingales*

The concepts developed in this section allow us to define, at least in an intuitive sense, semimartingales. Any process $X$ which can be decomposed in the form

$$X = V + M, \tag{8.29}$$

where $V$ is a continuous process this is of bounded variation and $M$ is a martingale is called a *semimartingale*. Certainly processes generated by the stochastic differential equation (8.25) are semimartingales since

$$x(t) = \left[ x(0) + \int_0^t \mu(x, s) ds \right] + \left[ \int_0^t \sigma(x, s) dz(s) \right] \equiv V + M.$$

Given our discussion in Chap. 5 we see that $V \equiv [x(0) + \int_0^t \mu(x, s) ds]$ is of bounded variation, whilst we have demonstrated that $M \equiv \int_0^t \sigma(x, s) dz(s)$ is a martingale.

## 8.2 Changes of Measure and Girsanov's Theorem

Quite frequently we will be working with a diffusion process having drift/diffusion $(\mu, \sigma)$ with the uncertainty being drawn from a probability distribution or measure $\mathbb{P}$ (i.e. the diffusion is defined on $(\Omega, F, \mathbb{P})$ to use the correct technical language), that is

$$dx = \mu dt + \sigma dz, \quad x(0) = x_0, \quad \text{under } \mathbb{P}, \tag{8.30}$$

where typically $x$ would be an asset price or some interest rate. Application of the continuous hedging argument will lead us to consider the same quantity, $x$, having a different drift coefficient which we designate $(\mu - \lambda\sigma)$. Usually $\lambda$ would involve the market price of risk of $x$. Since there is now a different drift term we are in fact dealing with a different diffusion process which we denote $y$ if we remain under the original probability distribution or measure $\mathbb{P}$, that is

$$dy = (\mu - \lambda\sigma)dt + \sigma dz, \quad y(0) = x_0, \quad \text{under } \mathbb{P}. \tag{8.31}$$

However in many applications we wish to remain with the original process $x$. One way to do this is to change the original probability measure $\mathbb{P}$ to a new one $\tilde{\mathbb{P}}$ such that under $\tilde{\mathbb{P}}$ the original process $x$ has drift $(\mu - \lambda\sigma)$, that is

$$dx = (\mu - \lambda\sigma)dt + \sigma dz, \quad x(0) = x_0, \quad \text{under } \tilde{\mathbb{P}}. \tag{8.32}$$

Figure 8.3 illustrates these two different perspectives on the process $x$.

We know the calculation of the prices of contingent claims can be reduced to calculating expectations of future payoffs. Hence the above change of measure technique poses the problem of how to adjust the calculation of the expectation operator when we change the drift coefficient of the underlying diffusion process. Girsanov's change of measure theorem provides the solution to this problem.

We use $\mathbb{E}_t[f(x(\tau))]$ to denote the expectation formed at $t$ of some function $f$ of $x$ at time $\tau(> t)$ under the measure $\mathbb{P}$ (i.e. the dynamics of $x$ governed by (8.30)).

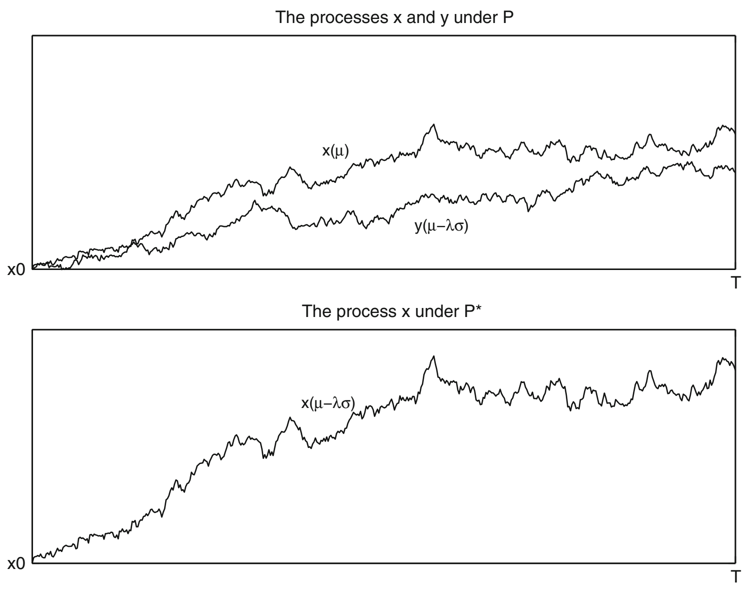


**Fig. 8.3** Illustrating the change of measure result

Similarly $\tilde{\mathbb{E}}_t[f(x(\tau)]$ denotes the expectation operation under the measure $\tilde{\mathbb{P}}$ (i.e. the dynamics of $x$ governed by (8.32)). Recall that

$$\mathbb{E}_t[f(x(\tau))] = \int f(x(\tau))d\mathbb{P}(\tau) = \int f(x_\tau)p(x_\tau, \tau \mid x_t, t)dx_\tau, \tag{8.33}$$

and

$$\tilde{\mathbb{E}}_t[f(x(\tau))] = \int f(x(\tau))d\tilde{\mathbb{P}}(\tau) = \int f(x_\tau)\tilde{p}(x_\tau, \tau \mid x_t, t)dx_\tau, \tag{8.34}$$

where $\mathbb{P}(\tau)$ and $\tilde{\mathbb{P}}(\tau)$ are the cumulative density functions from $x(t)$ to $x(\tau)$ and $p(x_\tau, \tau \mid x_t, t)$ and $\tilde{p}(x_\tau, \tau \mid x_t, t)$ are the corresponding conditional transition density functions from $x(t)$ to $x(\tau)$. We note that these latter are respectively the solutions to the Kolmogorov backward equations associated with the stochastic differential equations (8.30) and (8.32).

The idea of the Cameron-Martin-Girsanov formula is to find a stochastic variable $\xi(t)$ such that

$$d\tilde{\mathbb{P}} = \xi d\mathbb{P}, \tag{8.35}$$

i.e.

$$\tilde{p}(x_\tau, \tau \mid x_t, t)dx_\tau = \xi(\tau)p(x_\tau, \tau \mid x_t, t)dx_\tau.$$

For then

$$\begin{aligned}\mathbb{E}_t[\xi(\tau)f(x(\tau))] &= \int f(x(\tau))\xi(\tau)d\mathbb{P}(\tau) = \int f(x(\tau))\xi(\tau)p(x_\tau, \tau \mid x_t, t)dx_\tau\\ &= \int f(x(\tau))\tilde{p}(x_\tau, \tau \mid x_t, t)dx_\tau = \int f(x(\tau))d\tilde{\mathbb{P}}(\tau)\\ &= \tilde{\mathbb{E}}_t[f(x(\tau))].\end{aligned} \tag{8.36}$$

Two requirements on the (yet to be specified) function $\xi(t)$ are that

(i) Prob.$\{\xi > 0\} = 1$, so that probabilities remain positive under $\tilde{\mathbb{P}}$,

and

(ii) $\mathbb{E}_0[\xi(t)] = \int_0^t \xi(s)d\mathbb{P}(s) = \int_0^t \xi_s p(\xi_s, s \mid \xi_0, 0)d\xi_s = 1$.

This last result in turn ensures that

$$\int d\tilde{\mathbb{P}}(t) = \int \tilde{p}(\xi_t, t \mid \xi_0, 0)d\xi_t = 1,$$

so that $\tilde{\mathbb{P}}$ indeed qualifies as a probability distribution.

We now construct a quantity $\xi(t)$ fulfilling the above requirements. We use the quantity $\lambda(t)$ appearing in Eq. (8.32) to define the Ito diffusion process $u(t)$ given by

$$du = -\frac{1}{2}\lambda^2(t)dt - \lambda(t)dz(t), \quad u(0) = 0. \tag{8.37}$$

Then let $\xi(t)$ be defined by

$$\xi(t) = e^{u(t)} = \exp\left(-\int_0^t \frac{1}{2}\lambda^2(s)ds - \int_0^t \lambda(s)dz(s)\right), \tag{8.38}$$

so that $\xi(0) = e^{u(0)} = e^0 = 1$. Note that since $\xi(t)$ is the exponential of a random variable $u(t)$, it must always be positive i.e.

$$\text{Prob.}\{\xi > 0\} = 1. \tag{8.39}$$

Thus our first condition, (i) above, on $\xi(t)$ is satisfied.

By Ito's lemma (see our discussion in Sect. 8.1.2 on the exponential martingale) $\xi(t)$ satisfies the stochastic differential equation

$$d\xi(t) = -\xi(t)\lambda(t)dz(t), \tag{8.40}$$

so that

$$\mathbb{E}[d\xi(t)] = 0,$$

which implies that $\xi(t)$ is a martingale. It follows that

$$\mathbb{E}_0[\xi(t)] = \xi(0) = 1, \tag{8.41}$$

which is our second requirement, (ii) above, on the function $\xi(t)$. Here $\mathbb{E}_0$ is the unconditional expectation operator calculated at $t = 0$.

In fact $\xi(t)$ is a special example of the exponential martingale which we encountered in Sect. 8.1.3. The simulation of four paths of (8.40) for $\lambda(t) = 0.2$ is displayed in Fig. 8.4. The figure shows how typical sample paths wander around the initial value $\xi(0) = 1$, as is implied by Eq. (8.41).

The quantity $\xi(t)$ is known as the **Radon–Nikodym derivative** of the measure $\tilde{\mathbb{P}}$ with respect to the measure $\mathbb{P}$ and sometimes (8.35) is written

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \xi. \tag{8.42}$$

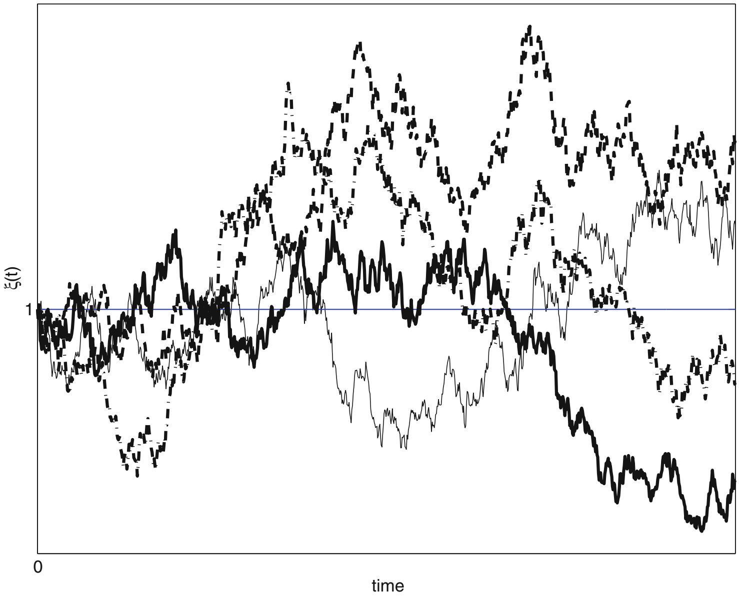


**Fig. 8.4** Some typical sample paths of the exponential martingale process

The measures $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are said to be equivalent which means events having zero probability under one measure have zero probability under the other, or more formally

$$\tilde{\mathbb{P}}(A) = 0 \quad \text{if and only if} \quad \mathbb{P}(A) = 0, \tag{8.43}$$

where A is a set belonging to the sample space $\Omega$. In other words if $x(t)$ is a price process then the set of prices attainable under $\mathbb{P}$ remains attainable under $\tilde{\mathbb{P}}$, albeit with a different probability. The main result we require in our subsequent applications is the following theorem.

**Change of Measure Theorem** *Suppose that $x$ is a $(\mu, \sigma)$ diffusion process under $\mathbb{P}$. Given the stochastic process $\xi$ and the probability measure $\tilde{\mathbb{P}}$ as defined by* (8.38) *and* (8.42) *respectively then $x$ is a $(\mu - \lambda\sigma, \sigma)$ diffusion process under $\tilde{\mathbb{P}}$.*

The proof of this theorem is somewhat technical and we refer the reader to Harrison (1990) for details. Versions of this theorem in a more general setting are known as **Girsanov's theorem**. The names of Cameron and Martin are connected with earlier versions.

In our applications we rarely need to calculate directly the Radon–Nikodym derivative $\xi(t)$ in Eq. (8.42). Rather we interpret the change of measure theorem in the following way. Firstly let $z(t)$ and $\tilde{z}(t)$ respectively denote Wiener processes under the probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$. Then we can view the diffusion process $x$ in two ways:

$$(i) \qquad dx = \mu dt + \sigma dz, \tag{8.44}$$

where under $\mathbb{P}$, $\quad dz \sim N(0, dt)$,

$$(ii) \qquad dx = (\mu - \lambda\sigma)dt + \sigma d\tilde{z}, \tag{8.45}$$

where under $\tilde{\mathbb{P}}$, $\quad d\tilde{z} \sim N(0, dt)$. Note that we can rewrite Eq. (8.45) in the following way:

$$dx = \mu dt + \sigma(d\tilde{z}(t) - \lambda(t)dt) = \mu dt + \sigma\left(d\tilde{z}(t) - d\int_0^t \lambda(s)ds\right),$$

i.e.

$$dx = \mu dt + \sigma d\left(\tilde{z}(t) - \int_0^t \lambda(s)ds\right). \tag{8.46}$$

Contrasting (8.46) with (8.44) we see that the stochastic quantity $\tilde{z}(t) - \int_0^t \lambda(s)ds$ must be a Wiener process under $\mathbb{P}$. That is we have the relationship

$$z(t) = \tilde{z}(t) - \int_0^t \lambda(s)ds \tag{8.47}$$

between the Wiener processes $z(t)$ and $\tilde{z}(t)$ under the two different probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ that are related via (8.42). It is this result which we shall mostly use in our applications.

It is also possible to state the change of measure theorem in a form such that the drift adjustment term $\lambda\sigma$ appears in the diffusion process under $\mathbb{P}$ and the "clean" drift term $\mu$ appears in the process under $\tilde{\mathbb{P}}$. From this perspective a diffusion process $x$ could be viewed in the following two ways:

$$(i) \qquad dx = (\mu + \lambda\sigma)dt + \sigma dz, \tag{8.48}$$

where under $\mathbb{P}$, $\quad dz \sim N(0, dt)$, or

$$(ii) \qquad dx = \mu dt + \sigma d\tilde{z}, \tag{8.49}$$

where under $\tilde{\mathbb{P}}$, $d\tilde{z} \sim N(0, dt)$.

**Table 8.1** Summarising the relation between the processes $z(t)$, $\tilde{z}(t)$ and the measures $\mathbb{P}$, $\tilde{\mathbb{P}}$

| | $\mathbb{P}$ | $\tilde{\mathbb{P}}$ |
|---|---|---|
| $z(t)$ | Wiener<br>$\mathbb{E}(dz(t)) = 0$ | Not wiener<br>$\tilde{\mathbb{E}}(dz(t)) = -\lambda dt \neq 0$ |
| $\tilde{z}(t)$ | Not wiener<br>$\mathbb{E}(d\tilde{z}(t)) = \lambda dt \neq 0$ | Wiener<br>$\tilde{\mathbb{E}}(d\tilde{z}(t)) = 0$ |

Here $z(t) = \tilde{z}(t) - \int_0^t \lambda(s)ds$, or $dz(t) = d\tilde{z}(t) - \lambda(t)dt$

Now we could manipulate Eq. (8.48) as

$$dx = \mu dt + \sigma d\left(z(t) + \int_0^t \lambda(s)ds\right). \tag{8.50}$$

Contrasting with Eq. (8.49) we see that $z(t) + \displaystyle\int_0^t \lambda(s)ds$ can be equated to the Wiener process $\tilde{z}(t)$ under $\tilde{\mathbb{P}}$ i.e.

$$\tilde{z}(t) = z(t) + \int_0^t \lambda(s)ds, \tag{8.51}$$

which not surprisingly is a re-expression of (8.47). It is from this perspective that we shall usually use the change of measure theorem in our applications. We shall see that conditions which guarantee an arbitrage free economy will result in diffusion processes of the form (8.48).

In Table 8.1 we summarise the key results of this section and indeed the ones mostly required for applications.

## 8.3 Girsanov's Theorem for Vector Processes

In later applications when considering options on multiple assets we shall require a multi-dimensional (or vector version) of Girsanov's theorem. To keep the notation simple we first give the two-dimensional version and then extend the result to a vector system.

Consider the two processes $x_1$ and $x_2$ given by

$$dx_1 = \mu_1 dt + \sigma_{11} dw_1 + \sigma_{12} dw_2, \tag{8.52}$$

$$dx_2 = \mu_2 dt + \sigma_{21} dw_1 + \sigma_{22} dw_2, \tag{8.53}$$

where $w_1$ and $w_2$ are independent Wiener processes under the measure $\mathbb{P}$. With the quantities $\lambda_1(t)$ and $\lambda_2(t)$ (which in our applications will be market prices of risk associated with $w_1$ and $w_2$ respectively) define the process

$$\xi(t) = \exp\left(-\frac{1}{2}\int_0^t (\lambda_1^2(s) + \lambda_2^2(s))ds - \int_0^t (\lambda_1(s)dw_1(s) + \lambda_2(s)dw_2(s))\right). \tag{8.54}$$

Application of Ito's Lemma yields

$$d\xi(t) = -\xi(t)(\lambda_1(t)dw_1(t) + \lambda_2(t)dw_2(t)), \tag{8.55}$$

so that $\xi(t)$ is a martingale under $\mathbb{P}$. An argument exactly analogous to that leading to (8.42) can be used to show that here also

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \xi. \tag{8.56}$$

The processes

$$\tilde{w}_1(t) = w_1(t) + \int_0^t \lambda_1(s)ds, \tag{8.57}$$

$$\tilde{w}_2(t) = w_2(t) + \int_0^t \lambda_2(s)ds, \tag{8.58}$$

will be Wiener processes under $\tilde{\mathbb{P}}$. Furthermore the dynamics for $x_1$ and $x_2$ under $\tilde{\mathbb{P}}$ are given by

$$dx_1 = (\mu_1 - \lambda_1\sigma_{11} - \lambda_2\sigma_{12})dt + \sigma_{11}d\tilde{w}_1 + \sigma_{12}d\tilde{w}_2, \tag{8.59}$$

$$dx_2 = (\mu_2 - \lambda_1\sigma_{21} - \lambda_2\sigma_{22})dt + \sigma_{21}d\tilde{w}_1 + \sigma_{22}d\tilde{w}_2. \tag{8.60}$$

Alternatively, one may start with the dynamics under $\mathbb{P}$ specified as

$$dx_1 = (\mu_1 + \lambda_1\sigma_{11} + \lambda_2\sigma_{12})dt + \sigma_{11}dw_1 + \sigma_{12}dw_2, \tag{8.61}$$

$$dx_2 = (\mu_2 + \lambda_1\sigma_{21} + \lambda_2\sigma_{22})dt + \sigma_{21}dw_1 + \sigma_{22}dw_2, \tag{8.62}$$

and the dynamics under $\tilde{\mathbb{P}}$ then become

$$dx_1 = \mu_1 dt + \sigma_{11}d\tilde{w}_1 + \sigma_{12}d\tilde{w}_2, \tag{8.63}$$

$$dx_2 = \mu_2 dt + \sigma_{21}d\tilde{w}_1 + \sigma_{22}d\tilde{w}_2. \tag{8.64}$$

For full details on the multi-dimensional version of Girsanov's theorem we refer the reader to Oksendal (2003).

More generally consider the vector process

$$dx_i = \mu_i dt + \sum_{j=1}^{n} \sigma_{ij} dw_j, \qquad i = 1, 2, \ldots, n. \tag{8.65}$$

Using $d\boldsymbol{w}$ to denote the column vector $(dw_1, dw_2, \ldots, dw_n)$ and $\boldsymbol{\lambda}$ the column vector of corresponding market prices of risk we consider the process

$$\xi(t) = exp\left(-\frac{1}{2}\int_0^t \boldsymbol{\lambda}^T(s)\boldsymbol{\lambda}(s)ds - \int_0^t \boldsymbol{\lambda}^T(s)d\boldsymbol{w}(s)\right) \tag{8.66}$$

that satisfies

$$d\xi(t) = -\xi(t)\boldsymbol{\lambda}^T(t)d\boldsymbol{w}(t). \tag{8.67}$$

Hence $\xi(t)$ is a martingale under $\mathbb{P}$ and again we have Eq. (8.56) and that the processes

$$\tilde{w}_i(t) = w_i(t) + \int_0^t \lambda_i(s)ds \tag{8.68}$$

will be Wiener processes under $\tilde{\mathbb{P}}$. Note that in differential form and using vector notation (8.68) may be written

$$d\tilde{\boldsymbol{w}} = d\boldsymbol{w} + \boldsymbol{\lambda} dt. \tag{8.69}$$

The dynamics of $x_i$ under $\tilde{\mathbb{P}}$ are then given by

$$dx_i = \left(\mu_i - \sum_{j=1}^{n} \lambda_j \sigma_{ij}\right)dt + \sum_{j=1}^{n} \sigma_{ij} d\tilde{w}_j. \tag{8.70}$$

Alternatively, we may start with the dynamics under $\mathbb{P}$ specified as

$$dx_i = \left(\mu_i + \sum_{j=1}^{n} \lambda_j \sigma_{ij}\right)dt + \sum_{j=1}^{n} \sigma_{ij} dw_j, \tag{8.71}$$

under which $\tilde{\mathbb{P}}$ becomes

$$dx_i = \mu_i dt + \sum_{j=1}^{n} \sigma_{ij} d\tilde{w}_j. \tag{8.72}$$

Often we need to deal with the dynamics for the $x_i$ in the form

$$dx_i = \mu_i dt + \sum_{j=1}^{n} s_{ij} dz_j, \quad (i = 1, \ldots, n) \tag{8.73}$$

where under the measure $\mathbb{P}$ the $z_i$ are correlated Wiener processes so that

$$\mathbb{E}[dz_i dz_j] = \rho_{ij} dt, \quad (i \neq j). \tag{8.74}$$

Let

$$\boldsymbol{x} = (x_1, \cdots, x_n)^T, \qquad \boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)^T, \qquad \boldsymbol{s} = (s_{ij})_{n \times n}.$$

Then, the stochastic differential equations (8.72) can be written as

$$d\boldsymbol{x} = \boldsymbol{\mu} dt + \boldsymbol{s} d\boldsymbol{z}. \tag{8.75}$$

We know from Sect. 5.3 that we can find independent Wiener processes $w_i(t)$ such that

$$d\boldsymbol{z} = A d\boldsymbol{w}, \tag{8.76}$$

where $d\boldsymbol{z} = (dz_1, dz_2, \ldots, dz_n)^T$ and $d\boldsymbol{w} = (dw_1, dw_2, \ldots, dw_n)^T$. In this notation, we will have processes $\phi_i$ (the market prices of risk of the $z_i$) so that the processes for the $x_i$ under $\tilde{\mathbb{P}}$ should look like

$$d\boldsymbol{x} = (\boldsymbol{\mu} - \boldsymbol{s}\boldsymbol{\phi}) dt + \boldsymbol{s} d\tilde{\boldsymbol{z}}, \qquad \boldsymbol{\phi} = (\phi_1, \cdots, \phi_n)^T \tag{8.77}$$

and the relation between the $z_i$ and $\tilde{z}_i$ is given by

$$d\tilde{z}_i = dz_i + \phi_i dt, \quad (i = 1, 2, \ldots, n), \tag{8.78}$$

or

$$d\tilde{\boldsymbol{z}} = d\boldsymbol{z} + \boldsymbol{\phi} dt.$$

We use the same matrix $A$ to transform the correlated $d\tilde{\boldsymbol{z}}$ to the independent $d\tilde{\boldsymbol{w}}$, thus

$$A d\tilde{\boldsymbol{w}} = A d\boldsymbol{w} + \boldsymbol{\phi} dt,$$

where we set $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_n)^T$. Thus

$$d\tilde{\boldsymbol{w}} = d\boldsymbol{w} + A^{-1} \boldsymbol{\phi} dt, \tag{8.79}$$

which upon comparison with (8.69) indicates that the market prices of risk for the independent and correlated Wiener processes are related by

$$\boldsymbol{\lambda} = A^{-1}\boldsymbol{\phi}. \tag{8.80}$$

Substituting (8.76) (in the form $d\boldsymbol{w} = A^{-1}d\boldsymbol{z}$) and (8.80) into (8.66) obtain the expression for the Radon–Nikodym derivative in terms of the correlated Wiener processes $\boldsymbol{z}$ and their corresponding market prices of risk $\boldsymbol{\phi}$. In fact, we find that

$$\xi(t) = \exp\left(-\frac{1}{2}\int_0^t \boldsymbol{\phi}^T(s)(A^{-1})^T A^{-1}\boldsymbol{\phi}(s)ds - \int_0^t \boldsymbol{\phi}^T(s)(A^{-1})^T A^{-1}d\mathbf{z}(s)\right)$$

which reduces to

$$\xi(t) = \exp\left(-\frac{1}{2}\int_0^t \boldsymbol{\phi}^T(s)\rho^{-1}\boldsymbol{\phi}(s)ds - \int_0^t \boldsymbol{\phi}^T(s)\rho^{-1}d\boldsymbol{z}(s)\right). \tag{8.81}$$

In achieving the last simplification we have used the fact that $(d\boldsymbol{z})(d\boldsymbol{z})^T = \rho dt = AA^T dt$ and hence $AA^T = \rho$ where $\rho$ is the correlation matrix (see Eq. (5.33)). Note also that we have used the results from matrix algebra that

$$(A^{-1})^T = (A^T)^{-1} \quad \text{and} \quad (A^{-1})^T A^{-1} = (A^T)^{-1}A^{-1} = (AA^T)^{-1}.$$

## 8.4 Derivation of Black–Scholes Formula by Girsanov's Theorem

Recall our standard assumption on stock price dynamics, namely,

$$dS = \mu S dt + \sigma S dz. \tag{8.82}$$

For the Ito process (8.82), the transition probability density function $p(S_T, T \mid S, t)$ (the probability density of reaching stock price $S_T$ at $T$ given that the stock price is $S$ at time $t < T$) is the solution of the Kolmogorov backward equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 p}{\partial S^2} + \mu S\frac{\partial p}{\partial S} + \frac{\partial p}{\partial t} = 0, \qquad t \le T, \tag{8.83}$$

subject to the initial condition

$$p(S', t \mid S, T) = \delta(S - S'), \tag{8.84}$$

where $\delta$ is the Dirac delta function.

As we have described in Chap. 2 and shown in Problem 3.3 the solution to this Kolmogorov backward equation is

$$p(S_T, T \mid S, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma S_T} \exp\left[-\frac{\left\{\ln(S_T/S) - \left(\mu - \sigma^2/2\right)(T-t)\right\}^2}{2\sigma^2(T-t)}\right]. \tag{8.85}$$

Since the option price is given by

$$C = C(S, t),$$

and we have seen in Chap. 6 that by Ito's lemma the option price dynamics are given by

$$dC = \mu_c C dt + \sigma_c C dz, \tag{8.86}$$

where

$$\mu_c = \frac{(\theta + \mu S \Delta + \frac{1}{2}\sigma^2 S^2 \Gamma)}{C}, \qquad \sigma_c = \frac{\sigma S \Delta}{C}.$$

By the continuous hedging argument (see Chap. 7), the risk adjusted expected excess returns on the stock and the option are related by (see Eq. (7.14))

$$\frac{\mu_c - r}{\sigma_c} = \frac{\mu - r}{\sigma} = \lambda, \tag{8.87}$$

where $\lambda$, the market price of risk of the uncertainty associated with the stock price dynamics (i.e. $z(t)$) is defined in Sect. 7.2. Therefore, considering separately the two equations in (8.87), we have

$$\mu = r + \lambda\sigma, \tag{8.88}$$

$$\mu_c = r + \lambda\sigma_c. \tag{8.89}$$

Substituting (8.88) into (8.82) and (8.89) into (8.86) the stochastic differential equations followed by $S$ and $C$, in an arbitrage free economy, become

$$dS = (r + \lambda\sigma) S dt + \sigma S dz, \tag{8.90}$$

$$dC = (r + \lambda\sigma_c) C dt + \sigma_c C dz. \tag{8.91}$$

Note that these can both be written

$$dS = rSdt + \sigma Sd\left(z(t) + \int_0^t \lambda(u)du\right), \tag{8.92}$$

$$dC = rCdt + \sigma_c Cd\left(z(t) + \int_0^t \lambda(u)du\right). \tag{8.93}$$

Now define a new process

$$\tilde{z}(t) = z(t) + \int_0^t \lambda(u)du, \tag{8.94}$$

then Eqs. (8.92) and (8.93) can be written

$$dS = rSdt + \sigma Sd\tilde{z}(t), \tag{8.95}$$

$$dC = rCdt + \sigma_c Cd\tilde{z}(t). \tag{8.96}$$

Let $\mathbb{P}$ be the probability distribution underlying the Wiener process in (8.82) i.e. the one whose probability density function is given by (8.85), so that under $\mathbb{P}$

$$dz \sim N(0, dt).$$

Then under the probability distribution $\mathbb{P}$, we would have

$$d\tilde{z} \sim N(\lambda dt, dt). \tag{8.97}$$

Hence $\tilde{z}(t)$ would not be a Wiener process under $\mathbb{P}$ because of its non-zero drift term. However the change of measure theorem of the previous section tells us that if we define a new probability distribution $\tilde{\mathbb{P}}$ by

$$d\tilde{\mathbb{P}}(T) = \xi(t, T)d\mathbb{P}(T), \tag{8.98}$$

i.e.

$$\tilde{p}(\tilde{S}_T, T \mid S, t)d\tilde{S}_T = \xi(t, T)p(S_T, T \mid S, t)dS_T,$$

where[1]

$$\xi(t, T) = \exp\left(-\int_t^T \lambda(s)dz(s) - \frac{1}{2}\int_t^T \lambda^2(s)ds\right), \tag{8.99}$$

[1]Note that we use $\xi(t, T)$ to denote $\xi(T)/\xi(t)$ where $\xi(t)$ is defined in Eq. (8.38).

then with respect to $\tilde{\mathbb{P}}$, the process $\tilde{z}(t)$ is a Wiener process, i.e.

$$d\tilde{z}(t) \sim N(0, dt). \tag{8.100}$$

Thus under the measure $\tilde{\mathbb{P}}$, the $S$ and $C$ in Eqs. (8.95) and (8.96) can be regarded as Ito processes with drifts $(rS, rC)$ and diffusion coefficients $(\sigma S, \sigma_c C)$ respectively. Under the probability measure $\tilde{\mathbb{P}}$, the transition probability density function $\tilde{p}(\tilde{S}_T, T \mid S, t)$ is the solution of the Kolmogorov backward equation determined by the Ito process (8.95), namely,

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \tilde{p}}{\partial S^2} + rS\frac{\partial \tilde{p}}{\partial S} + \frac{\partial \tilde{p}}{\partial t} = 0, \quad t \leq T, \tag{8.101}$$

which is also subject to the initial condition

$$\tilde{p}(S', T|S, T) = \delta(S - S'). \tag{8.102}$$

Again using the results in Problem 3.3 the solution to (8.101) with initial condition (8.102) can be obtained as

$$\tilde{p}(S_T, T \mid S, t)dS_T = \frac{1}{\sqrt{2\pi(T-t)}\sigma} \exp\left[-\frac{\left\{\ln(S_T/S) - \left(r - \sigma^2/2\right)(T-t)\right\}^2}{2\sigma^2(T-t)}\right]\frac{dS_T}{S_T}. \tag{8.103}$$

From Eq. (8.96) we note that

$$d(Ce^{-rt}) = e^{-rt}\sigma_c C d\tilde{z}(t). \tag{8.104}$$

Since $d\tilde{z}(t)$ is a Wiener increment under the probability measure $\tilde{\mathbb{P}}$ it follows from (8.104) that

$$\tilde{\mathbb{E}}_t[d(Ce^{-rt})] = 0, \tag{8.105}$$

where $\tilde{\mathbb{E}}_t$ is the expectation operator under the equivalent probability measure $\tilde{\mathbb{P}}$. The last equation implies that $Ce^{-rt}$ is a martingale under $\tilde{\mathbb{P}}$. Hence

$$\tilde{\mathbb{E}}_t[C_T e^{-rT}] = C_t e^{-rt}, \qquad \text{for all} \quad T \geq t, \tag{8.106}$$

therefore

$$\begin{aligned} C(S,t) &= e^{-r(T-t)}\tilde{\mathbb{E}}_t[C(S_T,T)] \\ &= e^{-r(T-t)}\int_0^\infty \max[S_T - E, 0]\tilde{p}(S_T, T \mid S, t)dS_T. \end{aligned} \tag{8.107}$$

Equation (8.107) is precisely Eq. (3.15) (with $h(S_T)$ set equal to the payoff on a European call option). Now we have obtained the option pricing relationship without assuming the investors are risk neutral. The result that investors price the option as if they were risk neutral arises naturally as a result of the hedging argument.

One way to operationalize (8.107) is to use the expression for $\tilde{p}(S_T, T \mid S, t)$ given in (8.103). We are then dealing with the integral we evaluated in Appendix 3.1 to yield the Black–Scholes formula.

Note that in cases where we cannot calculate $\tilde{p}(S_T, T \mid S, t)$ explicitly, we can use Eq. (8.95) as the basis of numerical simulation by random drawings of $d\tilde{z}(t) \sim N(0, dt)$ (see Problem 8.1). As we shall see in later chapters this is the approach we need to adopt for many exotic options.

Finally we note that the same set of operations that led from (8.104) to (8.106) may also be applied to the stochastic differential equation (8.95) to yield

$$S_t e^{-rt} = \tilde{\mathbb{E}}_t[S_T e^{-rT}]. \tag{8.108}$$

Thus $S_t e^{-rt}$ is also a martingale under $\tilde{\mathbb{P}}$. Equation (8.108) may be re-expressed as

$$S_t = e^{-r(T-t)}\tilde{\mathbb{E}}_t[S_T]. \tag{8.109}$$

This equation simply states that under the risk-neutral-measure $\tilde{\mathbb{P}}$ the current stock price is the discounted (at the risk free rate) expected value of the stock price at the future time $T$.

## 8.5 The Pricing Kernel Representation

We may use the relation between the probability measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ in Eq. (8.98) to re-express the option pricing relationship as an expectation under the original measure $\mathbb{P}$.

We note that by use of Eq. (8.36) we may express the first line of Eq. (8.107) as

$$C(S,t) = e^{-r(T-t)}\mathbb{E}_t[\xi(t,T)C(S_T,T)], \tag{8.110}$$

which upon use of Eq. (8.99) may be written

$$C(S,t) = e^{-r(T-t)}\mathbb{E}_t[e^{-\int_t^T \lambda(s)dz(s)-\frac{1}{2}\int_t^T \lambda^2(s)ds}C(S_T,T)].$$
$$= \mathbb{E}_t\left[\frac{M(T)}{M(t)}C(S_T,T)\right], \tag{8.111}$$

where we set

$$M(t) = \exp\left(-rt - \int_0^t \lambda(s)dz(s) - \frac{1}{2}\int_0^t \lambda^2(s)ds\right). \tag{8.112}$$

Equation (8.111) may be expressed as

$$M(t)C(S,t) = \mathbb{E}_t[M(T)C(S_T,T)], \tag{8.113}$$

and in this form we see that the quantity $M(t)C(S,T)$ is a martingale under the original probability measure $\mathbb{P}$. The quantity $M(t)$ by which we have to adjust the option price to obtain a martingale relationship under $\mathbb{P}$ is sometimes referred to as the pricing kernel.

We note that the pricing kernel can be written

$$M(t) = e^{-rt}\xi(t), \tag{8.114}$$

from which by using of (8.40) we readily see that $M(t)$ satisfies the stochastic differential equation

$$dM = -M(rdt + \lambda dz). \tag{8.115}$$

The reader can readily verify that $M(t)S(t)$ is also a martingale under $\mathbb{P}$, so that

$$S(t) = \mathbb{E}_t\left[\frac{M(T)}{M(t)}S(T)\right]. \tag{8.116}$$

Equation (8.116) states that under the original measure $\mathbb{P}$, the stock price at time $t$ is the expectation of the stock price at future time $T$, multiplied by $\frac{M(T)}{M(t)}$, the stochastic discount factor between these two times.

## 8.6 The Feynman–Kac Formula

Most applications of the theory of stochastic processes that we have been discussing come down to calculating the value of contingent claims as the expectation of some function or functional of a stochastic process. We also know that it is possible to derive such values as the solutions of certain partial differential equations, as for example in the original Black–Scholes derivation. It is useful to know how these alternative expressions for contingent claim values are related. The technical result that allows us to express these expectations in terms of partial differential equations and to go back and forth between the different ways of representing the option price is the Feynman–Kac formula which we now discuss.

For later use we state the vector version of the Feynman–Kac formula. We use $\mathbf{x(t)} = (x_1(t), \cdots, x_n(t))$ to denote the vector of $n$ factors driven by set of $n$ Ito diffusion processes (the reader should refer again to Sect. 5.4)

$$dx_i(t) = \mu_i(\mathbf{x}, t)dt + \sum_{j=1}^{n} \sigma_{ij}(\mathbf{x}, t)dW_j(t), \tag{8.117}$$

for $i = 1, 2, \cdots, n$ on $t_0 \le t < T$, with $\mathbf{x}(t_0) = \mathbf{x}_0$. Here $\mu_i(\mathbf{x}, t)$ and $\sigma_{ij}(\mathbf{x}, t)$, $(j = 1, 2, \cdots, n)$ are the drift and diffusion coefficients associated with the process for $x_i(t)$. The $w(t) = (w_1(t), w_2(t), \cdots, w_n(t))$ is a vector of $n$ independent Wiener processes.

We recall from Sect. 5.4 that from the matrix $\sigma$ (whose elements are the $\sigma_{ij}$ in (8.117)) we can form the matrix

$$S = \sigma\sigma^\top = (s_{ij})_{n\times n}.$$

From the matrix $S$ and drift coefficients $\mu_i$ we may form the infinitesimal generator $\mathscr{K}$ for the process $\mathbf{x}$, namely

$$\mathscr{K} = \sum_{i=1}^{n} \mu_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij} \frac{\partial^2}{\partial x_i \partial x_j}. \tag{8.118}$$

Then the Kolmogorov backward equation associated with the above diffusion process generates the conditional transition probability density function $p(\mathbf{X}, T|\mathbf{x}, t)$ and is written

$$\mathscr{K} p + \frac{\partial p}{\partial t} = 0, \qquad t_0 \le t < T \tag{8.119}$$

subject to the initial condition[2]

$$p(\mathbf{X}, T \mid \mathbf{x}, T) = \delta(\mathbf{x} - \mathbf{X}). \tag{8.120}$$

[2]The notation $\delta(\mathbf{x} - \mathbf{X})$ should be interpreted as

$$\delta(x_1 - X_1)\delta(x_2 - X_2)\cdots\delta(x_n - X_n).$$

Let $\mathbb{E}_{\mathbf{x},t}$ denote the expectation operator[3] associated with the Ito diffusion process (8.117) when $x(t) = x$. If $f(\mathbf{x}(s))$ is some sufficiently well behaved function of the random variable $\mathbf{x}(s)$ (for $t \le s \le T$) then

$$\mathbb{E}_{\mathbf{x},t}\,[f(\mathbf{x}(s))] = \int f(\mathbf{x}(s))p(\mathbf{x}(s), s \mid \mathbf{x}, t)d\mathbf{x}(s). \tag{8.121}$$

Let

$$u(\mathbf{x}, t) = \mathbb{E}_{\mathbf{x},t}\,[f(\mathbf{x}(T))], \qquad (t_0 \le t < T), \tag{8.122}$$

then it turns out that $u(\mathbf{x}, t)$ also satisfies the Kolmogorov backward equation. In fact we can state:

**Proposition 8.1 (Kolmogorov's Formula)** *The function $u(\mathbf{x}, t)$ defined in* (8.122) *satisfies the partial differential equation*

$$\mathscr{K} u + \frac{\partial u}{\partial t} = 0, \tag{8.123}$$

*subject to the initial condition*

$$\lim_{t \to T} u(\mathbf{x}, t) = f(\mathbf{x}(T)). \tag{8.124}$$

*Proof* See Appendix 8.1. ■

We have in essence seen the basic idea of the proof of this result when we derived the partial differential equation (3.7) for the option price in Sect. 3.1. The above proposition allows us to handle the conditional expectation of a function of a stochastic process, however, in many problems in continuous time finance we need to deal with the conditional expectation of a functional of a diffusion process. For instance, in the term structure of interest rate modelling we encounter quantities of the type

$$v(\mathbf{x}, t) = \mathbb{E}_{\mathbf{x},t}\left[\exp\left(\lambda \int_t^T g[\mathbf{x}(s), s]ds\right)\right], \tag{8.125}$$

for $g$ a sufficiently well behaved function and $\lambda$ a constant. Typically $v(\mathbf{x}, t)$ would be the price at time $t$ of a pure discount bond maturing at time $T$.

[3] For the purposes of the discussion in this section it is useful to have a notation for the expectation operator that indicates both the time, $t$, as well as the initial value, $\mathbf{x}$, of the underlying stochastic process when expectations are formed. We shall not use this notation elsewhere.

**Proposition 8.2 (The Feynman–Kac Formula)** *The function* $v(\mathbf{x}, t)$ *defined in Eq.* (8.125) *satisfies the partial differential equation*

$$\frac{\partial v}{\partial t} + \mathscr{K} v + \lambda g(\mathbf{x}, t) v = 0, \tag{8.126}$$

*subject to the initial condition*

$$\lim_{t \to T} v(\mathbf{x}, t) = 1. \tag{8.127}$$

*Proof* See Appendix 8.2. ■

A more general version of the Feynman–Kac formula involves the quantity

$$v(\mathbf{x}, t) = \mathbb{E}_{\mathbf{x},t}\left[ f(\mathbf{x}(T)) \exp\left( \lambda \int_t^T g[\mathbf{x}(s), s] ds \right) \right], \tag{8.128}$$

where $g$ and $f$ are sufficiently well behaved functions and $\lambda$ is some constant. In later chapters we shall encounter such quantities when evaluating options within a stochastic interest rate environment.

**Proposition 8.3 (The Feynman–Kac Formula II)** *The function* $v(\mathbf{x}, t)$ *defined in Eq.* (8.128) *satisfies the partial differential equation*

$$\frac{\partial v}{\partial t} + \mathscr{K} v + \lambda g(\mathbf{x}, t) v = 0, \tag{8.129}$$

*subject to the initial condition*

$$\lim_{t \to T} v(\mathbf{x}, t) = f(\mathbf{x}(T)). \tag{8.130}$$

*Proof* See Appendix 8.3. ■

Of course Proposition 8.3 contains the previous two as special cases. Thus setting $\lambda = 0$ yields Proposition 8.1 and setting $g = 1$ yields Proposition 8.2. However we have preferred to state them as separate propositions as each proposition relates to specific problems that we encounter in stochastic finance. Thus Proposition 8.1 relates to the stock option problem, Proposition 8.2 to the bond pricing (or term structure of interest rates) problem and Proposition 8.3 to the interest rate option problem.

It may happen that we need to consider functionals involving integrals with respect to increments of the Wiener process. It is possible to obtain a generalization of the Feynman–Kac formula that allows us to deal with such functionals. Here we consider just the one-dimensional case.

We consider functions having the general form

$$\psi(x,t) = \mathbb{E}_{x,t}\left[f(x(T))\exp\left\{\lambda\int_t^T g(x(s),s)ds + \gamma\int_t^T h(x(s),s)dz(s)\right\}\right], \tag{8.131}$$

where $\lambda$ and $\gamma$ are constants. By appropriate manipulations we can reduce this function to one of the same form that occurs in Proposition 8.3.

**Proposition 8.4** *The function $\psi(x,t)$ defined in Eq.* (8.131) *satisfies the partial differential equation*

$$\frac{\partial\psi}{\partial t} + \mathscr{K}\psi + \gamma\sigma h\frac{\partial\psi}{\partial x} + \left(\lambda g + \frac{1}{2}\gamma^2 h^2\right)\psi = 0, \tag{8.132}$$

*subject to the initial condition*

$$\lim_{t\to T}\psi(x,t) = f(x(T)), \tag{8.133}$$

*Proof* See Appendix 8.4. ■

## 8.7 Appendix

### Appendix 8.1 Proof of Proposition 8.1

This proof is based on that given in Gihman and Skorohod (1979) and here we consider just the one-dimensional case so as to illustrate the essential ideas. We note that (see Fig. 8.5)

$$u(x,t) = \mathbb{E}_{x,t}f(x(T)) = \mathbb{E}_{x,t}\left[\mathbb{E}_{x_{t+\Delta t},t+\Delta t}f(x(T))\right] = \mathbb{E}_{x,t}u(x(t+\Delta t),t+\Delta t). \tag{8.134}$$

By Ito's Lemma

$$\begin{aligned} u(x(t+\Delta t),t+\Delta t) &= u(x,t+\Delta t) \\ &+ \int_t^{t+\Delta t}\left[\mu(x(s),s)\frac{\partial u}{\partial x}(x(s),t+\Delta t) + \frac{1}{2}\sigma^2(x(s),s)\frac{\partial^2 u}{\partial x^2}(x(s),t+\Delta t)\right]ds \\ &+ \int_t^{t+\Delta t}\sigma(x(s),s)\frac{\partial u}{\partial x}(x(s),t+\Delta t)dz(s). \end{aligned} \tag{8.135}$$

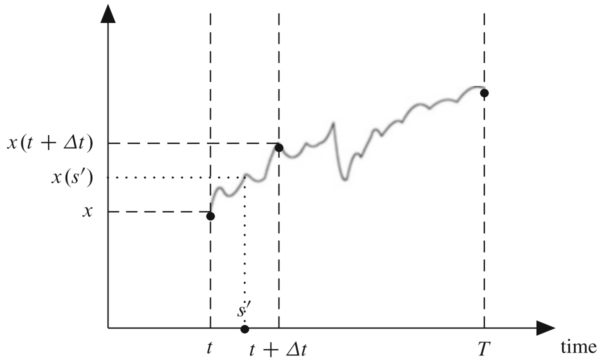


**Fig. 8.5** Successive initial points for $u(x, t)$

Applying the expectation operator $\mathbb{E}_{x,t}$ across the last equation and bearing in mind that the expectation of the stochastic integral on the right hand side is zero and that $\mathbb{E}_{x,t} u(x, t + \Delta t) = u(x, t + \Delta t)$ we obtain from (8.134) and (8.135) that

$$\begin{aligned} u(x,t) =& u(x, t + \Delta t) + \mathbb{E}_{x,t} \int_t^{t+\Delta t} \left[ \mu(x(s), s) \frac{\partial u}{\partial x}(x(s), t + \Delta t) \right. \\ & \left. + \frac{1}{2}\sigma^2(x(s), s) \frac{\partial^2 u}{\partial x^2}(x(s), t + \Delta t) \right] ds \end{aligned} \tag{8.136}$$

so that, on application of the mean value theorem for integrals

$$\begin{aligned} 0 =& u(x, t + \Delta t) - u(x, t) + \mathbb{E}_{x,t} \left[ \mu(x(s'), s') \frac{\partial u}{\partial x}(x(s'), t + \Delta t) \right. \\ & \left. + \frac{1}{2}\sigma^2(x(s), s) \frac{\partial^2 u}{\partial x^2}(x(s'), t + \Delta t) \right] \Delta t \end{aligned} \tag{8.137}$$

where $s' \in (t, t + \Delta t)$. Dividing the last equation by $\Delta t$ and passing to the limit $\Delta t \to 0$ we obtain

$$\frac{\partial u}{\partial t}(x, t) + \mu(x, t) \frac{\partial u}{\partial x}(x, t) + \frac{1}{2}\sigma^2(x, t) \frac{\partial^2 u}{\partial x^2}(x, t) = 0, \tag{8.138}$$

which is Eq. (8.123). The initial condition $\lim_{t \to T} u(x, t) = f(x)$ follows since the $\mathbb{E}_{x,t}$ and lim operations may be interchanged. Thus

$$\lim_{t \to T} u(x, t) = \lim_{t \to T} \mathbb{E}_{x,t} f(x(T)) = \mathbb{E}_{x,t} \lim_{t \to T} f(x(T)) = f(x).$$

We can use (8.138) to prove an important subsidiary result that will be useful in proving Proposition 8.2. Note that (8.134) may be written

$$\frac{\mathbb{E}_{x,t}u(x(t+\Delta t), t+\Delta t) - u(x, t+\Delta t)}{\Delta t} = \frac{-(u(x, t+\Delta t) - u(x,t))}{\Delta t}.$$

Taking the limit $\Delta t \to 0$ we obtain the result

$$\lim_{\Delta t \to 0} \frac{\mathbb{E}_{x,t}u(x(t+\Delta t), t+\Delta t) - u(x, t+\Delta t)}{\Delta t} = -\frac{\partial u}{\partial t},$$

which by use of (8.138) and the definition of the operator $\mathscr{K}$ becomes

$$\lim_{\Delta t \to 0} \frac{\mathbb{E}_{x,t}u(x(t+\Delta t), t+\Delta t) - u(x, t+\Delta t)}{\Delta t} = \mathscr{K}u. \tag{8.139}$$

The relationship (8.139) holds for any functional of the process $x$.

## Appendix 8.2 Proof of Proposition 8.2

We note first of all that by integration by parts

$$\begin{aligned}
&\lambda \int_{t'}^{t''} \exp\left[\lambda \int_s^T g(x(u), u)du\right] g(x(s), s)ds \\
&= \exp\left[\lambda \int_{t'}^T g(x(u), u)du\right] - \exp\left[\lambda \int_{t''}^T g(x(u), u)du\right].
\end{aligned} \tag{8.140}$$

Taking $t' < t''$ and applying the operator $\mathbb{E}_{x,t}$ across the last equation, we obtain (see Fig. 8.6)

$$\begin{aligned}
&v(x, t') - \mathbb{E}_{x,t'}\left(\exp\left[\lambda \int_{t''}^T g(x(u), u)du\right]\right) \\
&= \lambda \int_{t'}^{t''} \mathbb{E}_{x,t'}\left(g(x(s), s) \exp\left[\lambda \int_s^T g(x(u), u)du\right]\right)ds \\
&= \lambda \int_{t'}^{t''} \mathbb{E}_{x,t'}\big(g(x(s), s)\big) \cdot \mathbb{E}_{x(s),s}\left(\exp\left[\lambda \int_s^T g(x(u), u)du\right]\right)ds \\
&= \lambda \int_{t'}^{t''} \mathbb{E}_{x,t'}\Big(g(x(s), s) \cdot v(x(s), s)\Big)ds.
\end{aligned} \tag{8.141}$$

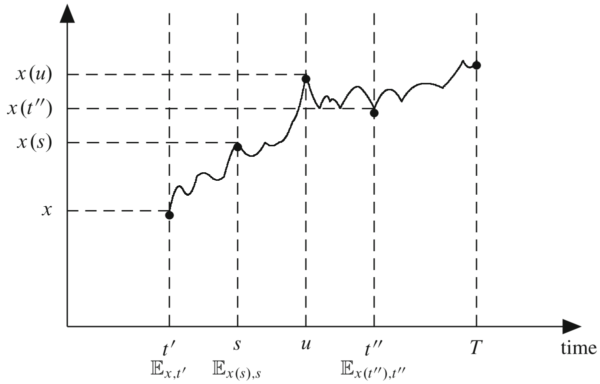


**Fig. 8.6** The telescoping of expectations for Eq. (8.141)

Note also that

$$\mathbb{E}_{x,t'}\left(\exp\left[\lambda\int_{t''}^{T} g(x(u),u)du\right]\right) = \mathbb{E}_{x,t'}\left(\mathbb{E}_{x(t''),t''}\exp\left[\lambda\int_{t''}^{T} g(x(u),u)du\right]\right)$$
$$= \mathbb{E}_{x,t'}v(x(t''),t''), \tag{8.142}$$

so that (8.141) becomes

$$v(x,t') - \mathbb{E}_{x,t'}v(x(t''),t'') = \lambda\int_{t'}^{t''} \mathbb{E}_{x,t'}\Big(g(x(s),s)v(x(s),s)\Big)ds. \tag{8.143}$$

The rest of the proof consists in setting $t' = t, t'' = t + h$ and considering the limit as $h \to 0$. Thus (8.142) becomes

$$v(x,t) - \mathbb{E}_{x,t}v(x(t+h),t+h) = \lambda\int_{t}^{t+h} \mathbb{E}_{x,t}[g(x(s),s)v(x(s),s)]ds, \tag{8.144}$$

or, adding and subtracting $v(x,t+h)$ on the left and applying the mean value theorem for integrals on the right of (8.144)

$$v(x,t) - v(x,t+h) + v(x,t+h) - \mathbb{E}_{x,t}v(x(t+h),t+h)$$
$$= \lambda g(x(t),t)\cdot v(x(t),t)\cdot h + 0(h). \tag{8.145}$$

Dividing through the last equation by $h$, taking the limit $h \to 0$, we obtain

$$-\lim_{h\to 0}\frac{\mathbb{E}_{x,t}v(x(t+h),t+h) - v(x,t+h)}{h} = \frac{\partial v}{\partial t} + \lambda g(x,t)v.$$

Applying the result (8.139) in the present context we have that

$$\lim_{h\to 0} \frac{\mathbb{E}_{x,t} v(x(t+h), t+h) - v(x, t+h)}{h} = \mathscr{K} v(x,t). \tag{8.146}$$

Hence we obtain the result that $v(x,t)$ satisfies the partial differential equation

$$\frac{\partial v}{\partial t} + \mathscr{K} v + \lambda g v = 0. \tag{8.147}$$

The initial condition $\lim_{t\to t} v(x,t) = 1$ follows fairly simply from the definition of $v(x,t)$.

## Appendix 8.3 Proof of Proposition 8.3

The proof of Proposition 8.3 turns out to be more convenient by defining a function $g$ such that (8.128) may be re-written as

$$v(x,t) = \mathbb{E}_{x,t}\left[\exp\left(\lambda \int_t^T g(x(s), s)ds + f(x(T))\right)\right].$$

It suffices to modify the first line to the proof of Proposition 8.2 (i.e. Eq. (8.140)) to

$$\begin{aligned} &\lambda \int_{t'}^{t''} \exp\left[\lambda \int_s^T g(x(u), u)du + f(x(T))\right] \cdot g(x(s), s)ds \\ &= \exp\left[\lambda \int_t^T g(x(u), u)du + f(x(T))\right] - \exp\left[\lambda \int_{t''}^T g(x(u), u)du + f(x(T))\right]. \end{aligned} \tag{8.148}$$

Applying the operator $\mathbb{E}_{x,t'}$ across this last equation, bearing in mind the definition of $v(x,t)$ and performing similar manipulations to those leading to Eq. (8.143), we find that

$$v(x,t') - \mathbb{E}_{x,t'} v(x(t''), t'') = \lambda \int_{t'}^{t''} \mathbb{E}_{x,t'}\Big[g(x(s), s) v(x(s), s)\Big] ds. \tag{8.149}$$

This last equation is the analogue in the present context of Eq. (8.143) in Proposition 8.2. By letting $t' = t, t'' = t + h$ and considering the limit as $h \to 0$ we will analogously find that $v(x,t)$ satisfies the partial differential equation

$$\frac{\partial v}{\partial t} + \mathscr{K} v + \lambda g(x,t) v(x,t) = 0.$$

The initial condition (8.130) follows easily by taking the limit $t \to T$ in (8.128).

## Appendix 8.4 Proof of Proposition 8.4

We define the function $H(x, t)$ satisfying

$$H_x(x,t) = \frac{\partial H(x,t)}{\partial x} = \frac{h(x,t)}{\sigma(x,t)}. \tag{8.150}$$

By application of Ito's lemma we find that

$$\begin{aligned} dH(x(s),s) &= [H_t(x(s),s) + \mathcal{K} H(x(s),s)]\, ds + \sigma(x(s),s) H_x(x(s),s) dz(s) \\ &= [H_t(x(s),s) + \mathcal{K} H(x(s),s)]\, ds + h(x(s),s) dz(s), \end{aligned} \tag{8.151}$$

by use of (8.151), where $H_t(x,t) = \partial H(x,t)/\partial t$. Hence integrating the last equation over the internal $[t, T]$ we obtain

$$\begin{aligned} \int_t^T h(x(s),s) dz(s) = H(x(T),T) - H(x,t) - \int_t^T & [H_t(x(s),s) \\ & + \mathcal{K} H(x(s),s)]\, ds. \end{aligned} \tag{8.152}$$

Substituting Eq. (8.152) into Eq. (8.131) we find that $\psi(x,t)$ may be expressed as

$$\begin{aligned} \psi(x,t) = e^{-\gamma H(x,t)} \mathbb{E}_{x,t} \Bigg[ \exp \Bigg\{ \lambda \int_t^T g(x(s),s) ds \\ - \gamma \int_t^T [H_t(x(s),s) + \mathcal{K} H(x(s),s)]\, ds \Bigg\} e^{\gamma H(x(T),T)} f(x(T)) \Bigg]. \end{aligned}$$

Finally we apply Proposition 8.3 to the function $\phi(x,t) \equiv e^{\gamma H(x,t)} \psi(x,t)$ to obtain

$$\begin{aligned} \frac{\partial}{\partial t} \left[ e^{\gamma H(x,t)} \psi(x,t) \right] + \mathcal{K} \left[ e^{\gamma H(x,t)} \psi(x,t) \right] \\ + [\lambda g(x,t) - \gamma H_t(x,t) - \gamma \mathcal{K} H(x,t)]\, e^{\gamma H(x,t)} \psi(x,t) = 0. \end{aligned} \tag{8.153}$$

Noting that

$$\begin{aligned} \mathcal{K} [e^{\gamma H(x,t)} \psi(x,t)] = e^{\gamma H(x,t)} \left[ \frac{1}{2} \sigma^2 \frac{\partial^2 \psi}{\partial x^2} + \gamma \sigma h \frac{\partial \psi}{\partial x} + \frac{1}{2} (\gamma \sigma^2 \frac{\partial^2 H}{\partial x^2} + \gamma^2 h^2) \psi \right] \\ + e^{\gamma H(x,t)} \left[ \mu \frac{\partial \psi}{\partial x} + \mu \gamma \frac{\partial H}{\partial x} \psi \right], \end{aligned}$$

and that

$$\mathscr{K}H(x,t) = \frac{1}{2}\sigma^2\frac{\partial^2 H}{\partial x^2} + \mu\frac{\partial H}{\partial x},$$

we find after some algebraic manipulations that $\psi(x,t)$ indeed satisfies the partial differential equation (8.132). Finally we note directly from (8.131) that

$$\lim_{t\to T}\psi(x,t) = f(x(T)),$$

which provides the initial condition (8.133).

## 8.8 Problems

**Problem 8.1** Consider the expression (8.107) for the price of a European call option, namely

$$C(S,t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[C(S_T,T)],$$

where $\tilde{\mathbb{E}}_t$ is generated according to the process

$$dS = rSdt + \sigma Sd\tilde{z}(t).$$

(i) By simulating $M$ paths for $S$ approximate the expectation with

$$\frac{1}{M}\sum_{i=0}^{M} C(S_T^{(i)},T),$$

where $i$ indicates the $i$th path. Take $r = 5\,\%$p.a., $\sigma = 20\,\%$p.a., $S = 100$, $E = 100$ and $T = 6$ months.

(ii) Compare graphically the simulated values for various $M$ with the true Black–Scholes value.

(iii) Instead of using discretisation to simulate paths for $S$, use instead the result in Eq. (6.16). We know from Problem 6.16 that this involves no discretisation error.

# Chapter 9
# The Partial Differential Equation Approach Under Geometric Brownian Motion

**Abstract** The Partial Differential Equation (PDE) Approach is one of the techniques in solving the pricing equations for financial instruments. The solution technique of the PDE approach is the Fourier transform, which reduces the problem of solving the PDE to one of solving an ordinary differential equation (ODE). The Fourier transform provides quite a general framework for solving the PDEs of financial instruments when the underlying asset follows a jump-diffusion process and also when we deal with American options. This chapter illustrates that in the case of geometric Brownian motion, the ODE determining the transform can be solved explicitly. It shows how the PDE approach is related to pricing derivatives in terms of integration and expectations under the risk-neutral measure.

## 9.1 Introduction

In this chapter we outline two ways in which the pricing equations of financial instruments may be solved. The first starts from the expression

$$\begin{aligned} C(S,t) &= e^{-r(T-t)}\tilde{\mathbb{E}}_t[C(S_T,T)] \\ &= e^{-r(T-t)}\int_0^{\infty} \max[S_T - E, 0]\tilde{p}(S_T, T \mid S_t)dS_T \end{aligned}$$

that gives the derivative value in terms of an expectation under the risk-neutral measure. To operationalise this expression we need to obtain an explicit expression for the conditional transition density function, so that the problem of derivative security evaluation reduces to an exercise in integration. It turns out that for the case when the underlying asset price is driven by geometric Brownian motion it is possible to obtain explicitly the conditional transition density function directly from an analysis of the stochastic differential equation.

The second approach seeks to solve the (parabolic) partial differential equation

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} = rC$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_9

that determines the derivative price. Different problems (e.g. European options, put options, digital options, exchange options or just simply the Kolmogorov equation) are characterized by different boundary conditions for this equation. The solution technique we propose to solve those partial differential equation is the Fourier transform. Essentially this technique reduces the problem of solving the partial differential equation first to one of solving an ordinary differential equation to obtain the transform and then to one of carrying out an integration to invert the transform. In the case of geometric Brownian motion this technique works very nicely because the ordinary differential equation determining the transform can be solved explicitly. Of course both approaches lead to the same final expression for the derivative price. It is important to see how they are related to each other. Each has its advantages in particular problems and we encourage the reader to always visualize the derivative pricing problem from both perspectives.

The Fourier transform perhaps has the advantage that it provides quite a general framework for solving the partial differential equations of financial instruments as it generalizes quite nicely to handle the partial differential equations we encounter when the underlying asset follows a jump-diffusion process and also when we come to deal with American options.

## 9.2 The Transition Density Function for Geometric Brownian Motion

Consider a derivative instrument written on an underlying asset whose price $x$ follows the diffusion process

$$dx = \mu(x,t)dt + \sigma(x,t)dz.$$

Furthermore suppose that the derivative instrument is of European type that pays off only at time $T$ and use $h(x_T, T)$ to denote the payoff function.

Under the further assumption of a constant risk-free rate of interest, $r$, we know from the discussion of Chap. 8 that the value at time $t$ of the derivative, that we denote $f(x,t)$, is given by

$$f(x,t) = e^{-r(T-t)} \int_0^\infty h(x_T, T)\tilde{p}(x_T, T|x,t)dx_T. \tag{9.1}$$

In (9.1) the density function $\tilde{p}(x_T, T|x,t)$ is the one that is the solution of the Kolmogorov backward equation associated with the asset price dynamics under the risk-neutral measure, namely

$$dx = rxdt + \sigma(x,t)d\tilde{z}. \tag{9.2}$$

The Kolmogorov backward equation associated with (9.2) may be written

$$\frac{\partial \tilde{p}}{\partial t} + rx\frac{\partial \tilde{p}}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 \tilde{p}}{\partial x^2} = 0,$$

subject to the terminal condition

$$\tilde{p}(x_T, T|x, T) = \delta(x_T - x).$$

In this chapter we focus on the case when the dynamics under the risk-neutral measure are geometric Brownian motion. This is the case when $\sigma(x,t) = \sigma x$ ($\sigma$ constant) so that (9.2) become

$$dx = rxdt + \sigma xd\tilde{z}. \tag{9.3}$$

We know from Sect. 6.3.2 that for the stochastic process defined by (9.3), if $y = \ln x$ then the quantity $y(T) - y(t)$ over the interval $t$ to $T$ is normally distributed with mean $(r - \frac{1}{2}\sigma^2)(T-t)$ and variance $\sigma^2(T-t)$. This means that the density function $\tilde{p}(x_T, T|x, t)$ is log-normal i.e.

$$\tilde{p}(x_T, T|x, t) = \frac{1}{\sqrt{2\pi(T-t)}\sigma x_T}\exp\left[-\frac{\left(\ln(\frac{x_T}{x}) - (r - \frac{1}{2}\sigma^2)(T-t)\right)^2}{2\sigma^2(T-t)}\right]. \tag{9.4}$$

Substituting this density function into (9.1) we obtain

$$f(x,t) = \frac{e^{-r(T-t)}}{\sqrt{2\pi(T-t)}\sigma}\int_0^\infty h(x_T, T)e^{-\frac{\left(\ln\frac{x_T}{x} - (r-\frac{1}{2}\sigma^2)(T-t)\right)^2}{2\sigma^2(T-t)}}\frac{dx_T}{x_T}. \tag{9.5}$$

The evaluation of the derivative then becomes an exercise in integration once the particular functional form for $h(x_T, T)$ is chosen. For example in the case of a European call option $h(x_T, T) = (x_T - E)^+$ and the integral in (9.5) reduces to the integral (3.16) whose evaluation has been detailed in Appendix 3.1. We shall consider other examples in later sections.

## 9.3 The Fourier Transform

Now we turn to the Fourier transform approach. The Fourier transform of the function $f(x)$ is defined as

$$\mathscr{F}[f(x)] = \bar{f}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx.$$

The Fourier inversion theorem states that

$$\mathscr{F}^{-1}[\bar{f}(\omega)] = f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{f}(\omega)e^{i\omega x}d\omega.$$

The Fourier convolution theorem states that for two functions $f$ and $g$,

$$\mathscr{F}\left[\int_{-\infty}^{\infty} f(x-u)g(u)du\right] = \bar{f}(\omega)\bar{g}(\omega),$$

which can alternatively be stated as

$$\mathscr{F}^{-1}[\bar{f}(\omega)\bar{g}(\omega)] = \int_{-\infty}^{\infty} f(x-u)g(u)du. \tag{9.6}$$

It is the result (9.6) that will generally be most useful in our applications.

We consider the partial differential equation for the case of one underlying asset that pays a continuously compounded dividend at the rate $q$ and whose dynamics are driven by geometric Brownian motion. From Chap. 10 we know that the derivative security price $f$ satisfies

$$\frac{\partial f}{\partial t} + (r(t) - q(t))x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(t)x^2\frac{\partial^2 f}{\partial x^2} - r(t)f = 0. \tag{9.7}$$

Note that we have adopted a notation that allows the risk-free rate $r$, continuous dividend yield and volatility $\sigma$ to all be functions of time. Equation (9.7) is to be solved on the interval $(t, T)$ subject to the final time condition

$$f(x_T, T) = h(x_T), \tag{9.8}$$

when $h(x_T)$ is some payoff function. By considering different functions $h$ we obtain different derivative securities, e.g.

$$\begin{aligned} h(x_T) &= (x_T - E)^+; && \text{European call option}\\ h(x_T) &= (E - x_T)^+; && \text{European put option}\\ h(x_T) &= E\mathscr{H}(x_T - E); && \text{European digital option,} \end{aligned}$$

where the Heaviside function $\mathscr{H}(x)$ is defined by[1]

$$\mathscr{H}(x) = \begin{cases} 0, & \text{for } x < 0,\\ 1, & \text{for } x \geq 0. \end{cases} \tag{9.9}$$

[1]Note that it is common in the literature to use instead of the Heaviside function notation, the indicator function notation

$$\mathbf{1}_{\{x\geq 0\}} = \begin{cases} 1, & \text{for } x \geq 0,\\ 0, & \text{otherwise.} \end{cases}$$

Before proceeding we first make the change of variable

$$y = \ln x, \qquad F(y,t) \equiv f(e^y, t), \tag{9.10}$$

in terms of which the partial differential equation (9.7) becomes

$$\frac{\partial F}{\partial t} + c(t)\frac{\partial F}{\partial y} + \frac{1}{2}\sigma^2(t)\frac{\partial^2 F}{\partial y^2} - rF = 0, \tag{9.11}$$

where we set $c(t) = r(t) - q(t) - \frac{1}{2}\sigma^2(t)$, the cost-of-carry. In terms of the variable $y$ the final time condition (9.8) becomes

$$F(y_T, T) = f(e^{y_T}, T) = h(e^{y_T}) \equiv H(y_T). \tag{9.12}$$

Define the Fourier transform of the solution to (9.11) $\bar{F}(\omega, t)$ by

$$\bar{F}(\omega, t) = \int_{-\infty}^{\infty} F(y,t)e^{-i\omega y}\,dy. \tag{9.13}$$

We note that

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\partial F(y,t)}{\partial y} e^{-i\omega y}\,dy &= [F(y,t)e^{-i\omega y}]_{-\infty}^{\infty} + i\omega \int_{-\infty}^{\infty} F(y,t)e^{-i\omega y}\,dy \\ &= [F(y,t)e^{-i\omega y}]_{-\infty}^{\infty} + i\omega \bar{F}(\omega, t).\end{aligned}$$

At this stage we shall tentatively assume that $\lim_{y\to\pm\infty} F(y,t)e^{-i\omega y} = 0$ so that

$$\int_{-\infty}^{\infty} \frac{\partial F}{\partial y}(y,t)e^{-i\omega y}\,dy = i\omega \bar{F}(\omega, t). \tag{9.14}$$

Next consider

$$\int_{-\infty}^{\infty} \frac{\partial^2 F}{\partial y^2}(y,t)e^{-i\omega y}\,dy = \left[\frac{\partial F}{\partial y}(y,t)e^{-i\omega y}\right]_{-\infty}^{\infty} + i\omega \int_{-\infty}^{\infty} \frac{\partial F}{\partial y}(y,t)e^{-i\omega y}\,dy.$$

Here also we tentatively assume that $\lim_{y\to\pm\infty} \frac{\partial F}{\partial y}(y,t)e^{-i\omega y} = 0$. So that, upon use of (9.14),

$$\int_{-\infty}^{\infty} \frac{\partial^2 F}{\partial y^2} e^{-i\omega y}\,dy = -\omega^2 \bar{F}(\omega, t). \tag{9.15}$$

Next we observe by differentiating (9.13) with respect to time that

$$\int_{-\infty}^{\infty} \frac{\partial F}{\partial t}(y,t)e^{-i\omega y}dy = \frac{\partial \bar{F}}{\partial t}(\omega,t). \tag{9.16}$$

Thus as a result of applying the Fourier transform to Eq. (9.11) we have obtained the ordinary differential equation

$$\frac{\partial \bar{F}}{\partial t}(\omega,t) = [\frac{1}{2}\sigma^2(t)\omega^2 + r(t) - i\omega c(t)]\bar{F}(\omega,t). \tag{9.17}$$

This equation must be solved for the transform $\bar{F}(\omega,t)$, subject to the boundary condition

$$\bar{F}(\omega,T) = \int_{-\infty}^{\infty} H(y)e^{-i\omega y}dy \equiv \bar{H}(\omega), \tag{9.18}$$

obtained by taking the Fourier transform of Eq. (9.12). Equation (9.17) can be rewritten

$$\frac{\partial}{\partial t}\left[\bar{F}(w,t)e^{-\int_0^t(\frac{1}{2}\sigma^2(s)\omega^2+r(s)-i\omega c(s))ds}\right] = 0. \tag{9.19}$$

Integrating (9.19) from $t$ to $T$, using the final time condition (9.18) and rearranging we obtain

$$\bar{F}(\omega,t) = \bar{H}(\omega)e^{-\int_t^T(\frac{1}{2}\sigma^2(s)\omega^2+r(s)-i\omega c(s))ds}. \tag{9.20}$$

For the purpose of subsequent manipulations it will be convenient to introduce the time averaged quantities[2]

$$\bar{\sigma}^2(T-t) = \frac{1}{T-t}\int_t^T \sigma^2(s)ds,$$
$$\bar{r}(T-t) = \frac{1}{T-t}\int_t^T r(s)ds,$$
$$\bar{c}(T-t) = \frac{1}{T-t}\int_t^T c(s)ds,$$

[2]Note that these may also be written

$$\bar{\sigma}^2(\theta) = \frac{1}{\theta}\int_0^\theta \sigma^2(T-u)du, \quad \bar{r}(\theta) = \frac{1}{\theta}\int_0^\theta r(T-u)du, \quad \bar{c}(\theta) = \frac{1}{\theta}\int_0^\theta c(T-u)du.$$

in terms of which (9.8) can be more succinctly written

$$\bar{F}(\omega, t) = \bar{H}(\omega)e^{-[\frac{1}{2}\bar{\sigma}^2(T-t)\omega^2+\bar{r}(T-t)-i\omega\bar{c}(T-t)](T-t)}. \tag{9.21}$$

Our next task is to use the Fourier inversion theorem to recover the function $F(y, t)$ from (9.8). By the Fourier inversion theorem

$$F(y, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{F}(\omega, t)e^{i\omega y}d\omega,$$

which by use of (9.21) becomes

$$\begin{aligned} F(y, t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{H}(\omega)e^{-(\frac{1}{2}\bar{\sigma}^2(T-t)\omega^2+\bar{r}(T-t)-i\omega\bar{c}(T-t))(T-t)+i\omega y}d\omega \\ &= \frac{e^{-\bar{r}(T-t)(T-t)}}{2\pi}\int_{-\infty}^{\infty} \bar{H}(\omega)\bar{K}(\omega)e^{i\omega y)}d\omega, \end{aligned} \tag{9.22}$$

where

$$\bar{K}(\omega) = e^{[-\frac{1}{2}\bar{\sigma}^2(T-t)\omega^2+i\omega\bar{c}(T-t)](T-t)}.$$

By the Fourier inversion theorem

$$K(y, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\bar{\sigma}^2(t)t\omega^2+i\omega(\bar{c}(t)t+y)}d\omega. \tag{9.23}$$

Using the result that

$$\int_{-\infty}^{\infty} e^{-p\omega^2-q\omega}d\omega = \sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}} \tag{9.24}$$

with $p = \frac{1}{2}\bar{\sigma}^2(t)t$ and $q = -i(\bar{c}(t)t + y)$, we obtain that

$$K(y, t) = \frac{1}{\sqrt{2\pi t}\bar{\sigma}(t)}e^{-\frac{(\bar{c}(t)t+y)^2}{2\bar{\sigma}^2(t)t}}. \tag{9.25}$$

Application of the convolution theorem to (9.22) yields

$$\begin{aligned} F(y, t) &= e^{-\bar{r}(T-t)(T-t)}\int_{-\infty}^{\infty} H(\xi)K(y-\xi, T-t)d\xi \\ &= \frac{e^{-\bar{r}(T-t)(T-t)}}{\sqrt{2\pi(T-t)}\bar{\sigma}(T-t)}\int_{-\infty}^{\infty} H(\xi)e^{-\frac{(\bar{c}(T-t)(T-t)+y-\xi)^2}{2\bar{\sigma}^2(T-t)(T-t)}}d\xi. \end{aligned}$$

Finally we use (9.10) to transform back to the original variable $x(= e^y)$ and correspondingly make the change of integration variable $u = e^\xi$, so that

$$f(x,t) = \frac{e^{-\bar{r}(T-t)(T-t)}}{\sqrt{2\pi(T-t)}\bar{\sigma}(T-t)} \int_0^\infty h(u) e^{-\frac{(\ln(u/x)-\bar{c}(T-t)(T-t))^2}{2\bar{\sigma}^2(T-t)(T-t)}} \frac{du}{u}. \tag{9.26}$$

## 9.4 Solutions for Specific Payoff Functions

In this section we consider various forms for the payoff function $h$, which yield various cases of interest.

### *9.4.1 The Kolmogorov Equation*

In Sect. 3.1 we stated the solution of the Kolmogorov equation, under both the historical and risk-neutral measures (see Eqs. (3.13) and (3.14)). We can now use (9.26) to obtain the results.

We note first of all that to obtain the Kolmogorov equation we interpret $f(x,t)$ in (9.7) as $\tilde{p}(x_T, T|x,t)$ and set $q(t) = 0$ in the first derivative term and set $r = 0$ in the $rf$ term to obtain the risk-neutral density, that is,

$$\frac{\partial \tilde{p}}{\partial t} + r(t)x\frac{\partial \tilde{p}}{\partial x} + \frac{1}{2}\sigma^2(t)x^2\frac{\partial^2 \tilde{p}}{\partial x^2} = 0.$$

Note that

$$h(u) = \delta(u - x_T). \tag{9.27}$$

With these changes (9.26) gives the risk-neutral conditional transition density

$$\begin{aligned} &\tilde{p}(x_T, T|x,t) \\ &= \frac{1}{\sqrt{2\pi(T-t)}\bar{\sigma}(T-t)} \int_0^\infty \delta(u-x_T) e^{-\frac{(\ln(u/x)-(\bar{r}(T-t)-\frac{1}{2}\bar{\sigma}^2(T-t))(T-t))^2}{2\bar{\sigma}^2(T-t)(T-t)}} \frac{du}{u} \\ &= \frac{1}{\sqrt{2\pi(T-t)}\bar{\sigma}(T-t)x_T} \exp\left[-\frac{(\ln(x_T/x)-(\bar{r}(T-t)-\frac{1}{2}\bar{\sigma}^2(T-t))(T-t))^2}{2\bar{\sigma}^2(T-t)(T-t)}\right]. \end{aligned} \tag{9.28}$$

We observe that (9.28) is precisely the same as (9.4) that was obtained directly from analysis of the stochastic differential equation (9.3) generating the sample paths of this conditional transition density function. In the case $r$ and $\sigma$ are constant (9.28) reduces to (3.14).

To obtain the historical conditional transition density $p(x_T, T|x,t)$ given in Eq. (3.13) we need to set $q = r - \mu$ in the first derivative term of (9.7) and still maintain $r = 0$ in the $rf$ term. We leave the details as an exercise for the reader.

### 9.4.2 *The European Digital Option*

A European digital option with exercise price $E$ pays $E$ if the underlying asset price at maturity is greater than $E$, and pays zero otherwise. Thus the payoff for this option is given by

$$h(x_T) = E\mathscr{H}(x_T - E)$$

and is illustrated in Fig. 9.1. Applying Eq. (9.26) the value of this option at time $t$ is given by

$$f(x,t) = \frac{e^{-\bar{r}(T-t)}}{\sqrt{2\pi(T-t)}\bar{\sigma}} \int_E^{\infty} E e^{-\frac{\{\ln(u/x)-\bar{c}(T-t)\}^2}{2\bar{\sigma}^2(T-t)}} \frac{du}{u}. \tag{9.29}$$

The integral involved in (9.29) is essentially $A_2$ (Eq. (3.29)) of Appendix 3.1 and we readily find that

$$f(x,t) = Ee^{-\bar{r}(T-t)}\mathscr{N}(d_2), \qquad d_2 = \frac{\ln(x/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

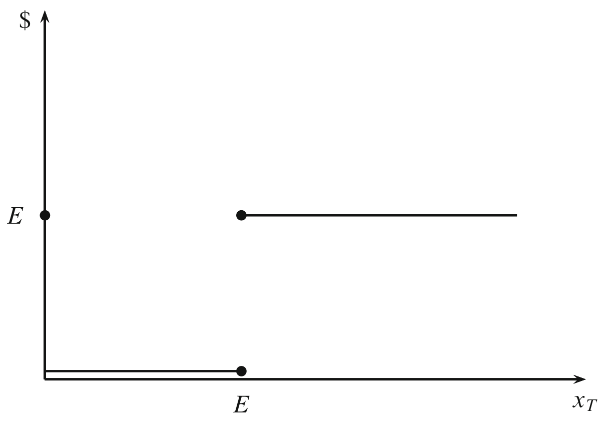


**Fig. 9.1** Payoff on a European digital option

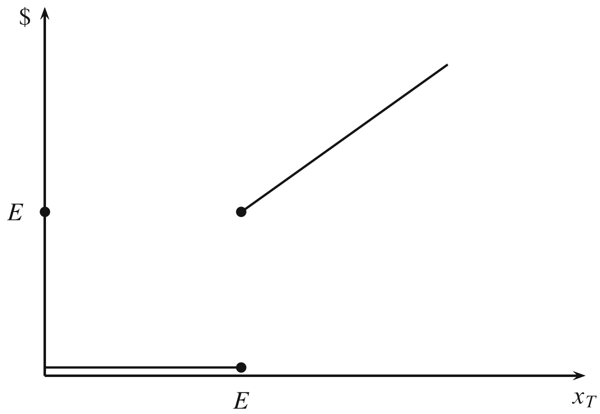


**Fig. 9.2** Payoff on a European all-or-nothing option

### 9.4.3 The European All-or-Nothing Option

A European all-or-nothing option with exercise price $E$ pays the value of the underlying asset if it is greater than $E$ at maturity, otherwise it pays nothing. The payoff for this option may be written

$$h(x_T) = x_T \mathscr{H}(x_T - E)$$

and is illustrated in Fig. 9.2. Applying Eq. (9.26) the value of an all-or-nothing European option at time $t$ is given by

$$f(x,t) = \frac{e^{-\bar{r}(T-t)}}{\sqrt{2\pi(T-t)}\bar{\sigma}} \int_E^\infty u e^{-\frac{\{\ln(u/x)-\bar{c}(T-t)\}^2}{2\bar{\sigma}^2(T-t)}} \frac{du}{u}.$$

This integral is essentially $A_1$ (Eq. (3.28)) of Appendix 3.1 and as a result

$$f(x,t) = x\mathscr{N}(d_1), \qquad d_1 = \frac{\ln(x/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}.$$

### 9.4.4 The European Call Option

Equation (9.26) reduces to the value of a European call option by applying the payoff condition that

$$h(u) = (u - E)^+,$$

so that

$$f(x,t) = \frac{e^{-\bar{r}(T-t)}}{\sqrt{2\pi(T-t)}\bar{\sigma}} \int_E^\infty (u-E)e^{-\frac{\{\ln(u/x)-\bar{c}(T-t)\}^2}{2\bar{\sigma}^2(T-t)}} \frac{du}{u}.$$

This is precisely the integral we evaluated in Appendix 3.1, so that

$$f(x,t) = x\mathscr{N}(d_1) - Ee^{-\bar{r}(T-t)}\mathscr{N}(d_2), \tag{9.30}$$

where

$$d_1 = \frac{\ln(x/E) + (\bar{r} + \bar{\sigma}^2/2)(T-t)}{\bar{\sigma}\sqrt{T-t}},$$
$$d_2 = d_1 - \bar{\sigma}\sqrt{T-t}.$$

In order to interpret each term of Eq. (9.30) we note that the payoff at maturity, $\max(S - E, 0)$, could also be constructed by holding a portfolio consisting of a short position in a European digital option and long position in a European all-or-nothing option. As we have seen in the previous sections; the value at time $t$ of a long position in a European all-or-nothing option with exercise price $E$ is $x\mathscr{N}(d_1)$, and this is the first term of Eq. (9.30). The value of a short position in a European digital option with exercise price $E$ and payoff $E$ is $-Ee^{-\bar{r}(T-t)}\mathscr{N}(d_2)$, and this is the second term of Eq. (9.30).

## 9.5 Interpreting the General Pricing Relation

The solution (9.28) to the risk-neutral Kolmogorov equation allows us to give a very simple economic interpretation to Eq. (9.26). In terms of the notation of Eq. (9.28) we see that the exponential quantity multiplying the payoff $h(u)$ in the integrand in (9.26) is in fact the risk-neutral conditional transition density $\tilde{p}(u, T|x, t)$. Thus Eq. (9.26) may be written

$$f(x,t) = e^{-\bar{r}(T-t)(T-t)} \int_0^\infty h(u)\tilde{p}(u, T|x, t)du. \tag{9.31}$$

Equation (9.31) has the discounted expected payoff interpretation that we have already encountered back at Eq. (3.5). There however our financial arguments were somewhat tentative. We have now arrived at the same result using precise financial and mathematical arguments.

Because $\tilde{p}(u, T|x, t)$ is associated with the boundary condition (9.27) it can be interpreted as the value at time $t$ (where the price is $x$) of an elementary claim that pays \$1 at time $T$ if the price is $u$, and \$0 if any other price occurs at $T$. It is often convenient to incorporate the discount factor with the $\tilde{p}(u, T|x, t)$ and define the function

$$G(u, T|x, t) = e^{-\bar{r}(T-t)(T-t)}\tilde{p}(u, T|x, t).$$

The function $G(u, T|x, t)$ can be interpreted as the discounted value at time $t$ (where the price is $x$) of an elementary claim that pays \$1 at time $T$ if the price is $u$, and \$0 if any other price occurs at $T$. In terms of this function the solution (9.31) may be written

$$f(x, t) = \int_0^{\infty} h(u)G(u, T|x, t)du, \tag{9.32}$$

which may be interpreted as follows: *If at time $T$ the price $u$ occurs the payoff on the option is $h(u)$; this is multiplied by the discounted value of the elementary claim that pays \$1 if this state*[3] *occurs to give the discounted value of the payoff in this particular state; the integral in* (9.32) *then sums over all possible states to give the discounted value of the total payoff.*

In the literature on the solution of partial differential equations, the function $G(u, T|x, t)$ is known as the Green's function, and (9.32) is the Green's function representation of the solution of the partial differential equation (9.7) with boundary condition (9.8). The Green's function technique reduces the problem of solving a partial differential equation subject to a given boundary condition to a two-pass process.

First solve the partial differential equation subject to the $\delta$-function boundary condition (in financial economic applications thus usually means solving the Kolmogorov equation associated with the problem at hand) to obtain the Green's function for this particular partial differential operator. In the second phase the integral operation (9.32) combines the Green's function with the boundary condition (payoff function) to give the solution. In this way the problem of solving the partial differential equation is reduced to an exercise in integration (often a very difficult one). For a more extensive discussion on Green's functions we refer the reader to Greenberg (1971).

We obtain yet another important interpretation of expression (9.31) by recalling that $\tilde{p}(u, T|x, t)du$ is the probability of observing a price in the interval $(u, u+du)$ at time $T$ given that the price was $x$ at time $t$. Thus the integral in (9.31) is the expected value (where probability evolves under the risk-neutral Kolmogorov equation) at time $t$ of the payoff to be received at time $T$. The discount factor $e^{-\bar{r}(T-t)(T-t)}$ converts this expected payoff into dollars at time $t$. In terms of the notation we used in Chap. 3 we could thus write (9.31) as

$$f(x, t) = e^{-\bar{r}(T-t)(T-t)}\tilde{\mathbb{E}}_t[h(x_T)], \tag{9.33}$$

[3]The state being the occurrence of the particular price $u$ at time $T$.

which is essentially (3.18) with some obvious changes of notation. We have shown in Chap. 8 how to arrive at (9.33) via martingale arguments.

## 9.6 Appendix

### Appendix 9.1 Transforming the Black–Scholes Partial Differential Equation to the Heat Equation

The early literature on option pricing obtained the solution to the Black–Scholes partial differential equation by transforming it to the heat equation and then applying known solutions for that problem. We reproduce this approach here to give another perspective on the solution of the Black–Scholes partial differential equation.

Consider the Black–Scholes partial differential equation

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC, \tag{9.34}$$

subject to the boundary condition

$$C(0, t; E, \sigma^2, r) = 0, \tag{9.35}$$

and final condition

$$C(S, T; E, \sigma^2, r) = \max[0, S - E]. \tag{9.36}$$

Transform the time variable $t$ to time-to-maturity $\tau$ given by $\tau = T - t$ in terms of which (9.34) becomes

$$-\frac{\partial C}{\partial \tau} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC, \tag{9.37}$$

and the conditions (9.35) and (9.36) become

$$\begin{aligned} C(0, \tau; E, \sigma^2, r) &= 0, \\ C(S, 0; E, \sigma^2, r) &= \max[0, S - E]. \end{aligned} \tag{9.38}$$

The condition (9.38) is now an initial condition. The following sequence of transformations reduce (9.37) to the classical heat equation.

First of all define a new function $F(S, \tau)$ by

$$C(S, \tau) = e^{-r\tau} F(S, \tau). \tag{9.39}$$

Since

$$\frac{\partial C}{\partial S} = e^{-r\tau}\frac{\partial F}{\partial S}, \quad \frac{\partial^2 C}{\partial S^2} = e^{-r\tau}\frac{\partial^2 F}{\partial S^2},$$

and

$$\frac{\partial C}{\partial \tau} = -re^{-r\tau}F + e^{-r\tau}\frac{\partial F}{\partial \tau},$$

the partial differential equation satisfied by $F$ is

$$-\frac{\partial F}{\partial \tau} + rS\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = 0.$$

Next introduce a new state variable $X$ related to the original state variable $S$ by

$$X = Se^{r\tau} \quad \text{or} \quad S = Xe^{-r\tau}$$

and put $F(Xe^{-r\tau}, \tau) \equiv f(X, \tau)$. Since

$$\frac{\partial f}{\partial S} = \frac{\partial f}{\partial X} \cdot \frac{\partial X}{\partial S} = e^{r\tau}\frac{\partial f}{\partial X} = \frac{X}{S}\frac{\partial f}{\partial X},$$

$$\frac{\partial^2 f}{\partial S^2} = e^{r\tau} \cdot e^{r\tau}\frac{\partial^2 f}{\partial X^2} = \left(\frac{X}{S}\right)^2 \frac{\partial^2 f}{\partial X^2},$$

and

$$\frac{\partial f}{\partial \tau} = \frac{\partial f}{\partial X} \cdot \frac{\partial X}{\partial \tau} + \frac{\partial f}{\partial \tau} = rX\frac{\partial f}{\partial X} + \frac{\partial f}{\partial \tau},$$

then the partial differential equation for $f$ becomes

$$-\frac{\partial f}{\partial \tau} + \frac{1}{2}\sigma^2 X^2 \frac{\partial^2 f}{\partial X^2} = 0.$$

Next set $\theta = \sigma^2 \tau$ so that

$$-\frac{\partial f}{\partial \theta} + \frac{1}{2}X^2 \frac{\partial^2 f}{\partial X^2} = 0.$$

Recall that this must be solved subject to the boundary condition

$$f(0, \theta) = 0,$$

and initial condition

$$f(X,0) = \max[0, X - E].$$

If we let

$$Y = X/E$$

and

$$f(EY,\theta) \equiv g(Y,\theta)$$

then we have for $g$ the partial differential equation

$$-\frac{\partial g}{\partial \theta} + \frac{1}{2}Y^2\frac{\partial^2 g}{\partial Y^2} = 0, \tag{9.40}$$

subject to

$$g(0,\theta) = 0,$$
$$g(Y,0) = \max[0, Y - 1].$$

Equation (9.40) could have been arrived at more readily by the original transformation

$$\theta = \sigma^2\tau, \qquad Y = \frac{S}{E}e^{r\tau}, \qquad g = \frac{e^{r\tau}}{E}C. \tag{9.41}$$

Finally we perform the change of variable

$$y = \ln Y + \frac{1}{2}\theta, \qquad \phi(y,\theta) = g(Y,\theta)/Y,$$

to arrive at

$$\frac{1}{2}\frac{\partial^2\phi}{\partial y^2} = \frac{\partial\phi}{\partial\theta}, \tag{9.42}$$

subject to

$$\begin{aligned} &\text{(a)} \quad |\,\phi\,| \le 1, \\ &\text{(b)} \quad \phi(y,0) = \max[0, 1 - e^{-y}]. \end{aligned} \tag{9.43}$$

The partial differential equation (9.42) for $\phi$ is the heat conduction equation in an infinite rod. The solution may be obtained either by the technique of separation of variables (see e.g. Greenberg 1978, pp. 533–536) or by use of the Fourier transform discussed in this chapter. The solution may be written

$$\phi(y,\theta) = \int_{-\infty}^{\infty} \phi(\zeta,0)\frac{e^{-(\zeta-y)^2/2\theta}}{\sqrt{2\pi\theta}}d\zeta.$$

After substituting in the expression for $\phi(\zeta,0)$ from (9.43) and reversing the transformations we arrive at the integral expression for the option price given in Appendix 3.1 and which as we saw there becomes the Black–Scholes formula.

# Chapter 10
# Pricing Derivative Securities: A General Approach

**Abstract** This chapter extends the hedging argument developed in Chap. 7 and the martingale approach developed in Chap. 8 by allowing derivative securities to depend on multiple sources of risks and multiple underlying factors, some are tradable and some are not tradable. It provides a general PDE and martingale approaches to pricing derivative securities.

## 10.1 Risk Neutral Valuation

A feature of the Black–Scholes option pricing model is that the theoretical option price is independent of the drift of the price process of the underlying asset. This is a reflection of the fact that the theoretical option price is preference free since it does not involve in any way investor's attitude to risk (i.e. an investor who is very cautious about risky investments would price the option in exactly the same way as an investor who is eager to gamble all on the riskiest of investment proposals). This feature reflects the fact that the option pricing formula is a relative pricing formula i.e. given the underlying asset price, the risk-free arbitrage argument tells us what the option on that underlying asset should be worth in an efficient market. The factors that do affect the option price are the exercise price (E), the current stock price (S), the stock price volatility ($\sigma$), the time to maturity ($T - t$) and the risk free rate of interest ($r$), all of which are observable in the market (either directly or indirectly).

Finance theory classifies investors as *risk averse, risk neutral* or *risk lover*. A risk averse investor would require an expected return in excess of the risk free rate in order to hold a risky asset, the premium would depend in part on the investor's degree of risk aversion. An investor who is satisfied with an expected return below the risk free rate is said to be a risk lover. An investor who is satisfied with an expected return equal to the risk-free rate on all risky assets is said to be risk-neutral.

The expected return, $\rho$, is the rate at which investors will discount the expected payoff on risky investments. This discount rate may be expressed as

$$\rho = r + \text{risk premium}.$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_10

To value at time $t$ a risky asset yielding some expected payoff at time $T$ in the future, the investor uses the discounting rule

$$\text{value of risky asset} = e^{-\rho(T-t)}\mathbb{E}_t[\text{payoff of risky asset at time } T],$$

where $\mathbb{E}_t$ denotes expectation conditional on information up to and including time $t$. Thus a European call option would be valued according to

$$\text{value of call option} = e^{-\rho(T-t)}\mathbb{E}_t[\max(0, S_T - E)],$$

and the expectation operator $\mathbb{E}_t$ would be calculated according to the log-normal distribution, involving in particular the expected stock return $\mu$.

The Black–Scholes option pricing model tells us that investors value the option as if they were risk-neutral i.e. with $\rho = r$ and $\mu = r$. This is *not* to assert that all investors are risk neutral. Certainly the values of $\rho$ and $\mu$ in the above valuation formula vary as investor's risk preferences vary. It turns out that in perfect capital markets with investors maximizing profit or utility, these two effects offset each other in such a way that apparently $\rho = r$ and $\mu = r$.

This observation has given rise to the risk neutral valuation principle, namely that investors value the options *as if* they were risk neutral and were expecting the underlying stock to yield a return equal to the risk free rate. In the following sections we shall see how this principle can be extended to yield a quite general framework for the pricing of derivative securities.

## 10.2 The Market Price of Risk

The theory of option pricing that we developed in Chaps. 6 and 8 worked out so neatly because we were dealing with one risk source (the Wiener process $w(t)$) and one traded factor (the common stock) which when combined with the option in a portfolio allowed us to hedge away the one risk source. In this chapter we want to allow for multiple sources of risk and multiple underlying factors. Some of the underlying factors may be traded (e.g. common stock, foreign currency) and some may not be traded (e.g. interest rates, inflation, stochastic volatility). We will find ourselves dealing with a situation in which the number of sources of risk is greater than the number of underlying traded factors. The number of traded factors is then insufficient to hedge away all sources of risk when combined in a portfolio with an option dependent on the underlying factors. We are now dealing with a so-called incomplete markets situation. In order to now hedge away all sources of risk we need to complete the market in some fashion. This is usually done by placing a sufficient number of options of different maturities in the hedging portfolio. It is then possible to hedge away all sources of risk, but the resulting option pricing equations involve a set of quantities which are the market prices of risk of some of the sources of risk, the number of these being equal to the number of non-traded underlying factors. The market prices of risk are difficult to specify and to estimate because they depend on investor preferences and attitudes to risk. One way to properly model the market price of risk would be to develop a dynamic general equilibrium model involving a representative investor. However we would obtain different expressions for the market price of risk depending on the utility function we use. This is a source of the non-uniqueness that we discuss later. For the moment we assume that the market prices of risk can somehow be estimated.

### *10.2.1 Tradable Asset*

Consider first of all the case of one *traded factor* whose stochastic dynamics are driven by one source of risk (uncertainty) represented by the Wiener process $z(t)$. If the factor's price $x$ (in return form) follows a diffusion process

$$\frac{dx}{x} = mdt + sdz,$$

then the market price of risk of the source of uncertainty is defined as

$$\lambda = \frac{m - r}{s},$$

which may be written as

$$m - r = \lambda s,$$

where

$$\begin{aligned} m &= \text{expected return from the factor,} \\ r &= \text{the instantaneous risk free rate,} \\ s &= \text{the volatility of the return of the factor,} \end{aligned}$$

If the traded asset earns a continuously compounded dividend at the rate $q$, then

$$m + q - r = \lambda s. \tag{10.1}$$

We interpret $s$ as the amount of risk associated with the uncertainty arising from the stochastic factor $z$ and $\lambda$ as the amount of compensation required by a rational investor for bearing an additional unit of risk associated with $z$. Then (10.1) can be interpreted as stating that the expected return on the risky factor equals the risk free rate $r$ plus a risk premium, $\lambda s$, as compensation for bearing the risk associated with the source of uncertainty $z$. The risk premium, $\lambda s$, consisting of the product of the cost of 1 unit of risk, $\lambda$, with the amount of risk, $s$.

Suppose the price of the factor $x$ is driven by several sources of uncertainty modelled by independent Wiener processes[1] $w_1, w_2, \ldots, w_M$, then there will be a market price of risk $\lambda_1, \lambda_2, \ldots, \lambda_M$ associated with each of these sources of uncertainty. In particular we assume $x$ follows the diffusion process

$$\frac{dx}{x} = mdt + \sum_{j=1}^{M} s_j dw_j. \tag{10.2}$$

Now we interpret $\lambda_j$ as the cost of 1 unit of risk associated with the source of uncertainty $w_j$ and $s_j$ as the corresponding amount of risk, so that $\lambda_j s_j$ is the risk premium required for bearing the risk associated with $w_j$. If the traded asset earns a continuously compounded dividend at the rate $q$, the expected excess return relationship now becomes

$$m + q - r = \sum_{j=1}^{M} \lambda_j s_j, \tag{10.3}$$

with the sum on the right-hand side representing the total risk premium required as compensation for bearing the risk associated with all of the sources of uncertainty $w_1, w_2, \ldots, w_M$. Equation (10.3) for the expected excess return on a risky factor has a similar appearance to the multifactor arbitrage pricing theory of Ross (1976), and is indeed closely related to it. It is important to stress that since $x$ is a traded factor then (10.3) is a relationship between returns and risk premia in a securities market which would determine the $\lambda_j$ as actual market prices of risk.

### *10.2.2 Non-tradable Asset*

The market price of risk of a *non-traded factor* is more subtle as by definition this factor is not traded and so there is no market in which a risk premium for holding it can be earned. More specifically, let $x$ be a factor which is the price[2] of a non-traded factor (e.g. an interest rate or volatility) and assume it follows the diffusion process

$$\frac{dx}{x} = mdt + sdz.$$

[1]Recall from Sect. 5.3 that we can always transform a set of correlated Wiener processes to a set of independent Wiener processes.

[2]The term price here is being used to in a broader sense to mean the level of $x$ in the units in which it is measured.

We formally define $\lambda$ as the risk premium for bearing the risk of this non-traded factor via

$$\lambda = \frac{m - r}{s},$$

or

$$m - r = \lambda s. \tag{10.4}$$

It is probably best to think of $\lambda s$ as a shadow risk premium for bearing the risk associated with the non-traded factor $x$, as there is no market in which this risk premium can be earned. Similarly, $m - r$ could be interpreted as a shadow excess return. Expressions for the shadow risk premium $\lambda s$ could be obtained in a dynamic general equilibrium model along the lines of Cox et al. (1985b).

In the case in which $x$ is driven by several sources of uncertainty (again modelled by the independent Wiener processes $z_j$), so that

$$\frac{dx}{x} = mdt + \sum_{j=1}^{M} s_j dw_j,$$

then (10.4) generalizes to

$$m - r = \sum_{j=1}^{M} \lambda_j s_j. \tag{10.5}$$

Now $\lambda_j$ is the market price of risk associated with the source of uncertainty $w_j$ and $\lambda_j s_j$ the corresponding shadow risk premium associated with this source of uncertainty. The summation on the right-hand side of Eq. (10.5) is the total shadow risk premium for bearing the risk associated with the non-traded factor $x$.

## 10.3 Pricing Derivative Securities Dependent on Two Factors

As we have stated above we will be interested in pricing derivative securities which are dependent on several underlying factors (e.g. bond options depending on a long term and a short term rate). In this section we shall go through the detail of the two factors case and then consider the general result in the following section.

Suppose the derivative security of interest depends on two underlying factors whose values[3] $x_1$, $x_2$ follow the diffusion processes

$$\begin{aligned}\frac{dx_1}{x_1} &= m_1 dt + s_1 dz_1,\\ \frac{dx_2}{x_2} &= m_2 dt + s_2 dz_2.\end{aligned} \tag{10.6}$$

We would like to allow for the possibility that the stochastic terms $dz_1$, $dz_2$ may be correlated. Thus $\mathbb{E}(dz_1) = \mathbb{E}(dz_2) = 0$, $\mathrm{var}(dz_1) = \mathrm{var}(dz_2) = dt$ and $\mathbb{E}(dz_1 dz_2) = \rho_{12} dt$. Later in our discussion we will need to apply Girsanov's Theorem in the presence of two Wiener processes. As this theorem is usually stated in terms of independent Wiener processes it will be more convenient to transform the diffusion processes for the prices $x_1, x_2$ to ones involving independent Wiener processes. We have already discussed the general procedure for transforming the Wiener processes driving a system of stochastic differential equations from dependent to independent ones. The general result required in the present context is outlined in Sect. 5.3.1.2. Using $w_1(t), w_2(t)$ to denote the independent Wiener processes, the diffusion processes for $x_1, x_2$ become

$$\begin{aligned}dx_1 &= m_1 x_1 dt + s_1 x_1 dw_1,\\ dx_2 &= m_2 x_2 dt + \rho_{12} s_2 x_2 dw_1 + \sqrt{1-\rho_{12}^2}\, s_2 x_2 dw_2.\end{aligned} \tag{10.7}$$

Since such a transformation from correlated to independent Wiener processes is always possible, we shall set up our general derivative pricing framework in terms of independent Wiener processes. The final pricing relationship can always be re-expressed in terms of the original variance-covariance structure by reversing the transformations. Thus, we assume that the factors $x_1, x_2$ are driven by the diffusion processes

$$\begin{aligned}\frac{dx_1}{x_1} &= m_1 dt + s_{11} dw_1 + s_{12} dw_2,\\ \frac{dx_2}{x_2} &= m_2 dt + s_{21} dw_1 + s_{22} dw_2,\end{aligned} \tag{10.8}$$

where $w_1(t), w_2(t)$ are independent Wiener processes and the coefficients $m_i, s_{ij}$ are possibly functions of $x_1, x_2$ and $t$. Whenever $x_i$ is interpreted as a traded asset, it is assumed to earn a continuously compounded dividend at the rate $q_i$.

[3]The term "value" should be broadly interpreted here. It could refer to prices in the case of traded factors, but could also to quantities such as inflation rates, level of volatility in the case of non-traded factors.

We denote by $\lambda_1, \lambda_2$, the market prices of risk associated with the risk sources $w_1, w_2$, respectively. Thus,

$$\left.\begin{array}{c}\text{the risk premium required}\\ \text{for bearing the risk}\\ \text{associated with factor } x_i\end{array}\right\} = \lambda_1 s_{i1} + \lambda_2 s_{i2}, \qquad (i = 1, 2).$$

Depending on whether $x_i$ is a traded or non-traded factor, the risk premium may be priced in a market or may be a shadow risk premium, in line with our earlier discussion. In particular, when $x_i$ is a traded factor and it earns a continuously compounded dividend at the rate $q_i$, then

$$m_i + q_i - r = \lambda_1 s_{i1} + \lambda_2 s_{i2}. \tag{10.9}$$

The price, $f$, of the derivative security of interest (e.g. an option) depends on the prices $x_1$, $x_2$ and time $t$, i.e.

$$f = f(x_1, x_2, t).$$

We know that by application of Ito's lemma (see Sect. 6.5.2) that the derivative security inherits the dynamics

$$\frac{df}{f} = \mu_f dt + \sigma_1 dw_1 + \sigma_2 dw_2, \tag{10.10}$$

where the expected return $\mu_f$ of the security is given by

$$f\mu_f = \theta + m_1 x_1 \Delta_1 + m_2 x_2 \Delta_2 + \mathfrak{D} f, \tag{10.11}$$

and the volatilities associated with each noise term are given by

$$\begin{aligned} f\sigma_1 &= s_{11} x_1 \Delta_1 + s_{21} x_2 \Delta_2, \\ f\sigma_2 &= s_{12} x_1 \Delta_1 + s_{22} x_2 \Delta_2. \end{aligned} \tag{10.12}$$

Here the quantities $\theta$, $\Delta_1$, $\Delta_2$, $\Gamma_{11}$, $\Gamma_{22}$ and $\Gamma_{12}$ are the various hedge ratios which measure the sensitivities of the price $f$ to changes in time and in prices of the underlying factors, i.e.

$$\begin{aligned} &\theta = \frac{\partial f}{\partial t}, \qquad \Delta_1 = \frac{\partial f}{\partial x_1}, \quad \Delta_2 = \frac{\partial f}{\partial x_2}, \\ &\Gamma_{11} = \frac{\partial \Delta_1}{\partial x_1} = \frac{\partial^2 f}{\partial x_1^2}, \quad \Gamma_{12} = \frac{\partial \Delta_1}{\partial x_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2} = \Gamma_{21}, \quad \Gamma_{22} = \frac{\partial \Delta_2}{\partial x_2} = \frac{\partial^2 f}{\partial x_2^2}, \\ &\mathfrak{D} f = \frac{1}{2}(s_{11}^2 + s_{12}^2) x_1^2 \Gamma_{11} + (s_{11} s_{21} + s_{12} s_{22}) x_1 x_2 \Gamma_{12} + \frac{1}{2}(s_{21}^2 + s_{22}^2) x_2^2 \Gamma_{22}. \end{aligned} \tag{10.13}$$

Equation (10.12) may be written in vector matrix notation as

$$\begin{bmatrix} f\sigma_1 \\ f\sigma_2 \end{bmatrix} = s^\top \begin{bmatrix} x_1\triangle_1 \\ x_2\triangle_2 \end{bmatrix}, \qquad s = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \tag{10.14}$$

### 10.3.1 Two Traded Assets

Consider the portfolio consisting of a position in one option, $Q_1$ of asset $x_1$ and $Q_2$ of asset $x_2$. The value of this portfolio is given by

$$V = Q_1 x_1 + Q_2 x_2 + f. \tag{10.15}$$

When calculating the change in the value of $V$ over $(t, t+dt)$ we need to keep in mind that the positions in $x_1$ and $x_2$ earn not only the capital gains $dx_1$ and $dx_2$ but also the dividends $q_1 x_1 dt$ and $q_2 x_2 dt$. Thus the instantaneous change in value of $V$ is given by

$$dV = Q_1(dx_1 + q_1 x_1 dt) + Q_2(dx_2 + q_2 x_2 dt) + df,$$

which by use of (10.10) becomes

$$\begin{aligned} dV = {} & [Q_1(m_1 + q_1)x_1 + Q_2(m_2 + q_2)x_2 + \mu_f f]dt \\ & + [Q_1 s_{11} x_1 + Q_2 s_{21} x_2 + \sigma_1 f]dw_1 \\ & + [Q_1 s_{12} x_1 + Q_2 s_{22} x_2 + \sigma_2 f]dw_2. \end{aligned} \tag{10.16}$$

We stress that we are using $\sigma_1, \sigma_2$ as defined by Eq. (10.12). We choose the proportions $Q_1$, $Q_2$ so as to eliminate the noise terms, i.e.

$$\begin{aligned} Q_1 s_{11} x_1 + Q_2 s_{21} x_2 + \sigma_1 f &= 0, \\ Q_1 s_{12} x_1 + Q_2 s_{22} x_2 + \sigma_2 f &= 0, \end{aligned} \tag{10.17}$$

which after use of the definitions (10.12) of $\sigma_1, \sigma_2$ result in (see Appendix 10.1)

$$Q_1 = -\Delta_1, \quad Q_2 = -\Delta_2. \tag{10.18}$$

This generalizes in a natural way the corresponding result in the single asset case in Chap. 6 namely that the investor takes positions in the underlying asset according to the delta with respect to the asset price. Also the investor takes a position in the underlying asset opposite to that of the option position. Substituting (10.18) into (10.15) and (10.16) we find that in the resulting riskless economy $V$ and $dV$ evolve according to

$$\begin{aligned} V &= -\Delta_1 x_1 - \Delta_2 x_2 + f, \\ dV &= [-\Delta_1(m_1 + q_1)x_1 - \Delta_2(m_2 + q_2)x_2 + \mu_f f]dt. \end{aligned} \tag{10.19}$$

Given that the change in portfolio value is now riskless, it follows that in order to avoid riskless arbitrage opportunities the portfolio must earn the riskless rate $r$, i.e. $V$ and $dV$ in Eq. (10.19) must satisfy

$$dV = rVdt,$$

which simplifies to

$$(\mu_f - r)f = \Delta_1(m_1 + q_1 - r)x_1 + \Delta_2(m_2 + q_2 - r)x_2. \tag{10.20}$$

Substituting Eq. (10.11) for $\mu_f$ and making use of the definitions in Eq. (10.13), Eq. (10.20) reduces to

$$\mathfrak{D}f + (r - q_1)x_1\frac{\partial f}{\partial x_1} + (r - q_2)x_2\frac{\partial f}{\partial x_2} + \frac{\partial f}{\partial t} - rf = 0. \tag{10.21}$$

To obtain a solution to (10.21) we would need to specify boundary conditions which would depend on the type of option we are trying to evaluate such as European, American etc.

We saw in Chap. 8 how it is possible to re-express the no riskless arbitrage condition (10.20) by use of martingale concepts and hence obtain the derivative security price as an expectation under a different probability measure. We now show how to extend the same argument to the case of the two traded assets $x_1$ and $x_2$.

First we need to consider what the excess return relationship (10.3) becomes for assets $x_1, x_2$ in the present context. Taking into account the dividends on each asset and the risk premium terms defined in (10.9) we have for each asset the expected excess return relationships

$$\begin{aligned} m_1 + q_1 - r &= \lambda_1 s_{11} + \lambda_2 s_{12}, \\ m_2 + q_2 - r &= \lambda_1 s_{21} + \lambda_2 s_{22}. \end{aligned} \tag{10.22}$$

We can use (10.22) to re-express the underlying asset price dynamics $x_1, x_2$ as

$$\begin{aligned} \frac{dx_1}{x_1} &= (r - q_1 + \lambda_1 s_{11} + \lambda_2 s_{12})dt + s_{11}dw_1 + s_{12}dw_2, \\ \frac{dx_2}{x_2} &= (r - q_2 + \lambda_1 s_{21} + \lambda_2 s_{22})dt + s_{21}dw_1 + s_{22}dw_2. \end{aligned} \tag{10.23}$$

Next we seek an expression for the derivative security price dynamics in the arbitrage free economy. To this end, we first substitute the expected excess return relationships (10.22) into the no-riskless arbitrage relationship (10.20) to obtain

$$\mu_f - r = (\lambda_1 s_{11} + \lambda_2 s_{12})\frac{\Delta_1 x_1}{f} + (\lambda_1 s_{21} + \lambda_2 s_{22})\frac{\Delta_2 x_2}{f}. \tag{10.24}$$

Using (10.12), (10.24) becomes

$$\mu_f - r = \lambda_1\sigma_1 + \lambda_2\sigma_2. \tag{10.25}$$

This equation may be interpreted as stating that the expected excess return on the derivative security is the sum of the risk premia ($\lambda_i\sigma_i$) associated with each source of uncertainty. Substituting (10.25) into Eq. (10.10) we find that the derivative security price dynamics follow the diffusion process

$$\frac{df}{f} = rdt + \sigma_1[dw_1 + \lambda_1 dt] + \sigma_2[dw_2 + \lambda_2 dt]. \tag{10.26}$$

Equation (10.26) suggests that we define adjusted Wiener processes

$$\tilde{w}_1(t) = w_1(t) + \int_0^t \lambda_1(u)du, \qquad \tilde{w}_2(t) = w_2(t) + \int_0^t \lambda_2(u)du,$$

in which case Eq. (10.26) becomes

$$\frac{df}{f} = rdt + \sigma_1 d\tilde{w}_1 + \sigma_2 d\tilde{w}_2. \tag{10.27}$$

After some algebraic manipulations we find that the asset price dynamics (10.23) can also be written in terms of the adjusted Wiener processes $\tilde{w}_1(t), \tilde{w}_2(t)$ as

$$\begin{aligned} \frac{dx_1}{x_1} &= (r - q_1)dt + s_{11}d\tilde{w}_1 + s_{12}d\tilde{w}_2, \\ \frac{dx_2}{x_2} &= (r - q_2)dt + s_{21}d\tilde{w}_1 + s_{22}d\tilde{w}_2. \end{aligned} \tag{10.28}$$

We note that $\tilde{w}_1(t), \tilde{w}_2(t)$ are not standard Wiener processes under the (so called historical) probability measure $\mathbb{P}$ as they have non-zero means in general. However, as we discussed in Chap. 8, we may apply Girsanov's theorem to obtain an equivalent probability measure $\tilde{\mathbb{P}}$ under which $\tilde{w}_1(t)$ and $\tilde{w}_2(t)$ become standard Wiener processes.

We note that $e^{rt}$ is the money market account at time $t$, i.e. it is the value at time $t$ of an account formed by investing a dollar at $t = 0$ and reinvesting continuously at the risk free rate $r$. The quantity $fe^{-rt}$ may be interpreted as the option value measured in units of the money market account. Applications of Ito's lemma and some simple manipulations of (10.27) reveal that

$$d(fe^{-rt}) = e^{-rt} f(\sigma_1 d\tilde{w}_1 + \sigma_2 d\tilde{w}_2). \tag{10.29}$$

Using $\tilde{\mathbb{E}}_t$ to denote expectation at time $t$ under the equivalent probability distribution $\tilde{\mathbb{P}}$, we have from (10.29) that

$$\tilde{\mathbb{E}}_t[d(fe^{-rt})] = 0. \tag{10.30}$$

This generalizes the result we obtained in Chap. 8, namely that the option price, measured in units of the money market account, is a martingale under the equivalent probability measure $\tilde{\mathbb{P}}$. From (10.30) we may derive the result

$$f_t = e^{-r(T-t)}\tilde{\mathbb{E}}_t[f_T]. \tag{10.31}$$

Application of the Feynman–Kac formula to their last expression would connect us back to the partial differential equation (10.21).

### *10.3.2 Two Traded Assets-Vector Notation*

In order to facilitate the discussion of the general case in Sect. 10.4 it is useful to reconsider the derivation of Eq. (10.20) in vector-matrix notation.

First we introduce the notation that if $v = v = (v_1, v_2)^\top$ is a column vector then $V = diag v$ indicates the matrix[4]

$$V = diag(v) = \begin{bmatrix} v_1 & 0 \\ 0 & v_2 \end{bmatrix}.$$

We also introduce the unit vector $\mathbf{1} = (1, 1)^\top$. Furthermore we write

$$x = (x_1, x_2)^\top, \quad f = (f_1, f_2)^\top, \quad m = (m_1, m_2)^\top, \quad q = (q_1, q_2)^\top,$$
$$\sigma = (\sigma_1, \sigma_2)^\top, \quad dw = (dw_1, dw_2)^\top,$$
$$s = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}, \quad Q = (Q_1, Q_2)^\top, \quad \Delta = (\Delta_1, \Delta_2)^\top.$$

With these notations the stochastic differential equation system (10.8) can be written

$$dx = X \cdot (m \cdot dt + s \cdot dw), \tag{10.32}$$

[4]It is useful to note the result that for any two vectors of the same length $v$ and $x$, there holds $diag(v) \cdot x = diag(x) \cdot v$.

whilst Eq. (10.10) can be written

$$df = f(\mu_f dt + \sigma^\top \cdot dw). \tag{10.33}$$

The hedging portfolio can be written as

$$V = Q^\top x + f, \tag{10.34}$$

and the change in $V$ is given by

$$dV = Q^\top (dx + Xq dt) + df. \tag{10.35}$$

Using (10.32), Eq. (10.35) can be written

$$\begin{aligned} dV &= Q^\top X \cdot (m+q) \cdot dt + Q^\top X \cdot s \cdot dw + f\mu_f dt + f\sigma^\top \cdot dw \\ &= [Q^\top X \cdot (m+q) + f\mu_f]dt + [Q^\top X \cdot s + f\sigma^\top]dw. \end{aligned} \tag{10.36}$$

We note that Eq. (10.12) may be re-expressed as[5]

$$f\sigma^\top = \triangle^\top \cdot X \cdot s, \tag{10.37}$$

so that (10.36) becomes

$$dV = [Q^\top X \cdot (m+q) + f\mu_f]dt + (Q+\triangle)^\top X \cdot s \cdot dw. \tag{10.38}$$

Clearly the $dw$ term is eliminated by setting

$$Q = -\triangle, \tag{10.39}$$

or (as in Sect. 10.3) $Q_1 = -\triangle_1$ and $Q_2 = -\triangle_2$. Thus now

$$dV = [-\triangle^\top X \cdot (m+q) + f\mu_f]dt. \tag{10.40}$$

But since this change in portfolio value is riskless it must also be the case that

$$dV = r \cdot V \cdot dt, \tag{10.41}$$

where now, by (10.34) and (10.39),

$$V = -\triangle^\top x + f = -\triangle^\top X\mathbf{1} + f.$$

[5]Take care to note that the matrix that appears on the right hand side of (10.12) is $s^T$.

Equating (10.40) and (10.41) and noting that $rx = r\,diag(x)\mathbf{1}$, we obtain

$$(\mu_f - r)f = \triangle^\top X \cdot (m + q - r\mathbf{1}). \tag{10.42}$$

When written componentwise Eq. (10.42) becomes Eq. (10.20). The expected excess return relationships (10.22) can be written

$$(m + q - r\mathbf{1}) = s\lambda, \tag{10.43}$$

and hence substituting (10.43) into (10.42) the excess return relation for the derivative security becomes

$$(\mu_f - r)f = \triangle^\top \cdot X \cdot s\lambda. \tag{10.44}$$

We may use (10.37) to eliminate $\triangle^\top \cdot X \cdot s$ in Eq. (10.44) to obtain (recall that $f$ is a scalar which cancels on both sides),

$$\mu_f - r = \sigma^\top \lambda, \tag{10.45}$$

which coordinate-wise becomes Eq. (10.25).

### *10.3.3 One Traded Asset and One Non-traded Asset*

In this subsection we consider the case where the first asset, $x_1$, is traded (e.g. common stock, foreign currency) and the second asset, $x_2$, is not traded (e.g. stochastic volatility, expected inflation, interest rate).

Now we have two risks ($x_1$ and $x_2$) but only one traded asset ($x_1$) to combine with the option in a hedging portfolio.[6] In order to hedge away the two risks we face, we need another traded asset to place in our hedging portfolio. One way to create an additional traded asset is to have two options (both written on $x_1$) of different maturity in the hedging portfolio. We use $f^{(1)}(x_1, x_2, t)$, $f^{(2)}(x_1, x_2, t)$ to denote the values at time $t$ of options (written on $x_1$) of maturity $T_1$, $T_2$ respectively.

We need to generalize the notation used in Eqs. (10.10)–(10.13) to express the dynamics of the option price. Application of Ito's Lemma to $f^{(i)}(x_1, x_2, t)$ ($i = 1, 2$) yields, for each option, the dynamics

$$\frac{df^{(i)}}{f^{(i)}} = \mu_f^{(i)} dt + \sigma_1^{(i)} dw_1 + \sigma_2^{(i)} dw_2, \tag{10.46}$$

[6] Note that the following argument needs to be modified in the special case $s_{12} = s_{22} = 0$. That is, there is only one noise term. This special case is treated in the next section. Such a case arises for instance in the stochastic volatility model of Hobson and Rogers (1998).

where

$$f^{(i)}\mu_f^{(i)} = \theta^{(i)} + m_1 x_1 \Delta_1^{(i)} + m_2 x_2 \Delta_2^{(i)} + \mathfrak{D} f^{(i)},$$
$$f^{(i)}\sigma_1^{(i)} = s_{11} x_1 \Delta_1^{(i)} + s_{21} x_2 \Delta_2^{(i)},$$
$$f^{(i)}\sigma_2^{(i)} = s_{12} x_1 \Delta_1^{(i)} + s_{22} x_2 \Delta_2^{(i)},$$

and

$$\theta^{(i)} = \frac{\partial f^{(i)}}{\partial t}, \qquad \Delta_1^{(i)} = \frac{\partial f^{(i)}}{\partial x_1}, \quad \Delta_2^{(i)} = \frac{\partial f^{(i)}}{\partial x_2},$$
$$\Gamma_{11}^{(i)} = \frac{\partial^2 f^{(i)}}{\partial x_1^2}, \quad \Gamma_{12}^{(i)} = \frac{\partial^2 f^{(i)}}{\partial x_1 \partial x_2} = \Gamma_{21}^{(i)}, \quad \Gamma_{22}^{(i)} = \frac{\partial^2 f^{(i)}}{\partial x_2^2},$$
$$\mathfrak{D} f^{(i)} = \frac{1}{2}(s_{11}^2 + s_{12}^2) x_1^2 \Gamma_{11}^{(i)} + (s_{11}s_{21} + s_{12}s_{22}) x_1 x_2 \Gamma_{12}^{(i)} + \frac{1}{2}(s_{21}^2 + s_{22}^2) x_2^2 \Gamma_{22}^{(i)}.$$

We apply the hedging argument using the vector notation already introduced in Sect. 10.3.2. It is now that we appreciate the advantages of the vector notation. The alternative (and more cumbersome) approach without the use of vector notation is laid out in Appendix 10.2. In vector notation Eq. (10.46) may be written

$$df = F \cdot (\mu_f \cdot dt + \sigma \cdot dw), \tag{10.47}$$

where

$$f = (f^{(1)}, f^{(2)})^\top, \quad \mu_f = (\mu_f^{(1)}, \mu_f^{(2)})^\top, \quad F = diag(f)$$

and $\sigma$ is the matrix

$$\sigma = \begin{pmatrix} \sigma_1^{(1)} & \sigma_2^{(1)} \\ \sigma_1^{(2)} & \sigma_2^{(2)} \end{pmatrix}.$$

We consider a portfolio consisting of 1 unit of traded asset $x_1$, $Q_1$ units of $f^{(1)}$ and $Q_2$ units of $f^{(2)}$. The value, $V$, of this portfolio, in vector notation, at time $t$ is given by

$$V = x_1 + Q^\top f,$$

whilst the change in $V$ is given by

$$dV = (dx_1 + q_1 x_1 dt) + Q^\top df. \tag{10.48}$$

Using the first of Eqs. (10.8) and (10.47) we may express (10.48) as

$$\begin{aligned} dV &= (m_1 + q_1)x_1 dt + x_1(s_{11}, s_{12})dw \\ &\quad + Q^\top \cdot F \cdot \mu_f \cdot dt + Q^\top \cdot F \cdot \sigma \cdot dw \\ &= [(m_1 + q_1)x_1 + Q^\top \cdot F \cdot \mu_f]dt \\ &\quad + [x_1(s_{11}, s_{12}) + Q^\top \cdot F \cdot \sigma]dw. \end{aligned} \tag{10.49}$$

The stochastic $dw$ term is eliminated by choosing $Q$ so that

$$Q^\top F\sigma = -x_1(s_{11}, s_{12}). \tag{10.50}$$

The dynamics of the now riskless portfolio is then given by

$$dV = [(m_1 + q_1)x_1 + Q^\top \cdot F \cdot \mu_f]dt. \tag{10.51}$$

The condition, $dV = rV\,dt$, that the portfolio can only earn the riskless rate becomes

$$(m_1 + q_1)x_1 + Q^\top F\mu_f = r(x_1 + Q^\top f). \tag{10.52}$$

Noting the result in footnote 4, Eq. (10.52) can be expressed in the form

$$(m_1 + q_1 - r)x_1 + Q^\top F(\mu_f - r\mathbf{1}) = 0. \tag{10.53}$$

At this point we recall the expected excess return relationship for the asset $x_1$ [see the first of Eq. (10.22)],

$$m_1 + q_1 - r = s_{11}\lambda_1 + s_{12}\lambda_2 = (s_{11}, s_{12})\lambda, \tag{10.54}$$

where

$$\lambda = (\lambda_1, \lambda_2)^\top.$$

Thus (10.53) may be written

$$x_1[s_{11}, s_{12}]\lambda + Q^\top F \cdot (\mu_f - r\,\mathbf{1}) = 0. \tag{10.55}$$

Using (10.50) to eliminate the first term in (10.55) we obtain

$$-Q^\top F\sigma\lambda + Q^\top F \cdot (\mu_f - r\,\mathbf{1}) = 0,$$

which may be re-expressed as

$$Q^\top F[-\sigma\lambda + \mu_f - r\mathbf{1}] = 0. \tag{10.56}$$

From Eq. (10.50) we have that

$$Q^\top F = -x_1(s_{11}, s_{12})\sigma^{-1},$$

which together with (10.56) implies that

$$x_1(s_{11}, s_{12})\sigma^{-1}[\sigma\lambda - \mu_f + r\mathbf{1}] = 0. \tag{10.57}$$

Equation (10.57) can hold for options of any arbitrary maturity only if the term in square brackets is equal to zero. Equation (10.57) hence implies that

$$\mu_f - r\mathbf{1} = \sigma\lambda.$$

Coordinate-wise this last equation states that

$$\mu_{f^{(1)}} - r = \sigma_1^{(1)}\lambda_1 + \sigma_2^{(1)}\lambda_2,$$
$$\mu_{f^{(2)}} - r = \sigma_1^{(2)}\lambda_1 + \sigma_2^{(2)}\lambda_2.$$

Since the maturities $T_1$ and $T_2$ were arbitrarily chosen and since each of the relationships in (10.58) is maturity specific, a similar relationship holds for an option of any maturity. Thus we can assert that for an option of any maturity there holds

$$\mu_f - r = \sigma_1\lambda_1 + \sigma_2\lambda_2. \tag{10.58}$$

Equation (10.58) merely states that for an option of arbitrary maturity the expected excess return equals the sum of the risk premia ($\lambda_i\sigma_i, i = 1, 2$) associated with each source of uncertainty.

Upon employing the expression for $\mu_f$, $\sigma_1$, $\sigma_2$ and re-arranging, Eq. (10.58) becomes

$$\theta + (m_1 - s_{11}\lambda_1 - s_{12}\lambda_2)x_1\Delta_1 + (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})x_2\Delta_2 + \mathfrak{D}f - rf = 0,$$

which upon use of (10.54) becomes

$$\theta + (r - q_1)x_1\Delta_1 + (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})x_2\Delta_2 + \mathfrak{D}f - rf = 0.$$

Replacing $\theta$, $\Delta_1$ and $\Delta_2$ by their partial derivative expressions we see that we have the pricing partial differential equation

$$\frac{\partial f}{\partial t} + (r - q_1)x_1\frac{\partial f}{\partial x_1} + (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})x_2\frac{\partial f}{\partial x_2} + \mathfrak{D}f - rf = 0. \tag{10.59}$$

In order to obtain the martingale representation we use (10.58) to replace $\mu_f$ in the stochastic differential equation for $f$ to obtain

$$\frac{df}{f} = rdt + \sigma_1(dw_1 + \lambda_1 dt) + \sigma_2(dw_2 + \lambda_2 dt),$$

which in terms of the adjusted Wiener processes

$$\begin{aligned} \tilde{w}_1(t) &= w_1(t) + \int_o^t \lambda_1(\tau)d\tau, \\ \tilde{w}_2(t) &= w_2(t) + \int_o^t \lambda_2(\tau)d\tau, \end{aligned} \tag{10.60}$$

may be written

$$\frac{df}{f} = rdt + \sigma_1 d\tilde{w}_1 + \sigma_2 d\tilde{w}_2. \tag{10.61}$$

The argument involved in the last subsection to obtain the martingale representation of the option price may also be applied to (10.61), so we obtain

$$f_t = e^{-r(T-t)}\tilde{\mathbb{E}}_t(f_T), \tag{10.62}$$

where $\tilde{\mathbb{E}}_t$ is the expectation under the equivalent measure $\tilde{\mathbb{P}}$.

In order to obtain the dynamics of $x_1, x_2$ under the equivalent measure we use (10.60) to replace the $dw_1$ and $dw_2$ in (10.8) with $d\tilde{w}_1 - \lambda_1 dt$ and $d\tilde{w}_2 - \lambda_2 dt$ respectively to obtain

$$\begin{aligned} \frac{dx_1}{x_1} &= (m_1 - \lambda_1 s_{11} - \lambda_2 s_{12})dt + s_{11}d\tilde{w}_1 + s_{12}d\tilde{w}_2, \\ \frac{dx_2}{x_2} &= (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})dt + s_{21}d\tilde{w}_1 + s_{22}d\tilde{w}_2. \end{aligned}$$

By use of the expected excess return relationship for $x_1$, we replace $m_1 - \lambda_1 s_{11} - \lambda_2 s_{12}$ by $r - q_1$ to finally obtain for the dynamics of the underlying factors under the equivalent measure

$$\frac{dx_1}{x_1} = (r - q_1)dt + s_{11}d\tilde{w}_1 + s_{12}d\tilde{w}_2, \tag{10.63}$$

$$\frac{dx_2}{x_2} = (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})dt + s_{21}d\tilde{w}_1 + s_{22}d\tilde{w}_2. \tag{10.64}$$

We highlight the presence of the risk premium term $\lambda_1 s_{21} + \lambda_2 s_{22}$, in the process (10.64) for the non-traded underlying factor $x_2$. Under the historical measure $\mathbb{P}$ the adjusted Wiener processes $\tilde{w}_1$, $\tilde{w}_2$ are not standard Wiener processes. Again we may apply Girsanov's theorem to obtain an equivalent probability measure $\tilde{\mathbb{P}}$ under which $\tilde{w}_1$, $\tilde{w}_2$ become standard Wiener processes. However in this non-traded asset case, which is an example of an incomplete market, the measure $\tilde{\mathbb{P}}$ is not unique. This reflects the fact that we may specify a number of ways in which the market may settle on a value for the risk premium, e.g. by appealing to utility maximization arguments, by minimization of variance of trading costs, etc. Some authors argue that the risk involved in $x_2$ would be diversified away in an efficient market so we may set $\lambda_1 s_{21} + \lambda_2 s_{22} = 0$. Discussion of these issues is beyond the scope of this book. For the moment, we assume the risk premium $\lambda_1 s_{21} + \lambda_2 s_{22}$ is somehow determined empirically and we use this value to choose a particular $\tilde{\mathbb{P}}$. We note that $\tilde{\mathbb{E}}_t$ could be calculated by simulating (10.63) and (10.64) over $t$ to $T$. Again application of the Feynman–Kac formula would yield the partial differential equation (10.59).

## 10.4 The General Case

Consider the general situation in which there are $n_t$ traded factors and $n_n$ non-traded factors with $n_t + n_n = n$ the total number of factors. In this discussion we shall assume that each factor has its own driving source of uncertainty, $z_i(t)$, $(i = 1, 2, \ldots, n)$. In general the sources of uncertainty may be correlated. Thus the diffusion processes for the factors may be written

$$\frac{dx_i}{x_i} = m_i dt + s_i dz_i(t), \quad (i = 1, 2, \ldots, n). \tag{10.65}$$

We assume that

$$\mathbb{E}[dz_i(t)dz_j(t)] = \rho_{ij} dt. \tag{10.66}$$

We shall adopt the convention that the first $n_t$ factors are the traded factors. For the reasons stated in Sect. 10.3 we prefer to express the factor dynamics in terms of independent Wiener processes. Using the transformation procedure detailed in Sect. 5.3 we can re-express the factor dynamics in terms of independent Wiener processes $w_i(t)$, $i = 1, 2, \ldots, n$ as

$$\frac{dx_i}{x_i} = m_i dt + \sum_{j=1}^{n} s_{ij} dw_j(t). \tag{10.67}$$

The relationship between the $s_{ij}$ of (10.67) and the $s_i$ and $\rho_{ij}$ of Eqs. (10.65) and (10.66) can be determined in the manner outlined in Sect. 5.3. If $\lambda_i$ denotes the market price of risk associated with the risk source $w_i(t)$ then

$$\left.\begin{array}{c}\text{the risk premium required for bearing}\\ \text{the risk associated with factor } x_i\end{array}\right\} = \sum_{j=1}^{n} \lambda_j s_{ij}. \tag{10.68}$$

As we shall consider a portfolio containing the $n_t$ traded assets we need a convenient notation to keep track of the dynamics of just this subset of the underlying factors. We denote by $x_{tr}$ the column vector of traded factors $(x_1, x_2, \dots, x_{n_t})^\top$, $m_{tr}$ the column vector $(m_1, m_2, \dots, m_{n_t})^\top$, $s_{tr}$ the $(n_t \times n)$ matrix with elements $s_{ij}(i = 1, \dots, n_t; j = 1, \dots, n)$. With this notation we can write the dynamics of $x_{tr}$ as

$$dx_{tr} = X_{tr} \cdot (m_{tr} \cdot dt + s_{tr} \cdot dw), \qquad X_{tr} = diag(x_{tr}). \tag{10.69}$$

In order to hedge away the $n_n$ non-traded risks we need to introduce $(n - n_t + 1)$ traded options of maturities $T_1, T_2, \dots, T_{n-n_t+1}$. We use $f^{(l)}(x_1, \dots, x_n, t)(l = 1, \dots, n - n_t + 1)$ to denote these options, which are assumed to depend (possibly) on all the state variables. By application of Ito's Lemma the dynamics of each option are given by

$$\frac{df^{(l)}}{f^{(l)}} = \mu_f^{(l)} dt + \sum_{j=1}^{n} \sigma_j^{(l)} dw_j, \tag{10.70}$$

where

$$f^{(l)}\mu_f^{(l)} = \theta^{(l)} + \sum_{j=1}^{n} m_j x_j \Delta_j^{(l)} + \mathscr{D} f^{(l)}, \qquad f^{(l)}\sigma_j^{(l)} = \sum_{k=1}^{n} s_{kj} x_k \Delta_k^{(l)},$$

and

$$\theta^{(l)} = \frac{\partial f^{(l)}}{\partial t}, \quad \Delta_j^{(l)} = \frac{\partial f^{(l)}}{\partial x_j}, \quad \Gamma_{ij}^{(l)} = \frac{\partial^2 f^{(l)}}{\partial x_i \cdot \partial x_j},$$
$$\mathscr{D} f^{(l)} = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \sum_{k=1}^{n} s_{ik} s_{jk} \right) x_i x_j \Gamma_{ij}^{(l)}.$$

We note that in vector notation (10.70) may be written

$$df = F \cdot (\mu_f \cdot dt + \sigma \cdot dw), \tag{10.71}$$

where

$$f = (f^{(1)}, f^{(2)}, \dots, f^{(n-n_t+1)})^\top, \quad F = diag(f),$$
$$\mu_f = (\mu_{f^{(1)}}, \mu_{f^{(2)}}, \dots, \mu_{f^{(n-n_t+1)}})^\top$$

and $\sigma$ is the $(n - n_t + 1) \times n$ matrix whose $(l, j)$th element is $\sigma_j^{(l)}(j = 1, \dots, n; l = 1, \dots, n - n_t + 1)$.

We now form a portfolio consisting of $Q_i$ units of the traded underlying asset $x_i$ for $i = 1, \dots, n_t$ and $Q_{n_t+l}$ units of the traded options $f^{(l)}$ for $l = 1, \dots, n - n_t + 1$.

The value of this portfolio at time $t$ is given by

$$V = \sum_{i=1}^{n_t} Q_i x_i + \sum_{l=1}^{n-n_t+1} Q_{n_t+l} f^{(l)}. \tag{10.72}$$

It is convenient to partition $Q$ according to

$$Q_{tr} = (Q_1, \ldots, Q_{n_t})^\top, \qquad Q_f = (Q_{n_t+1}, \ldots, Q_{n+1})^\top$$

so that

$$V = Q_{tr}^\top x_{tr} + Q_f^\top f. \tag{10.73}$$

Using (10.67), incorporating the continuously compounded dividend and (10.69) we see that the instantaneous change in $V$ may be written

$$\begin{aligned} dV = &\; Q_{tr}^\top dx_{tr} + Q_f^\top df \\ = &\; [Q_{tr}^\top X_{tr}(m_{tr} + q_{tr}) + Q_f^\top F\mu_f]dt \\ &+ [Q_{tr}^\top X_{tr} \cdot s_{tr} + Q_f^\top \cdot F \cdot \sigma]dw. \end{aligned}$$

In order to render the portfolio riskless the vector $Q$ has to be chosen so that

$$Q_{tr}^\top X_{tr} s_{tr} + Q_f^\top F\sigma = 0. \tag{10.74}$$

The hedging portfolio is now riskless and evolves according to

$$dV = [Q_{tr}^\top X_{tr}(m_{tr} + q_{tr}) + Q_f^\top F\mu_f]dt. \tag{10.75}$$

Following the now standard argument, the riskless hedging portfolio can only earn the risk-free interest rate, a condition which by use of (10.73) becomes

$$dV = r[Q_{tr}^\top x_{tr} + Q_f^\top f]dt. \tag{10.76}$$

Equations (10.75) and (10.76) imply that the vector $Q$ also has to be chosen so that[7]

$$Q_{tr}^\top X_{tr}(m_{tr} + q_{tr}) + Q_f^\top F\mu_f = rQ_{tr}^\top x_{tr} + rQ_f^\top f. \tag{10.77}$$

The last equation can be expressed as

$$Q_{tr}^\top X_{tr}(m_{tr} + q_{tr} - r\mathbf{1}) + Q_f^\top F(\mu_f - r\mathbf{1}) = 0. \tag{10.78}$$

---

[7]We again make use of the result in footnote 4.

At this point it is convenient to set $Q^\top := (Q_{tr}^\top, Q_f^\top)$, so that (10.78) can be rewritten as

$$Q^\top \begin{pmatrix} X_{tr}(m_{tr} + q_{tr} - r\mathbf{1}) \\ F(\mu_f - r\mathbf{1}) \end{pmatrix} = 0. \tag{10.79}$$

We note that in terms of the vector $Q$ we can rewrite Eq. (10.74) as

$$Q^\top \begin{pmatrix} [X_{tr}s_{tr}]_{n_t \times n} \\ [F\sigma]_{(n_n+1)\times n} \end{pmatrix}_{(n+1)\times n} = 0. \tag{10.80}$$

Equations (10.79) and (10.80) constitute an $(n + 1)$-dimensional linear homogeneous equation system

$$Q^\top \begin{pmatrix} X_{tr} & 0 \\ 0 & F \end{pmatrix}_{(n+1)\times(n+1)} \begin{pmatrix} (m_{tr} + q_{tr} - r\mathbf{1}) & s_{tr} \\ (\mu_f - r\mathbf{1}) & \sigma \end{pmatrix}_{(n+1)\times(n+1)} = 0.$$

Recall that $Q$(of size $(n + 1) \times 1$) represent the hedging portfolio, so the linear system must allow non-trivial solutions for $Q$. The condition for this is the rank degeneracy of the matrix

$$\text{Rank} \quad \begin{pmatrix} (m_{tr} + q_{tr} - r\mathbf{1}) & s_{tr} \\ (\mu_f - r\mathbf{1}) & \sigma \end{pmatrix} \leq n. \tag{10.81}$$

The right hand part of the matrix, consisting of $s_{tr}$, $\sigma$ comes from the volatility structure of the factors. Usually we do not impose restrictions on the volatility structure so this part should have rank $n$. Therefore, the no-arbitrage condition (10.81) requires that the vectors of the excess returns, $(m_{tr} + q_{tr} - r\mathbf{1})$, $(\mu_f - r\mathbf{1})$ must be represented as a linear combination of the vectors of the volatility, which is given by

$$\begin{pmatrix} (m_{tr} + q_{tr} - r\mathbf{1}) \\ (\mu_f - r\mathbf{1}) \end{pmatrix} = \begin{pmatrix} s_{tr} \\ \sigma \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{10.82}$$

Thus, as usual, the no-arbitrage condition turns out to be a restriction between the excess returns and the asset volatility structures as given in (10.82). It follows from (10.82) that

$$m_{tr} + q_{tr} - r\mathbf{1} = s_{tr}\lambda. \tag{10.83}$$

and

$$\mu_f - r\mathbf{1} = \sigma\lambda. \tag{10.84}$$

Equation (10.83) is simply the expected excess return relation for the underlying assets, which componentwise reads

$$m_i + q_i - r = \sum_{j=1}^{n} s_{ij}\lambda_j, \quad (i = 1, 2, \ldots, n_t).$$

Coordinate-wise Eq. (10.84) states that for the option of maturity $T_l$ the no riskless arbitrage condition is

$$\mu_{f^{(l)}} - r = \sum_{j=1}^{n} \sigma_j^{(l)} \lambda_j. \tag{10.85}$$

Since the maturity $T_l$ is arbitrary the relation (10.85) must hold for an option of any maturity, that is

$$\mu_f - r = \sum_{j=1}^{n} \sigma_j \lambda_j. \tag{10.86}$$

Upon making use of the expressions for $\mu_f$ and $\sigma_j$ defined below Eq. (10.70), the condition (10.86) reduces to the parabolic partial differential equation

$$\theta + \sum_{j=1}^{n} m_j x_j \Delta_j + \mathscr{D} f - rf = \sum_{j=1}^{n} \lambda_j (\sum_{k=1}^{n} s_{kj} x_k \Delta_k). \tag{10.87}$$

Consider separately the term

$$\begin{aligned} Z &\equiv \sum_{j=1}^{n} m_j x_j \Delta_j - \sum_{j=1}^{n} \lambda_j (\sum_{k=1}^{n} s_{kj} x_k \Delta_k) \\ &= \sum_{j=1}^{n} m_j x_j \Delta_j - \sum_{k=1}^{n} x_k \Delta_k (\sum_{j=1}^{n} \lambda_j s_{kj}) \end{aligned} \tag{10.88}$$

From (10.83) we have that for $k = 1, \ldots, n_t$

$$\sum_{j=1}^{n} \lambda_j s_{kj} = m_k + q_k - r,$$

hence (10.88) can be re-written

$$Z = \sum_{j=1}^{n_t} m_j x_j \Delta_j + \sum_{j=n_t+1}^{n} m_j x_j \Delta_j - \sum_{k=1}^{n_t} x_k \Delta_k (m_k + q_k - r)$$

$$- \sum_{k=n_{t+1}}^{n} x_k \Delta_k (\sum_{j=1}^{n} \lambda_j s_{kj})$$

$$= \sum_{k=1}^{n_t} x_k \Delta_k (r - q_k) + \sum_{k=n_t+1}^{n} x_k \Delta_k \left( m_k - \sum_{j=1}^{n} \lambda_j s_{kj} \right). \tag{10.89}$$

Using (10.89) in (10.87), the pricing partial differential equation becomes

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{n_t} (r - q_i) x_i \frac{\partial f}{\partial x_i} + \sum_{i=n_t+1}^{n} (m_i - \sum_{j=1}^{n} \lambda_j s_{ij}) x_i \frac{\partial f}{\partial x_i} + \mathfrak{D} f - r f = 0, \tag{10.90}$$

where

$$\mathfrak{D} f \equiv \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\sum_{k=1}^{n} s_{ik} s_{jk}) x_i x_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Using (10.86) to eliminate $\mu_f$ the dynamics for $f$ turn out to be

$$\frac{df}{f} = (r + \sum_{j=1}^{n} \sigma_j \lambda_j) dt + \sum_{j=1}^{n} \sigma_j dw_j.$$

Introducing the modified Wiener processes

$$\tilde{w}_j(t) = w_j(t) + \int_0^t \lambda_j(\tau) d\tau, \quad (j = 1, \ldots, n), \tag{10.91}$$

the dynamics for $f$ can be written

$$\frac{df}{f} = r dt + \sum_{j=1}^{n} \sigma_j d\tilde{w}_j.$$

Under the original historical measure $\mathbb{P}$ the $\tilde{w}$ are not standard Wiener processes, but by Girsanov's theorem we can find an equivalent measure $\tilde{\mathbb{P}}$ such that the $\tilde{w}$ are standard Wiener processes. So we obtain the generalization of (10.62)

$$f_t = e^{-r(T-t)} \tilde{\mathbb{E}}_t(f_T). \tag{10.92}$$

Using (10.91) to express the factor dynamics in terms of the $d\tilde{w}_i(t)$ we obtain

$$\frac{dx_i}{x_i} = (m_i - \sum_{j=1}^{n} \lambda_j s_{ij}) dt + \sum_{j=1}^{n} s_{ij} d\tilde{w}_j(t). \tag{10.93}$$

For the $n_t$ traded factors we may use the expected excess return relation

$$m_i + q_i - r = \sum_{j=1}^{n} \lambda_j s_{ij},$$

to eliminate the market price of risk factors from the dynamic for these factors. So the factor dynamics may finally be expressed in the form

$$\frac{dx_i}{x_i} = (r - q_i)dt + \sum_{j=1}^{n} s_{ij} d\tilde{w}_j, \tag{10.94}$$

for $(i = 1, 2, \ldots, n_t)$ and

$$\frac{dx_i}{x_i} = (m_i - \sum_{j=1}^{n} \lambda_j s_{ij})dt + \sum_{j=1}^{n} s_{ij} d\tilde{w}_j,$$

for $(i = n_t + 1, n_t + 2, \ldots, n)$. Simulating this stochastic dynamic system would be one way that the expectation operator $\tilde{\mathbb{E}}_t$ in (10.92) could be calculated.

The comments at the end of Sect. 10.3 concerning the non-uniqueness of the measure $\tilde{\mathbb{P}}$ apply equally. The non-uniqueness is here expressed through the need to choose the market prices of risk $\lambda_1, \lambda_2, \ldots, \lambda_n$. Again we assume here that these have been chosen by one of the procedures discussed at the end of Sect. 10.3.

## 10.5 Appendix

### Appendix 10.1 Derivation of $Q_1 = -\Delta_1$ and $Q_2 = -\Delta_2$

Consider Eq. (10.17)

$$Q_1 s_{11} x_1 + Q_2 s_{21} x_2 + \sigma_1 f = 0,$$
$$Q_1 s_{12} x_1 + Q_2 s_{22} x_2 + \sigma_2 f = 0,$$

and substitute the definitions (10.12) of $\sigma_1, \sigma_2$. Then we obtain the equations

$$s_{11} x_1 (Q_1 + \Delta_1) + s_{21} x_2 (Q_2 + \Delta_2) = 0,$$
$$s_{12} x_1 (Q_1 + \Delta_1) + s_{22} x_2 (Q_2 + \Delta_2) = 0,$$

or

$$\begin{bmatrix} s_{11} x_1 & s_{21} x_2 \\ s_{12} x_1 & s_{22} x_2 \end{bmatrix} \begin{bmatrix} Q_1 + \Delta_1 \\ Q_2 + \Delta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We note that the determinant

$$\begin{vmatrix} s_{11}x_1 & s_{21}x_2 \\ s_{12}x_1 & s_{22}x_2 \end{vmatrix} = (s_{11}s_{22} - s_{12}s_{21})x_1x_2 = \mathscr{G}x_1x_2 \neq 0$$

given our assumption that $\mathscr{G} = (s_{11}s_{22} - s_{12}s_{21}) \neq 0$. Therefore,

$$Q_1 + \Delta_1 = 0,$$
$$Q_2 + \Delta_2 = 0,$$

i.e.

$$Q_1 = -\Delta_1,$$
$$Q_2 = -\Delta_2.$$

## Appendix 10.2 Alterative Derivation for One Traded and One Non-traded Asset

We consider a portfolio consisting of 1 unit of traded asset $x_1$, $Q_1$ units of $f^{(1)}$ and $Q_2$ units of $f^{(2)}$. The value of this portfolio at time $t$ is given by

$$V = x_1 + Q_1 f^{(1)} + Q_2 f^{(2)}. \tag{10.95}$$

The instantaneous change in the value of $V$ is given by

$$\begin{aligned} dV =& [(m_1 + q_1)x_1 + Q_1 f^{(1)}\mu_{f^{(1)}} + Q_2 f^{(2)}\mu_{f^{(2)}}]dt \\ &+ [s_{11}x_1 + Q_1 f^{(1)}\sigma_1^{(1)} + Q_2 f^{(2)}\sigma_1^{(2)}]dw_1 \\ &+ [s_{12}x_1 + Q_1 f^{(1)}\sigma_2^{(1)} + Q_2 f^{(2)}\sigma_2^{(2)}]dw_2. \end{aligned} \tag{10.96}$$

We choose the portfolio weights $Q_1$, $Q_2$ of the two options so as to eliminate the noise terms in (10.96) i.e. we solve simultaneously the two equations

$$\begin{aligned} Q_1 f^{(1)}\sigma_1^{(1)} + Q_2 f^{(2)}\sigma_1^{(2)} &= -s_{11}x_1, \\ Q_1 f^{(1)}\sigma_2^{(1)} + Q_2 f^{(2)}\sigma_2^{(2)} &= -s_{12}x_1, \end{aligned} \tag{10.97}$$

to obtain

$$\begin{aligned} Q_1 &= -\frac{x_1}{\mathscr{S} f^{(1)}}[s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}], \\ Q_2 &= \frac{x_1}{\mathscr{S} f^{(2)}}[s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}], \end{aligned} \tag{10.98}$$

where

$$\mathscr{S} \equiv \sigma_1^{(1)}\sigma_2^{(2)} - \sigma_2^{(1)}\sigma_1^{(2)}.$$

With this choice of portfolio weights the instantaneous change in $V$ becomes

$$dV = [(m_1 + q_1)x_1 + Q_1 f^{(1)}\mu_{f^{(1)}} + Q_2 f^{(2)}\mu_{f^{(2)}}]dt. \tag{10.99}$$

By a now familiar argument, the absence of arbitrage opportunities implies that $dV$ and $V$ with $Q_1, Q_2$ chosen according to (10.98) must satisfy

$$dV = rVdt,$$

which results in

$$(m_1 + q_1 - r)x_1 + Q_1 f^{(1)}(\mu_f^{(1)} - r) + Q_2 f^{(2)}(\mu_f^{(2)} - r) = 0, \tag{10.100}$$

which upon use of (10.98) for $Q_1, Q_2$ becomes

$$m_1 + q_1 - r = \frac{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}{\mathscr{S}}(\mu_{f^{(1)}} - r) - \frac{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}}{\mathscr{S}}(\mu_{f^{(2)}} - r). \tag{10.101}$$

Further re-arrangement yields

$$\frac{(m_1 + q_1 - r)\mathscr{S}}{(s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)})(s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)})} = \frac{\mu_{f^{(1)}} - r}{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}} - \frac{\mu_{f^{(2)}} - r}{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}. \tag{10.102}$$

We note that[8]

$$\frac{\mathscr{S}}{(s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)})(s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)})} = \frac{1}{\mathscr{G}}\left[\frac{s_{22}\sigma_1^{(1)} - s_{21}\sigma_2^{(1)}}{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}} - \frac{s_{22}\sigma_1^{(2)} - s_{21}\sigma_2^{(2)}}{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}\right],$$

so that (10.102) can be written

$$\frac{m_1 + q_1 - r}{\mathscr{G}}\left[\frac{s_{22}\sigma_1^{(1)} - s_{21}\sigma_2^{(1)}}{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}} - \frac{s_{22}\sigma_1^{(2)} - s_{21}\sigma_2^{(2)}}{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}\right]$$

[8] Recall that $\mathscr{G} = det(s) = s_{11}s_{22} - s_{12}s_{21}$.

$$= \frac{\mu_f^{(1)} - r}{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}} - \frac{\mu_f^{(2)} - r}{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}.$$

Algebraic manipulations of this last equation yield

$$\begin{aligned} &\frac{\mu_f^{(1)} - r - \frac{m_1+q_1-r}{\mathcal{G}}(s_{22}\sigma_1^{(1)} - s_{21}\sigma_2^{(1)})}{s_{11}\sigma_2^{(1)} - s_{12}\sigma_1^{(1)}} \\ &\qquad = \frac{\mu_f^{(2)} - r - \frac{m_1+q_1-r}{\mathcal{G}}(s_{22}\sigma_1^{(2)} - s_{21}\sigma_2^{(2)})}{s_{11}\sigma_2^{(2)} - s_{12}\sigma_1^{(2)}}. \end{aligned} \tag{10.103}$$

The quantity on the left-hand side of (10.103) involves only the option of maturity $T_1$ whilst the quantity on the right-hand side involves only the option of maturity $T_2$. Since the option maturities may be chosen arbitrarily it must be the case that the quantity in question must be the same for an option of any maturity and equal to a common factor which we set to $(\lambda_1 s_{21} + \lambda_2 s_{22})/\mathcal{G}$.[9] Thus we conclude from (10.103) that for an option of any maturity $T$

$$\mu_f - r = \frac{m_1 + q_1 - r}{\mathcal{G}}(s_{22}\sigma_1 - s_{21}\sigma_2) + \frac{\lambda_1 s_{21} + \lambda_2 s_{22}}{\mathcal{G}}(s_{11}\sigma_2 - s_{12}\sigma_1). \tag{10.104}$$

Upon employing the expression for $\mu_f$, $\sigma_1$, $\sigma_2$ and re-arranging, the last equation reduces to

$$\theta + (r - q_1)x_1\Delta_1 + (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})x_2\Delta_2 + \mathfrak{D}f - rf = 0, \tag{10.105}$$

or replacing $\theta$, $\Delta_1$ and $\Delta_2$ by their partial derivative expressions we see that we have the pricing partial differential equation

$$\frac{\partial f}{\partial t} + (r - q_1)x_1\frac{\partial f}{\partial x_1} + (m_2 - \lambda_1 s_{21} - \lambda_2 s_{22})x_2\frac{\partial f}{\partial x_2} + \mathfrak{D}f - rf = 0. \tag{10.106}$$

We re-arrange (10.104) by expressing the expected excess return $(m_1 + q_1 - r)$ as the risk premium [see Eq. (10.22)] $\lambda_1 s_{11} + \lambda_2 s_{12}$ to obtain

$$\mu_f - r = \frac{\lambda_1 s_{11} + \lambda_2 s_{12}}{\mathcal{G}}(s_{22}\sigma_1 - s_{21}\sigma_2) + \frac{\lambda_1 s_{21} + \lambda_2 s_{22}}{\mathcal{G}}(s_{11}\sigma_2 - s_{12}\sigma_1). \tag{10.107}$$

[9] This choice for the common factor seems quite unintuitive. We could simply set this common factor equal to $\lambda'$, then we would find that the right hand side of (10.108) could only be interpreted as the appropriate risk premia terms if $\lambda'$ were chosen in the way indicated.

A little algebra reduces this last equation to

$$\mu_f - r = \lambda_1\sigma_1 + \lambda_2\sigma_2. \tag{10.108}$$

This is (10.58) which can then be used to obtain the martingale representation.

## 10.6 Problems

**Problem 10.1** An asset price $x_1$ is driven by the diffusion process

$$dx_1 = m_1 x_1 dt + s_{11}(x_1, x_2, t)dw_1. \tag{10.109}$$

The factor $x_2$ appearing in the $s_{11}$ function is an exponentially declining weighted average of past Wiener increment (shock) terms, i.e.

$$x_2(t) = \int_0^t e^{-\kappa(t-s)} s_{21}(x_2, s)dw_1(s). \tag{10.110}$$

Here $\kappa > 0$ and $s_{21}$ is a given function. By taking the differential of Eq. (10.110) show that $x_2(t)$ satisfies the stochastic differential equation

$$dx_2(t) = -\kappa x_2(t)dt + s_{21}(x_2, t)dw_1(t). \tag{10.111}$$

An option is written on the asset $x_1$ and has maturity $T$ and payoff $h(x_1(T), T)$. Write down the value of this option as a partial differential equation. Be careful to specify the boundary condition.

**Problem 10.2** An asset price is driven by the diffusion process

$$dx = mxdt + \sigma xdz_1.$$

The quantity $v = \sigma^2$ is driven by the diffusion process

$$dv = \kappa(\bar{v} - v)dt + s(v)dz_2,$$

where $s(v)$ is some function and the Wiener increments $dz_1, dz_2$ are correlated according to $\mathbb{E}[dz_1\, dz_2] = \rho dt$. Express the stochastic differential system for $x$ and $\sigma$ in terms of independent Wiener increments.

An option written on the asset $x$ has payoff $h(x(T), T)$ where $T$ is the maturity date. Obtain the partial differential equation satisfied by the option price and specify its boundary condition.

*Note*: The quantity $v$ is *not* a traded asset.

**Problem 10.3** Fill in all the missing details in Appendix 10.2. In particular be sure you can derive Eqs. (10.102)–(10.104).

# Chapter 11
# Applying the General Pricing Framework

**Abstract** This chapter applies the general pricing framework developed in Chap. 10 to some standard one factor examples including stock options, currency options, futures options and a two factor model of exchange option.

## 11.1 Introduction

The key to using the general framework of the previous chapter is to interpret the excess return of each underlying factor to its market price of risk and its volatility

$$m_i + q_i - r = \sum_{j=1}^{n} \lambda_j s_{ij}, \tag{11.1}$$

where $(m_i + q_i - r)$ is the excess return of factor $i$. In calculating the excess return we must account for any income, or costs, associated with holding the asset underlying price $x_i$. This is captured by the $q_i$ term which may be either a continuously compounded dividend or cost. In the next two sections we show how the general pricing structure of Sect. 10.4 may be applied, once Eq. (11.1) has been appropriately interpreted for the situation at hand. Section 11.2 considers some standard one-factor examples, whilst Sect. 11.3 considers some two-factor examples.

## 11.2 One-Factor Examples

### *11.2.1 Stock Options*

In this case we have one underlying factor, the price $x$ of the stock, which we assume here pays no dividend. It follows the diffusion process

$$\frac{dx}{x} = mdt + sdw,$$


C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_11

and from (11.1) the excess expected return from holding the stock is

$$m - r = \lambda s,$$

from which

$$m - \lambda s = r. \tag{11.2}$$

Substituting this last expression into the general pricing Eq. (10.90) with $n_t = n = 1$, $n_n = 0$ we obtain

$$\frac{\partial f}{\partial t} + rx\frac{\partial f}{\partial x} + \frac{1}{2}s^2x^2\frac{\partial^2 f}{\partial x^2} = rf, \tag{11.3}$$

which is of course the Black–Scholes partial differential equation which we obtained earlier.

If the stock does pay a continuously compounded dividend $q$ then the expected excess return relation becomes

$$m + q - r = \lambda s$$

so that (11.2) is replaced by

$$m - \lambda s = r - q. \tag{11.4}$$

The partial differential equation (11.3) then becomes

$$\frac{\partial f}{\partial t} + (r - q)x\frac{\partial f}{\partial x} + \frac{1}{2}s^2x^2\frac{\partial^2 f}{\partial x^2} = rf.$$

### *11.2.2 Foreign Currency Options*

Here we have one factor $x$ which is the exchange rate (domestic currency/unit of foreign currency) that is the price of a unit of the foreign currency. The foreign currency yields continuously the risk free rate in the foreign country, which we denote by $r_f$. If the diffusion process followed by the exchange rate is written as

$$\frac{dx}{x} = mdt + sdw,$$

the expected return from holding the foreign currency is $(m + r_f)$ and hence the relationship (11.1) for expected excess return as it applies to foreign currency becomes

$$(m + r_f) - r = \lambda s,$$

and hence

$$m - \lambda s = r - r_f. \tag{11.5}$$

Upon substituting this into the general pricing Eq. (10.90) with $n = 1$ yields

$$\frac{\partial f}{\partial t} + (r - r_f)x\frac{\partial f}{\partial x} + \frac{1}{2}s^2x^2\frac{\partial^2 f}{\partial x^2} = rf, \tag{11.6}$$

which is the equation obtained by Garman and Kohlhagen (1983) for the pricing of a foreign currency option. In the case of a European call option with exercise exchange rate $E$ on the foreign currency its solution turns out to be

$$f(x,t) = x\, e^{-r_f(T-t)}\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2), \tag{11.7}$$

where

$$d_1 = \frac{\ln(x/E) + (r - r_f + s^2/2)(T-t)}{s\sqrt{T-t}},$$
$$d_2 = d_1 - s\sqrt{T-t}.$$

It should also be noted that from (10.92) the foreign currency option value can also be expressed as

$$f(x,t) = \tilde{\mathbb{E}}_t[f(x_T, T)]. \tag{11.8}$$

From (10.94) with $q_i = r_f$ the dynamics for the exchange rate under the equivalent measure $\tilde{\mathbb{P}}$ is given by

$$\frac{dx}{x} = (r - r_f)dt + sd\tilde{w} \tag{11.9}$$

It is the conditional transition density function associated with (11.9) that is required to calculate the $\tilde{\mathbb{E}}_t$ in (11.8).

### *11.2.3 Futures Options*

In this case the factor $x$ is the price of a futures contract on an underlying asset whose price is $S$. The derivative security in this case is an option on the futures contract. Simple arbitrage arguments can be used to show that the relationship between the futures price $x$ and the price $S$ of the asset underlying the futures contract, is

$$x = Se^{\alpha(T^*-t)}, \tag{11.10}$$

where $T^*$ is the maturity date of the futures contract. Here

$$\alpha = \text{risk free rate} - \text{yield on the asset}. \tag{11.11}$$

For example if the underlying asset were a commodity, then $\alpha$ would be, the risk-free rate plus storage costs minus the convenience yield. If the price of the underlying asset follows the diffusion process

$$\frac{dS}{S} = \mu dt + \sigma dw, \tag{11.12}$$

then a straight forward application of Ito's lemma reveals that $x$ follows the diffusion process

$$\frac{dx}{x} = (\mu - \alpha)dt + \sigma dw \equiv \mu_x dt + \sigma dw. \tag{11.13}$$

It is well known that a futures price can be regarded as the price of a security paying a continuous dividend yield at the risk free rate $r$ which means we set $\alpha = r$ in (11.13) (see Hull 2000). Thus applying (11.1) here yields $(r + \mu_x) - r = \lambda\sigma$, i.e.

$$\mu_x - \lambda\sigma = 0,$$

and so the pricing equation becomes

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} = rf.$$

In the case of a European futures option the solution to this partial differential equation is

$$f = e^{-r(T-t)}[x\mathcal{N}(d_1) - E\mathcal{N}(d_2)],$$

where

$$d_1 = \frac{\ln(x/E) + (\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t},$$

which is Black (1976) model. Note that formally Black's model could be obtained from the foreign currency option model by setting $r_f = r$. From (10.92) the futures option value can also be expressed as

$$f(x,t) = \tilde{\mathbb{E}}_t[f(x_T, T)], \tag{11.14}$$

where by use of (10.94) with $q_i = r$, the dynamics of the futures price under the equivalent measure $\tilde{\mathbb{P}}$ are

$$\frac{dx}{x} = \sigma d\tilde{w}. \tag{11.15}$$

The transition probability density function associated with (11.15) is used to calculate $\tilde{\mathbb{E}}_t$ in (11.14). Incidentally we note that (11.15) indicates that the futures price is a Martingale under the equivalent measure $\tilde{\mathbb{P}}$.

## 11.3 Options on Two Underlying Factors

As an application of options on two underlying assets, both of which are traded, we consider exchange options which were first studied by Margrabe (1978). These are the most basic of a class of multi-asset options (digital options, quotient options, foreign equity options, quanto options etc.) which derive their value from the correlation structure between two underlying traded assets. For much more information and details about multi-asset options, we refer the reader to Zhang (1997).

We consider the framework and notation of Sect. 10.3.1 dealing with two traded underlying assets. A European exchange option to pay the second asset in exchange for the first has payoff given by

$$f_T = \max[x_1(T) - x_2(T), 0].$$

This payoff differs from that of a standard European call option on $x_1$ in that the exercise price $E$ is replaced by $x_2(T)$, the value at maturity of the second asset. Alternatively, it could be regarded as a European put option on $x_2$ with exercise price equal to $x_1(T)$. An investor might be interested in this type of option if $x_1$ and $x_2$ were negatively correlated and the investor wished to have some insurance against $x_2$ performing "badly".

The price of the exchange option satisfies the partial differential equation (10.21) subject to the terminal condition

$$f(x_1, x_2, T) = \max[x_1(T) - x_2(T), 0].$$

Alteratively, we may use the martingale representation (10.31) so that

$$f(x_1, x_2, t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[\max(x_1(T) - x_2(T), 0)], \tag{11.16}$$

where we recall that $\tilde{\mathbb{E}}_t$ is calculated using the distribution generated by the stochastic differential equations (10.28), which may be re-expressed as (see Sect. 6.3.2 and Problem 6.7)

$$\begin{aligned} d(\ln x_1) &= [r - q_1 - \frac{1}{2}(s_{11}^2 + s_{12}^2)]dt + s_{11}d\tilde{w}_1 + s_{12}d\tilde{w}_2, \\ d(\ln x_2) &= [r - q_2 - \frac{1}{2}(s_{21}^2 + s_{22}^2)]dt + s_{21}d\tilde{w}_1 + s_{22}d\tilde{w}_2. \end{aligned} \tag{11.17}$$

We shall concentrate on the case represented by the stochastic differential equations (10.7) so that (11.17) becomes

$$\begin{aligned} d(\ln x_1) &= (r - q_1 - \frac{s_1^2}{2})dt + s_1 d\tilde{w}_1, \\ d(\ln x_2) &= (r - q_2 - \frac{s_2^2}{2})dt + \rho s_2 d\tilde{w}_1 + \sqrt{1-\rho^2} s_2 d\tilde{w}_2. \end{aligned}$$

Thus, the probability density function for the joint distribution of $\ln[x_1(T)/x_1(t)]$, $\ln[x_2(T)/x_2(t)]$ is[1]

$$\tilde{\pi}[v_1(T), v_2(T), T] = \frac{1}{2\pi s_1 s_2 \sqrt{1-\rho^2}(T-t)} \exp\left[-\frac{u_1^2 - 2\rho u_1 u_2 + u_2^2}{2(1-\rho^2)}\right], \tag{11.18}$$

where $v_i(T) \equiv \ln[x_i(T)/x_i(t)]$ (as we are expressing the distribution of relative prices we have dropped the notation for conditioning on time $t$) and

$$u_i \equiv \frac{v_i(T) - (r - q_i - \frac{1}{2}s_i^2)(T-t)}{s_i\sqrt{T-t}}, \qquad \text{for} \qquad i = 1, 2.$$

Equation (11.18) is the bivariate normal distribution for the logarithm of the relative prices (see Fig. 11.1). Equation (11.18) may also be expressed as

$$\tilde{\pi}[v_1(T), v_2(T), T] = \tilde{\pi}_1[v_1(T), T]\tilde{\pi}_2[v_2(T), T \mid v_1(T), T],$$

where

$$\tilde{\pi}_1[v_1(T), T] = \frac{1}{s_1\sqrt{2\pi(T-t)}} \exp\left(-\frac{u_1^2}{2}\right), \tag{11.19}$$

[1] See Problem 6.7. Note that the density function there denoted as $p(y_1)(T), y_2(T), T|y_1(t), y_2(t), (t)$ is here denoted as $\tilde{\pi}\ [v_1(T), v_2(T), T]$.

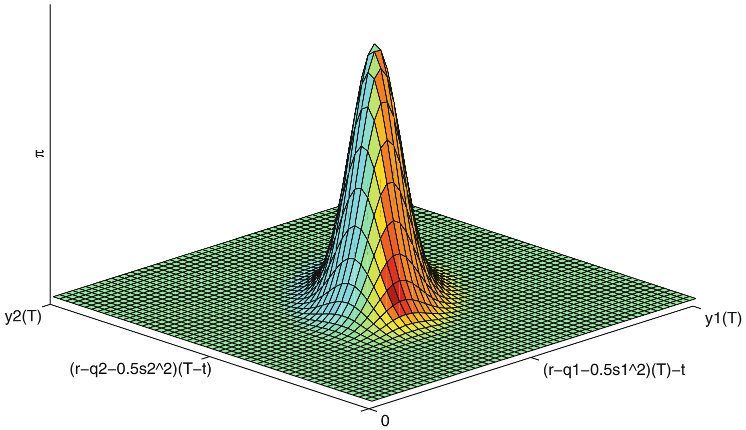


**Fig. 11.1** Bivariate normal distribution

and

$$\tilde{\pi}_2[v_2(T), T \mid v_1(T), T] = \frac{1}{s_2\sqrt{1-\rho^2}\sqrt{2\pi(T-t)}} \exp\left(-\frac{(u_2-\rho u_1)^2}{2(1-\rho^2)}\right), \tag{11.20}$$

or, alternatively as

$$\tilde{\pi}[v_1(T), v_2(T), T] = \tilde{\pi}_2[v_2(T), T]\tilde{\pi}_1[v_1(T), T \mid v_2(T), T],$$

where now (with a slight abuse of notation)

$$\tilde{\pi}_2[v_2(T), T] = \frac{1}{s_2\sqrt{2\pi(T-t)}} \exp\left(-\frac{u_2^2}{2}\right),$$

and

$$\tilde{\pi}_1[v_1(T), T \mid v_2(T), T] = \frac{1}{s_1\sqrt{1-\rho^2}\sqrt{2\pi(T-t)}} \exp\left(-\frac{(u_1-\rho u_2)^2}{2(1-\rho^2)}\right).$$

We shall use $\tilde{\pi}[x_1(T), x_2(T), T]$ to denote the corresponding probabilities in terms of the original asset prices $x_1, x_2$. In terms of these distributions, Eq. (11.16) becomes

$$f(x_1, x_2, T) = e^{-r(T-t)} \int_0^\infty \left( \int_0^{x_1} (x_1 - x_2)\tilde{\pi}(x_1, x_2, T)\frac{dx_2}{x_2} \right) \frac{dx_1}{x_1}, \tag{11.21}$$

where the region of integration is illustrated in Fig. 11.2.

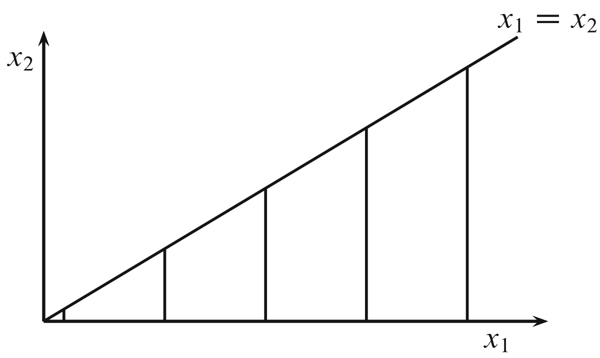


**Fig. 11.2** Region of integration for the exchange option

Since the integration with respect to $x_2$ is performed, whilst $x_1$ is held constant, it is more convenient to use the conditional distributions (11.19), (11.20) (after converting to $\ln S$). Thus, we may express (11.21) as

$$f(x_1, x_2, T) = e^{-r(T-t)} \int_0^\infty \left( \int_0^{x_1} (x_1 - x_2)\tilde{\pi}_2(x_2, T \mid x_1, T)\frac{dx_2}{x_2} \right) \tilde{\pi}_1(x_1, T)\frac{dx_1}{x_1}$$
$$\equiv e^{-r(T-t)} \left( \int_0^\infty J_1(x_1, T)\tilde{\pi}_1(x_1, T)dx_1 - \int_0^\infty J_2(x_1, T)\tilde{\pi}_1(x_1, T)\frac{dx_1}{x_1} \right), \tag{11.22}$$

where

$$J_1(x_1, T) \equiv \int_0^{x_1} \tilde{\pi}_2(x_2, T \mid x_1, T)\frac{dx_2}{x_2}, \tag{11.23}$$

and

$$J_2(x_1, T) \equiv \int_0^{x_1} \tilde{\pi}_2(x_2, T \mid x_1, T)dx_2. \tag{11.24}$$

Our integration task is simplified by noting that the integrals $J_1$, $J_2$ are essentially the integrals $A_1$, $A_2$ that we evaluated in Appendix 3.1. The same changes of variables, completing the square etc., need to be applied to the evaluation of $J_1$ and $J_2$. In Appendix 11.1 we show that (setting $\tau = T - t$)

$$J_1(x_1, T) = \mathcal{N}\left( \frac{(1-\overline{\rho})\ln x_1 + \ln\left(\frac{x_1(t)^{\overline{\rho}}}{x_2(t)}\right) - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}} \right), \tag{11.25}$$

when

$$\overline{\rho} = \rho\frac{s_2}{s_1}, \qquad \gamma \equiv q_2 + \frac{s_2^2\rho^2}{2} + \overline{\rho}\left(r - q_1 - \frac{s_1^2}{2}\right), \tag{11.26}$$

and

$$J_2(x_1, T) = x_2(t)(\frac{x_1}{x_1(t)})^{\overline{\rho}} e^{(r-\gamma)\tau}$$
$$\mathcal{N}\left(\frac{(1-\overline{\rho})\ln x_1 + \ln\left(\frac{x_1(t)^{\overline{\rho}}}{x_2(t)}\right) - \left(r - \gamma + \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right). \tag{11.27}$$

Substituting (11.25) and (11.27) into (11.22) we have

$$f(x_1, x_2, T) = e^{-r\tau}B_1 - e^{-\gamma\tau}B_2, \tag{11.28}$$

where

$$B_1 \equiv \frac{1}{\sqrt{2\pi\tau}s_1}\int_0^{\infty}\mathcal{N}\left(\frac{(1-\overline{\rho})\ln x_1 + \ln\left(\frac{x_1(t)^{\overline{\rho}}}{x_2(t)}\right) - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)$$
$$\exp\left[-\frac{\left\{\ln\left(\frac{x_1}{x_1(t)}\right) - (r - q_1 - \frac{s_1^2}{2})\tau\right\}^2}{2s_1^2\tau}\right]dx_1,$$

and we show in Appendix 11.2 that

$$B_1 = x_1(t)e^{(r-q_1)\tau}\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\mathcal{N}(\alpha_1 + \beta z)dz,$$

where

$$\alpha_1 = \frac{\ln\left(\frac{x_1(t)}{x_2(t)}\right) + (1-\overline{\rho})(r - q_1 + \frac{s_1^2}{2})\tau - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}},$$

$$\beta = \frac{s_1}{s_2}\frac{1-\overline{\rho}}{\sqrt{1-\rho^2}},$$

and that

$$B_2 = x_2(t)e^{\overline{\rho}(r-q_1+\frac{(\overline{\rho}-1)s_1^2}{2})\tau}\int_{-\infty}^{\infty}\frac{e^{-z^2/2}}{\sqrt{2\pi}}\mathcal{N}(\alpha_2+\beta z)dz,$$

where

$$\alpha_2 = \frac{\ln\left(\frac{x_1(t)}{x_2(t)}\right)+(1-\overline{\rho})(r-q_1+s_1^2(\overline{\rho}-\frac{1}{2}))\tau-\left(r-\gamma+\frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}.$$

Finally we use the result that (see Appendix 11.2)

$$\int_{-\infty}^{\infty}\frac{e^{-z^2/2}}{\sqrt{2\pi}}\mathcal{N}(\alpha+\beta z)dz = \mathcal{N}\left(\frac{\alpha}{\sqrt{1+\beta^2}}\right),$$

from which we obtain

$$B_1 = x_1(t)e^{(r-q_1)\tau}\mathcal{N}\left(\frac{\alpha_1}{\sqrt{1+\beta^2}}\right),$$

$$B_2 = x_2(t)e^{\overline{\rho}(r-q_1+\frac{(\overline{\rho}-1)s_1^2}{2})\tau}\mathcal{N}\left(\frac{\alpha_2}{\sqrt{1+\beta^2}}\right).$$

The value of the exchange option (see (11.28)) is given by

$$f(x_1,x_2,T) = x_1(t)e^{-q_1\tau}\mathcal{N}\left(\frac{\alpha_1}{\sqrt{1+\beta^2}}\right)-x_2(t)e^{-q_2\tau}\mathcal{N}\left(\frac{\alpha_2}{\sqrt{1+\beta^2}}\right).$$

## 11.4 Appendix

### Appendix 11.1 The Integrals $J_1$ and $J_2$

#### *A. Evaluation of $J_1(x_1, T)$*

From Eq. (11.23) we have

$$J_1(x_1,T) = \int_0^{x_1}\tilde{\pi}_2(x_2,T|x_1,T)\frac{dx_2}{x_2}.$$

By use of (11.20) (expressed in terms of $x_2$, $x_1$) we have (setting $\tau = T - t$)

$$J_1(x_1,T) = \frac{1}{\sqrt{2\pi\tau}s_2\sqrt{1-\rho^2}} \int_0^{x_1} \exp\left[-\frac{\left\{\ln\left(\frac{x_2(T)}{x_2(t)}\right) - \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau\right\}^2}{2s_2^2(1-\rho^2)\tau}\right]\frac{dx_2(T)}{x_2(T)},$$

where to simplify notation we have set

$$\delta\tau = \frac{s_2^2\rho^2\tau}{2} - \rho\frac{s_2}{s_1}\left[\ln\frac{x_1(T)}{x_1(t)} - (r - q_1 - \frac{s_1^2}{2})\tau\right].$$

Since $J_1$ represents the area under the log-normal curve from $(0, x_1)$, we may also express it as

$$J_1(x_1,T) = 1 - \frac{1}{\sqrt{2\pi\tau}s_2\sqrt{1-\rho^2}} \int_{x_1}^{\infty} \exp\left[-\frac{\left\{\ln\left(\frac{x_2(T)}{x_2(t)}\right) - \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau\right\}^2}{2s_2^2(1-\rho^2)\tau}\right]\frac{dx_2(T)}{x_2(T)}.$$

Now we are dealing precisely with the integral $A_2$ in Eq. (3.29) of Appendix. 3.1, so that after appropriate identifications we may write

$$\begin{aligned} J_1(x_1,T) &= 1 - \mathcal{N}\left(\frac{\ln\left(\frac{x_2(t)}{x_1}\right) + \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right) \\ &= \mathcal{N}\left(\frac{\ln\left(\frac{x_1}{x_2(t)}\right) - \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right) \\ &= \mathcal{N}\left(\frac{(1-\overline{\rho})\ln x_1 + \ln\left(\frac{x_1(t)^{\overline{\rho}}}{x_2(t)}\right) - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right), \end{aligned}$$

where

$$\overline{\rho} = \rho\frac{s_2}{s_1}, \qquad \gamma \equiv q_2 + \frac{s_2^2\rho^2}{2} + \overline{\rho}\left(r - q_1 - \frac{s_1^2}{2}\right).$$

## *B. Evaluation of $J_2(x_1, T)$*

From Eq. (11.24) we have

$$J_2(x_1, T) = \int_0^{x_1} \tilde{\pi}_2(x_2, T | x_1, T) dx_2$$

which by use of (11.20) may be expressed

$$J_2(x_1, T) = \frac{1}{\sqrt{2\pi\tau} s_2 \sqrt{1-\rho^2}} \int_0^{x_1} \exp\left[ -\frac{\left\{ \ln\left(\frac{x_2(T)}{x_2(t)}\right) - \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau \right\}^2}{2s_2^2(1-\rho^2)\tau} \right] dx_2(T),$$

or, in terms of the log price variable $y_2 = \ln\left(\frac{x_2(T)}{x_2(t)}\right)$

$$\begin{aligned} J_2(x_1, T) &\\ = \frac{x_2(t)}{\sqrt{2\pi\tau} s_2 \sqrt{1-\rho^2}} &\int_{-\infty}^{\ln\left(\frac{x_1}{x_2(t)}\right)} \exp\left[ -\frac{\left\{ y_2 - \left(r - q_2 - \delta - \frac{s_2^2(1-\rho^2)}{2}\right)\tau \right\}^2}{2s_2^2(1-\rho^2)\tau} \right] e^{y_2} dy_2 \\ = \frac{x_2(t) e^{(r-q_2-\delta)\tau}}{\sqrt{2\pi\tau} s_2 \sqrt{1-\rho^2}} &\int_{-\infty}^{\ln\left(\frac{x_1}{x_2(t)}\right)} \exp\left[ -\frac{\left\{ y_2 - \left(r - q_2 - \delta + \frac{s_2^2(1-\rho^2)}{2}\right)\tau \right\}^2}{2s_2^2(1-\rho^2)\tau} \right] dy_2, \end{aligned}$$

where we have employed the same completion of the square technique as was used to simplify $A_1$ in Appendix 3.1. Setting

$$z = \frac{y_2 - \left(r - q_2 - \delta + \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2 \sqrt{1-\rho^2} \sqrt{\tau}},$$

we then obtain

$$J_2(x_1, T) = \frac{x_2(t) e^{(r-q_2-\delta)\tau}}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{1}{2}z^2} dz,$$

where

$$d \equiv \frac{\ln\left(\frac{x_1}{x_2(t)}\right) - \left(r - q_2 - \delta + \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}.$$

Recalling the definition of $\mathcal{N}(d)$, we can finally write

$$J_2(x_1, T) = x_2(t)e^{(r-q_2-\delta)\tau}\mathcal{N}(d).$$

Reexpressing this so as to highlight the dependence on $x_1$,

$$J_2(x_1, T) = x_2(t)(\frac{x_1}{x_1(t)})^{\overline{\rho}}e^{(r-\gamma)\tau}$$
$$\mathcal{N}\left(\frac{(1-\overline{\rho})\ln x_1 + \ln\left(\frac{x_1(t)^{\overline{\rho}}}{x_2(t)}\right) - \left(r - \gamma + \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right).$$

## Appendix 11.2 The Integrals $B_1$ and $B_2$

### *The Integral $B_1$*

With a slight re-arrangement of the argument of the $\mathcal{N}$ function, $B_1$ can be written

$$B_1 = \frac{1}{\sqrt{2\pi\tau}s_1}\int_0^\infty \mathcal{N}\left(\frac{(1-\overline{\rho})\ln\left(\frac{x_1}{x_1(t)}\right) + \ln\left(\frac{x_1(t)}{x_2(t)}\right) - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)$$
$$\exp\left[\frac{-\left\{\ln\left(\frac{x_1}{x_1(t)}\right) - (r - q_1 - \frac{s_1^2}{2})\tau\right\}^2}{2s_1^2\tau}\right] dx_1.$$

Setting

$$z = \frac{\ln\frac{x_1}{x_1(t)} - (r - q_1 - \frac{s_1^2}{2})\tau}{s_1\sqrt{\tau}}$$

we may re-express $B_1$ as

$$B_1 = \frac{x_1(t)}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{z^2}{2}+s_1\sqrt{\tau}z+(r-q_1-\frac{s_1^2}{2})\tau}\mathcal{N}\left(\frac{(1-\overline{\rho})s_1\sqrt{\tau}z + \ln\left(\frac{x_1(t)}{x_2(t)}\right) + b_1}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right) dz,$$

where

$$b_1 = (1-\overline{\rho})(r - q_1 - \frac{s_1^2}{2})\tau - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau.$$

Completing the square in the exponent we can write

$$B_1 = x_1(t)e^{(r-q_1)\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-s_1\sqrt{\tau})^2}\mathcal{N}\left(\frac{(1-\overline{\rho})s_1\sqrt{\tau}z + \ln\left(\frac{x_1(t)}{x_2(t)}\right) + b_1}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)dz,$$

which after a further change of variable becomes

$$B_1 = x_1(t)e^{(r-q_1)\tau}\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}\mathcal{N}(\alpha_1 + \beta z)dz,$$

where

$$\alpha_1 = \frac{\ln\left(\frac{x_1(t)}{x_2(t)}\right) + (1-\overline{\rho})(r - q_1 + \frac{s_1^2}{2})\tau - \left(r - \gamma - \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}},$$

$$\beta = \frac{s_1}{s_2}\frac{1-\overline{\rho}}{\sqrt{1-\rho^2}}.$$

## *The Integral $B_2$*

$$B_2 = \frac{1}{\sqrt{2\pi\tau}s_1}\int_0^{\infty}\frac{x_2(t)}{x_1(t)^{\overline{\rho}}}x_1^{\overline{\rho}-1}\mathcal{N}\left(\frac{(1-\overline{\rho})\ln\left(\frac{x_1}{x_1(t)}\right) + \ln\left(\frac{x_1(t)}{x_2(t)}\right) - \left(r - \gamma + \frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)$$
$$\exp\left[-\frac{\left\{\ln\left(\frac{x_1}{x_1(t)}\right) - (r - q_1 - \frac{s_1^2}{2})\tau\right\}^2}{2s_1^2\tau}\right]dx_1.$$

We make the change of variable

$$z = \frac{\ln\frac{x_1}{x_1(t)} - (r - q_1 - \frac{s_1^2}{2})\tau}{s_1\sqrt{\tau}},$$

to obtain

$$B_2 = \frac{x_2(t)e^{\overline{\rho}(r-q_1-\frac{s_1^2}{2})\tau}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}+\overline{\rho}s_1\sqrt{\tau}z}$$

$$\mathcal{N}\left(\frac{(1-\overline{\rho})s_1\sqrt{\tau}z + \ln\left(\frac{x_1(t)}{x_2(t)}\right) + (1-\overline{\rho})(r-q_1-\frac{s_1^2}{2})\tau - \left(r-\gamma+\frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)dz.$$

Completing the square and simplifying,

$$B_2 = x_2(t)e^{\overline{\rho}(r-q_1+\frac{(\overline{\rho}-1)s_1^2}{2})\tau} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}(z-\overline{\rho}s_1\sqrt{\tau})^2}}{\sqrt{2\pi}}$$

$$\mathcal{N}\left(\frac{(1-\overline{\rho})s_1\sqrt{\tau}z + \ln\left(\frac{x_1(t)}{x_2(t)}\right) + (1-\overline{\rho})(r-q_1-\frac{s_1^2}{2})\tau - \left(r-\gamma+\frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}}\right)dz.$$

A further change of variable allows us to write

$$B_2 = x_2(t)e^{\overline{\rho}(r-q_1+\frac{(\overline{\rho}-1)s_1^2}{2})\tau} \int_{-\infty}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \mathcal{N}(\alpha_2+\beta z)dz,$$

where

$$\alpha_2 = \frac{\ln\left(\frac{x_1(t)}{x_2(t)}\right) + (1-\overline{\rho})(r-q_1+s_1^2(\overline{\rho}-\frac{1}{2}))\tau - \left(r-\gamma+\frac{s_2^2(1-\rho^2)}{2}\right)\tau}{s_2\sqrt{1-\rho^2}\sqrt{\tau}},$$

$$\beta = \frac{s_1}{s_2}\frac{1-\overline{\rho}}{\sqrt{1-\rho^2}}.$$

## 11.5 Problems

**Problem 11.1** Solve Eq. (11.6) in the case of a European foreign currency option to obtain the solution (11.7). Do this in two ways. First, by applying the Fourier transform techniques of Chap. 9. Note that by setting $q = r_f$ you can take a lot of the results set out there. Second, obtain the solution to the Kolmogorov backward equation associated with (11.9) and then use (11.8) and integration. To implement this approach you should use (6.23) (appropriately re-interpreted) to obtain $\tilde{p}(x_T, T|x, t)$. The integration may be done by using the results in Appendix 3.1. These will have to be extended slightly to accommodate the $r_f$.

# Chapter 12
# Jump-Diffusion Processes

**Abstract** This chapter considers jump-diffusion processes to allow for price fluctuations to have two components, one consisting of the usual increments of a Wiener process, the second allows for "large" jumps from time-to-time. We introduce Poisson jump process with either absolute or proportional jump sizes through the stochastic integrals and provide solutions when both the stock price and Poisson jump size are log-normal. We also extend Ito's lemma for the jump-diffusion processes.

## 12.1 Introduction

In the derivation of the Black–Scholes model we listed the "ideal market conditions" which are assumed to hold. One important assumption is that concerning the asset price dynamics, namely that the price follows a stochastic process with a continuous sample path. Essentially this means that the asset price changes satisfy a local Markov property, that in a short interval of time, the asset price changes only by a small amount. The technical statement of this property is the Lindeberg condition which we discussed in Sect. 2.4, and leads to the modelling of the uncertain stochastic term by the increments of a Wiener process (i.e. the $dz$ term in the stochastic differential equation). Implicit in this approach to modelling the asset price dynamics is the notion that the factors causing the random fluctuations around the average trend (e.g. changes in market conditions, changes in general economic conditions) cause only *marginal* changes in the asset price. This assumption is important as it underlies the continuous hedging argument. Of course in real markets continuous hedging is not possible. However Merton and Samuelson (1974) have been able to demonstrate that the Black–Scholes continuous trading solution is a reasonable approximation to the more realistic discrete time trading solution, provided that the asset price dynamics are generated by a stochastic process with continuous sample paths.

It is therefore of some interest to consider the effect of relaxing the continuous sample path assumption and to allow for random fluctuations which have more than a marginal effect on the price of the underlying asset. The stochastic process that allows us to incorporate this type of effect is the jump process, which we have



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_12

already discussed in Chap. 2. This process allows the random fluctuations of the asset price to have two components, one consisting of the usual increments of a Wiener process, the second allows for "large" jumps in the asset price from time-to-time.

There is some empirical evidence to suggest that jump processes may be more appropriate for describing the dynamics of foreign exchange rates (see e.g. Akgiray and Booth 1988) and also for the dynamics of interest rates (see e.g. Ahn and Thompson 1988). Furthermore, Merton (1982) shows that in a continuous trading environment asset price dynamics can always be described by a mixture of diffusion processes and Poisson jump processes.

## 12.2 Mathematical Description of the Jump Process

We have already shown in Sect. 2.6 how to modify the Fokker–Planck and Kolmogorov equations to incorporate jump process terms. However, in order to extend the hedging portfolio approach, we need a sample path description of mixed jump-diffusion processes. Stochastic differential equations incorporating both continuous diffusion and Poisson jump elements are discussed by Kushner (1967) and Gihman and Skorohod (1972).

The arrival of the "events" causing a "large" price jump is assumed to follow a Poisson process. In particular, the arrivals of the "events" are independently and identically distributed, so that the probability of an event occurring during a time interval of length $\Delta t$ can be written

$$\text{Prob}\left\{\begin{matrix}\text{event occurs once in the}\\ \text{time interval}(t, t+\Delta t)\end{matrix}\right\} = \lambda\Delta t + o(\Delta t),$$

$$\text{Prob}\left\{\begin{matrix}\text{event occurs more than once}\\ \text{in the time interval}(t, t+\Delta t)\end{matrix}\right\} = o(\Delta t),$$

$$\text{Prob}\left\{\begin{matrix}\text{event does not occur in the}\\ \text{time interval}(t, t+\Delta t)\end{matrix}\right\} = 1-\lambda\Delta t + o(\Delta t),$$

where $\lambda$ is the mean number of arrivals of events per unit time.

There are two ways in which we may expound the description of jump processes. One is to describe the jump in absolute terms (this is useful if we are focusing on prices or interest rates), the other is to describe the jump in proportional terms (this is more useful if our focus is on returns). Since both points of view are useful we shall describe both. We start our exposition in terms of jump sizes being measured in absolute terms and then consider the situation of jumps being measured in relative terms as a special case. Since this special case is relevant for the stock option pricing framework we have developed in previous chapters we in fact enter into quite a deal of detail, particularly concerning the price distribution when returns are governed by both diffusion and jump stochastic components.

### 12.2.1 Absolute Jumps

Consider the diffusion process generated by the stochastic differential equation

$$dx = \mu(x,t)dt + \sigma(x,t)dw \tag{12.1}$$

to which we wish to add Poisson jumps. Here $\mu$ is the instantaneous expected return on the asset, $\sigma^2$ is the instantaneous variance of the return (conditional on no arrival of Poisson jump events) and, $dw$ is the increment of a Wiener process $w$ under the physical measure $\mathbb{P}$.

The jumps are modelled as arriving at random times $t_i$ with a jump intensity $\lambda$. If we use $n(dt)$ to denote the number of jumps occurring in the interval $dt$, then

$$\begin{aligned} Pr[n(dt) = 0] &= 1 - \lambda dt + o(dt), \\ Pr[n(dt) = 1] &= \lambda dt + o(dt), \\ Pr[n(dt) > 1] &= o(dt). \end{aligned}$$

Furthermore if $N$ is the number of jumps up to time $t$ then

$$Pr[N = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}. \tag{12.2}$$

We shall use $\mathbb{Q}_j$ to indicate the probability measure governing the jump arrival times. Let the size of the jump be $y$, which is drawn from a distribution whose density function we denote by $g(y)$ and the associated measure $\mathbb{Q}_y$. The behaviour of the process at a typical jump point $t_i$ is displayed in Fig. 12.1a for absolute jumps and Fig. 12.1b for proportional jumps. Thus, for the absolute jumps that we are currently considering

$$x(t_i^+) = x(t_i^-) + y.$$

In order to express the jump process within the stochastic differential equation framework it is convenient to introduce the Poisson increment $dN$ which we define as

$$dN = \begin{cases} 1, \text{ with prob.} & \lambda dt, \\ 0, \text{ with prob.} & (1 - \lambda dt). \end{cases}$$

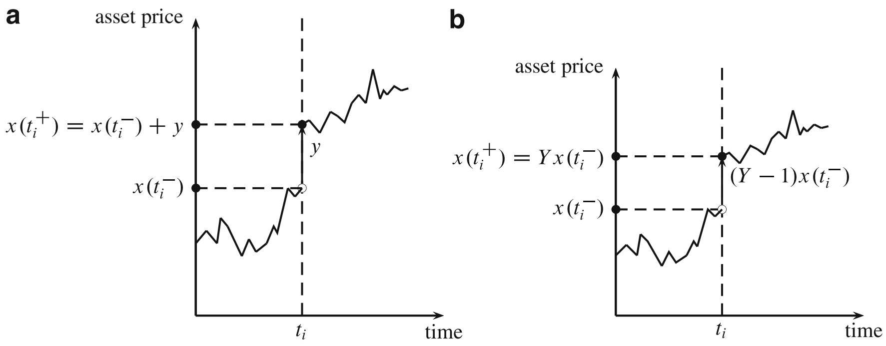


**Fig. 12.1** Sample path description of a jump-diffusion process. (**a**) An absolute jump. (**b**) A proportional jump

In other words, $N(t + dt) - N(t)$ is drawn from a Poisson distribution with mean $\lambda dt$.

In setting up the notation for the stochastic differential equation with Poisson jumps it is convenient to do so in such a way that the expected change in $dx$ over the infinitesimal interval $dt$ remains $\mu(x,t)dt$, as it is in (12.1). But clearly the jump process will alter the mean change in $dx$. To determine this mean change due to jumps we note that when a jump occurs the mean jump size is

$$k \equiv \mathbb{E}^{\mathbb{Q}_y}[y] = \int yg(y)dy,$$

where we write $\mathbb{E}^{\mathbb{Q}_y}$ to denote expectation with respect to the jump size measure $\mathbb{Q}_y$. However the jumps only occur with probability $\lambda dt$, so that the mean change in $x$ over an infinitesimal interval $dt$ due just to jumps is given by (note that we are assuming the independence of $\mathbb{Q}_y$ and $\mathbb{Q}_j$)

$$\mathbb{E}^{\mathbb{Q}_y}\mathbb{E}^{\mathbb{Q}_j}[dx] = \mathbb{E}^{\mathbb{Q}_y}[(1-\lambda dt)0 + \lambda\,(dt)\,y] = \lambda k dt, \tag{12.3}$$

where we write $\mathbb{E}^{\mathbb{Q}_y}$ to denote expectation with respect to the jump-arrival measure $\mathbb{Q}_j$. So if we want the mean change in $dx$, over $dt$, to be $\mu(x,t)dt$ we need to compensate for the mean jump component in (12.3). Keeping the above facts in mind, we write the stochastic differential equation for a combined jump diffusion process as

$$dx = (\mu(x,t) - \lambda k)dt + \sigma(x,t)dw + ydN. \tag{12.4}$$

Following Cox and Ross (1976a) this stochastic differential equation can be interpreted as

$$dx = \begin{cases} (\mu(x,t) - \lambda k)dt + \sigma(x,t)dw, & \text{with prob.} \quad (1 - \lambda dt), \\ (\mu(x,t) - \lambda k)dt + \sigma(x,t)dw + y, & \text{with prob.} \quad \lambda dt. \end{cases} \tag{12.5}$$

From the latter interpretation and the assumption of the independence of $\mathbb{P}$, $\mathbb{Q}_y$ and $\mathbb{Q}_j$ we calculate

$$\begin{aligned} \mathbb{E}[dx] =& \mathbb{E}^{\mathbb{Q}_y}\mathbb{E}^{\mathbb{Q}_j}\mathbb{E}^{\mathbb{P}}[dx] = \mathbb{E}^{\mathbb{Q}_y}\mathbb{E}^{\mathbb{Q}_j}[(\mu(x,t) - \lambda k)dt + ydN] \\ =& (\mu(x,t) - \lambda k)dt + \mathbb{E}^{\mathbb{Q}_y}(y)\mathbb{E}^{\mathbb{Q}_j}(dN) = \mu(x,t)dt, \end{aligned} \tag{12.6}$$

giving the desired property that the mean of the increment $dx$ remains the same as under the pure diffusion process. It is also sometimes convenient to write (12.4) as

$$dx = \mu(x,t)dt + \sigma(x,t)dw + (ydN - \lambda kdt). \tag{12.7}$$

Since

$$\mathbb{E}(ydN - \lambda kdt) = \mathbb{E}^{\mathbb{Q}_y}(y)\mathbb{E}^{\mathbb{Q}_j}(dN) - \lambda kdt = k\lambda dt - \lambda kdt = 0,$$

we see that under the representation (12.7) both stochastic terms are martingales.

Of course, just as in Chaps. 4 and 6, the stochastic differential equation notations in (12.4), (12.5) and (12.7) are just convenient shorthand notations. These equations need to be properly mathematically defined in terms of stochastic integrals. Thus we need to extend to Poisson jump processes the concept of a stochastic integral. To this end we focus just on the pure jump component and define a process $J(t)$ which is the contribution to $x(t)$ arising purely from the jump components. Using $N(t)$ to denote the number of jumps that have occurred up to time $t$ and referring to Fig. 12.2 (and assuming $J(0) = 0$) we see that, for a particular path,

$$J(t) = \sum_{i=1}^{N(t)} y_i,$$

where $y_i$ denotes the change in $x$ at jump time $t_i$, i.e.

$$y_i = x(t_i^+) - x(t_i^-).$$

The stochastic properties of $J(t)$ are determined by the Poisson arrival process and the jump size distribution of $y$. If the jump sizes are bounded then the definition of the stochastic integral with respect to jump processes is much simpler than the corresponding definition of the stochastic integral with respect to the Wiener process, where the principal difficulty lies in the fact that the changes in the Wiener

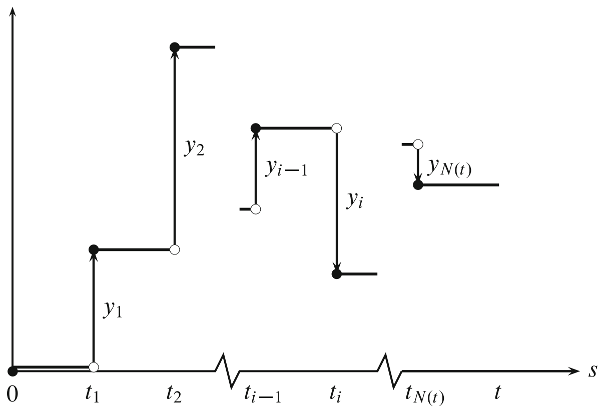


**Fig. 12.2** A typical sample path of the pure jump process

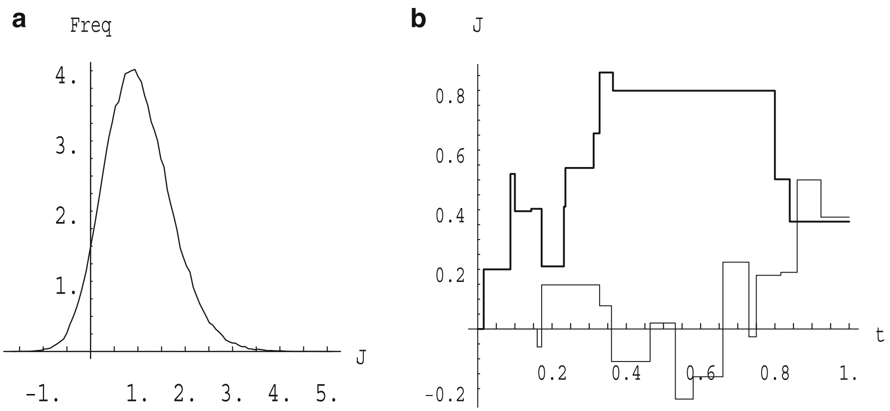


**Fig. 12.3** (**a**) The simulated distribution of the jump-component $J(t)$ at $t = 1$ with L = 250 subdivisions of [0,1]. (**b**) Two simulated sample paths when $y$ is normally distributed

process are of unbounded variation. If we were to limit our attention to jumps of finite size we could define the stochastic integral with respect to jump processes by using the Riemann-Stieltjes integral. However we do want to allow for the situation where $y$ could be drawn from a distribution of unbounded variation (e.g. $y$ could be drawn from a normal or log-normal distribution). So we would still need to appeal to the mean-square limit definition of the stochastic integral. Thus we shall write

$$J(t) = \int_0^t y(s)dN(s), \tag{12.8}$$

and the stochastic integral (12.8) should be understood in the sense of the least squares limit. To illustrate the stochastic nature of the quantity $J(t)$ we display in Fig. 12.3 the distribution of $J(t)$ at $t = 1$ for the case where the jump sizes are drawn

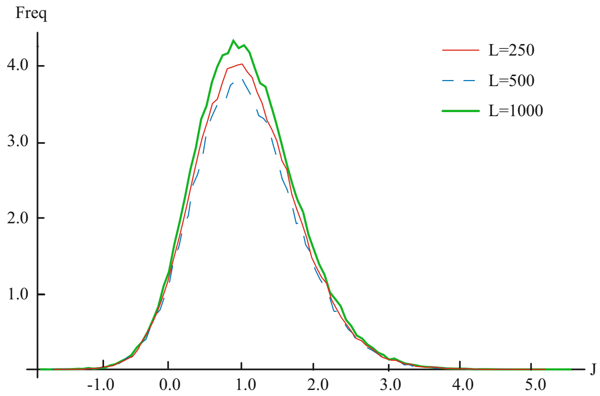


**Fig. 12.4** The simulated distributions of $J(t)$ for varying values of $L$

from a normal distribution.[1] We also display in Fig. 12.3 two simulated sample paths for $J(t)$ when $y$ was drawn from a normal distribution. The distribution for $J(t)$ has been generated using a particular subdivision of the interval $(0, t)$. To reinforce the notion that the stochastic integral (12.8) is the limit of such distributions, we show in Fig. 12.4 how the distribution evolves as the number of subdivisions $L$ takes the values 250, 500 and 1,000. Here we again focus on the case where the jump sizes are normally distributed as in Fig. 12.3.

To perform the manipulations of stochastic calculus for jump-diffusion processes we need to be able to calculate the mean and the variance of the quantity $J(t)$. Thus

$$\mathbb{E}[J(t)] = \mathbb{E}^{\mathbb{Q}_y}\mathbb{E}^{\mathbb{Q}_j}\sum_{i=1}^{N(t)} y_i. \tag{12.9}$$

In performing the calculation in (12.9) we note that for all $i$, $\mathbb{E}^{\mathbb{Q}_y}[y_i] = k$. We also note that the jump times are independent and that conditional on the $(i-1)^{\mathrm{st}}$ jump

$$\mathbb{E}^{\mathbb{Q}_j}[y_i] = \left(1 - \int_{t_{i-1}}^{t_i} \lambda(s)ds\right) \cdot 0 + \left(\int_{t_{i-1}}^{t_i} \lambda(s)ds\right) \cdot y_i = \left(\int_{t_{i-1}}^{t_i} \lambda(s)ds\right) \cdot y_i.$$

[1]The distribution has been obtained by simulating 100,000 paths of the Poisson process up to $t = 1$. The normal distribution has mean of 0.1 and standard deviation of 0.2. The value of $\lambda$ used was 10 and $\Delta t = 0.004$ (i.e. 250 subdivisions of the time interval).

Then taking the expectation with respect to $\mathbb{Q}_y$ we have, conditional on the $(i-1)^{\text{st}}$ jump

$$\mathbb{E}[y_i] = k \int_{t_{i-1}}^{t_i} \lambda(s) ds \cdot .$$

Summing all such contributions we see that

$$\mathbb{E}[J(t)] = \mathbb{E}\left[\int_0^t y(s) dN(s)\right] = k \int_0^t \lambda(s) ds.$$

It follows that the quantity

$$\tilde{J}(t) = \int_0^t \Big( y(s) dN(s) - k\lambda(s) ds \Big)$$

is a martingale. The quantity $\int_0^t k\lambda(s)ds$ is known as the compensator. Thus Eq. (12.7) should be more correctly written as

$$x(t) = x(0) + \int_0^t \mu(x,s) ds + \int_0^t \sigma(x,s) dw(s) + \int_0^t \left(y(s) dN(s) - k\lambda(s) ds\right),$$

and may be more simply interpreted as

$$x(t) = x(0) + \int_0^t \mu(x,s) ds + \int_0^t \sigma(x,s) dw(s) + (\sum_{i=1}^{N(t)} y_i - k \int_0^t \lambda(s) ds).$$

Alternatively the compensator may be incorporated into the drift term so that Eq. (12.4) should be interpreted as

$$x(t) = x(0) + \int_0^t \left(\mu(x,s) - k\lambda(s)\right) ds + \int_0^t \sigma(x,s) dw(s) + \int_0^t y(s) dN(s),$$

or more simply as

$$x(t) = x(0) + \int_0^t \left(\mu(x,s) - k\lambda(s)\right) ds + \int_0^t \sigma(x,s) dw(s) + \sum_{i=1}^{N(t)} y_i .$$

### 12.2.2 Proportional Jumps

Now consider the special case when the underlying diffusion process (12.1) is expressed in return form i.e.

$$\frac{dx}{x} = \mu dt + \sigma dw \tag{12.10}$$

and $\mu$ and $\sigma$ may be functions of $x$ and $t$. Note also that in this subsection we assume $\mu$, $\sigma$ and $\lambda$ are constants in order to obtain the particular distributional results derived below.

If there is the arrival of just one jump event at time $t_i$ then now it is assumed that

$$x(t_i^+) = Y_i x(t_i^-),$$

or alternatively,

$$x(t_i^+) - x(t_i^-) = (Y_i - 1)x(t_i^-),$$

where $Y_i$ is drawn at time $t_i$ from a jump-size probability distribution of $Y$ with measure $\mathbb{Q}_Y$ and the set of $Y$ from successive events are assumed to be independently and identically distributed. In between the arrival of the price jump events, the asset price follows the continuous diffusion process (12.10). The asset price dynamics may be written

$$\frac{dx}{x} = (\mu - \lambda k)dt + \sigma dw + (Y - 1)dN, \tag{12.11}$$

where

$$k = \mathbb{E}^{\mathbb{Q}_Y}(Y - 1).$$

Thus $(Y - 1)$ is, as we have seen, the (random) percentage change in the asset price if the Poisson jump event occurs and $\mathbb{E}^{\mathbb{Q}_Y}$ is the expectation operator taken over the probability distribution of the random variable $Y$. If we denote this distribution by $G(Y)$ then

$$k = \int (Y - 1)G(Y)dY = \int YG(Y)dY - 1.$$

The stochastic integral equation (12.11) may be written

$$x(t) = x(0) + \int_0^t (\mu - \lambda k)x(s)ds + \int_0^t \sigma x(s)dw(s) + \sum_{i=1}^{N(t)} (Y_i - 1)x(t_i^-). \tag{12.12}$$

Equation (12.11) can be interpreted in long hand form as

$$\frac{dx}{x} = \begin{cases} (\mu - \lambda k)dt + \sigma dw, & \text{with prob.} \quad (1 - \lambda dt), \\ (\mu - \lambda k)dt + \sigma dw + (Y - 1), & \text{with prob.} \quad \lambda dt. \end{cases}$$

The sample path $x$ will be continuous most of the time with finite jumps of differing signs and magnitudes occurring at discrete points in time. Under the assumption that the parameters $\mu$, $\lambda$, $k$ and $\sigma$ are constant then the random variable $x(t)$, conditional on $x(0)$, can be written

$$\frac{x(t)}{x(0)} = \exp\left[\left(\mu - \frac{\sigma^2}{2} - \lambda k\right) t + \sigma(w(t) - w(0))\right] X_N, \tag{12.13}$$

where $w(t)$ is a standard Wiener process and the $X$ process is defined by[2]

$$X_N = 1, \qquad \text{if} \qquad N(t) = 0,$$

and

$$X_N = \prod_{i=1}^{N(t)} Y_i, \qquad \text{if} \quad N(t) \geq 1.$$

The $Y_i$ are assumed to independently and identically distributed, and $N(t)$, the number of jumps in $(0, t)$, is again drawn from the Poisson distribution with parameter $\lambda t$.

Equation (12.13) is easily derived by noting that, illustrated in Fig. 12.5, between two jump times $t_{i-1}^{+}$ and $t_i^{-}$ the path for $x$ is driven by the diffusion process

$$\frac{dx}{x} = (\mu - \lambda k)dt + \sigma dw,$$

which may also be written

$$d(\ln x) = (\mu - \lambda k - \frac{1}{2}\sigma^2)dt + \sigma dw,$$

whose solution for $t_{i-1}^{+} \leq t \leq t_i^{-}$ can be written (see Eq. (6.16))

$$x(t) = x(t_{i-1}^{+}) \exp[(\mu - \frac{\sigma^2}{2} - \lambda k)(t - t_{i-1}) + \sigma(w(t) - w(t_{i-1}))].$$

[2]Note that for notation simplicity we write $X_N$ instead of the more strictly correct $X_{N(t)}$.

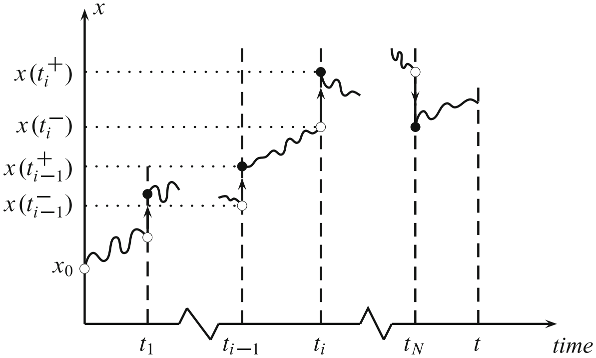


**Fig. 12.5** Diffusion movements between a sequence of jump times

Thus

$$x(t_i^-) = x(t_{i-1}^+)\exp[(\mu - \frac{\sigma^2}{2} - \lambda k)(t_i - t_{i-1}) + \sigma(w(t_i) - w(t_{i-1}))].$$

After the jump at $t_i$ we have that

$$x(t_i^+) = Y_i x(t_i^-),$$

so that

$$x(t_i^+) = x(t_{i-1}^+)\exp[(\mu - \frac{\sigma^2}{2} - \lambda k)(t_i - t_{i-1}) + \sigma(w(t_i) - w(t_{i-1}))]Y_i. \quad (12.14)$$

Applying (12.14) successively from $t = 0$ to a time $t$ between jump times $t_n$ and $t_{n+1}$ we obtain (12.13).

If we make the further assumption that

$$\frac{dY}{Y} = \gamma dt + \delta dW_Y,$$

where $\gamma$ and $\delta$ are constants and $W_Y$ is an independent Wiener process. Then $Y_i$ are log-normally distributed, hence $x(t)/x(0)$ will be log-normally distributed. More specifically, if[3]

$$\ln Y_i \sim \phi(\gamma - \frac{1}{2}\delta^2, \delta^2), \quad (12.15)$$

[3]For later use note that the density function for $Y$ which we denote $G(Y)$, is given by

$$G(Y)dY = \frac{1}{\sqrt{2\pi}\delta}\exp\left[-\frac{1}{2}\left(\frac{\ln Y - (\gamma - \delta^2/2)}{\delta}\right)^2\right]\frac{dY}{Y}$$

which follows from (6.23).

then the distribution of $\ln[x(t)/x(0)]$ is

$$\ln\left(\frac{x(t)}{x(0)}\right) \sim \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}\phi\left((\mu-\lambda k-\sigma^2/2)t+n\left(\gamma-\frac{\delta^2}{2}\right),\quad \sigma^2 t+n\delta^2\right). \tag{12.16}$$

This last expression tells us that, the log return distribution is stationary over time and is described as a Poisson mixture of normal distributions.

To derive (12.16) we note that by taking logs Eq. (12.13) becomes

$$\ln\left(\frac{x(t)}{x(0)}\right) = \left(\mu-\frac{\sigma^2}{2}-\lambda k\right)t+\sigma(w(t)-w(0))+\sum_{i=1}^{N(t)}\ln Y_i.$$

Since $w(t)-w(0)$ and each of the $\ln Y_i$ are normally distributed it follows that $\ln(x(t)/x(0))$ is normally distributed. Furthermore, conditional on $n$ jumps having occurred up to time $t$, $\ln(x(t)/x(0))$ has mean and variance given by

$$\mathbb{E}_0\left[\ln\left(\frac{x(t)}{x(0)}\right)\right] = \left(\mu-\lambda k-\frac{\sigma^2}{2}\right)t+n\left(\gamma-\frac{\delta^2}{2}\right),$$

and

$$\text{var}_0\left[\ln\left(\frac{x(t)}{x(0)}\right)\right] = \sigma^2 t+n\delta^2.$$

Equation (12.16) then follows by using the result (12.2) and summing over all possible numbers of jumps.

In Fig. 12.6 we illustrate (on the left hand side) some typical sample paths when the $Y_i$ are log-normally distributed.[4] For the same simulated path of the Wiener process we compare the $\lambda = 0$ case (the pure diffusion case) with the case $\lambda = 2$. On the right hand side we display the outcome of the $(Y-1)xdq$ process on the simulated path. To appreciate the significance of the size of $\lambda$ (the jump arrival rate) note that a value of $\lambda = 2$ corresponds to two jumps per year on average.

It can be shown that the distribution (12.16) is leptokurtic (i.e. has fat tails) and therefore, captures an important feature of asset price movements that is not well captured by the simple geometric Brownian motion process. In Fig. 12.7 we show the effect on the distributions for $\ln(x(t)/x(0))$ and $x(t)/x(0)$ as $\lambda$ increases from $\lambda = 0$. The parameter values used are displayed in Table 12.1, and have been chosen

[4] For this simulation we took $\mu = 0.15$, $\sigma = 0.20$, $\gamma = 0.05$, $\delta = 0.02$. The value of $k(= e^{\gamma}-1)$ was 0.05. The value of $\Delta t$ was 0.01.

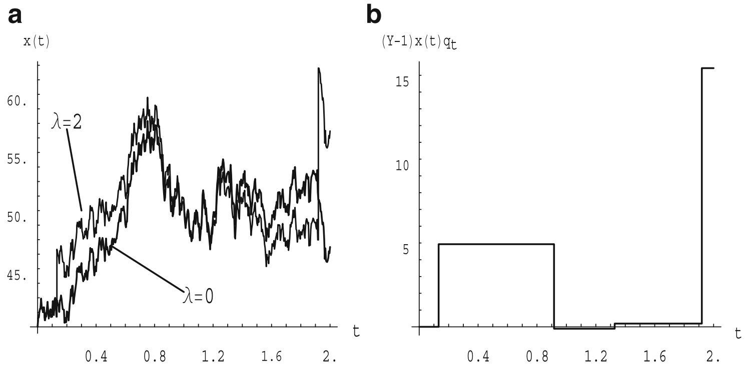


**Fig. 12.6** Simulating the jump diffusion process; (**a**) comparing the pure diffusion process ($\lambda = 0$) with a jump-diffusion process ($\lambda = 2$), (**b**) the pure jump component of the simulated ($\lambda = 2$) path in (**a**)

to maintain $\text{var}_0[x_t] = 0.2318$, where $\text{var}_0[x_t]$ denotes the overall variance[5] (as opposed to the variance conditional on $n$ jumps). The fat tails are clearly evident as is the skewness and central peakedness as the jump intensity increases. These are both features observed in financial market data.

Finally we note that as a stochastic integral equation, Eq. (12.11) may be written

$$x(t) = x(0) + \int_0^t (\mu - \lambda k)\, x(s) ds + \int_0^t \sigma x(s) dw(s) + \int_0^t (Y(s) - 1)\, x(s) dN(s),$$

or more simply as

$$x(t) = x(0) + \left( \int_0^t (\mu - \lambda k)\, x(s) ds \right) + \int_0^t \sigma x(s) dw(s) + \sum_{i=1}^{N(t)} (Y_i - 1) x(t_i). \tag{12.17}$$

Written in this form the process for $x$ is rather difficult to analyse. As we have seen in this section it is simpler to work with the stochastic integral equation behind the solution (12.13) which is

$$\ln x(t) = \ln x(0) + \int_0^t \left( \mu - \frac{\sigma^2}{2} - \lambda k \right) ds + \int_0^t \sigma dw(s) + \sum_{i=1}^{N(t)} \ln Y_i. \tag{12.18}$$

[5]Note that for the geometric jump-diffusion process (12.12)

$$\text{var}_t[x_T] = x_t^2 e^{2\mu(T-t)} \left[ \exp((\sigma^2 - \lambda(2e^{\gamma} - e^{2\gamma+\delta^2} - 1))(T-t)) - 1 \right].$$

See Problem 12.5.

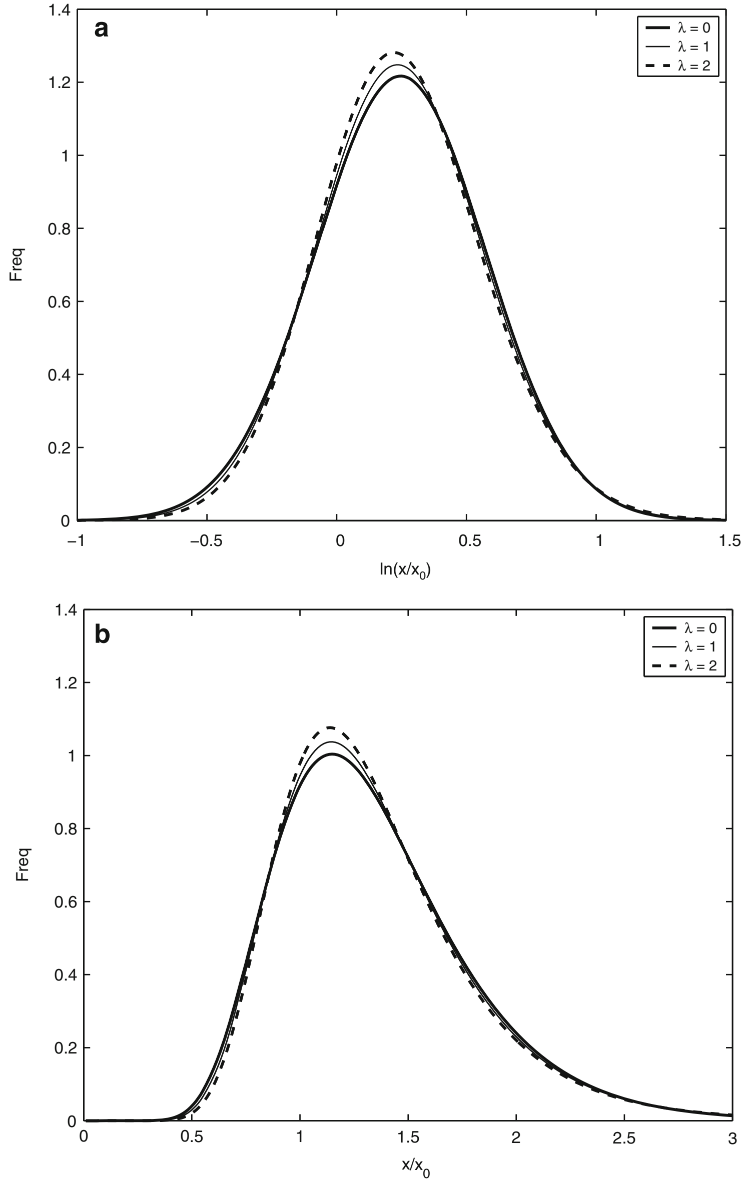


**Fig. 12.7** Effect of increasing values of $\lambda$ on the distribution at $t = 2$; (**a**) for the returns $\ln(x(t)/x(0))$, (**b**) for the relative price $x(t)/x(0)$

In the next section we will show how (12.18) can be formally derived from (12.17) an application of Ito's lemma for jumps.

**Table 12.1** Parameter values used to generate the distributions in Fig. 12.7

| $\mu$ | $\sigma$ | $\lambda$ | $\gamma$ | $\delta$ |
|---|---|---|---|---|
| 0.15 | 0.2318 | 0 | – | – |
| 0.15 | 0.2 | 1 | 0.05 | 0.10 |
| 0.15 | 0.1621 | 2 | 0.05 | 0.10 |

These have been selected so that for all distributions $\text{var}_0[x_t] = 0.2318$

### 12.2.3 A General Process of Dependent Jump Size

Both the absolute and proportional jump sizes of the previous two subsections can be generalised to allow the case when the jump size is some function of the process $x$. In particular we shall assume that at a jump time $t_i$

$$x(t_i^+) = x(t_i^-) + y_i h(x(t_i^-)), \tag{12.19}$$

where $h$ is some well behaved function of the process $x$. In the case of absolute jumps of Sect. 12.2.1 we have $h(x) = 1$. In the case of proportional jumps of Sect. 12.2.2 we set $h(x) = x$. If in addition we impose the restriction that at a jump time $x(t_i^+)$ cannot become negative then we would also require that the support for the distribution of $y$ be $(-1, \infty)$. In Sect. 12.2.2 this was indeed the case since there the support for $Y$ was $(0, \infty)$ and, $y$ and $Y$ are here related by $y = Y - 1$.

From (12.19) we calculate that

$$\text{The mean jump size at time } t_i = \mathbb{E}^{\mathbb{Q}_y}[yh(x(t_i^-))] = h(x(t_i^-))k,$$

where we still use $k$ to denote $\mathbb{E}^{\mathbb{Q}_y}[y]$. Now in this more general situation (12.7) becomes

$$dx = \mu(x,t)dt + \sigma(x,t)dw + (yh(x)dN - \lambda kh(x)dt), \tag{12.20}$$

which may be interpreted as the stochastic integral equation

$$x(t) = x(0) + \int_0^t \mu(x,s)ds + \int_0^t \sigma(x,s)dw(s) + \int_0^t \Big(yh(x(s))dN(s) - \lambda kh(x(s))ds\Big), \tag{12.21}$$

or more simply as

$$x(t) = x(0) + \int_0^t \mu(x,s)ds + \int_0^t \sigma(x,s)dw(s) + \sum_{i=1}^{N(t)} \left(y_i - k\int_0^t \lambda(s)ds\right) h(x(t_i)). \tag{12.22}$$

Equations (12.21) and (12.22) are easily re-arranged to give the forms where the compensator is included in the drift term.

## 12.3 Ito's Lemma for Jump-Diffusion Processes

Suppose $v$ depends on $x$ and $t$, i.e.

$$v = v(x, t),$$

where $x$ is driven by the jump-diffusion stochastic process (12.20). We seek the stochastic process driving $v$. The derivation we sketch out below is based loosely on Kushner (1967) and Gihman and Skorohod (1972).

We first observe that in between the Poisson jump events $x$ follows the continuous diffusion given by the upper part of (12.5). Thus at times $t \neq t_i$ (a typical jump time), we calculate, the change in $v$ using the same approach as in Sect. 6.2.2. Thus

$$\begin{aligned} dv &= \frac{\partial v}{\partial t}dt + \frac{\partial v}{\partial x}dx + \frac{1}{2}\frac{\partial^2 v}{\partial x^2}(dx)^2 \\ &= \left(\frac{\partial v}{\partial t} + (\mu - \lambda k h(x))\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2}\right)dt + \sigma\frac{\partial v}{\partial x}dw. \end{aligned} \tag{12.23}$$

The change in $v$ brought about by the jump at a typical jump time, $t = t_i$, is given by (see Fig. 12.8)

$$v(t_i^+) - v(t_i^-) = v(x + y_i h(x), t_i) - v(x, t_i).$$

At any time $t$, for the change in $v$ brought about by the Poisson jump process we may write

$$v(t^+) - v(t^-) = [v(x + yh(x), t) - v(x, t)]dN(t). \tag{12.24}$$

Figure 12.1 illustrates the change in $v$ at a typical jump time. It is convenient to define the expected change in $v$ over the jump distribution, namely

$$k_v = \mathbb{E}^{Q_y}[v(x + yh(x), t) - v(x, t)] = \int [v(x + yh(x), t) - v(x, t)]g(y)dy.$$

The changes in $v$ in (12.23) and (12.24) occur with probabilities $(1 - \lambda dt)$ and $\lambda dt$ respectively, thus calculating the expected change in $v$ with respect to both the jump time distribution and the jump size distribution we have

$$\begin{aligned} \mathbb{E}[dv] = \mathbb{E}^{\mathbb{Q}_y}\mathbb{E}^{\mathbb{Q}_j}\mathbb{E}^{\mathbb{P}}[dv] &= \left(\frac{\partial v}{\partial t} + (\mu - \lambda k h(x))\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2}\right)dt(1 - \lambda dt) + k_v\lambda dt \\ &= \left(\frac{\partial v}{\partial t} + (\mu - \lambda k h(x))\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2} + \lambda k_v\right)dt. \end{aligned}$$

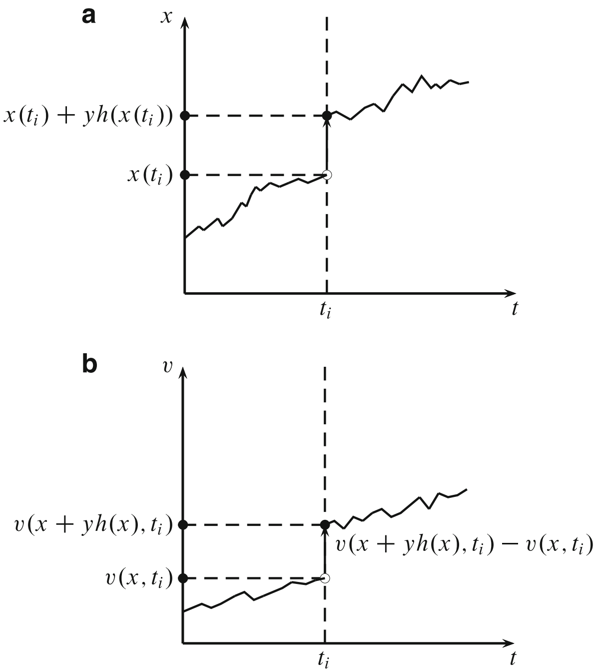


**Fig. 12.8** (**a**) The change in $x$ and (**b**) change in $v$, at a typical jump time

Thus we see that if we define

$$\mu_v = \frac{\partial v}{\partial t} + (\mu - \lambda k h(x))\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2} + \lambda k_v, \tag{12.25}$$

and

$$\sigma_v = \sigma\frac{\partial v}{\partial x}.$$

then

$$\mathbb{E}[dv] = \mu_v dt,$$

in analogy with (12.6). Furthermore using the same logic as that behind the representation (12.5) we can write

$$dv = \begin{cases} (\mu_v - \lambda k_v)dt + \sigma_v dw, & \textit{with prob.}\ (1\text{-}\lambda dt), \\ (\mu_v - \lambda k_v)dt + \sigma_v dw + [v(x + yh(x), t) - v(x,t)], & \textit{with prob.}\quad \lambda dt, \end{cases}$$

which in analogy with (12.4) may be written

$$dv = (\mu_v - \lambda k_v)dt + \sigma_v dw + [v(x + yh(x), t) - v(x, t)]dN. \tag{12.26}$$

We may express (12.26) in a martingale representation by rewriting it in the form (compare with Eq. (12.20))

$$dv = \mu_v dt + \sigma_v dw + ([v(x + yh(x), t) - v(x, t)]dN - \lambda k_v dt). \tag{12.27}$$

Equation (12.26) (or alternatively (12.27)) is the expression of Ito's lemma for jump-diffusion processes.

A simple (but important for future applications) example of the use of Ito's Lemma is to derive the jump-diffusion stochastic differential equation followed by $\ln x$ where $x$ is the geometric jump-diffusion process (12.11). This extends to the jump-diffusion setting the example in Sect. 6.3.2. We set

$$u = \ln x$$

to denote the log of the process $x$. Hence the function $h(x)$ in (12.19) is simply $x$, so the jump in $x$ can be denoted $x + yx$, or $x + (Y - 1)x = Yx$ which is the form we shall use. We calculate

$$\frac{\partial u}{\partial t} = 0, \quad \frac{\partial u}{\partial x} = \frac{1}{x}, \quad \frac{\partial^2 u}{\partial x^2} = -\frac{1}{x^2}.$$

The quantity of (12.25) $\mu_v$ in the current context becomes

$$\mu_v = (\mu x - \lambda k x) \cdot \frac{1}{x} + \frac{1}{2}\sigma^2 x^2 \left(-\frac{1}{x^2}\right) + \lambda k_v = \mu - \lambda k - \frac{1}{2}\sigma^2 + \lambda k_v,$$

where

$$\begin{aligned} k_v &= \int [\ln((1 + y)x) - \ln x]\, g(y)dy = \int [\ln(Yx) - \ln x]\, G(Y)dY \\ &= \int \ln(Y)G(Y)d(Y), \end{aligned}$$

and

$$\sigma_v = \sigma x \cdot \frac{1}{x} = \sigma.$$

Thus applying Eq. (12.26) the stochastic differential equation for $u$ or ($\ln x$) becomes

$$d(\ln x) = (\mu - \frac{\sigma^2}{2} - \lambda k)dt + \sigma dw + \ln Y dN. \tag{12.28}$$

Integrating from 0 to $t$ we obtain

$$\ln x(t) = \ln x(0) + \int_0^t (\mu - \frac{\sigma^2}{2} - \lambda k)ds + \int_0^t \sigma dw(s) + \int_0^t \ln Y(s)dN(s). \tag{12.29}$$

Bearing in mind that

$$\int_0^t \ln Y(s)dN(s) = \sum_{i=1}^{N(t)} \ln Y_i,$$

where $N(t)$ is the number of jumps that have occurred up to time $t$, we see that Eq. (12.29) is equivalent to Eq. (12.18).

## 12.4 Appendix

### Appendix 12.1 Kolmogorov Equation and Feynman–Kac Formula for Processes with Jumps

In this appendix we merely summarise the key results concerning the Kolmogorov equation and Feynman–Kac formula for jump-diffusion (sometimes called Wiener-Poisson) stochastic differential equations. We refer the reader to Gihman and Skorohod (1972) for details of proofs.

Consider the jump-diffusion stochastic differential equation (see Eq. (12.7))

$$dx = \mu(x,t)dt + \sigma(x,t)dw + (ydN - \lambda k dt). \tag{12.30}$$

The partial differential operator for the transition probability density function $p(x_T, T|x,t)(T \geq t)$ associated with (12.30) (the so-called infinitesimal generator) is given by

$$\mathscr{K} p = \mu(x,t)\frac{\partial p}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 p}{\partial x^2} + \int \left[ p(x+y) - p(x) - \lambda y \frac{\partial p}{\partial x} \right] g(y)dy,$$

which can also be written

$$\mathscr{K} p = (\mu(x,t) - \lambda k)\frac{\partial p}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 p}{\partial x^2} + \int \left[ p(x+y) - p(x) \right] g(y)dy.$$

The Kolmogorov equation for $p$ is

$$\frac{\partial p}{\partial t} + \mathscr{K} p = 0,$$

subject to the usual boundary condition

$$\lim_{t \to T} p(x_T, T|x, t) = \delta(x - x_T).$$

### *The Feynman–Kac Formula*

Set

$$v(t, x) = \mathbb{E}_{t,x}\left[ f(x_T) \exp\left\{ \int_t^T \gamma g(s, x_s) ds \right\} \right]$$

where $f$ and $g$ are sufficiently well-behaved functions, $\gamma$ is a constant, and $x$ is generated by the jump-diffusion process (12.30). Then the Feynman–Kac formula in this situation states that $v(t, x)$ satisfies

$$\frac{\partial v}{\partial t} + Av + \gamma g(t, x) v(t, x) = 0$$

subject to the boundary condition

$$\lim_{t \to T} v(t, x) = f(x_T).$$

Note that for the applications in Chap. 13, $\gamma = 0$.

## 12.5 Problems

**Problem 12.1**

(a) Show that

$$\mathbb{E}_0\left[ \int_0^t f(s) dN^2(s) \right] = \int_0^t f(s) \lambda(s) ds,$$

for a suitably well behaved function $f$.

(b) Hence, evaluate

$$\mathbb{E}_0\left[ \int_0^t \beta_i(u) dN_i(u) \int_0^t \beta_j(u) dN_j(u) \right]$$

for well-behaved functions $\beta_i(t), \beta_j(t)$ and independent jumps $N_i, N_j$.

**Problem 12.2** Consider the independent Poisson processes $N_i(t)$ and $N_j(t)$ with intensities $\lambda_i(t)$ and $\lambda_j(t)$ respectively. For well-behaved functions $\beta_i(t)$ and $\beta_j(t)$, prove that

$$\mathbb{E}_0\left[\int_0^t \beta_i(u)[dN_i(u)-\lambda_i(u)du]\int_0^t \beta_j(u)[dN_j(u)-\lambda_j(u)du]\right] = \begin{cases} \int_0^t \lambda_i(s)\beta_i^2(s)ds, & i=j, \\ 0, & i\neq j. \end{cases}$$

**Problem 12.3** In the case when $\mu$, $\sigma$ and $\lambda$ are time dependent show that Eq. (12.13) becomes

$$\frac{x(t)}{x(0)} = \exp\left[\int_0^t \left(\mu(s) - \frac{\sigma(s)^2}{2} - \lambda(s)k\right)ds + \int_0^t \sigma(s)dw(s)\right]X_N.$$

**Problem 12.4** Consider the mean-reverting jump-diffusion equation

$$dx = \beta(\bar{x} - \lambda k/\beta - x)dt + \sigma dw + ydN.$$

Use the same idea as was used to derive Eq. (12.13) (that is, solve the underlying diffusion equation between jump times) to show, for a path along which $n$ jumps have occurred up to time $t$, that

$$x(t) = x(0)e^{-\beta t} + \int_0^t \beta(\bar{x}-\lambda k/\beta)e^{-\beta(t-s)}ds + \int_0^t \sigma e^{-\beta(t-s)}dw(s) + \sum_{i=1}^{n} e^{-\beta(t_n-t_i)}y_i.$$

**Problem 12.5** For the geometric jump-diffusion process (12.13), show that

$$\mathbb{E}_t[x_T] = x_t e^{\mu(T-t)},$$

and

$$\mathrm{var}_t[x_T] = x_t^2 e^{2\mu(T-t)}\left[\exp\left\{(\sigma^2 - \lambda(2e^{\gamma} - e^{2\gamma+\delta^2} - 1))(T-t)\right\} - 1\right].$$

# Chapter 13
# Option Pricing Under Jump-Diffusion Processes

**Abstract** This chapter extends the hedging argument of option pricing developed for continuous diffusion processes previously to the situations when the underlying asset price is driven by the jump-diffusion stochastic differential equations. By constructing hedging portfolios and employing the capital asset pricing model, we provide an option pricing integro-partial differential equations and a general solution. We also examine alternative ways to construct the hedging portfolio and to price option when the jump sizes are fixed.

## 13.1 Introduction

Now let us turn to the problem of developing the hedging argument under the assumption that the underlying asset price $x$ is driven by the jump-diffusion stochastic differential equation (12.11). To develop a hedging argument we need to know the dynamics of the option price. If the option price $f$ is given by

$$f = f(x,t),$$

then application of the results (12.26) implies that

$$\frac{df}{f} = (\mu_f - \lambda k_f)dt + \sigma_f dw + (Y_f - 1)dN,$$

where[1]

$$\begin{aligned} f\mu_f &= \theta + (\mu - \lambda k)x\Delta + \frac{1}{2}\sigma^2 x^2 \Gamma + \lambda f k_f, \\ f\sigma_f &= \sigma x \Delta, \\ f k_f &= \mathbb{E}^{Q_Y}[f(xY,t) - f(x,t)] = \int [f(xY,t) - f(x,t)]G(Y)dY, \end{aligned}$$

[1] We recall the definitions $\theta = \frac{\partial f}{\partial t}$, $\Delta = \frac{\partial f}{\partial x}$, $\Gamma = \frac{\partial^2 f}{\partial x^2}$.



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_13

and

$$Y_f - 1 \equiv (f(xY,t) - f(x,t))/f(x,t)$$

is the random variable percentage change in the option price. If the Poisson event for the asset occurs and the proportional jump size takes on the value $Y$, then the Poisson event for the option occurs and the proportional jump size in the option value is given by

$$Y_f = \frac{f(xY,t)}{f(x,t)},$$

which is a nonlinear relationship connecting the random variables $Y_f$ and $Y$.

## 13.2 Constructing a Hedging Portfolio

Consider a portfolio which contains the asset, the option on the asset and the riskless asset with return $r$ per unit time in the proportions $\pi_x$, $\pi_f$, and $\pi_r$, so that

$$\pi_x + \pi_f + \pi_r = 1.$$

If $V$ is the value of the portfolio then the return dynamics of the portfolio are given by

$$\begin{aligned}\frac{dV}{V} &= \pi_x \frac{dx}{x} + \pi_f \frac{df}{f} + \pi_r dr \\ &= \pi_x[(\mu - \lambda k)dt + \sigma dw + (Y-1)dN] \\ &\quad + \pi_f[(\mu_f - \lambda k_f)dt + \sigma_f dw + (Y_f - 1)dN] + \pi_r r dt.\end{aligned}$$

Collecting terms and using $\pi_r = 1 - \pi_x - \pi_f$ we obtain

$$\frac{dV}{V} = (\mu_V - \lambda k_V)dt + \sigma_V dw + (Y_V - 1)dN, \tag{13.1}$$

where

$$\begin{aligned}\mu_{\mathrm{V}} &= \pi_x(\mu - r) + \pi_f(\mu_f - r) + r, \\ \sigma_{\mathrm{V}} &= \pi_x\sigma + \pi_f\sigma_f, \\ Y_{\mathrm{V}} - 1 &= \pi_x(Y-1) + \pi_f[f(xY,t) - f(x,t)]/f(x,t), \\ k_{\mathrm{V}} &= \mathbb{E}^{Q_Y}[Y_V - 1].\end{aligned} \tag{13.2}$$

Here $(Y_V - 1)$ is the random variable percentage change in the portfolio's value if the Poisson jump event occurs.

When the asset price follows a diffusion process the hedging portfolio is rendered riskless by choosing the portfolio proportions $\pi_x$, $\pi_f$ such that

$$\pi_x \sigma + \pi_f \sigma_f = 0. \tag{13.3}$$

However, this choice of portfolio weights in the case of a jump-diffusion process, while eliminating the $\sigma_V$ term will not eliminate the jump risk (i.e. the $Y_V - 1$ term). In fact, there is no choice of $\pi_x$ and $\pi_f$ which eliminates the jump risk term (i.e. makes $Y_V = 1$).

Let us nevertheless determine the return characteristics of the portfolio when the Black–Scholes hedge is followed. Letting $\pi_x^*$ and $\pi_f^*$ denote the values of $\pi_x$, $\pi_f$ satisfying (13.3) and $V^*$ the corresponding portfolio value we have from (13.1)

$$\frac{dV^*}{V^*} = (\mu_V^* - \lambda k_V^*)dt + (Y_V^* - 1)dN. \tag{13.4}$$

The portfolio return has thus been reduced to a pure jump process, and could also be written

$$\frac{dV^*}{V^*} = \begin{cases} (\mu_V^* - \lambda k_V^*)dt, & \text{if the Poisson jump event does not occur,} \\ (\mu_V^* - \lambda k_V^*)dt + (Y_V^* - 1) & \text{if the Poisson jump event occurs.} \end{cases} \tag{13.5}$$

Equation (13.5) tells us that most of the time the portfolio return will be predictable and earn $(\mu_V^* - \lambda k_V^*)$. However every $(1/\lambda)$ units of time, on average, the portfolio return takes an unexpected jump.

It is possible to say something about the qualitative characteristics of the portfolio return. Note first of all that

$$Y_V^* - 1 = \pi_f^* \frac{f(xY, t) - f(x, t) - f_x(x, t)(xY - x)}{f(x, t)}.$$

Since the option price is a strictly convex function of the asset price it follows that

$$\frac{f(xY, t) - f(x, t)}{xY - x} > f_x(x, t),$$

for $Y > 1$, and

$$\frac{f(xY, t) - f(x, t)}{xY - x} < f_x(x, t),$$

for $Y < 1$. Thus for all values of $Y$, it follows that

$$f(xY, t) - f(x, t) - f_x(x, t)(xY - x) > 0.$$

Hence

$$\mathrm{sign}(Y_V^* - 1) = \mathrm{sign}(\pi_f^*).$$

Suppose an investor is long the stock and short the option (i.e. $\pi_f^* < 0$) then most of the time he or she would earn more than the expected return on the hedge $\mu_V^*$, since $k_V^* < 0$. The investor will however suffer losses when the asset price jumps from time to time. These losses occur at such a frequency so as to, on average, offset the excess return $-\lambda k_V^*$. If we define as a "quiet" period, that period in between the arrival of Poisson jump events, and if we assume that the jump events are related to asset specific information then the above argument shows that during quiet periods writers of options will tend to make what appear to be positive excess returns. Purchasers of options on the other hand would make negative excess returns and therefore appear as "losers". However, at the arrival (relatively infrequently) of Poisson jump events, the options writers will suffer loss and the buyers appear as "winners". Since the arrival of the Poisson events is random, there is no systematic way of exploiting this understanding of the dynamics. The reverse argument applies when the investor is short the asset and long the option (i.e. $\pi_f^* > 0$).

## 13.3 Pricing the Option

The clue to pricing the option in the presence of jump-diffusion processes is the alternative approach used by Black–Scholes employing the Capital Asset Pricing model.

We have already stressed that the Poisson jump events are asset specific. It follows that the jump component of the asset's return represents non-systematic risk. It also follows that, since the only uncertainty in the $V^*$ portfolio of the previous section is the Poisson jump component, then its risk is uncorrelated with the market, i.e. it contains only non-systematic risk. From modern portfolio theory we have the result that portfolios containing only non-systematic risk have a beta factor of zero. Furthermore, if the CAPM describes security returns then the return on a zero beta portfolio must equal the riskless rate. It follows that

$$\mu_V^* = r,$$

or, from (13.2) that

$$\pi_x^*(\mu - r) + \pi_f^*(\mu_f - r) = 0,$$

which when combined with

$$\pi_x^* \sigma + \pi_f^* \sigma_f = 0,$$

yields

$$\frac{\mu - r}{\sigma} = \frac{\mu_f - r}{\sigma_f}. \tag{13.6}$$

After applying the definitions of $\mu_f$ and $\sigma_f$ in the last equation, we obtain the following equation for the option price

$$\frac{\partial f}{\partial t} + (r - \lambda k)x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} - rf + \lambda \mathbb{E}^{Q_Y}[f(xY, t) - f(x, t)] = 0. \tag{13.7}$$

Because of the expectation operator $\mathbb{E}^{Q_Y}$, Eq. (13.7) is an integro-partial differential equation and solution techniques for it require a degree of complexity beyond those for the Black–Scholes partial differential equation.

We may use (13.6) to obtain a martingale representation of the price. Using an argument familiar from Chaps. 8 and 10, if we use $\phi$ to denote the market price of risk associated with the risk factor $dw$ then (13.6) may be interpreted as

$$\mu = r + \phi\sigma,$$
$$\mu_f = r + \phi\sigma_f.$$

Thus in the absence of riskless arbitrage opportunities the stochastic differential equations for $x$ and $f$ may be written

$$\frac{dx}{x} = (r - \lambda k + \phi\sigma)dt + \sigma dw + (Y - 1)dN,$$
$$\frac{df}{f} = (r - \lambda k_f + \phi\sigma_f)dt + \sigma_f dw + (Y_f - 1)dN.$$

Or alternatively as

$$\frac{dx}{x} = rdt + \sigma d\tilde{w} + [(Y - 1)dN - \lambda k dt], \tag{13.8}$$

$$\frac{df}{f} = rdt + \sigma_f d\tilde{w} + [(Y_f - 1)dN - \lambda k_f dt], \tag{13.9}$$

where

$$\tilde{w}(t) = w(t) + \int_0^t \phi(s)ds.$$

Under the original measure $\mathbb{P}$, $\tilde{w}$ will not be a standard Wiener process, but application of Girsanov's theorem for processes involving jumps (see Bremaud 1981) allows us to assert that it is possible to obtain an equivalent measure $\tilde{\mathbb{P}}$ under which $\tilde{w}$ is a standard Wiener process and $N$ remains a jump process with jump intensity $\lambda$.

We note that (13.9) may be written

$$d(fe^{-rt}) = e^{-rt}\sigma_f f d\tilde{w} + e^{-rt} f[(Y_f - 1)dN - \lambda k_f dt]$$

so that under $\tilde{\mathbb{P}}$ the quantity $fe^{-rt}$, the option price measured in units of the money market account $e^{rt}$, is a martingale, i.e.

$$f(x,t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[f(x_T, T)],$$

where $\tilde{\mathbb{E}}_t$ is the expectation operator under $\tilde{\mathbb{P}}$.

We note that one way to calculate $\tilde{\mathbb{E}}_t$ would be to simulate the jump-diffusion process (13.8) for $x$. Application of the Feynman–Kac formula for jump-diffusion processes (see Appendix 12.1) would yield the integro-partial differential equation (13.7). Thus we have established the link between the martingale viewpoint and the integro-partial differential equation viewpoint.

## 13.4 General Form of the Solution

Recall that in Eq. (13.7), $t$ is the current time. If we switch the time variable to $\tau = T - t =$ time-to-maturity, then Eq. (13.7) becomes

$$-\frac{\partial f}{\partial \tau} + (r - \lambda k)x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} - rf + \lambda \mathbb{E}^{Q_Y}[f(xY,\tau) - f(x,\tau)] = 0. \tag{13.10}$$

To fully appreciate the nature of the pricing equation (13.10), recall that $G(Y)$ is the probability density function for the random variable $Y$ then (13.10) may be written

$$\begin{aligned} -\frac{\partial f}{\partial \tau} + (r - \lambda k)x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} - rf \\ + \lambda \int_{-\infty}^{\infty} [f(xY,\tau) - f(x,\tau)]G(Y)dY = 0, \end{aligned} \tag{13.11}$$

where

$$k = \int_{-\infty}^{\infty} YG(Y)dY - 1.$$

This type of equation may be classed as a mixed integro-partial differential equation. Whilst the solution of such equations is in general quite difficult, it turns out that the general form of the solution may be expressed in a convenient form even before we specify the density function $G(Y)$.

In the situation when the underlying asset is common stock equation (13.10) must be solved subject to the boundary condition

$$f(0,\tau) = 0, \tag{13.12}$$

and the initial condition

$$f(x,0) = \max[0, x - E], \tag{13.13}$$

where $E$ is the exercise price of the option. Let $M(x,\tau; E, \sigma^2, r)$ denote the solution to (13.10) in the absence of the jump component, i.e. when $\lambda = 0$. Thus $M$ would be the Black–Scholes solution given by

$$M(x,\tau; E,\sigma^2, r) = x\mathcal{N}(d_1) - Ee^{-r\tau}\mathcal{N}(d_2), \tag{13.14}$$

where

$$d_1 = \frac{\ln(x/E) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \qquad d_2 = d_1 - \sigma\sqrt{\tau}.$$

Define the random variable $X_n = \prod_{i=1}^{n} Y_i$ as one having the same distribution as the product of $n$ independently identically distributed random variables, each identically distributed as the random variable price change $Y$. It is assumed $X_0 = 1$. Define $\mathbb{E}^n$ to be the expectation operator over the distribution of $X_n$ (Fig. 13.1).

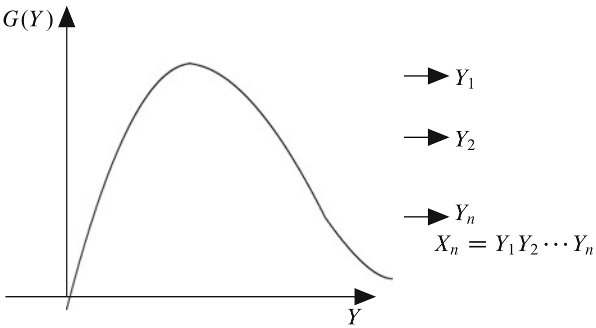


**Fig. 13.1** Constructing the random variable $X_n$

We show in Appendix 13.1 that the solution to (13.10) subject to the boundary and initial conditions (13.12), (13.13) can be written[2]

$$f(x,\tau) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \mathbb{E}^n[M(xX_n e^{-\lambda k\tau}, \tau; E, \sigma^2, r)]. \qquad (13.15)$$

To apply the solution (13.15) we need to specify the probability distribution of the random variable $Y$. Let us consider in particular the case when $Y$ follows a log-normal distribution $\ln Y \sim \phi(\gamma - \delta^2/2, \delta^2)$. It follows that

$$\gamma = \ln(1+k),$$

and that $X_n$ has a log-normal distribution with

$$\mathbb{E}^n[X_n] = e^{n\gamma}, \qquad \mathrm{var}[\ln X_n] = n\delta^2.$$

If we let

$$M_n(x,\tau) = M(x,\tau; E, v_n^2, r_n),$$

where

$$v_n^2 = \sigma^2 + \frac{n\delta^2}{\tau}, \qquad r_n = r - \lambda k + \frac{n\gamma}{\tau},$$

then the solution (13.15) reduces to

$$f(x,\tau) = \sum_{n=0}^{\infty} \frac{e^{-\lambda'\tau}(\lambda'\tau)^n}{n!} M_n(x,\tau),$$

where $\lambda' = \lambda(1+k)$. The quantity $M_n(x,\tau)$ is the value of the option, conditional on knowing that exactly $n$ Poisson jumps will occur during the life of the option. The option price is then the expectation of all such values where the expectation is taken over the Poisson distribution (with parameter $\lambda'\tau$) that $n$ jumps will occur during the life of the option.

In Figs. 13.2 and 13.3 we show the effect on the option price and on delta of increasing values of $\lambda$. Here we have used the parameter values $T = 1$, $E = 1$, $r = 0.05$, $\sigma = 0.2$, $\gamma = 0$ and $\delta = 0.25$.

[2]The forms of the solution given here are from the original Merton (1976) paper. He only demonstrates that these solutions indeed satisfy the integro-partial differential equation (13.11) and relevant boundary conditions. Theory on uniqueness of solutions guarantees that this is indeed “the solution”. Appendix 13.1 reproduces (modulo some notational changes) Merton’s calculations. However this approach gives us no systematic method to solve the integro-partial differential equations encountered in the jump-diffusion case. In Chap. 14 we outline the use of the Fourier transform technique as one such systematic approach.

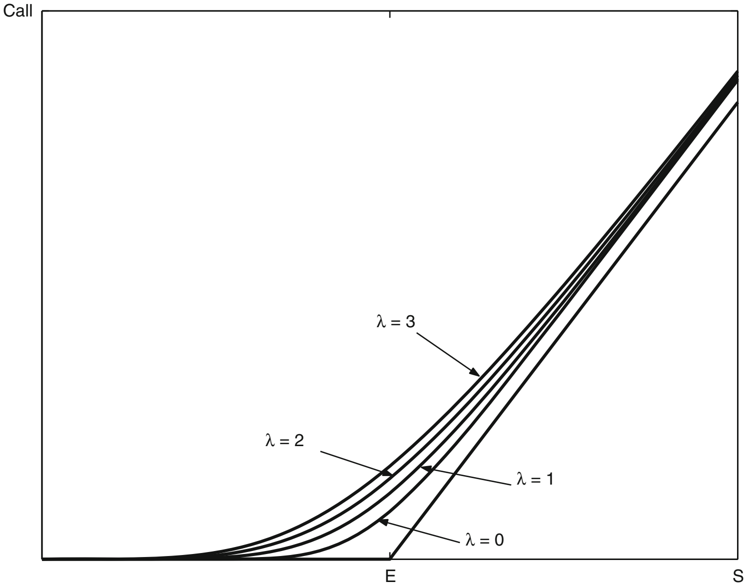


**Fig. 13.2** Effect of increasing values of $\lambda$ on the option price

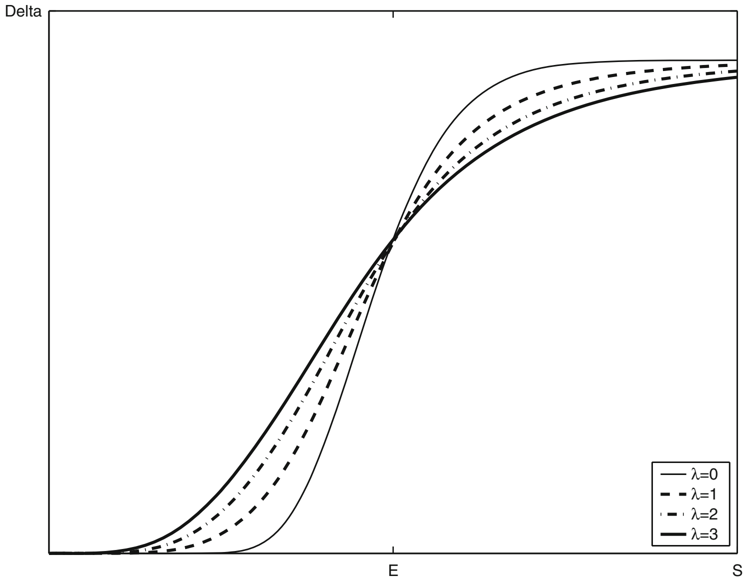


**Fig. 13.3** Effect of increasing values of $\lambda$ on the delta

## 13.5 Alternative Ways of Completing the Market

As we have seen in the previous sections the incorporation of jumps into the diffusion process governing the dynamics of the asset price introduces an additional source of risk. Namely the risk associated with the Poisson stochastic process governing the jump part of the process followed by the asset price. In order to successfully apply the hedging argument we need some way of hedging this additional risk. The way proposed by Merton in Sect. 13.3 is one way to do this. However other ways are also possible and these usually involve introducing some additional hedging instruments into the hedging portfolio. Such a procedure of introducing a sufficient number of traded instruments to hedge away the number of risk factors is known as "completing the market".

One way of completing the market is to introduce additional options into the hedging portfolio, an approach which was developed by Jones (1984). It is also possible to complete the market by using interest rate market instruments as in Jarrow and Madan (1995).

Here we follow the approach of Jones (1984) and introduce several options into the hedging portfolio (for example, options with different strike prices). For instance we may introduce two options on the stock under consideration. Since we have a finite number of hedging instruments we can only hedge a finite number of "jump risks". Hence in this approach we have to restrict the type of jumps that can occur. In the case of the availability of two options as hedging instruments we allow jumps to have only two amplitudes, as shown in Fig. 13.4.

Hence we write the stock price process as

$$\frac{dx}{x} = \mu dt + \sigma dw + k_1 dN_1 + k_2 dN_2, \tag{13.16}$$

where

$$Pr(dN_i = 1) = \lambda_i dt, \qquad Pr(dN_i = 0) = 1 - \lambda_i dt$$

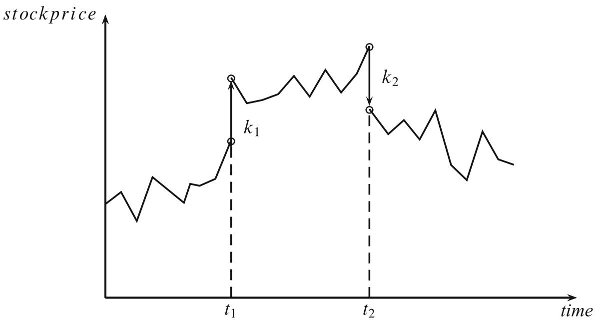


**Fig. 13.4** A finite number of fixed jump sizes

for $i = 1, 2$, and $k_1, k_2$ measure the proportional price jumps in the case of Poisson events.

Let $g$ and $h$ represent the prices of two options written on the stock and assume that the option dynamics contain the same kind of risks as the stock itself. Then the option price dynamics may be written

$$\frac{dg}{g} = \mu_g dt + \sigma_g dw + k_{g_1} dN_1 + k_{g_2} dN_2, \tag{13.17}$$

$$\frac{dh}{h} = \mu_h dt + \sigma_h dw + k_{h_1} dN_1 + k_{h_2} dN_2, \tag{13.18}$$

where the coefficients $\mu$, $\sigma$, $k$ represent expected return, volatility and proportional price jumps for each option. All coefficients are assumed to be functions of $x, g, h$ and time $t$.

We note from (13.16)–(13.18) that the unconditional expected returns are given by

$$\mathbb{E}\left[\frac{dx}{x}\right] = (\mu + k_1\lambda_1 + k_2\lambda_2)dt,$$

$$\mathbb{E}\left[\frac{dg}{g}\right] = (\mu_g + k_{g_1}\lambda_1 + k_{g_2}\lambda_2)dt,$$

$$\mathbb{E}\left[\frac{dh}{h}\right] = (\mu_h + k_{h_1}\lambda_1 + k_{h_2}\lambda_2)dt.$$

Let $f$ be the price of any other option on the stock having an expiry date earlier than that of options $g$ and $h$. We form a hedging portfolio consisting of the three options, the stock and the risk-free asset. We assume that the price of option $f$ is a function $f(x, g, h, \tau)$ of the stock price, the other two option prices and its time-to-maturity $\tau$ in general.

By an application of Ito's Lemma in several variables (see Sect. 6.5) and Ito's Lemma for jump processes (Sect. 12.3) the dynamics of the option $f$ are given by

$$\frac{df}{f} = \mu_f dt + \sigma_f dw + k_{f_1} dN_1 + k_{f_2} dN_2,$$

where

$$\mu_f \equiv \frac{1}{f}\left(\mathscr{D}f + \mu x\frac{\partial f}{\partial x} + \mu_g g\frac{\partial f}{\partial g} + \mu_h h\frac{\partial f}{\partial h} - \frac{\partial f}{\partial \tau}\right), \tag{13.19}$$

$$\mathscr{D}f \equiv \frac{1}{2}\sigma^2 x^2\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\sigma_g^2 g^2\frac{\partial^2 f}{\partial g^2} + \frac{1}{2}\sigma_h^2 h^2\frac{\partial^2 f}{\partial h^2}$$

$$+ \sigma x \sigma_g g \frac{\partial^2 f}{\partial x \partial g} + \sigma x \sigma_h h \frac{\partial^2 f}{\partial x \partial h} + \sigma_g g \sigma_h h \frac{\partial^2 f}{\partial g \partial h}, \tag{13.20}$$

$$\sigma_f \equiv \frac{1}{f}\left(\sigma x \frac{\partial f}{\partial x} + \sigma_g g \frac{\partial f}{\partial g} + \sigma_h h \frac{\partial f}{\partial h}\right), \tag{13.21}$$

$$k_{f_i} \equiv \frac{1}{f}[f(xY_i, gY_{g_i}, hY_{h_i}, \tau) - f(x, g, h, \tau)], \qquad (i = 1, 2), \tag{13.22}$$

where,

$$Y_i = (k_i + 1), \qquad Y_{g_i} = k_{g_i} + 1, \qquad Y_{h_i} = k_{h_i} + 1, \qquad (i = 1, 2).$$

We note that all coefficients are functions of the stock price, the first two option prices and time. The dynamics of $x, g, h$ and $f$ each contain the three risk terms $dz$, $dN_1$ and $dN_2$. The stock $x$ and options $g, h$ span the three risk dimensions that they have in common with the option $f$. Hence by forming a hedge of $x$, $g$ and $h$ we can cancel any risk due to $f$. This reflects the redundancy of $f$ since it can be viewed as an instrument which duplicates a return pattern already available via a dynamic portfolio strategy.

Consider the hedging portfolio and suppose that the weights of the risky asset $x$, options $g, h, f$ and riskless asset $r$ are $\pi$, $\pi_g$, $\pi_h$, $\pi_f$, $\pi_r$ respectively (so that $\pi_r \equiv -(\pi + \pi_h + \pi_g + \pi_f)$ since the weights sum to zero). If $V$ denotes the value of the hedging portfolio then

$$\begin{aligned} \frac{dV}{V} &= \left[\pi(\mu - r) + \pi_g(\mu_g - r) + \pi_h(\mu_h - r) + \pi_f(\mu_f - r)\right] dt \\ &+ [\pi\sigma + \pi_g\sigma_g + \pi_h\sigma_h + \pi_f\sigma_f]dw \\ &+ [\pi k_1 + \pi_g k_{g_1} + \pi_h k_{h_1} + \pi_f k_{f_1}]dN_1 \\ &+ [\pi k_2 + \pi_g k_{g_2} + \pi_h k_{h_2} + \pi_f k_{f_2}]dN_2 \,. \end{aligned}$$

The portfolio will be riskless if

$$\pi\sigma + \pi_g\sigma_g + \pi_h\sigma_h + \pi_f\sigma_f = 0, \tag{13.23}$$

$$\pi k_1 + \pi_g k_{g_1} + \pi_h k_{h_1} + \pi_f k_{f_1} = 0, \tag{13.24}$$

$$\pi k_2 + \pi_g k_{g_2} + \pi_h k_{h_2} + \pi_f k_{f_2} = 0. \tag{13.25}$$

The return on the hedging portfolio would then be

$$\frac{dV}{V} = \left[\pi(\mu - r) + \pi_g(\mu_g - r) + \pi_h(\mu_h - r) + \pi_f(\mu_f - r)\right] dt.$$

Following a now standard argument, this return must be zero so that

$$\pi(\mu - r) + \pi_g(\mu_g - r) + \pi_h(\mu_h - r) + \pi_f(\mu_f - r) = 0. \tag{13.26}$$

The four simultaneous Eqs. (13.23)–(13.26) in the weights $(\pi, \pi_g, \pi_h, \pi_f)$ may be written in matrix form as

$$\begin{bmatrix} \mu - r & \mu_g - r & \mu_h - r & \mu_f - r \\ \sigma & \sigma_g & \sigma_h & \sigma_f \\ k_1 & k_{g_1} & k_{h_1} & k_{f_1} \\ k_2 & k_{g_2} & k_{h_2} & k_{f_2} \end{bmatrix} \begin{bmatrix} \pi \\ \pi_g \\ \pi_h \\ \pi_f \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{13.27}$$

Using standard results in linear algebra (13.27) implies that there must exist quantities $\xi$, $\gamma_1$, $\gamma_2$ such that

$$\mu - r = \xi\sigma + \gamma_1 k_1 + \gamma_2 k_2, \tag{13.28}$$

$$\mu_g - r = \xi\sigma_g + \gamma_1 k_{g_1} + \gamma_2 k_{g_2}, \tag{13.29}$$

$$\mu_h - r = \xi\sigma_h + \gamma_1 k_{h_1} + \gamma_2 k_{h_2}, \tag{13.30}$$

$$\mu_f - r = \xi\sigma_f + \gamma_1 k_{f_1} + \gamma_2 k_{f_2}. \tag{13.31}$$

Making use of (13.31) and substituting (13.28)–(13.30) and (13.21), we find that the option price $f$ must satisfy

$$\begin{aligned} &\mathscr{D}f + (r + \gamma_1 k_1 + \gamma_2 k_2)x\frac{\partial f}{\partial x} + (r + \gamma_1 k_{g_1} + \gamma_2 k_{g_2})g\frac{\partial f}{\partial g} \\ &\quad + (r + \gamma_1 k_{h_1} + \gamma_2 k_{h_2})h\frac{\partial f}{\partial h} - (r + \gamma_1 k_{f_1} + \gamma_2 k_{f_2})f - \frac{\partial f}{\partial \tau} = 0. \end{aligned} \tag{13.32}$$

Note that Eqs. (13.28)–(13.31) extend the familiar interpretation of the no-riskless arbitrage condition. First we interpret $\xi$ as the market price of risk associated with the uncertainty due to the continuous diffusion part of the asset price process and $\gamma_i$ as the market price of risk associated with the $i$th jump component. Then Eqs. (13.28)–(13.31) assert that in equilibrium the expected return on each risky asset equals the risk free rate plus the sum of the market price of each risk component times the amount of associated risk.

A considerable simplification of the option pricing equation (13.32) is possible if we assume that all parameters are functions of the stock price and time alone i.e. $f(x, g, h, \tau) = f(x, \tau)$. Then Eq. (13.32) reduces to

$$\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} + (r + \gamma_1 k_1 + \gamma_2 k_2)x\frac{\partial f}{\partial x} - (r + \gamma_1 k_{f_1} + \gamma_2 k_{f_2})f - \frac{\partial f}{\partial \tau} = 0, \tag{13.33}$$

where we recall that

$$k_{f_i} = \frac{1}{f}[f((k_i + 1)x, \tau) - f(x, \tau)], \quad (i = 1, 2).$$

If we assume that all parameters are constant then Eq. (13.33) may be solved in a way similar to that used to solve Merton's equation (13.11) and the solution turns out to be

$$f(x, \tau) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \left[ e^{-(\gamma+\delta)\tau} \frac{(\gamma_1 \tau)^m}{m!} \frac{(\gamma_2 \tau)^n}{n!} \right] \times$$
$$M[xY_1^m Y_2^n e^{-(\gamma_1 k_1 + \gamma_2 k_2)\tau}, \tau; E, \sigma^2, r], \tag{13.34}$$

where $Y_i = k_i + 1$ for $i = 1, 2$. Suppose we maintain our assumption that all parameters are functions of stock price and time only. Then in the argument leading up to Eq. (13.34) the roles of $f$, $g$ and $h$ can be interchanged. It follows that $g$ and $h$ must also satisfy an equation like (13.34).

If we assume knowledge of $\sigma, k_1, k_2$ is already available, then we have two unknown parameters $\gamma_1, \gamma_2$. Using market values of $g, h$ we may solve $g(x, \tau; \gamma_1, \gamma_2) = g_{market}$ and $h(x, \tau; \gamma_1, \gamma_2) = h_{market}$ to obtain $\widehat{\gamma_1}, \widehat{\gamma_2}$, which may then be used to price the option $f$.

## 13.6 Large Jumps

In this section we restrict our attention to binomial jumps. That is we assume $Y_1 = k_1 + 1, Y_2 = k_2 + 1$ satisfy $Y_1 Y_2 = 1$. In this case, if we define

$$k_2 = 1/Y_1 - 1,$$
$$k_{f_2} = \frac{1}{f}(f(x/Y_1, \tau) - f(x, \tau)),$$

then Jones (1984) shows that the option pricing formula (13.34) specialises to

$$f(x, \tau) = \sum_{n=-\infty}^{\infty} (\gamma_1/\gamma_2)^{n/2} e^{-\upsilon\tau} I_n(2\tau\sqrt{\gamma_1\gamma_2}) M(xY_1^n, e^{-\upsilon\psi\tau}, \tau; E, \sigma^2, r),$$

where

$$I_n(z) = \sum_{j=1}^{\infty} \frac{z^{n+2}}{j!(n+j)!}$$

is a modified Bessel function of the first kind of integer order $n$,

$$v \equiv \gamma_1 + \gamma_2 = \text{the probability of a jump},$$

and

$$\psi \equiv (\gamma_1 k_1 + \gamma_2 k_2)/v = \text{the expected jump amplitude}.$$

These last two results are derived in Feller (1966).

We wish to consider the limiting case in which the jump amplitude becomes large, but at the same time the expected jump amplitude remains constant. In such a case the expected returns on the stock remains finite. If we define $\chi \equiv \ln Y_1 = -\ln Y_2$, then we can define the conditional probabilities for upward versus downward jumps as

$$\gamma_1/v = (\psi + 1 - e^{-\chi})/2\sinh\chi, \qquad \gamma_2/v = (e^{\chi} - \psi - 1)/2\sinh\chi.$$

Note that

$$\lim_{\chi\to\infty} \frac{\gamma_1}{v} = 0,$$

whilst

$$\lim_{\chi\to\infty} \frac{\gamma_2}{v} = 1.$$

These results indicate that large positive jumps are “rare” compared to large negative jumps.

The jump magnitude becoming large is captured by considering $\chi \to \infty$. In this case Jones (1984) shows that the conditional expected upward jump in the option price satisfies

$$\lim_{\chi\to\infty}\left[\frac{\gamma_1}{v}(f(xe^{\chi},\tau) - f(x,\tau))\right] = (\psi + 1)x,$$

and that the conditional expected downward jump satisfies

$$\lim_{\chi\to\infty}\left[\frac{\gamma_2}{v}(f(xe^{\chi},\tau) - f(x,\tau))\right] = f(0,\tau) - f(x,\tau) = -f.$$

Note that in the present notation the partial differential equation (13.33) for $f$ may be written

$$\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} + (r - v\chi)x\frac{\partial f}{\partial x} - \frac{\partial f}{\partial \tau} + \gamma_1[f(xe^{\chi}, \tau) - f(x,\tau)] + \gamma_2[f(x, e^{-\chi}, \tau) - f(x,\tau)] - rf = 0.$$

Taking the limit as $\chi \to \infty$ we obtain the partial differential equation

$$\frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} + (\eta - \theta)x\frac{\partial f}{\partial x} - \eta f + \frac{\partial f}{\partial \tau} + \theta x = 0, \tag{13.35}$$

where

$$\eta \equiv r + v \quad \text{and} \quad \theta \equiv v(\psi + 1).$$

The solution to (13.35) turns out to be

$$f(x,\tau) = x[1 - e^{-\theta\tau}N(-b_1)] - Ed^{-\eta\tau}N(b_2),$$

where

$$b_1 \equiv \frac{\ln(x/E) + (\eta - \theta + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}, \qquad b_2 \equiv b_1 - \sigma\sqrt{\tau}.$$

## 13.7 Appendix

### Appendix 13.1 The Solution of the Integro-Partial Differential Equation

To simplify the notation put

$$P_n(\tau) = \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}, \qquad V_n = xX_n e^{-\lambda k\tau}.$$

We note the derivatives

$$\begin{aligned}\frac{dP_n(\tau)}{d\tau} &= -\lambda\frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} + \lambda\frac{e^{-\lambda\tau}(\lambda\tau)^{n-1}}{(n-1)!}\\ &= \begin{cases} -\lambda P_n(\tau) + \lambda P_{n-1}(\tau), & (n > 0),\\ -\lambda P_n(\tau), & (n = 0),\end{cases}\end{aligned}$$

and

$$\frac{\partial V_n}{\partial \tau} = -\lambda k x X_n e^{-\lambda k \tau} = -\lambda k V_n.$$

Using the above notation the proposed solution (13.15) may be written

$$f(x,\tau) = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{M(V_n,\tau;E,\sigma^2,r)\}. \tag{13.36}$$

We shall simply show that (13.36) satisfies the integro-partial differential equation (13.10) and the associated boundary and initial conditions (13.12) and (13.13). Observe that

$$\begin{aligned}\frac{\partial f}{\partial x} &= \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\left\{\frac{\partial}{\partial x}M(V_n,\tau;E,\sigma^2,r)\right\}\\ &= \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\left\{\frac{\partial V_n}{\partial x}M^{(1)}(V_n,\tau;E,\sigma^2,r)\right\}\\ &= \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{X_n e^{-\lambda k\tau}M^{(1)}(V_n,\tau;E,\sigma^2,r)\}.\end{aligned} \tag{13.37}$$

Here $M^{(1)}$ indicates the first partial derivative of $M$ with respect to its first argument. Upon multiplying through by $x$ the last equation reads

$$x\frac{\partial f}{\partial x} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{V_n M^{(1)}(V_n,\tau;E,\sigma^2,r)\}.$$

Differentiating (13.37) again with respect to $x$ we obtain

$$\frac{\partial^2 f}{\partial x^2} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\left\{(X_n e^{-\lambda k\tau})^2 M^{(11)}(V_n,\tau;E,\sigma^2,r)\right\},$$

which after multiplication by $x^2$ becomes

$$x^2\frac{\partial^2 f}{\partial x^2} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{V_n^2 M^{(11)}(V_n,\tau;E,\sigma^2,r)\},$$

where $M^{(11)}$ indicates the second partial derivative of $M$ with respect to its first argument. Finally

$$\frac{\partial f}{\partial \tau} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\left\{\frac{d}{d\tau}M(V_n,\tau;E,\sigma^2,r)\right\} + \sum_{n=0}^{\infty}\frac{dP_n(\tau)}{d\tau}\mathbb{E}^n\{M(v_n,\tau;E,\sigma^2,r)\}. \tag{13.38}$$

Since

$$\frac{d}{d\tau}M(V_n,\tau;E,\sigma^2,r) = \frac{dV_n}{d\tau}M^{(1)}(V_n,\tau;E,\sigma^2,r) + M^{(2)}(V_n,\tau;E,\sigma^2,r),$$

Eq. (13.38) becomes

$$\frac{\partial f}{\partial \tau} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{-\lambda k V_n M^{(1)} + M^{(2)}\} + \sum_{n=0}^{\infty}(-\lambda)P_n(\tau)\mathbb{E}^n\{M\}$$
$$+\sum_{n=1}^{\infty}\lambda P_{n-1}(\tau)\mathbb{E}^n\{M\},$$

where $M$, $M^{(1)}$ and $M^{(2)}$ are all evaluated at $(V_n,\tau;E,\sigma^2,r)$. Upon rearranging, the last expression can be written as

$$\frac{\partial f}{\partial \tau} = -\lambda f - \lambda k\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{V_n M^{(1)}\} + \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{M^{(2)}\}$$
$$+\lambda\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{M(V_{n+1},\tau;E,\sigma^2,r)\}.$$

Now

$$\frac{-\partial f}{\partial \tau} + (r-\lambda k)x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2\frac{\partial^2 f}{\partial x^2} - rf$$
$$= \lambda f + \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{\lambda k V_n M^{(1)} + M^{(2)}\} + (r-\lambda k)\sum_{n-0}^{\infty} P_n(\tau)\mathbb{E}^n\{V_n M^{(1)}\}$$
$$+\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\left\{\frac{1}{2}\sigma^2 V_n^2 M^{(11)}\right\} - r\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{M\}$$
$$-\lambda\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^{n+1}\{M(V_{n+1},\tau;E,\sigma^2,r)\}$$

$$= \lambda f - \lambda \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^{n+1}\{M(V_{n+1}, \tau; E, \sigma^2, r)\}$$

$$+ \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n \left\{ M^{(2)} + rV_n M^{(1)} + \frac{1}{2}\sigma^2 V_n^2 M^{(11)} - rM \right\}. \tag{13.39}$$

The expression in the curly bracket in the third term of (13.39) is zero since $M(V_n, \tau; E, \sigma^2, r)$ is the solution of

$$M^{(2)} + rV_n M^{(1)} + \frac{1}{2}\sigma^2 V_n M^{(11)} - rM = 0.$$

Thus (13.39) reduces to

$$\frac{-\partial f}{\partial \tau} + (r - \lambda k)x\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 f}{\partial x^2} - rf$$
$$= \lambda f - \lambda \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^{n+1}\{M(V_{n+1}, \tau; E, \sigma^2, r)\}. \tag{13.40}$$

The final step in the proof is to show that the term on the right-hand side of (13.40) equals

$$\lambda \mathbb{E}^{Q_Y}[f(x, \tau) - f(xY, \tau)],$$

where we write $Q_Y$ to indicate clearly that expectations are being taken with respect to the distribution of the random variable $Y$. Replacing $x$ by $xY$ in (13.36) and applying the operator $\mathbb{E}^{Q_Y}$ we have

$$\mathbb{E}^{Q_Y}\{f(xY, \tau)\} = \mathbb{E}^{Q_Y}\left[\sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^n\{M(YV_n, \tau; E, \sigma^2, r)\}\right]. \tag{13.41}$$

Given the definition of $X_n$ as the product of $n$ independent drawings from the distribution of $Y$ and $\mathbb{E}^n$ as the expectation operator over the distribution of $X_n$ it should be clear that

$$\mathbb{E}^{Q_Y}\mathbb{E}^n M(YV_n, \ldots.) = \mathbb{E}^{n+1} M(V_{n+1}, \ldots.).$$

Thus (13.41) becomes

$$\mathbb{E}^{Q_Y}\{f(xY, \tau)\} = \sum_{n=0}^{\infty} P_n(\tau)\mathbb{E}^{n+1}\{M(V_{n+1}, \tau; E, \sigma^2, r)\}.$$

The summation on the right-hand side above is the same as the summation in the second term on the right-hand side of (13.41), so that this last equation may be written

$$\frac{-\partial f}{\partial \tau}+(r-\lambda k)x\frac{\partial f}{\partial x}+\frac{1}{2}\sigma^2x^2\frac{\partial^2 f}{\partial x^2}-rf=\lambda f(x,\tau)-\lambda\mathbb{E}^{Q_Y}\{f(xY,\tau)\}, \qquad (13.42)$$

which may be rearranged to

$$\frac{-\partial f}{\partial \tau}+(r-\lambda k)x\frac{\partial f}{\partial x}+\frac{1}{2}\sigma^2x^2\frac{\partial^2 f}{\partial x^2}-rf+\lambda\mathbb{E}^{Q_Y}[f(xY,\tau)-f(x,\tau)]=0,$$

which is Eq. (13.10).

We have thus shown that Eq. (13.15) is the general form of the solution. It remains only to show that this form of the solution also satisfies the boundary and initial conditions. Since $x = 0$, implies $V_n = 0$ and given that

$$M(0,\tau;E,\sigma^2,r)=0,$$

it follows that

$$f(0,\tau)=0,$$

indicating that the boundary condition (13.13) is satisfied by the solution (13.15). To show that the initial condition (13.13) is satisfied requires a little more analysis. Note first of all that

$$M(V_n,0;E,\sigma^2,r)=\max[0,V_n-E],$$

and so

$$\begin{aligned}\mathbb{E}^n\{M(V_n,0;E,\sigma^2,r)\} &= \mathbb{E}^n\{\max[0,V_n-E]\}\\ &\le \mathbb{E}^n\{V_n\}=\mathbb{E}^n\{xX_n\}=x\mathbb{E}^n\{X_n\}=x(1+k)^n.\end{aligned}$$

The last equality follows from the definition of $k$ as $k = \mathbb{E}^{Q_Y}(Y-1)$ and the fact that $\mathbb{E}^n$ is the expectation over the distribution of $n$ independent drawings from the distribution of $Y$. Now

$$\begin{aligned}f(x,0) &= \lim_{\tau\to 0}\sum_{n=0}^{\infty}P_n(\tau)\mathbb{E}^n\{M(V_n,\tau;E,\sigma^2,r)\}\\ &= P_0(\tau)\mathbb{E}^0\{M(V_0,0:E,\sigma^2,r)\}+\lim_{\tau\to 0}\sum_{n=1}^{\infty}P_n(\tau)\mathbb{E}^n\{M(V_n,\tau;E,\sigma^2,r)\}.\end{aligned}$$

Since $P_0(\tau) = 1$ and

$$\mathbb{E}^0\{M(V_0, 0; E, \sigma^2, r)\} = \mathbb{E}^0\{M(x, 0; E, \sigma^2, r)\} = \max[0, x - E],$$

we have

$$f(x, 0) = \max[0, x - E] + \lim_{\tau \to 0} \sum_{n=1}^{\infty} P_n(\tau)\mathbb{E}^n\{M(V_n, \tau; E, \sigma^2, r)\}.$$

Thus we need to show that the summation term on the right-hand side is zero. To show this proceed as follows:

$$\begin{aligned}
&\lim_{\tau \to 0} \sum_{n=1}^{\infty} P_n(\tau)\mathbb{E}^n\{M(V_n, \tau; E, \sigma^2, r)\} \\
&\leq \lim_{\tau \to 0} x e^{-\lambda\tau} \sum_{n=1}^{\infty} \frac{[(1+k)\lambda\tau]^n}{n!} \qquad \text{(using (13.43))} \\
&= \lim_{\tau \to 0} x e^{-\lambda\tau} [e^{(1+k)\lambda\tau} - 1] \\
&= 0.
\end{aligned}$$

Thus we have shown that $f(x, 0) = \max[0, x - E]$ which is the final step in the demonstration that Eq. (13.15) is the general form of the solution.

# Chapter 14
# Partial Differential Equation Approach Under Geometric Jump-Diffusion Process

**Abstract** In this chapter we consider the solution of the integro-partial differential equation that determines derivative security prices when the underlying asset price is driven by a jump-diffusion process. We take the analysis as far as we can for the case of a European option with a general pay-off and the jump-size distribution is left unspecified. We obtain specific results in the case of a European call option and when the jump size distribution is log-normal. We illustrate two approaches to the problem. The first is the Fourier transform technique that we have used in the case that the underlying asset follows a diffusion process. The second is the direct approach using the expectation operator expression that follows from the martingale representation. We also show how these two approaches are connected.

## 14.1 The Integro-Partial Differential Equation

Consider the integro-partial differential equation (13.11) that prices a derivative security when the underlying asset price is driven by a jump-diffusion process, viz.

$$-\frac{\partial f}{\partial \tau}+(r-\lambda k)\,x\frac{\partial f}{\partial x}+\frac{1}{2}\sigma^2x^2\frac{\partial^2 f}{\partial x^2}-rf+\lambda\int_0^\infty[f(xY,\tau)-f(x,\tau)]\,G(Y)\,dY=0, \tag{14.1}$$

where we recall that $\tau$ represents time-to-maturity and $G(Y)$ is the density function for the log-normal distribution (see (12.15) and the associated footnote). Equation (14.1) may be rearranged into the form

$$\frac{\partial f}{\partial \tau}=\frac{1}{2}\sigma^2x^2\frac{\partial^2 f}{\partial x^2}+(r-\lambda k)x\frac{\partial f}{\partial x}-(r+\lambda)f+\lambda\int_0^\infty f(xY,\tau)G(Y)dY. \tag{14.2}$$

We recall the boundary condition

$$f(0,\tau)=0, \tag{14.3}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_14

which is applicable to all European style options. Here we consider a general initial condition

$$f(x,0) = h(x). \tag{14.4}$$

We recall that $x \geq 0$ and $Y \geq 0$. We transform to the logarithmic variables $u = \ln x$, in terms of which we define the transformed price function. Thus

$$f(x,\tau) = F(\ln x, \tau) = F(u,\tau). \tag{14.5}$$

In terms of the transformed variables (14.2) becomes[1]

$$\frac{\partial F}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 F}{\partial u^2} + (r - \lambda k - \frac{\sigma^2}{2})\frac{\partial F}{\partial u} - (r+\lambda)F + \lambda\int_0^\infty F(u + \ln Y, \tau)G(Y)dY, \tag{14.6}$$

which must be solved subject to the boundary condition

$$F(-\infty, \tau) = 0$$

and initial condition

$$F(u,0) = f(x,0) = h(e^u) \equiv H(u).$$

## 14.2 The Fourier Transform

Following the solution procedure established in Chap. 9, we consider the Fourier transform of $F(u,\tau)$, namely

$$\hat{F}(\xi,\tau) = \int_{-\infty}^{\infty} F(u,\tau)e^{-iu\xi}du. \tag{14.7}$$

**Proposition 14.1** *The Fourier transform* (14.7) *is given by*

$$\hat{F}(\xi,\tau) = \hat{F}(\xi,0)e^{[-\frac{\sigma^2\xi^2}{2} + (r-\lambda k - \frac{\sigma^2}{2})i\xi - (r+\lambda) + \lambda A(\xi)]\tau}. \tag{14.8}$$

[1] We note that

$$\frac{\partial f}{\partial x} = \frac{1}{x}\frac{\partial F}{\partial u}, \qquad \frac{\partial^2 f}{\partial x^2} = -\frac{1}{x^2}\frac{\partial F}{\partial u} + \frac{1}{x^2}\frac{\partial^2 F}{\partial u^2} = \frac{1}{x^2}\left(\frac{\partial^2 F}{\partial u^2} - \frac{\partial F}{\partial u}\right),$$

and $f(xY,\tau) = F(\ln(xY),\tau) = F(u + \ln Y, \tau)$.

*where*

$$A(\xi) = \int_0^\infty G(Y) e^{i\xi \ln Y} dY. \tag{14.9}$$

*Proof* As in Chap. 9 we apply the Fourier transform operation to the integro-partial differential equation (14.6), note that in doing so we again make use of the results (9.14) and (9.15). We thus obtain for $\hat{F}(\xi, \tau)$ the ordinary differential equation

$$\begin{aligned} \frac{\partial \hat{F}}{\partial \tau} &= \left[ \frac{-\sigma^2 \xi^2}{2} + (r - \lambda k - \frac{\sigma^2}{2}) i\xi - (r + \lambda) \right] \hat{F} \\ &\quad + \int_{-\infty}^{\infty} \left[ \lambda \int_0^\infty F(u + \ln Y, \tau) e^{-iu\xi} G(Y) dY \right] du. \end{aligned} \tag{14.10}$$

Consider the integral term in (14.10), which is the new feature brought in by the jump process, it can be rearranged as

$$\begin{aligned} &\int_{-\infty}^{\infty} \left[ \int_0^\infty \lambda F(u + \ln Y, \tau) e^{-iu\xi} G(Y) dY \right] du \\ &= \lambda \int_0^\infty \int_{-\infty}^{\infty} F(u + \ln Y, \tau)\, G(Y)\, e^{-iu\xi}\, du\, dY \\ &= \lambda \int_0^\infty \int_{-\infty}^{\infty} F(Z, \tau) G(Y) e^{-i\xi(Z - \ln Y)} dZ dY \qquad \text{(by setting } Z = u + \ln Y) \\ &= \lambda \int_0^\infty G(Y) e^{i\xi \ln Y} dY \int_{-\infty}^{\infty} F(Z, \tau) e^{-i\xi Z} dZ \\ &= \lambda A(\xi) \hat{F}(\xi, \tau), \end{aligned}$$

where we set

$$A(\xi) = \int_0^\infty G(Y) e^{i\xi \ln Y} dY.$$

Hence the ordinary differential equation (14.10) becomes

$$\frac{\partial \hat{F}}{\partial \tau} = \left[ \frac{-\sigma^2 \xi^2}{2} + (r - \lambda k - \frac{\sigma^2}{2}) i\xi - (r + \lambda) + \lambda A(\xi) \right] \hat{F},$$

whose solution is easily obtained and is given in (14.8). ■

To obtain the solution we then apply the Fourier inversion theorem to the function $\hat{F}(\xi, \tau)$. The result turns out to be

**Proposition 14.2** *The solution to the integro-partial differential equation (14.2) turns out to be*

$$f(x,\tau) = e^{-(\lambda+r)\tau} \int_{-\infty}^{\infty} H(Z)K(Z,x,\tau)dZ, \tag{14.11}$$

*where*

$$K(Z,x,\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(\frac{1}{2}\sigma^2\xi^2-\lambda A(\xi))\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z]}d\xi. \tag{14.12}$$

*Proof* Applying the Fourier inversion theorem to $\hat{F}(\xi,\tau)$ defined in (14.8) we obtain

$$\begin{aligned} F(u,\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{F}(\xi,0)e^{[-\frac{\sigma^2\xi^2}{2}+(r-\lambda k-\frac{\sigma^2}{2})i\xi-(r+\lambda)+\lambda A(\xi)]\tau}e^{iu\xi}d\xi \\ &= \frac{1}{2\pi}e^{-(r+\lambda)\tau}\int_{-\infty}^{\infty} \hat{F}(\xi,0)e^{[-\frac{\sigma^2\xi^2}{2}+\lambda A(\xi)]\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+u]}d\xi. \end{aligned} \tag{14.13}$$

By substituting $u = \ln x$, we can express (14.13) in terms of the original variables,

$$f(x,\tau) = \frac{1}{2\pi}e^{-(r+\lambda)\tau}\int_{-\infty}^{\infty} \hat{F}(\xi,0)e^{-(\frac{1}{2}\sigma^2\xi^2-\lambda A(\xi))\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x]}d\xi. \tag{14.14}$$

Recall that

$$\hat{F}(\xi,0) = \int_{-\infty}^{\infty} F(u,0)e^{-iu\xi}du = \int_{-\infty}^{\infty} f(e^u,0)e^{-iu\xi}du. \tag{14.15}$$

Upon substituting (14.15) into (14.14) we obtain

$$\begin{aligned} &f(x,\tau) \\ &= \frac{1}{2\pi}e^{-(r+\lambda)\tau}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(e^Z,0)e^{-iZ\xi}dZ\right]e^{-(\frac{1}{2}\sigma^2\xi^2-\lambda A(\xi))\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x]}d\xi \\ &= \frac{e^{-(r+\lambda)\tau}}{2\pi}\int_{-\infty}^{\infty} f(e^Z,0)\left(\int_{-\infty}^{\infty} e^{-(\frac{1}{2}\sigma^2\xi^2-\lambda A(\xi))\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z]}d\xi\right)dZ \\ &= e^{-(\lambda+r)\tau}\int_{-\infty}^{\infty} H(Z)K(Z,x,\tau)dZ, \end{aligned}$$

where we set

$$K(Z,x,\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(\frac{1}{2}\sigma^2\xi^2-\lambda A(\xi))\tau+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z]}d\xi.$$

■

The kernel function $K(Z, x, \tau)$ in (14.12) is analogue to the kernel $K(y, t)$ of (9.23) for the corresponding problem without jumps. The essential difference is the $A(\xi)$ term which depends on the jump size distribution. The next step in the solution procedure is to perform the integration in (14.12) and obtain the analogue of (9.25).

Some standard manipulations of probability density functions and some well known results on integrals involving exponential functions allow us to express the kernel function $K(Z, x, \tau)$ as a Poisson weighted sum of expectations over the jump size distribution. It is interesting to observe that this result can be obtained without specifying the nature of the jump-size distribution.

**Proposition 14.3** *The kernel function $K(Z, x, \tau)$ may be expressed as*

$$K(Z, x, \tau) = \frac{1}{\sigma\sqrt{2\pi\tau}}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\mathbb{E}^n\left\{e^{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}}\right\}$$

*where for a sequence of jump $Y_1, Y_2, \ldots, Y_n$ we have $\theta_n = xY_1Y_2\ldots Y_n$, $\theta_0 = 1$ and we define the n-fold expectation operator $\mathbb{E}^n$ by*

$$\mathbb{E}^n(x) \equiv \int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty} xG(Y_1)G(Y_2)\cdots G(Y_n)dY_1dY_2\cdots dY_n.$$

*Note that*

$$\mathbb{E}^0(x) \equiv x.$$

*Proof* Expanding $e^{\lambda\tau A(\xi)}$ in a Taylor series, the expression (14.12) reduces to

$$K(Z, x, \tau) = \frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2\tau\xi^2-i[Z-\ln x-(r-\lambda k-\frac{\sigma^2}{2})\tau]\xi}A(\xi)^n d\xi. \tag{14.16}$$

Note that

$$\begin{aligned}A(\xi)^n &= \left(\int_0^{\infty} G(Y)e^{i\xi\ln Y}dY\right)^n\\ &= \int_0^{\infty} G(Y_1)e^{i\xi\ln Y_1}dY_1\int_0^{\infty} G(Y_2)e^{i\xi\ln Y_2}dY_2\cdots\int_0^{\infty} G(Y_n)e^{i\xi\ln Y_n}dY_n.\end{aligned}$$

Hence

$$\begin{aligned}K(Z, x, \tau) =& \frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2\tau\xi^2-i[Z-\ln x-(r-\lambda k-\frac{\sigma^2}{2})\tau]\xi}\\ &\left\{\int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty} G(Y_1)G(Y_2)\cdots G(Y_n)e^{i\xi\ln(Y_1Y_2\cdots Y_n)}dY_1dY_2\cdots dY_n\right\}d\xi\end{aligned}$$

$$=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty}I(\theta_n,\tau)G(Y_1)G(Y_2)\cdots G(Y_n)I(\theta_n,\tau)$$
$$\times\, dY_1 dY_2\cdots dY_n,$$

where

$$I(\theta_n,\tau)\equiv\int_{-\infty}^{\infty}e^{-\frac{1}{2}\sigma^2\tau\xi^2-i[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]\xi}d\xi$$

with

$$\theta_n\equiv xY_1Y_2\cdots Y_n.$$

Consider $I(\theta_n,\tau)$, which may be expressed as

$$I(\theta_n,\tau)=\int_{-\infty}^{\infty}e^{-p\xi^2-q\xi}d\xi,$$

where

$$p=\frac{1}{2}\sigma^2\tau\quad\text{and}\quad q=i[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau].$$

Recalling the result

$$\int_{-\infty}^{\infty}e^{-p\xi^2-q\xi}d\xi=\sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}},$$

we have

$$I(\theta_n,\tau)=\sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}}=\sqrt{\frac{2\pi}{\sigma^2\tau}}\exp\left\{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}\right\}.$$

Hence

$$K(Z,x,\tau)=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_0^{\infty}\int_0^{\infty}\cdots\int_0^{\infty}G(Y_1)G(Y_2)\cdots G(Y_n)$$
$$\sqrt{\frac{2\pi}{\sigma^2\tau}}\exp\{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}\}dY_1dY_2\cdots dY_n$$
$$=\frac{1}{\sigma\sqrt{2\pi\tau}}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\mathbb{E}^n\left\{e^{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}}\right\},\qquad(14.17)$$

where

$$\mathbb{E}^n(x) \equiv \int_0^\infty \int_0^\infty \cdots \int_0^\infty x\, G(Y_1)G(Y_2)\cdots G(Y_n) dY_1 dY_2 \cdots dY_n. \tag{14.18}$$

Note that

$$\mathbb{E}^0(x) \equiv x \quad \text{and} \quad \theta_0 \equiv x.$$

■

Propositions 14.2 and 14.3 may be brought together to express the European call option value as a Poisson weighted sums of $n$-fold expectations of Black–Scholes values conditional on $n$-jumps. The actual result is

**Proposition 14.4** *In the case of a European call option Propositions 14.2 and 14.3 imply that*

$$f(x,\tau) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \mathbb{E}^n \left\{ M\left(xX_n e^{-\lambda k\tau}, \tau; X, \sigma^2, r\right)\right\} \tag{14.19}$$

*where $M$ is the Black–Scholes pricing function*

$$M(x,\tau;X,\sigma^2,\tau) = x\mathcal{N}(d_1) - Xe^{-r\tau}\mathcal{N}(d_2), \tag{14.20}$$

$$d_1 = \frac{\ln(x/X) + (r+\sigma^2/2)\tau}{\sigma\sqrt{\tau}}, \quad d_2 = d_1 - \sigma\sqrt{\tau}.$$

*Here $X_n \equiv Y_1 Y_2 \ldots Y_n$, will $X_0 \equiv 1$ and $\mathbb{E}^n$ is the $n$-fold expectation operator defined in Proposition 14.3.*

*Proof* Substituting the simplified expression for the kernel function (14.17) into (14.11) and using the European call option payoff $H(Z) = e^Z - X$ and interchanging integration and expectation operations we find that the call option price, $f(x,\tau)$ is given by

$$\begin{aligned} f(x,\tau) =& \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \mathbb{E}^n \Bigg\{ \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}} \int_{\ln X}^{\infty} (e^Z \\ & - X) \exp[\frac{-[Z - \ln\theta_n - (r - \lambda k - \frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}] dZ \Bigg\} \\ =& \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \mathbb{E}^n\{f_1(x,\tau) - f_2(x,\tau)\}, \end{aligned}$$

where

$$f_1(x,\tau) = \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}} \int_{\ln X}^{\infty} e^Z e^{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}} dZ$$

and

$$f_2(x,\tau) = \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}} \int_{\ln X}^{\infty} X e^{\frac{-[Z-\ln\theta_n-(r-\lambda k-\frac{\sigma^2}{2})\tau]^2}{2\sigma^2\tau}} dZ.$$

*First consider* $f_1(x,\tau)$. By the change of variable

$$\alpha = (Z - \ln\theta_n - (r - \lambda k - \frac{\sigma^2}{2})\tau)/\sigma\sqrt{\tau}$$

and setting

$$\hat{D} = \frac{\ln X - \ln\theta_n - (r - \lambda k - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}},$$

$f_1(x,\tau)$ is simplified to

$$f_1(x,\tau) = \frac{e^{-r\tau}}{\sigma\sqrt{2\pi\tau}} \int_{\hat{D}}^{\infty} e^Z e^{-\frac{\alpha^2}{2}} \sigma\sqrt{\tau} d\alpha = \theta_n e^{-\lambda k\tau} \frac{1}{\sqrt{2\pi}} \int_{\hat{D}}^{\infty} e^{-\frac{(\alpha-\sigma\sqrt{\tau})^2}{2}} d\alpha.$$

Making the further change of variable

$$\omega = \alpha - \sigma\sqrt{\tau}$$

and let

$$\hat{d}_1 = \frac{\ln\frac{\theta_n}{X} + (r - \lambda k + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$$

we obtain

$$f_1(x,\tau) = \theta_n e^{-\lambda k\tau} N(\hat{d}_1).$$

*Next consider* $f_2(x,\tau)$. Let

$$\alpha = \frac{Z - \ln\theta_n - (r - \lambda k - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}},$$

we obtain

$$f_2(x,\tau) = Xe^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\hat{d}_2}^{\infty} e^{-\frac{-\alpha^2}{2}}\,d\alpha,$$

where

$$\hat{d}_2 = \frac{\ln\frac{\theta_n}{X} + (r - \lambda k - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}.$$

Thus

$$f_2(x,\tau) = Xe^{-r\tau}N(\hat{d}_2).$$

Putting the above results together we find that

$$f(x,\tau) = \sum_{n=0}^{\infty}\frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}\mathbb{E}^n\{\theta_n e^{-\lambda k\tau}N(\hat{d}_1) - Xe^{-r\tau}N(\hat{d}_2)\}.$$

Finally, define

$$X_n \equiv Y_1Y_2\cdots Y_n, \qquad X_0 \equiv 1$$

and

$$M(x,\tau;X,\sigma^2,r) = xN(d_1) - Xe^{-r\tau}N(d_2)$$

with

$$d_1 = \frac{\ln(\frac{x}{X}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \qquad d_2 = d_1 - \sigma\sqrt{\tau}.$$

Then the value of the call option can be written as

$$f(x,\tau) = \sum_{n=0}^{\infty}\frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}\mathbb{E}^n\{M(xX_n e^{-\lambda k\tau},\tau;X,\sigma^2,r)\}.$$

■

To generate specific models we need to specify the jump-size distribution. The most common assumption is that it is either log-normally distributed as discussed in Sect. 12.2 or follows a multi-nomial distribution as in Sect. 13.5. In the following sections we shall derive in detail the call option price when the jump-size is log-normally distributed and sketch briefly the derivation where the jump-size is given by a multi-nomial distribution.

## 14.3 Evaluating the Kernel Function Under a Log-Normal Jump Distribution

**Proposition 14.5** *For the case in which jump-size distribution is log-normally distributed the kernel function becomes*

$$K(Z, x, \tau) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!\sqrt{n+\tau}} \exp\left[-\frac{(Z-b_n)}{2\sigma^2(n+\tau)}\right] \tag{14.21}$$

*where*

$$b_n = \ln x + (r - \lambda k - \sigma^2/2)\tau + n(\gamma - \delta^2/2).$$

*Proof* Recall that when $G(y)$ is log-normally distributed the density function is given by[2]

$$G(y)dy = \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\ln y - (\gamma - \delta^2/2))^2}{\delta^2}} \frac{dy}{y}.$$

It follows in this case that the quantity $A(\xi)$ in (14.9) assumes the form

$$A(\xi) = \int_0^\infty \frac{1}{y\delta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln y - (\gamma - \delta^2/2)}{\delta}\right)^2 + i\xi \ln y} dy,$$

which by use of the transformation $x = \ln y$ becomes

$$A(\xi) = \int_{-\infty}^\infty \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x - (\gamma - \delta^2/2))^2}{\delta^2} + i\xi x} dx.$$

The further transformation $Z = (x - (\gamma - \delta^2/2))/\delta$ allows us to write

$$\begin{aligned} A(\xi) &= \int_{-\infty}^\infty \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{Z^2}{2} + i\xi(\delta Z + \gamma - \delta^2/2)} \delta dZ \\ &= \frac{1}{\sqrt{2\pi}} e^{i\xi(\gamma - \delta^2/2) - \frac{\xi^2\delta^2}{2}} \int_{-\infty}^\infty e^{-\frac{(Z - i\xi\delta)^2}{2}} dZ. \end{aligned} \tag{14.22}$$

Then by setting $q = (Z - i\xi\delta)/\sqrt{2}$, we finally obtain

$$A(\xi) = e^{i\xi(\gamma - \delta^2/2) - \frac{\xi^2\delta^2}{2}} \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-q^2} dq = e^{i\xi(\gamma - \delta^2/2) - \frac{\xi^2\delta^2}{2}}. \tag{14.23}$$

[2] See (12.15).

Substituting (14.23) into (14.16) and rearranging slightly we obtain

$$K(Z,x,\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\frac{\tau}{2}\sigma^2\xi^2+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z]}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}e^{\frac{-\xi^2\sigma^2 n}{2}+i\xi n(\gamma-\delta^2/2)}d\xi.$$

Some further re-arrangement allows us to write

$$\begin{aligned}K(Z,x,\tau)&=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(n+\tau)\sigma^2\xi^2+i\xi[(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z+n(\gamma-\delta^2/2)]}d\xi\\&=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(n+\tau)\sigma^2\xi^2+i\xi\alpha}d\xi,\end{aligned}$$

where

$$\alpha=(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x-Z+n(\gamma-\delta^2/2). \tag{14.24}$$

Thus

$$\begin{aligned}K(Z,x,\tau)&=\frac{1}{2\pi}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\int_{-\infty}^{\infty}e^{-\frac{1}{2}[(n+\tau)\sigma^2\xi^2-2i\xi\alpha]}d\xi\\&=\frac{1}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!\sigma\sqrt{n+\tau}}e^{-\frac{\alpha^2}{2\sigma^2(n+\tau)}},\end{aligned} \tag{14.25}$$

where the last equality has been obtained by use of the result (9.24) (with $p=(n+\tau)\frac{\sigma^2}{2}$ and $q=-ix$). ■

It is of interest to observe that (14.21) generalizes the kernel function in (9.25) in a very natural way. It is the weighted sum of the kernels conditional on $n$ jumps having occurred, the weights being the Poisson jump probabilities.

## 14.4 Option Valuation Under a Log-Normal Jump Distribution

Propositions 14.2 and 14.3 may be combined to yield the option value under a lognormal jump distribution. The result is

**Proposition 14.6** *In the case of a log-normal jump size distribution the European option value is given by*

$$f(x,\tau)=\sum_{n=0}^{\infty}\frac{e^{-\lambda'\tau}(\lambda'\tau)^n}{n!}M(x,\tau;K,v_n^2,r_n) \tag{14.26}$$

*where M is the Black–Scholes pricing function defined in Proposition 14.4 and*

$$\lambda' = \lambda(1+k), \quad v_n^2 = \sigma^2 + n\delta^2/\tau \quad \textit{and} \quad r_n = r - \lambda k + n\gamma/\tau.$$

*Proof* We substitute the expression (14.25) for the kernel function into (14.11). Consider a European Call option for which $h(z) = (e^z - K)^+$, then (keeping in mind that $\alpha$ is a function of $Z$)

$$\begin{aligned} f(x,\tau) &= e^{-(\lambda+r)\tau} \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!\sigma\sqrt{2\pi(n+\tau)}} \int_{-\infty}^{\infty} max(e^Z - K, 0) e^{\frac{-\alpha^2}{2\sigma^2(n+\tau)}} dZ \\ &= e^{-(\lambda+r)\tau} \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!\sigma\sqrt{2\pi(n+\tau)}} \int_{\ln K}^{\infty} (e^Z - K) e^{\frac{-\alpha^2}{2\sigma^2(n+\tau)}} dZ \\ &= e^{-(\lambda+r)\tau} \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!\sigma\sqrt{2\pi(n+\tau)}} I, \end{aligned} \tag{14.27}$$

where

$$I = \int_{\ln K}^{\infty} (e^Z) e^{\frac{-\alpha^2}{2\sigma^2(n+\tau)}} dZ - K \int_{\ln K}^{\infty} e^{\frac{-\alpha^2}{2\sigma^2(n+\tau)}} dZ. \tag{14.28}$$

To evaluate $I$, recall that from (14.24) that

$$\alpha = (r - \lambda k - \frac{1}{2}\sigma^2)\tau + \ln x - Z + n(\gamma - \delta^2/2).$$

In (14.28) make the change of variable

$$u = \frac{(r - \lambda k - \frac{\sigma^2}{2})\tau + \ln x - Z + n(\gamma - \delta^2/2)}{\sigma\sqrt{n+\tau}}, \tag{14.29}$$

and set

$$d_2 = \frac{(r - \lambda k - \frac{\sigma^2}{2})\tau + \ln \frac{x}{K} + n(\gamma - \delta^2/2)}{\sigma\sqrt{n+\tau}}.$$

Then

$$\begin{aligned} I &= (-\sigma\sqrt{n+\tau})\left[\int_{d_2}^{-\infty} e^Z e^{-\frac{u^2}{2}} du - K \int_{d_2}^{-\infty} e^{-\frac{1}{2}u^2} du\right] \\ &= \sigma\sqrt{n+\tau}\left(\int_{-\infty}^{d_2} e^{-\frac{u^2}{2} + (r-\lambda k - \frac{\sigma^2}{2})\tau + \ln x + n(\gamma - \delta^2/2) - \sigma\sqrt{n+\tau}u} du - K \int_{-\infty}^{d_2} e^{-\frac{1}{2}u^2} du\right). \end{aligned}$$

Rearranging

$$I = \sigma\sqrt{n+\tau}\Big(e^{(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x+n(\gamma-\delta^2/2)}\int_{-\infty}^{d_2} e^{-\frac{1}{2}(u^2+2\sigma(\sqrt{n+\tau})u)}du$$

$$-K\int_{-\infty}^{d_2} e^{-\frac{1}{2}u^2}du\Big)$$

$$= \sigma\sqrt{n+\tau}\Big(e^{(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x+n(\gamma-\delta^2/2)+\frac{1}{2}\sigma^2(n+\tau)}\int_{-\infty}^{d_2} e^{-\frac{1}{2}(u+\sigma\sqrt{n+\tau})^2}du$$

$$-K\int_{-\infty}^{d_2} e^{-\frac{1}{2}u^2}du\Big).$$

Making the further change of variable $V = u + \sigma\sqrt{n+\tau}$, and defining

$$d_1 = d_2 + \sigma\sqrt{n+\tau} = \frac{\ln\frac{x}{K} + n(\gamma - \delta^2/2 + \sigma^2) + (r - \lambda k + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{n+\tau}},$$

we obtain

$$I = \sigma\sqrt{n+\tau}\left(e^{(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x+n(\gamma-\delta^2/2)+\frac{1}{2}\sigma^2(n+\tau)}\int_{-\infty}^{d_1} e^{-\frac{V^2}{2}}dV - K\int_{-\infty}^{d_2} e^{-\frac{u^2}{2}}du\right).$$

Hence

$$f(x,\tau) = e^{-(\lambda+r)\tau}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}(e^{(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x+n(\gamma-\delta^2/2)+\frac{1}{2}\sigma^2(n+\tau)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_1} e^{-\frac{V^2}{2}}dV$$

$$- K\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2} e^{-\frac{u^2}{2}}du)$$

$$= e^{-(\lambda+r)\tau}\sum_{n=0}^{\infty}\frac{(\lambda\tau)^n}{n!}\Big(e^{(r-\lambda k-\frac{\sigma^2}{2})\tau+\ln x+n(\gamma-\delta^2/2)+\frac{1}{2}\sigma^2(n+\tau)}\mathcal{N}(d_1)$$

$$- K\mathcal{N}(d_2)\Big),$$

which simplifies to (14.26). Note that $e^{n\gamma} = e^{n\log(1+k)} = (1+k)^n$ and the $e^{-\lambda\tau}$ and $e^{-\lambda k\tau}$ terms may be grouped to yield $e^{-\lambda'\tau}$. ■

## 14.5 Using the Expectation Operator to Evaluate the Option Under Log-Normal Jumps

The expression for the European call option price derived in the previous section was obtained by using the expression for the kernel function under log-normally distributed jumps, substituting this into the Fourier inversion formula (14.11) and performing the integration. It is also possible to proceed directly from the expression (14.19) in terms of expectation operators (which makes no assumption about the jump size distribution), and calculate this expression when the jump-size distribution is log-normal. These calculations are laid out in this section.

**Proposition 14.7** *In the case of log-normally distributed jump sizes the expression* (14.26) *in Proposition 14.6 may be obtained directly from the expectation representation in Proposition 14.4 operator.*

*Proof* From the definition of the function $M$, (14.19) may be written

$$f(x,\tau) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \mathbb{E}^n\{(xX_n e^{-\lambda k\tau} N(d_1) - Ke^{-r\tau} N(d_2)\}, \tag{14.30}$$

where

$$d_1 = \frac{\ln\left(\frac{xX_n e^{-\lambda k\tau}}{K} + \left(r + \frac{\sigma^2}{2}\right)\tau\right)}{\sigma\sqrt{\tau}}, \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{\tau}.$$

We note first that

$$\mathbb{E}^n\{g(X_n)\} = \int_0^\infty \ldots \int_0^\infty g(X_n)G(Y_1)\ldots G(Y_n)dY_1\ldots dY_n, \tag{14.31}$$

where

$$X_n = Y_1Y_2\ldots Y_n, \quad (n = 1, 2\ldots), \qquad X_0 = 1, \quad \text{and} \quad \mathbb{E}^0\{g(X_0)\} = g(X_0).$$

Since

$$\ln Y \sim N(\gamma - \delta^2/2, \delta^2),$$

it follows that

$$\sum_{i=1}^{n} \ln Y_i = \ln X_n \sim N(n(\gamma - \delta^2/2), n\delta^2).$$

Thus

$$G(X_n) = \frac{1}{X_n \delta \sqrt{2\pi n}} \exp\left[-\frac{1}{2}\left(\frac{\ln X_n - n(\gamma - \delta^2/2)}{\delta\sqrt{n}}\right)^2\right].$$

As a result (14.31) may be written

$$\begin{aligned}\mathbb{E}^n\{g(X_n)\} &= \int_0^\infty g(X_n) G(X_n) dX_n \\ &= \int_0^\infty g(X_n) \frac{1}{X_n \delta \sqrt{2\pi n}} e^{-\frac{1}{2}\left(\frac{\ln X_n - n(\gamma - \delta^2/2)}{\delta\sqrt{n}}\right)^2} dX_n. \end{aligned} \tag{14.32}$$

Next we note that the expectation operation in (14.30) may be rewritten

$$\mathbb{E}^n\left\{xX_n e^{-\lambda k\tau} N(d_1) - Ke^{-r\tau} N(d_2)\right\} = xe^{-\lambda k \tau} A_1 - Ke^{-r\tau} A_2 \tag{14.33}$$

where

$$A_1 = \mathbb{E}^n\{X_n N(d_1)\} \quad \text{and} \quad A_2 = \mathbb{E}^n\{N(d_2)\}.$$

In Appendix 14.1 we evaluate $A_1$ and $A_2$, given that $X_n$ is log-normally distributed, and find that

$$A_1 = e^{n\gamma} N(\hat{d}_1), \quad \text{and} \quad A_2 = N(\hat{d}_2), \tag{14.34}$$

where

$$\hat{d}_1 = \frac{\ln \frac{x}{K} + \left(r_n + \frac{v_n^2}{2}\right)\tau}{v_n\sqrt{\tau}},$$

where $r_n \equiv r - \lambda k + n\gamma/\tau$, $v_n^2 = \sigma^2 + n\delta^2/\tau$, and $\hat{d}_2 = \hat{d}_1 - v_n\sqrt{\tau}$. Using the results (14.33) and (14.34) the price of the call option becomes

$$\begin{aligned} f(x,\tau) &= \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}\left[xe^{-\lambda k\tau} e^{n\gamma} N(\hat{d}_1) - Ke^{-r\tau} N(\hat{d}_2)\right] \\ &= \sum_{n=0}^{\infty} \frac{e^{-\lambda(k+1)\tau}(\lambda\tau)^n}{n!} e^{n\gamma}\left[xN(\hat{d}_1) - Ke^{-(r-\lambda k + \frac{n\gamma}{\tau})\tau} N(\hat{d}_2)\right] \\ &= \sum_{n=0}^{\infty} \frac{e^{-\lambda(k+1)\tau}(\lambda\tau)^n (k+1)^n}{n!}\left[xN(\hat{d}_1) - Ke^{-r_n\tau} N(\hat{d}_2)\right]. \end{aligned}$$

This result it may be more succinctly written

$$f(x,\tau)=\sum_{n=0}^{\infty}\frac{e^{-\lambda'\tau}(\lambda\tau)^n}{n!}M_n(x,\tau)$$

where we set

$$\lambda'=\lambda(1+k),\quad M_n(x,\tau)=M(x,\tau;K,v_n^2,r_n),$$

and

$$M(x,\tau;K,\sigma^2,r)=xN(d_1)-Ke^{-r\tau}N(d_2),$$

with

$$d_1=\frac{\ln\frac{x}{K}+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}},\quad \text{and}\quad d_2=d_1-\sigma\sqrt{\tau}.$$

■

## 14.6 Appendix

### Appendix 14.1 Calculating the $A_1$ and $A_2$

*The Evaluation* $A_1$. Applying the result (14.32) we find that

$$A_1=\mathbb{E}^n\{X_nN(d_1)\}=\frac{1}{2\pi\delta\sqrt{n}}\int_0^{\infty}\int_{-\infty}^{d_1}e^{-\frac{1}{2}\left(\frac{\ln X_n-n(\gamma-\delta^2/2)}{\delta\sqrt{n}}\right)^2-\frac{u^2}{2}}dudX_n.$$

Introduce the changes of variable $y=\ln X_n$ and $v=u-\ln X_n/(\sigma\sqrt{\tau})$ we readily find that

$$A_1=\frac{1}{2\pi\delta\sqrt{n}}\int_{-\infty}^{\infty}\int_{-\infty}^{\bar{d}_1}\exp\left\{-\frac{1}{2}\left(v+\frac{y}{\sigma\sqrt{\tau}}\right)^2\right\}$$
$$\times\exp\left\{-\frac{1}{2}\left(\frac{y-n(\gamma-\frac{\delta^2}{2})}{\delta\sqrt{n}}\right)^2+y\right\}dvdy,$$

where

$$\bar{d}_1 \equiv \frac{\ln \frac{x}{K} + \left(r - \lambda k + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}.$$

Since $\bar{d}_1$ does not depend on $y$, we can readily change the order of the integration, to obtain

$$\begin{aligned}
A_1 &= \frac{1}{2\pi\delta\sqrt{n}} \int_{-\infty}^{\bar{d}_1} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(v^2 + \frac{2vy}{\sigma\sqrt{\tau}} + \frac{y^2}{\sigma^2\tau}\right)\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\left(\frac{y^2 - 2ny(\gamma - \frac{\delta^2}{2}) + n^2(\gamma - \frac{\delta^2}{2})^2}{\delta^2 n}\right) + y\right\} dydv \\
&= \frac{1}{2\pi\delta\sqrt{n}} \int_{-\infty}^{\bar{d}_1} \exp\left\{-\frac{1}{2}\left(v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}\right)\right\} \\
&\qquad \times \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(y^2\left[\frac{\sigma^2\tau + \delta^2 n}{\sigma^2\tau\delta^2 n}\right] + y\left[\frac{2v\delta^2 - 2\gamma\sigma\sqrt{\tau} - \sigma\sqrt{\tau}\delta^2}{\sigma\sqrt{\tau}\delta^2}\right]\right)\right\} dydv.
\end{aligned}$$

Setting

$$\alpha = \frac{\sigma^2\tau + \delta^2 n}{\sigma^2\tau\delta^2 n}, \quad \beta_1 = \frac{2v\delta^2 - 2\gamma\sigma\sqrt{\tau} - \sigma\sqrt{\tau}\delta^2}{\sigma\sqrt{\tau}\delta^2} \tag{14.35}$$

and completing the square in $y$, we have

$$\begin{aligned}
A_1 = &\frac{1}{2\pi\delta\sqrt{n}} \int_{-\infty}^{\bar{d}_1} \exp\left\{-\frac{1}{2}\left(v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}\right.\right. \\
&\qquad \left.\left. - \frac{\beta_1^2}{4\alpha}\right)\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(\alpha\left(y + \frac{\beta_1}{2\alpha}\right)^2\right)\right\} dydv.
\end{aligned}$$

The two integrations are now independent. Applying the change of variable, $z = \sqrt{\alpha}\left(y + \frac{\beta_1}{2\alpha}\right)$, to the second integral

$$A_1 = \frac{1}{\delta\sqrt{2\pi n\alpha}} \exp\left\{-\frac{1}{2}\left(\frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}\right)\right\} \int_{-\infty}^{\bar{d}_1} \exp\left\{-\frac{1}{2}\left[v^2 - \frac{\beta_1^2}{4\alpha}\right]\right\} dv.$$

Thus[3] we obtain

$$A_1 = \frac{1}{\delta\sqrt{2\pi n\alpha}} \exp\left\{-\frac{1}{2}\frac{n(\gamma-\frac{\delta^2}{2})^2}{\delta^2}\right\} \exp\left\{\frac{1}{2}\frac{n(\sigma^2\tau+\delta^2 n)(\gamma+\frac{\delta^2}{2})^2}{\delta^2(\sigma^2\tau+\delta^2 n)}\right\}$$
$$\times \int_{-\infty}^{\bar{d}_1} \exp\left\{-\frac{1}{2}\left[\frac{\left(\sigma\sqrt{\tau}v+n(\gamma+\frac{\delta^2}{2})\right)^2}{\sigma^2\tau+\delta^2 n}\right]\right\} dv.$$

By making the change of variable $\xi = \left(\sigma\sqrt{\tau}v + n(\gamma+\frac{\delta^2}{2})\right)/(\sqrt{\sigma^2\tau+\delta^2 n})$ and setting

$$\hat{d}_1 = \frac{\ln\frac{x}{K} + \left(r_n + \frac{v_n^2}{2}\right)\tau}{v_n\sqrt{\tau}},$$

where $r_n \equiv r - \lambda k + n\gamma/\tau$ and $v_n^2 = \sigma^2 + n\delta^2/\tau$, we obtain

$$A_1 = \frac{1}{\delta\sqrt{2\pi n\alpha}}\frac{\sqrt{\sigma^2\tau+\delta^2 n}}{\sigma\sqrt{\tau}} d\xi \exp\left\{-\frac{1}{2}\left(\frac{n(\gamma-\frac{\delta^2}{2})^2}{\delta^2} - \frac{n(\gamma+\frac{\delta^2}{2})}{\delta^2}\right)\right\}$$
$$\times \int_{-\infty}^{\hat{d}_1} \exp\left\{-\frac{\xi^2}{2}\right\} d\xi.$$

Using the definition of $\alpha$ in (14.35),

$$\begin{aligned} A_1 &= \frac{\sqrt{\sigma^2\tau}+\delta^2 n}{\delta\sigma\sqrt{n\tau}}\frac{\sigma\delta\sqrt{n\tau}}{\sqrt{\sigma^2\tau+\delta^2 n}} \exp\left\{-\frac{n}{2\delta^2}\left(\gamma^2-\delta^2\gamma+\frac{\delta^4}{4}-\gamma^2-\delta^2\gamma-\frac{\delta^4}{4}\right)\right\} \\ &\quad\times\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\hat{d}_2} e^{-\frac{\xi^2}{2}} d\xi \\ &= e^{n\gamma}\mathcal{N}(\hat{d}_1). \end{aligned} \tag{14.36}$$

[3]Referring to (14.35) we see that

$$\begin{aligned} v^2 - \frac{\beta_1^2}{4\alpha} &= v^2 - \frac{n\left(v^2\delta^4 - 2v\delta^2\sigma\sqrt{\tau}(\gamma+\frac{\delta^2}{2}) + \sigma^2\tau(\gamma+\frac{\delta^2}{2})^2\right)}{\delta^2(\sigma^2\tau+\delta^2 n)} \\ &= \frac{\sigma^2\tau v^2 + 2n\sigma\sqrt{\tau}(\gamma+\frac{\delta^2}{2})v}{\sigma^2\tau+\delta^2 n} - \frac{\sigma^2\tau n(\gamma+\frac{\delta^2}{2})^2}{\delta^2(\sigma^2\tau+\delta^2 n)} \\ &= \frac{\left(\sigma\sqrt{\tau}v+n(\gamma+\frac{\delta^2}{2})\right)^2}{\sigma^2\tau+\delta^2 n} - \frac{n^2(\gamma+\frac{\delta^2}{2})^2}{\sigma^2\tau+\delta^2 n} - \frac{\sigma^2\tau n(\gamma+\frac{\delta^2}{2})^2}{\delta^2(\sigma^2\tau+\delta^2 n)}. \end{aligned}$$

*The evaluation of* $A_2$. By definition and use again of the result (14.32)

$$A_2 = \mathbb{E}^n\{N(d_2)\}$$
$$= \frac{1}{\delta\sqrt{2\pi n}}\int_0^\infty \int_{-\infty}^{d_2} \frac{1}{X_n} e^{-\frac{1}{2}\left(\frac{\ln X_n - n(\gamma - \frac{\delta^2}{2})}{\delta\sqrt{n}}\right)^2 - \frac{u^2}{2}} du dX_n.$$

Applying the change of variable $y = \ln X_n, \quad v = u - \ln X_n/\sigma\sqrt{\tau}$, and setting

$$\bar{d}_2 = \frac{\ln \frac{x}{K} + (r - \lambda k - \sigma^2/2)\tau}{\sigma\sqrt{\tau}},$$

then

$$A_2$$
$$= \frac{1}{2\pi\delta\sqrt{n}}\int_{-\infty}^{\infty}\int_{-\infty}^{\bar{d}_2} \exp\left\{-\frac{1}{2}\left(v + \frac{y}{\sigma\sqrt{\tau}}\right)^2\right\}\exp\left\{-\frac{1}{2}\left(\frac{y - n(\gamma - \frac{\delta^2}{2})}{\delta\sqrt{n}}\right)^2\right\} dv dy$$
$$= \frac{1}{2\pi\delta\sqrt{n}}\int_{-\infty}^{\bar{d}_2}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(v + \frac{y}{\sigma\sqrt{\tau}}\right)^2\right\}\exp\left\{-\frac{1}{2}\left(\frac{y - n(\gamma - \frac{\delta^2}{2})}{\delta\sqrt{n}}\right)^2\right\} dy dv$$
$$= \frac{1}{2\pi\delta\sqrt{n}}\int_{-\infty}^{\bar{d}_2} \exp\left\{-\frac{1}{2}\left(v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}\right)\right\}$$
$$\times \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(y^2\left(\frac{1}{\sigma^2\tau} + \frac{1}{\delta^2 n}\right) + y\left(\frac{2v}{\sigma\sqrt{\tau}} - \frac{2(\gamma - \frac{\delta^2}{2})}{\delta^2}\right)\right)\right\} dy dv.$$

Setting

$$\beta_2 = \frac{2\left(v\delta^2 - \sigma\sqrt{\tau}(\gamma - \frac{\delta^2}{2}\right)}{\sigma\sqrt{\tau}\delta^2}$$

we can write

$$A_2 = \frac{1}{2\pi\delta\sqrt{n}}\int_{-\infty}^{\bar{d}_2} \exp\left\{-\frac{1}{2}\left(v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}\right)\right\}$$
$$\times \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}(\alpha y^2 + \beta_2 y)\right\} dy dv$$

$$= \frac{1}{2\pi\delta\sqrt{n}} \int_{-\infty}^{\bar{d}_2} \exp\left\{ -\frac{1}{2}\left( v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2} - \frac{\beta_2^2}{4\alpha} \right) \right\}$$
$$\times \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2}\left( \alpha \left( y + \frac{\beta_2}{2\alpha} \right)^2 \right) \right\} dydv.$$

Setting $z = \sqrt{\alpha}\left( y + \frac{\beta_2}{2\alpha} \right)$,

$$A_2 = \frac{1}{2\pi\delta\sqrt{n}} \int_{-\infty}^{\bar{d}_2} \exp\left\{ -\frac{1}{2}\left( v^2 + \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2} - \frac{\beta_2^2}{4\alpha} \right) \right\} \int_{-\infty}^{\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{\alpha}} dz$$
$$= \frac{1}{\delta\sqrt{2\pi n\alpha}} \exp\left\{ -\frac{1}{2}\left( \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2} \right) \right\} \int_{-\infty}^{\bar{d}_2} \exp\left\{ -\frac{1}{2}\left[ v^2 - \frac{\beta_2^2}{4\alpha} \right] \right\} dv.$$

So that[4]

$$A_2 = \frac{1}{\delta\sqrt{2\pi n\alpha}} \exp\left\{ -\frac{1}{2}\left( \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2} \right) \right\}$$
$$\times \int_{-\infty}^{\bar{d}_2} \exp\left\{ -\frac{1}{2}\left[ \left( \frac{\sigma\sqrt{\tau}v + n(\gamma - \frac{\delta^2}{2})}{\sqrt{\sigma^2\tau + \delta^2 n}} \right)^2 \right] \right\} \exp\left\{ -\frac{1}{2}\left[ \frac{-n(\gamma - \frac{\delta^2}{2})^2}{\delta^2} \right] \right\} dv$$
$$= \frac{1}{\delta\sqrt{2\pi n}} \int_{-\infty}^{\bar{d}_2} \exp\left\{ -\frac{1}{2}\left( \frac{\sigma\sqrt{\tau}v + n(\gamma - \frac{\delta^2}{2})}{\sqrt{\sigma^2\tau + \delta^2 n}} \right)^2 \right\} dv.$$

Making the change of variable $\xi = (\sigma\sqrt{\tau}v + n(\gamma - \frac{\delta^2}{2}))/\sqrt{\sigma^2\tau + \delta^2 n}$ and defining

$$\hat{d}_2 = \frac{\ln\frac{x}{K} + \left( r_n - \frac{1}{2}v_n^2 \right)\tau}{v_n\sqrt{\tau}},$$

then we finally have

$$A_2 = \mathcal{N}(\hat{d}_2). \tag{14.37}$$

[4]We note that

$$v^2 - \frac{\beta_2^2}{4\alpha} = \frac{\sigma^2\tau v^2 + 2\sigma\sqrt{\tau}n(\gamma - \frac{\delta^2}{2})v}{\sigma^2\tau + \delta^2 n} - \frac{\sigma^2\tau n(\gamma - \frac{\delta^2}{2})^2}{\delta^2(\sigma^2\tau + \delta^2 n)}$$
$$= \frac{\left( \sigma\sqrt{\tau}v + n(\gamma - \frac{\delta^2}{2}) \right)^2}{\sigma^2\tau + \delta^2 n} - \frac{n(\gamma - \frac{\delta^2}{2})^2}{\delta^2}.$$

# Chapter 15
# Stochastic Volatility

**Abstract** Empirical studies show that the volatility of asset returns are not constant and the returns are more peaked around the mean and have fatter tails than implied by the normal distribution. These empirical observations have led to models in which the volatility of returns follows a diffusion process. In this chapter, we introduce some stochastic volatility models and consider option prices under stochastic volatility. In particular, we consider the solutions of the option pricing when volatility follows a mean-reverting diffusion process. We also introduce the Heston model, one of the most popular stochastic volatility models.

## 15.1 Introduction

The Black–Scholes model for the price of options on common stock that we derived in Chaps. 6 and 8, as well as, the Garman–Kohlhagen formula and Black's model in Chap. 10, all relied upon the assumption that $\sigma$, the volatility of returns on the underlying asset is constant. This means that the transition probability density function for the underlying asset is log-normal, which in turn enables us to perform analytically the integral expressions for the option price, or to solve analytically the pricing partial differential equations. However, it is well known from many empirical studies (e.g. Blattberg and Gonedes 1974; Scott 1987) that volatility is not constant. It is a feature of many financial assets that returns are more peaked around the mean and have fatter tails than would be implied by a normal distribution for returns.

Figure 15.1 plots the distribution of daily returns on the $S\&P$ 500 from May 1994 to 1999 and compares this with a normal distribution based on the mean and variance of the same data set. Figure 15.2 plots the distribution of the percentage changes in daily $US\$/£$ exchange rate over the same period. Both Figs. 15.1 and 15.2 illustrate the point made earlier about fat tails and the peaking around the mean.

Empirical studies of implied volatility (the volatility that makes the Black–Scholes model yield observed market option prices) also clearly show that the assumption of constant volatility is far for being realised. Figure 15.3 shows the



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_15

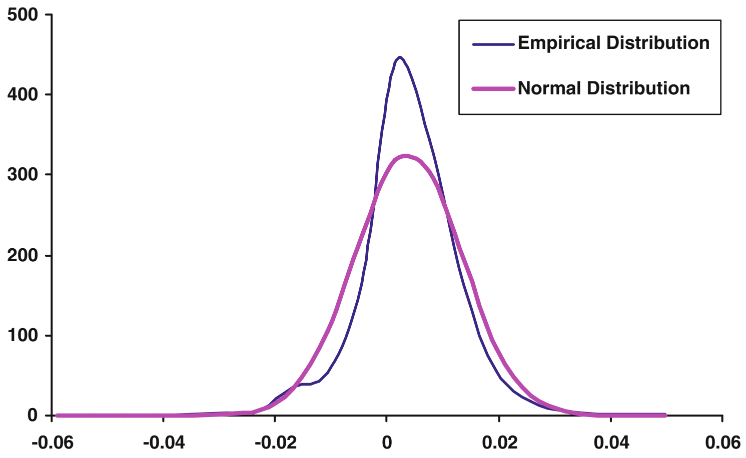


**Fig. 15.1** Distribution of daily returns on the $S\&P$ 500

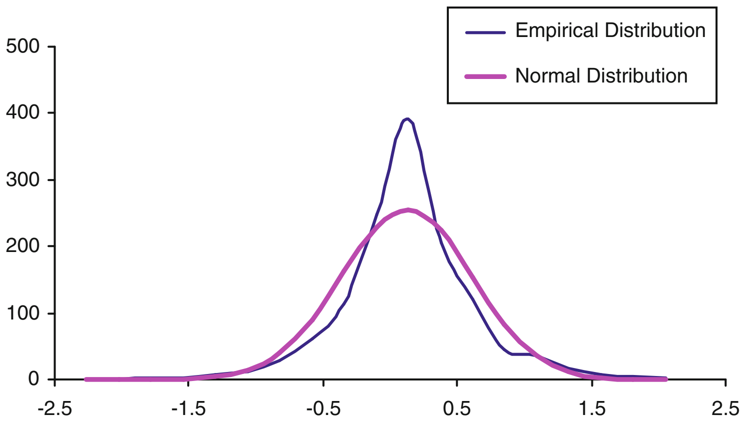


**Fig. 15.2** Distribution of percentage change in daily $US\$/£$ exchange rate

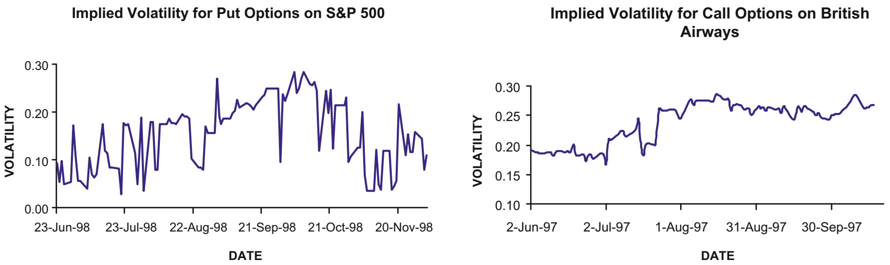


**Fig. 15.3** Implied volatilities for options

implied volatility for put options on the $S\&P$ 500 and for call options on British Airways. From both of these figures we see that implied volatility moves around quite a lot. Indeed it seems to have the characteristics of a mean reverting stochastic process.

The fat tails indicate a much larger probability of extreme movements in returns and the peaking at the centre more of a gathering around the mean that is allowed for by the normal distribution. These empirical observations have led a number of researchers including Johnson and Shanno (1987), Scott (1987), Hull and White (1987), Wiggins (1987) and Heston (1993) to consider models in which the volatility (or the variance) of returns follows a diffusion process. It is hoped thereby to capture some of the empirical features observed in financial markets.

To make the point that such an approach can indeed reproduce the fat tails and peakedness observed empirically, consider the stock price process

$$\frac{dS}{S} = \mu dt + \sigma dz_s, \tag{15.1}$$

where $\sigma$ itself follows the diffusion process

$$d\sigma = k(\overline{\sigma} - \sigma)dt + \delta dz_\sigma, \tag{15.2}$$

and $\mathbb{E}(dz_s dz_\sigma) = \rho\, dt$.

Using the values $\mu = 0.15$, $k = 1$, $\overline{\sigma} = 0.2$, $\delta = 0.2$ and $\rho = 0$ we have simulated (15.1) and (15.2) 10,000 times. Figure 15.4 illustrates the sample paths of $S$ for the case of constant $\sigma(= \overline{\sigma})$ and when $\sigma$ is stochastic according to (15.2). The wider distribution of sample paths in the stochastic volatility case suggest that the distribution is fat tailed. This is more clearly illustrated in Fig. 15.5 where the two distributions are compared. In addition, the peakedness of the distribution under stochastic volatility is also evident.

In Fig. 15.6 we illustrate a sample path for $\sigma$ and it is interesting to compare it with the pattern of observed implied volatilities in Fig. 15.3.

Some authors have suggested modelling the variance $v \equiv \sigma^2$ as a stochastic process, for example Heston (1993) and Stein and Stein (1991). In the context of (15.2), a version of the model considered by Heston can be obtained by setting $\overline{\sigma} = 0$ and using of Ito's Lemma to obtain for $v$ the stochastic differential equation

$$dv = (\delta^2 - 2kv)dt + 2\delta\sqrt{v}dz_\sigma. \tag{15.3}$$

The fact that this is a mean reverting stochastic process with a volatility dependent on $\sqrt{v}$ allows particular solution techniques to be employed when evaluating options. This is a point to which we return in a later section.

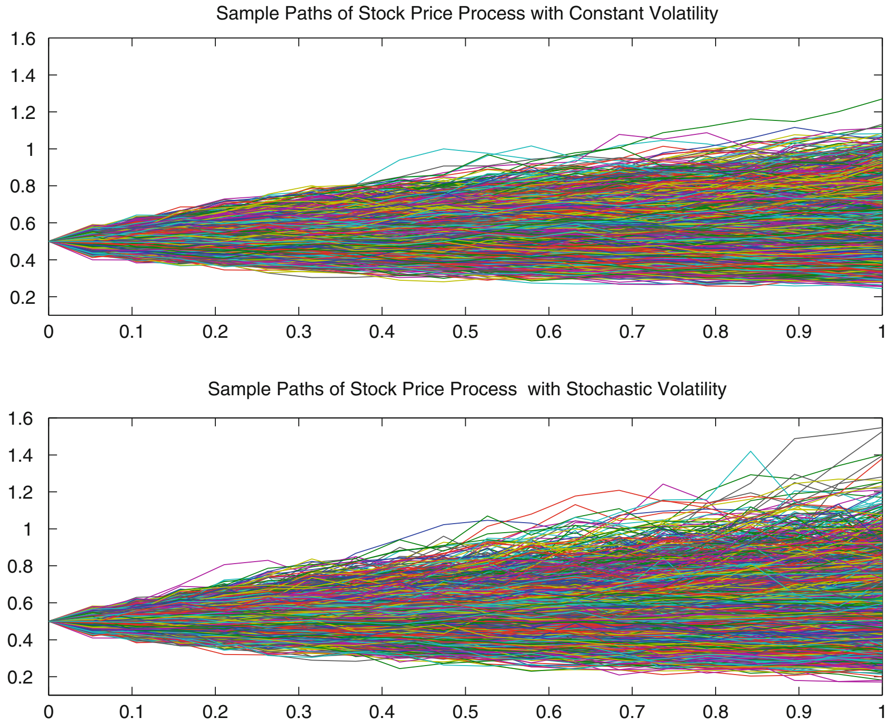


**Fig. 15.4** Sample paths of 10,000 simulations with constant (*upper*) and stochastic (*lower*) volatilities

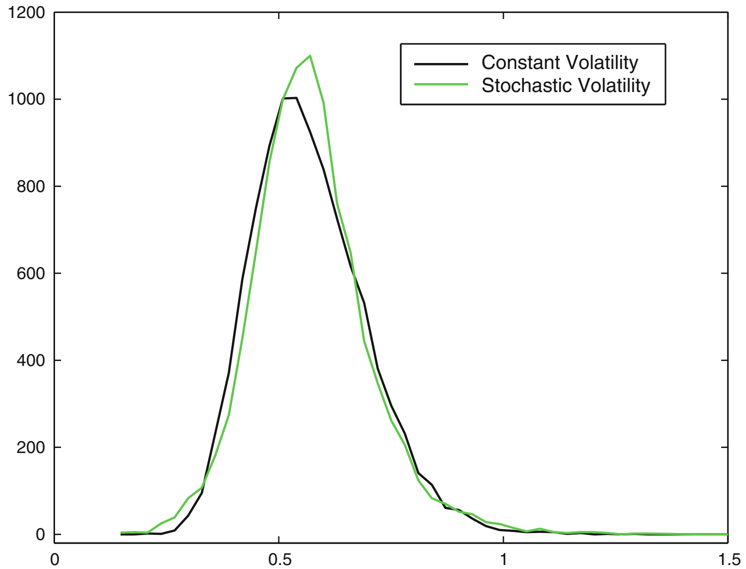


**Fig. 15.5** Comparison of constant and stochastic volatility distributions

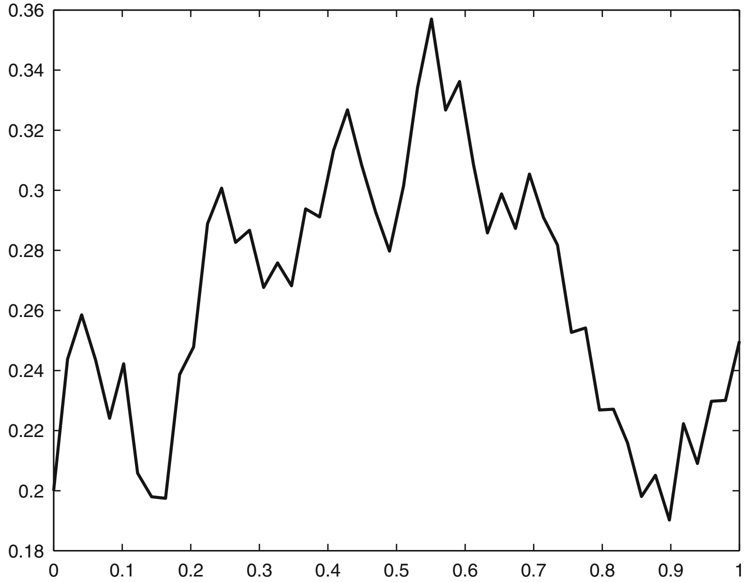


**Fig. 15.6** Sample path for $\sigma$

## 15.2 Modelling Stochastic Volatility

In our discussion of stochastic volatility in this chapter we maintain the assumption that the underlying asset price (typically this will be common stock, an index, foreign exchange, futures contracts or commodities) follows the now familiar log-normal diffusion process

$$\frac{dS}{S} = \mu dt + \sigma dz_s. \tag{15.4}$$

We use the subscript $s$ in $dz_s$ to indicate the source of uncertainty impinging on the asset market. We assume the volatility, $\sigma$, is driven by the diffusion process

$$d\sigma = a(\sigma, t)dt + b(\sigma, t)dz_\sigma, \tag{15.5}$$

whose drift and diffusion coefficients $a$ and $b$ respectively depend only on $\sigma$ itself and time $t$. Of course, in a more general treatment these coefficients could also be allowed to depend on the asset price $S$, however all stochastic volatility models developed to date do not make this assumption, so we shall retain the general forms specified in (15.5). The Wiener process $z_\sigma$ (the source of uncertainty impinging on the asset price volatility) is possibly correlated with $z_s$ (the source of uncertainty impinging on the asset price), that is we assume

$$\mathbb{E}(dz_s dz_\sigma) = \rho\, dt. \tag{15.6}$$

As in Sect. 5.3.1, it will be more convenient to work in terms of independent Wiener processes whose increments we denote $dw_s$ and $dw_\sigma$. We know from (5.42) that

$$dz_s = dw_s,$$
$$dz_\sigma = \rho\, dw_s + \sqrt{1-\rho^2}\, dw_\sigma,$$

and so the stochastic dynamic system (15.4)–(15.5) may alternatively be written as

$$\frac{dS}{S} = \mu dt + \sigma dw_s, \tag{15.7}$$

$$d\sigma = a(\sigma,t)dt + \rho\, b(\sigma,t)dw_s + \sqrt{1-\rho^2}\, b(\sigma,t)\, dw_\sigma. \tag{15.8}$$

By an appropriate choice of the functions $a, b$ and the correlation coefficient $\rho$, the specification (15.7)–(15.8) encompasses most of the stochastic volatility models developed in the literature.

Some authors prefer to model the variance $v(\equiv \sigma^2)$ as a diffusion process. By Ito's lemma (see (6.78)) the stochastic differential equation for $v$ implied by (15.8) is

$$dv = (2\sqrt{v}a + b^2)dt + 2\rho\sqrt{v}\, b\, dw_s + 2\sqrt{1-\rho^2}\sqrt{v}\, b\, dw_\sigma. \tag{15.9}$$

We shall return to processes for the variance when we consider the Heston model in Sect. 15.5. Let us now consider some specific forms for the coefficients $a$ and $b$ in (15.8). Almost invariably in the stochastic volatility literature, the drift coefficient $a(\sigma,t)$ has been chosen as

$$a(\sigma,t) = k(\overline{\sigma} - \sigma), \quad (k > 0), \tag{15.10}$$

which implies a mean reverting process for $\sigma$. Such behaviour is suggested by the time series behaviour of the implied volatility in Fig. 15.3. The mean reverting form of (15.8) is

$$d\sigma = k(\overline{\sigma} - \sigma)dt + \rho\, b(\sigma,t)dw_s + \sqrt{1-\rho^2}\, b(\sigma,t)\, dw_\sigma. \tag{15.11}$$

It is instructive to rewrite (15.11) in the form (by use of Ito's lemma),

$$d(\sigma e^{kt}) = k\overline{\sigma}e^{kt}dt + e^{kt}\rho\, b(\sigma,t)dw_s + e^{kt}\sqrt{1-\rho^2}\, b(\sigma,t)\, dw_\sigma,$$

so that

$$\sigma(t) = \overline{\sigma} + (\sigma_0 - \overline{\sigma})e^{-kt} + \rho \int_0^t e^{-k(t-\tau)} b(\sigma, \tau)\, dw_s(\tau) + \sqrt{1-\rho^2} \int_0^t e^{-k(t-\tau)} b(\sigma, \tau)\, dw_\sigma(\tau). \tag{15.12}$$

First note that, if $b = 0$ so that volatility is not stochastic, then $\sigma$ is a time varying deterministic function (which tends to $\overline{\sigma}$ fairly quickly). We would then be back to a model equivalent to one with a constant volatility (see Chap. 7). When $b \neq 0$ and we have a proper stochastic volatility model, the volatility can be viewed as having three components. The first component is a deterministic function of time. The second is an integral over past shocks impinging on the asset price. The third term is an integral over past shocks impinging on the volatility. Both integrals are weighted with the exponentially declining term, $e^{-k(t-\tau)}$, as well as, a function of past values of the volatility via the $b(\sigma, \tau)$ term. These integral terms show clearly the path dependence of the volatility $\sigma$, which is transmitted to the diffusion process for $S$. Thus the stochastic process for $S$ is non-Markovian. However, the form of this non-Markovian process is such that it can be written as a two-dimensional Markovian system. This is yet another example of a non-Markovian process that can be re-expressed as a Markovian system by a suitable expansion of the state space. It is also instructive to consider a discretised version of (15.12). Before discretising we set $\sigma_0 = \overline{\sigma}$ and note that by a simple change of variable

$$\int_0^t e^{-k(t-\tau)} b(\sigma, \tau) dw(\tau) = -\int_t^0 e^{-ku} b(\sigma, t-u) dw(t-u). \tag{15.13}$$

Thus using the simplest possible discretisation scheme and taking $n\Delta\tau = t$ and $\alpha_i = e^{-ki\Delta\tau}$ we can write

$$\int_0^t e^{-k(t-\tau)} b(\sigma, \tau) dw(\tau) \simeq -\sum_{i=1}^{n} \alpha_i b(\sigma_{t-i}, (n-i)\Delta\tau) \sqrt{\Delta\tau}\, \tilde{\xi}_{t-i}, \tag{15.14}$$

where $\tilde{\xi}_{t-i} \sim N(0, 1)$ is the shock term over the time interval commencing at $(t - i\Delta\tau)$, as shown in Fig. 15.7.

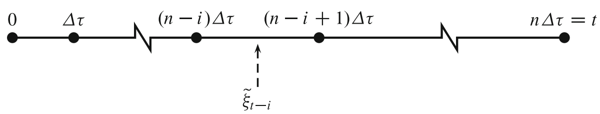


**Fig. 15.7** Time line for discretisation of the integral in (15.13)

With these notations the discretisation of (15.12) would become

$$\sigma_t = \sigma_0 - \rho \sum_{i=1}^{n} \alpha_i b(\sigma_{t-i}, (n-i)\Delta\tau)\sqrt{\Delta\tau}\,\tilde{\xi}_{S,N-i} \\ - \sqrt{1-\rho^2}\sum_{i=1}^{n} \alpha_i b(\sigma_{t-i}, (n-i)\Delta\tau)\sqrt{\Delta\tau}\tilde{\xi}_{\sigma,N-i}, \tag{15.15}$$

and $\tilde{\xi}_{S,t-i}$, $\tilde{\xi}_{\sigma,t-i}$ represent two sets of independent $N(0,1)$ shock terms. Thus, in discretised form, we have for the asset price process

$$\ln \frac{S_t}{S_{t-\Delta t}} = (\mu - \frac{1}{2}\sigma_t^2)\Delta t + \sigma_{t-\Delta t}\sqrt{\Delta t}\,\tilde{\xi}_{S,t}, \tag{15.16}$$

with $\sigma_t$ given by (15.15), which as we see depends on past values of $\sigma_t$ and past values of the shock terms. This representation of the asset price is somewhat suggestive of the ARCH and GARCH representations of the econometrics literature (see for example Bollerslev 1986), however it is the variance $\sigma_t^2$ rather than $\sigma_t$ which is modelled in this literature. In the ARCH and GARCH literature the dependence on past values of $\sigma_t$ and the shock terms occurs additively rather than multiplicatively as in (15.15). If we were to model $\sigma_t$ using the GARCH $(p, q)$ process then (15.15) would be replaced by

$$\sigma_t^2 = \sigma_0 + \sum_{i=1}^{q} \gamma_i \xi_{S,t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2. \tag{15.17}$$

We note that the squared past volatility shocks, $\xi_{\sigma,t-i}^2$, do not appear. The specifications (15.15) and (15.17) for $\sigma_t$ quite clearly differ in many respects but they both share the common feature of modelling volatility as a function of past shocks and past levels of volatility. It remains an open empirical question, as to, which of these two different specifications is best from the point of view of option pricing. Nelson (1991) has shown how ARCH models may be interpreted as approximations to diffusion processes.

## 15.3 Option Pricing Under Stochastic Volatility

We assume that the underlying asset price $S$ and volatility $\sigma$ are driven by the diffusion processes (15.7)–(15.8). We consider the problem of pricing a derivative security dependent on $S$ and $\sigma$ and whose value we denote $f(S, \sigma, t)$. We are now precisely in the situation of one traded asset ($S$) and one non-traded asset ($\sigma$) which we analysed is Sect. 10.3.3. We know from that analysis that the partial differential equation (10.59) determines the derivative price. Using $q$ to denote the continuous dividend yield on the underlying asset, and $\lambda_s$ and $\lambda_\sigma$ to denote respectively the market prices of risk associated with the sources of uncertainty $dw_s$ and $dw_\sigma$, that partial differential (10.59) in the current context becomes

$$\frac{\partial f}{\partial t}+(r-q)S\frac{\partial f}{\partial S}+(a(\sigma,t)-\rho b(\sigma,t)\lambda_s-\sqrt{1-\rho^2}b(\sigma,t)\lambda_\sigma)\frac{\partial f}{\partial \sigma}+\mathfrak{D}f-rf=0, \tag{15.18}$$

where

$$\mathfrak{D}f=\frac{1}{2}\sigma^2S^2\frac{\partial^2 f}{\partial S^2}+\rho b\sigma S\frac{\partial^2 f}{\partial S\partial\sigma}+\frac{1}{2}b^2\frac{\partial^2 f}{\partial\sigma^2}. \tag{15.19}$$

Alternatively, using (10.62), the derivative value may be calculated from

$$f(S,\sigma,t)=e^{-r(T-t)}\tilde{\mathbb{E}}_t[f(S_T,\sigma_T,T)], \tag{15.20}$$

where $f(S_T,\sigma_T,T)$ is the pay-off function at maturity $T$. Here $\tilde{\mathbb{E}}_t$ is calculated from the conditional density function generated by the two-dimensional diffusion process

$$\begin{aligned}\frac{dS}{S}&=(r-q)dt+\sigma d\tilde{w}_s,\\ d\sigma&=(a-\lambda_s\rho b-\lambda_\sigma\sqrt{1-\rho^2}b)dt+\rho bd\tilde{w}_s+\sqrt{1-\rho^2}bd\tilde{w}_\sigma,\end{aligned} \tag{15.21}$$

and $\tilde{w}_s$, $\tilde{w}_\sigma$ represent Wiener processes under the equivalent probability measure $\tilde{\mathbb{P}}$.

In principle at least the stochastic volatility option pricing problem may be considered solved, as all we need to do is to solve the two-dimensional partial differential equation (15.18) or use Monte-Carlo simulation to evaluate (15.20). In practice both of these tasks present formidable computational challenges and are still the object of ongoing research. For this reason, most of the developments in this area have focused on considering specifications of $a(\sigma,t)$ and $b(\sigma,t)$ that allow explicit solutions of the partial differential (15.18). We consider one such solution in the following section.

Before considering specific solutions we need to consider further the market prices of risk, $\lambda_s$ and $\lambda_\sigma$. The first market price of risk, $\lambda_s$, is associated with the shocks impinging on the asset price. It is common to assume that, as it is related to a traded asset, this risk could be diversified away if investors held sufficiently diversified portfolios. Hence, we set $\lambda_s=0$. The second market price of risk, $\lambda_\sigma$, cannot be diversified away to easily as generally volatility is a non-traded asset,[1] so it remains in the analysis. For reasons of tractability one common form that is

[1] Of course in recent years there has been the growth of the so-called VIX options. There are options on the so-called VIX implied volatility index. Within this framework volatility has become a traded factor. We shall discuss VIX option in a later section.

assumed is

$$\lambda_\sigma = \lambda\,\sigma, \tag{15.22}$$

where $\lambda$ is a constant. The assumption (15.22) also leads to the partial differential equation for the option price having coefficients that are linear in $\sigma$ which is convenient for some solution techniques. It implies that the risk premium demanded for bearing volatility risk is proportional to the level of volatility. We shall use (15.22) in much of the subsequent analysis.

## 15.4 The Mean Reverting Volatility Case

In this section we consider the solution of the pricing partial differential equation (15.18) in the case where volatility is driven by the mean reverting diffusion process (15.11) with the diffusion coefficient $b(\sigma, t)$, a constant, which we denote by $b$. Thus under the equivalent probability measure $\tilde{\mathbb{P}}$ and assuming $\lambda_s = 0$ and $\lambda_\sigma = \lambda\sigma$, the two-dimensional diffusion process (15.21) assumes the form

$$\begin{aligned} \frac{dS}{S} &= (r-q)dt + \sigma d\tilde{w}_s, \\ d\sigma &= [k\overline{\sigma} - (k + \lambda b\sqrt{1-\rho^2})\sigma]dt + \rho b d\tilde{w}_s + \sqrt{1-\rho^2}\, b d\tilde{w}_\sigma. \end{aligned} \tag{15.23}$$

The partial differential equation (15.18) under specification (15.23) thus becomes

$$\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + [k\overline{\sigma} - (k + \lambda b\sqrt{1-\rho^2})\sigma]\frac{\partial f}{\partial \sigma} + \mathfrak{D}f - rf = 0, \tag{15.24}$$

with $\mathfrak{D}f$ still given by (15.19) but $b$ taken as a constant.

We note that it is also possible to express the two-dimensional diffusion process (15.4) and (15.5) (with drift coefficient $a(\sigma, t) = k(\overline{\sigma} - \sigma)$ and constant diffusion coefficient), in terms of independent Wiener processes under the risk neutral measure in the form (in conditions (5.34)–(5.36) set $a_{11} = \sqrt{1-\rho^2}, a_{12} = \rho, a_{21} = 0, a_{22} = 1$)

$$\frac{dS}{S} = (r-q)dt + \sigma\sqrt{1-\rho^2}\, d\tilde{w}_s + \sigma\rho\, d\tilde{w}_\sigma, \tag{15.25}$$

$$d\sigma = [k\overline{\sigma} - (k + \lambda b)\sigma]dt + b d\tilde{w}_\sigma. \tag{15.26}$$

Note that with this specification we would have $s_{21} = 0$ in (10.64), and hence the change in the drift term in (15.26). Furthermore, since the $s_{21}$ terms involves the market price of risk $\lambda_s$, there is then no longer any need to make any assumption about the value of this factor since it no longer appears explicitly in the stochastic dynamics. Using this formulation, the diffusion process for $\sigma$ evolves independently of the diffusion process for $S$, because of the independence of the Wiener processes $\tilde{w}_s$ and $\tilde{w}_\sigma$. This fact will be significant in simplifying calculations at a later point.

We also note for later reference, that by application of Ito's lemma

$$d(\ln S) = [(r-q) - \frac{1}{2}\sigma^2]dt + \sigma\sqrt{1-\rho^2}d\tilde{w}_s + \sigma\rho d\tilde{w}_\sigma, \tag{15.27}$$

$$d(\sigma^2) = [b^2 + 2k\overline{\sigma}\sigma - 2(k+\lambda b)\sigma^2]dt + 2b\sigma d\tilde{w}_\sigma. \tag{15.28}$$

We shall consider the case where the derivative security is a European call option so that the payoff function is given by

$$f(S_T, \sigma_T, T) = (S_T - E)^+. \tag{15.29}$$

Substituting the payoff function into (15.20) we have

$$\begin{aligned} f(S,\sigma,t) &= e^{-r(T-t)}\tilde{\mathbb{E}}_t[(S_T - E)^+] \\ &= e^{-r(T-t)}\mathbb{E}_t[(S_T - E)\cdot \mathbb{1}_{\{S_T > E\}}], \end{aligned} \tag{15.30}$$

where $\mathbb{1}_{\{S_T>E\}}$ is the indicator function. The solution of (15.24) can be written

$$f(S,\sigma,t) = \tilde{\mathbb{E}}_t[e^{-r(T-t)}S_T\mathbb{1}_{\{S_T>E\}}] - e^{-r(T-t)}E\,\tilde{\mathbb{E}}_t[\mathbb{1}_{\{S_T>E\}}]. \tag{15.31}$$

In order to employ the solution technique of characteristic functions (see Appendix 15.1) it is more convenient to express the indicator function in terms of $\ln S_T$, in terms of which (15.31) becomes

$$f(S,\sigma,t) = \tilde{\mathbb{E}}_t[e^{-r(T-t)}S_T\mathbb{1}_{\{\ln S_T>\ln E\}}] - e^{-r(T-t)}E\,\tilde{\mathbb{E}}_t[\mathbb{1}_{\{\ln S_T>\ln E\}}]. \tag{15.32}$$

Using the technique of characteristic functions we show in Appendix 15.2 that the solution can be expressed as

$$f(S,\sigma,t) = Se^{-q(T-t)}\mathscr{I}_1 - Ee^{-r(T-t)}\mathscr{I}_2, \tag{15.33}$$

where

$$\mathscr{I}_1 = \frac{1}{2} - \frac{1}{2\pi}\int_0^\infty \frac{(\phi_1(-u)e^{iux} - \phi_1(u)e^{-iux})}{iu}du, \tag{15.34}$$

$$\mathscr{I}_2 = \frac{1}{2} - \frac{1}{2\pi}\int_0^\infty \frac{(\phi(-u)e^{iux} - \phi_1(u)e^{-iux})}{iu}du, \tag{15.35}$$

$$\phi_1(u) = \frac{\psi(1+iu)}{Se^{-q(T-t)}}, \quad \phi(u) = \frac{\psi(iu)}{e^{-r(T-t)}}, \tag{15.36}$$

with

$$\psi(\zeta) = \mathscr{K}^* \exp\left(\frac{1}{2}D(t)\sigma^2 + B(t)\sigma + C(t)\right) \tag{15.37}$$

$$\mathscr{K}^* = \exp\left[\zeta \ln(S_t e^{(r-q)(T-t)}) - r(T-t) - \frac{\zeta\rho}{2b}\sigma_t^2 - \frac{\zeta\rho b}{2}(T-t)\right]. \tag{15.38}$$

and

$$\begin{aligned}
D(t) &= \frac{(k+\lambda b)}{b^2} - \frac{\gamma_1 \sin h(\gamma_1\tau) + \gamma_2 \cos h(\gamma_1\tau)}{b^2 m},\\
B(t) &= \frac{(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)(1-\cos h(\gamma_1\tau)) - (k\bar{\sigma}\gamma_1\gamma_2 - \gamma_3)\sin h(\gamma_1\tau)}{\gamma_1 b^2 m},\\
C(t) &= -\frac{1}{2}\ln m + \frac{[(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)^2 - \gamma_3^2(1-\gamma_2^2)]\sin h(\gamma_1\tau)}{2\gamma_1^3 b^2 m}\\
&\quad + \frac{(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)\gamma_3(m-1)}{\gamma_1^3 b^2 m}\\
&\quad + \frac{\tau}{2\gamma_1^2 b^2}[(k+\lambda b)\gamma_1^2(b^2 - (k+\lambda b)k\bar{\sigma}) + \gamma_3^2],\\
\gamma_1 &= \sqrt{2b^2 s_1 + (k+\lambda b)^2},\\
\gamma_2 &= \frac{k+\lambda b - 2b^2 s_3}{\gamma_1},\\
\gamma_3 &= (k+\lambda b)k\bar{\sigma} - s_2 b^2,\\
m &= \cos h(\gamma_1\tau) + \gamma_2 \sin h(\gamma_1\tau),\\
\tau &= T - t.
\end{aligned}$$

The key to evaluating (15.33) is the numerical evaluations of the complex integrals $\mathscr{I}_1$ and $\mathscr{I}_2$ defined by Eqs. (15.34) and (15.35). Most computer packages (e.g. MATLAB) have tools to do complex integration and these quantities can in fact be readily calculated. The important point is although these quantities at first glance seem complex, because of the way that they combine complex conjugate quantities they turn out in fact to be real.

## 15.5 The Heston Model

One of the most popular stochastic volatility models is that of Heston (1993). Heston models the stochastic dynamics of the variance $v$ rather than the standard deviation (volatility) $\sigma$. In terms of the historical measure $\mathbb{P}$ the dynamics of $S$ and $v$ are assumed to be of the form

$$\frac{dS}{S} = \mu dt + \sqrt{v} dz_s, \tag{15.39}$$

$$dv = k_v(\bar{v} - v)dt + \sigma_v \sqrt{v} dz_v, \tag{15.40}$$

where

$$\mathbb{E}(dz_s dz_v) = \rho dt. \tag{15.41}$$

There are two important motivations for modelling the variance $v$ with the process (15.40), where we highlight the $\sqrt{v}$ term in front of the $dz_v$. The first is that the $\sqrt{v}$ ensures that the process for $v$ remains away from zero. In fact this square-root process, as it has become known, seems to have been first investigated by Feller (1951) who showed that $v > 0$ is ensured if

$$2k_v\bar{v} > \sigma_v^2, \tag{15.42}$$

a condition that typically will be satisfied by most data sets. The condition (15.42) is sometimes known as the Feller condition. The second reason for the popularity of the Heston model is that it is possible to calculate the characteristic function of the joint distribution of $S$ and $v$, which in turn allows semi-analytical formulas to be obtained for derivative prices. This result also builds on the work of Feller (1951) who solved the Kolmogorov backward equation associated with (15.40) using transform methods to obtain an expression for the density function. We give more discussion of the square root process in Chap. 22, where it is also used to model interest rate processes (again because it can guarantee positive rates).

Transforming to independent Wiener processes as in (15.7), (15.8) the dynamics for $S$ and $v$ become

$$\frac{dS}{S} = \mu dt + \sqrt{v} dw_s, \tag{15.43}$$

$$dv = k_v(\bar{v} - v)dt + \sigma_v \sqrt{v} \rho dw_s + \sigma_v \sqrt{v}\sqrt{1-\rho^2} dw_v. \tag{15.44}$$

In order to price an option written on the stock $S$, as in Sect. 15.3, we again appeal to the general results in Chap. 10 and adapt them to the current situation. We note that now the derivative price should be denoted as $f(S, v, t)$. We assume that the market price of risk associated with the source of uncertainty $dw_s$ is the constant $\lambda_s$, whilst the market price of risk associated with the uncertainty $dw_v$ is of the form $\lambda_v\sqrt{v}$ where $\lambda_v$ is constant. Following the same argument on used in Sect. 15.3 we find that the option price satisfies the partial differential equation

$$\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + (k_v(\bar{v}-v) - \lambda_s\sigma_v\sqrt{v}\rho - \lambda_v\sigma_v v\sqrt{1-\rho^2})\frac{\partial f}{\partial v} + \mathfrak{D} f - rf = 0, \tag{15.45}$$

where

$$\mathfrak{D}f = \frac{1}{2}vS^2\frac{\partial^2 f}{\partial S^2} + \sigma_v v\rho S\frac{\partial^2 f}{\partial S\partial v} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 f}{\partial v^2}. \tag{15.46}$$

Alternatively, again using (10.61) the derivative price may also be expressed as

$$f(S, v, t) = e^{-r(T-t)}\tilde{\mathbb{E}}_t[f(S_T, v_T, T)] \tag{15.47}$$

where $f(S_T, v_T, T)$ is the payoff at maturity (typically independent of $v_T$) and $\tilde{\mathbb{E}}_t$ is calculated using the conditional density function generated by the two dimensional diffusion process

$$\frac{dS}{S} = (r-q)dt + \sqrt{v}d\tilde{w}_s, \tag{15.48}$$

$$\begin{aligned} dv =& [k_v\bar{v} - \rho\sigma_v\lambda_s\sqrt{v} - (k_v + \sigma_v\sqrt{1-\rho^2}\lambda_v)v]dt \\ &+ \sigma_v\rho\sqrt{v}d\tilde{w}_s + \sigma_v\sqrt{v}\sqrt{1-\rho^2}d\tilde{w}_v. \end{aligned} \tag{15.49}$$

The term $\rho\sigma_v\lambda_s\sqrt{v}$ in the drift makes it difficult to do any analytical calculations on the systems (15.48) and (15.49), for instance it would not be possible to obtain the characteristic function with this term present. We could argue as in Sect. 15.3 that since the stock is a traded asset its risk could be diversified away and so we may set $\lambda_s = 0$. Alternatively we could arrange the transformation from correlated to independent Wiener process so that the stochastic differential equation for $S$ has both $d\tilde{w}_s$ and $d\tilde{w}_v$ terms and the stochastic differential equation for $v$ contains only the term $d\tilde{w}_v$. In this case we need make no assumption about $\lambda_s$ since it no longer appears explicitly in the stochastic dynamics and the dynamics for $S$ and $v$ then become

$$\frac{dS}{S} = (r-q)dt + \sqrt{1-\rho^2}\sqrt{v}d\tilde{w}_s + \rho\sqrt{v}d\tilde{w}_v, \tag{15.50}$$

$$dv = [k_v\bar{v} - (k_v + \sigma_v\lambda_v)v]dt + \sigma_v\sqrt{v}d\tilde{w}_v. \tag{15.51}$$

One of the reasons for the popularity of the Heston model is that it admits semi-analytical forms for the prices of simple put and call options. This in turn is a consequence of the fact that it is possible to find the characteristic equation for the joint distribution of $S$ and $v$ given by (15.48), (15.49) or (15.50), (15.51). We give full details in Appendix 15.3.

The dynamics of $v$ under $\tilde{\mathbb{P}}$ in (15.51) can be written

$$dv = \tilde{k}_v[\tilde{\bar{v}} - v]dt + \sigma_v\sqrt{v}d\tilde{w}_v$$

where

$$\tilde{k}_v \equiv k_v + \sigma_v \lambda_v,$$
$$\tilde{\tilde{v}} = k_v \bar{v}/(k_v + \sigma_v \lambda_v).$$

The new parameters $\tilde{k}_v$, $\tilde{\tilde{v}}$ are essentially the parameters $k_v$, $\bar{v}$ appearing in (15.40) but adjusted by the market price of volatility risk $\lambda_v$. If we adopt the view that we allow calibration of the Heston model to market data to determine the $\tilde{k}_v$ and $\tilde{\tilde{v}}$ then this procedure is really equivalent to setting $\lambda_v = 0$ and re-interpreting the $k_v$ and $\bar{v}$ that are found in the calibration procedure as $\tilde{k}_v$ and $\tilde{\tilde{v}}$. This is the justification for often setting $\lambda_v = 0$ in the Heston model.

In the case of a European call option the payoff is given by

$$f(S_T, v_T, T) = (S_T - E)^+.$$

Substituting this latter expression into (15.47) and using the results referred to in Appendix 15.3 it can be shown that the option price may be written as

$$f(S, v, t) = Se^{-q(T-t)} P_1 - Ee^{-r(T-t)} P_2, \tag{15.52}$$

where $P_1$ and $P_2$ are given by

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty R\left[\frac{e^{-iu \ln E} \varphi_j(u)}{iu}\right] du, \tag{15.53}$$

with

$$\varphi_2(u) = \exp(B(T-t, u) + A(T-t, u)v + iu \ln S), \tag{15.54}$$
$$\varphi_1(u) = e^{-(r-q)(T-t) - \ln S} \varphi_2(u - i). \tag{15.55}$$

The $A$ and $B$ functions appearing in (15.54) are given by

$$A(\tau, u) = -iu(1 - iu)e_- / \gamma(u), \tag{15.56}$$
$$B(\tau, u) = -iu(r-q)\tau + \frac{k_v \bar{v}}{\sigma_v^2}(k_v - iu\rho\sigma_v - d(u))\tau + \frac{2k_v \bar{v}}{\sigma_v^2} \ln\left(\frac{2d(u)}{\gamma(u)}\right), \tag{15.57}$$

where

$$e_{\pm} = 1 \pm e^{-d(u)\tau},$$
$$d(u) = \sqrt{(\rho\sigma_v iu - k_v)^2 + iu\sigma_v^2(1 - iu)},$$
$$\gamma(u) = d(u)e_{+} + (k_v - iu\rho\sigma_v)e_{-}.$$

It turns out that the quantities $P_1$ and $P_2$ are the probabilities that $S_T > E$ under the historical (or spot) and risk-neutral probability measures respectively. The quantities $\varphi_1$ and $\varphi_2$ are the characteristic functions respectively of the distribution of $\ln S_T$ at time $t$ under the two measures.

In the foregoing sections we have only been able to give the very basic notions of stochastic volatility modelling—the hedging argument underlying the derivation of the pricing equation, handling the market incompleteness issue through choice of the functional form of the market price of volatility risk, the basic notions of the solution methodology employed, and some of the standard models used.

We have not discussed the important calibration issues, which indeed are driving much of the intense research activity that is ongoing in this area. In terms of theoretical developments various authors have included jump processes either in the stock price dynamics, using the modelling framework of Chaps. 12, 13, or both the stock price dynamics and volatility dynamics (see Bates 1996; Cont and Tankov 2004). However most recent developments have been in models based on diffusion processes, with the inclusion of more stochastic factors driving volatility being an important consideration. We give some of the basic notions of these developments in some of the problems. An important recent development has been the market for VIX (volatility index) options which started trading on the CBOE in 2006. A good survey of many of these recent issues can be found in Gatheral (2008).

## 15.6 Appendix

### Appendix 15.1 Characteristic Functions

Let $f(x)$ be a probability density function. The characteristic function $\phi(\zeta)$ is defined as

$$\phi(\zeta) = \int_{-\infty}^{\infty} e^{i\zeta x} f(x)dx.$$

It follows by definition that

$$\phi(\zeta) = \mathbb{E}[e^{i\zeta x}].$$

The original probability density function may be recovered from the characteristic function via the Fourier inversion formula

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\zeta x} \phi(\zeta) d\zeta. \tag{15.58}$$

In many of our applications we are also interested in calculating the cumulative density function $F(x) = P_r(X \leq x)$, i.e.

$$F(x) = \int_{-\infty}^{x} f(s) ds.$$

It can be shown by integrating (15.58) (see Lamperti 1996) that the Fourier inversion to obtain $F(x)$ becomes

$$F(x) = \frac{1}{2} + \frac{1}{2\pi} \int_{0}^{\infty} \frac{\phi(-u)e^{iux} - \phi(u)e^{-iux}}{iu} du.$$

## Appendix 15.2 Expressing the Option Price in Terms of Characteristic Functions

The expression (15.32) involves the calculation of two expectations. We make the change of variable $x = lnS$ and write

$$I_1 = \tilde{\mathbb{E}}_t[e^{x_T} \mathbb{1}_{\{x_T > \ln E\}}], \tag{15.59}$$

and

$$I_2 = \tilde{\mathbb{E}}_t[\mathbb{1}_{\{x_T > \ln E\}}]. \tag{15.60}$$

We use $\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)$ to denote the conditional probability density function underlying the operator $\tilde{\mathbb{E}}_t$. This density function could be obtained by solving the Kolmogorov backward equation associated with the stochastic differential system (15.27) and (15.28), with $x$ replacing $lnS$ of course. This would in fact be a difficult calculation. The point of the manipulations laid out in this and the next section is to show how the quantities (15.59) and (15.60) may be calculated without directly calculating the function $\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)$.

We note first of all that $I_1$ is of the general form

$$\int_{-\infty}^{\infty} e^{x_T} g(x_T) \tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t) dx_T. \tag{15.61}$$

The task of performing the integration in (15.61) would be much simpler if we could find another conditional probability density function $\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)$ such that the integral in (15.61) reduces to an integral of the form

$$\int_{-\infty}^{\infty} g(x_T)\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T.$$

Such an outcome can be achieved by defining $\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)$ as the conditional probability density function satisfying the condition

$$\frac{\int_{-\infty}^{\infty} e^{x_T} g(x_T)\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T}{\int_{-\infty}^{\infty} e^{x_T}\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T} = \int_{-\infty}^{\infty} g(x_T)\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T. \tag{15.62}$$

It is necessary to divide by the term $\int_{-\infty}^{\infty} e^{x_T}\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T$ on the left hand side to guarantee that $\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)$ satisfies the requirement that

$$\int_{-\infty}^{\infty} \tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T = 1 \tag{15.63}$$

when $g(x_T)$ is set equal to 1. In (15.59) the function $g(x_T)$ is $\mathbb{1}_{(x_T > lnE)}$, but for the moment let us take

$$g(x_T) = e^{iux_T}, \tag{15.64}$$

this with a view to finding the alternative conditional probability density function $\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)$. In this case (15.62) reads

$$\frac{\int_{-\infty}^{\infty} e^{(1+iu)x_T}\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T}{\int_{-\infty}^{\infty} e^{x_T}\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T} = \int_{-\infty}^{\infty} e^{iux_T}\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T. \tag{15.65}$$

The right-hand side of (15.65) is in fact the characteristic function[2] of the density function $\tilde{\pi}_1$, and we use $\phi_1(u)$ to denote it. If we could calculate $\phi_1(u)$, then by the Fourier inversion theorem we could in principle calculate the density function $\tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)$. Then by setting $g(x_T) = \mathbb{1}_{(x_T > \ln E)}$ in (15.62) we reduce the calculation of $I_1$ to the calculation of

$$\tilde{\mathbb{E}}_t[e^{x_T}] = \int_{-\infty}^{\infty} e^{x_T}\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T \tag{15.66}$$

[2]See Appendix 15.1.

and of

$$\tilde{\mathbb{E}}_t^{(1)}[1_{(x_T > lnE)}] = \int_{lnE}^{\infty} \tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T, \tag{15.67}$$

where $\tilde{\mathbb{E}}_t^{(1)}$ denotes expectation with respect to the density function $\tilde{\pi}_1$. We shall see in the next subsection that the integrals in (15.66) and (15.67) can be conveniently handled.

To calculate $I_2$ we note that it is of the general form

$$\int_{-\infty}^{\infty} g(x_T)\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T. \tag{15.68}$$

By again setting $g(x_T) = e^{iux_T}$ this last integral would be the characteristic function of $\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)$, which we denote as $\phi(u)$, then by the Fourier inversion theorem we could in principle obtain $\tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)$. The calculation of $I_2$ would also be an integral of the form

$$\int_{lnE}^{\infty} \tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T. \tag{15.69}$$

## *Solving for the Option Price*

The characteristic functions $\phi_1(u)$ and $\phi(u)$ discussed in the last subsection can both be expressed in terms of the function

$$\psi(\zeta) = \tilde{\mathbb{E}}_t[e^{-r(T-t)+\zeta \ln S_T}], \tag{15.70}$$

where $\zeta$ is a complex number. We note that

$$\psi(0) = \tilde{\mathbb{E}}_t[e^{-r(T-t)}] = e^{-r(T-t)},$$
$$\psi(1) = \tilde{\mathbb{E}}_t[e^{-r(T-t)}S_T] = Se^{-q(T-t)}.$$

The last equality follows from (15.25) and manipulations similar to those yielding (15.66). Thus

$$\phi_1(u) = \frac{\psi(1+iu)}{\psi(1)} = \frac{\psi(1+iu)}{Se^{-q(T-t)}}, \tag{15.71}$$

$$\phi(u) = \frac{\psi(iu)}{\psi(0)} = \frac{\psi(iu)}{e^{-r(T-t)}}, \tag{15.72}$$

respectively. These density functions are expressed in terms of the logarithm of the asset price $x = \ln S$. We shall use $\tilde{\mathbb{E}}_t^{(1)}$ and $\tilde{\mathbb{E}}_t$ to denote expectations with respect to these probability density functions. Using the results discussed in the last subsection we can write

$$\begin{aligned}\tilde{\mathbb{E}}_t[e^{-r(T-t)}S_T\mathbb{1}_{(x_T>\ln E)}] &= Se^{-q(T-t)}\tilde{\mathbb{E}}_t^{(1)}[\mathbb{1}_{(x_T>\ln E)}]\\ &= Se^{-q(T-t)}\int_{\ln E}^{\infty}\tilde{\pi}_1(x_T,\sigma_T,T|x,\sigma,t)dx_T,\end{aligned} \tag{15.73}$$

and

$$\tilde{\mathbb{E}}_t[\mathbb{1}_{(x_T>\ln E)}] = \int_{\ln E}^{\infty}\tilde{\pi}(x_T,\sigma_T,T|x,\sigma,t)dx_T. \tag{15.74}$$

In order to evaluate the integrals in (15.73)–(15.74) we apply the Fourier inversion technique referred to in the previous section. To do this we require the characteristic functions $\phi_1(u)$, $\phi(u)$, which in turn require us to calculate the function $\psi(\zeta)$. First of all integrate the stochastic differential equations (15.27), (15.28) to obtain

$$\begin{aligned}\ln S_T = \ln S_t &+ (r-q)(T-t) - \frac{1}{2}\int_t^T\sigma^2(u)du\\ &+ \sqrt{1-\rho^2}\int_t^T\sigma(u)d\tilde{w}_s(u) + \rho\int_t^T\sigma(u)d\tilde{w}_\sigma(u),\end{aligned} \tag{15.75}$$

and

$$\begin{aligned}\sigma_T^2 - \sigma_t^2 = b^2(T-t) &+ 2k\overline{\sigma}\int_t^T\sigma(u)du - 2(k+\lambda b)\int_t^T\sigma^2(u)du\\ &+ 2b\int_t^T\sigma(u)d\tilde{w}_\sigma(u).\end{aligned} \tag{15.76}$$

Thus

$$\begin{aligned}\Psi &\equiv -r(T-t) + \zeta\ln S_T\\ &= -r(T-t) + \zeta\ln(S_te^{(r-q)(T-t)}) - \frac{\zeta}{2}\int_t^T\sigma^2(u)du\\ &+ \zeta\sqrt{1-\rho^2}\int_t^T\sigma(u)d\tilde{w}_s(u) + \zeta\rho\int_t^T\sigma(u)d\tilde{w}_\sigma(u).\end{aligned} \tag{15.77}$$

and

$$\psi(\zeta) = \tilde{\mathbb{E}}_t(e^{\Psi})$$
$$= \mathscr{K}\tilde{\mathbb{E}}_t\left[\exp\left(-\frac{\zeta}{2}\int_t^T \sigma^2(u)du + \zeta\rho\int_t^T \sigma(u)d\tilde{w}_\sigma(u)\right.\right.$$
$$\left.\left.+\zeta\sqrt{1-\rho^2}\int_t^T \sigma(u)d\tilde{w}_s(u)\right)\right] \tag{15.78}$$

where

$$\mathscr{K} \equiv \exp[\zeta \ln(S_t e^{(r-q)(T-t)}) - r(T-t)].$$

The first two stochastic integrals in the exponent in (15.78) are determined from the diffusion process for $\sigma$ which evolves independently of the diffusion process for $S$, thus we may write

$$\psi(\zeta) = \mathscr{K}\tilde{\mathbb{E}}_t\left\{\exp\left(-\frac{\zeta}{2}\int_t^T \sigma^2(u)du + \zeta\rho\int_t^T \sigma(u)d\tilde{w}_\sigma(u)\right)\right.$$
$$\left.\tilde{\mathbb{E}}_t\left[\exp(\zeta\sqrt{1-\rho^2}\int_t^T \sigma(u)d\tilde{w}_s(u))\right]\right\}.$$

Since[3]

$$\tilde{\mathbb{E}}_t\left[\exp\left(\zeta\sqrt{1-\rho^2}\int_t^T \sigma(u)d\tilde{w}_s(u)\right)\right] = \exp\left(\frac{1}{2}\zeta^2(1-\rho^2)\int_t^T \sigma^2(u)du\right),$$

then we can write

$$\psi(\zeta) = \mathscr{K}\tilde{\mathbb{E}}_t\left[\exp\left(\frac{\zeta^2(1-\rho^2)-\zeta}{2}\int_t^T \sigma^2(u)du + \zeta\rho\int_t^T \sigma(u)d\tilde{w}_\sigma(u)\right)\right]. \tag{15.79}$$

Next we use (15.76) to remove $\int_t^T \sigma(u)d\tilde{w}_\sigma(u)$ from the last expression. Thus

$$\psi(\zeta) = \mathscr{K}^*\tilde{\mathbb{E}}_t\left[\exp\left(-s_1\int_t^T \sigma^2(u)du - s_2\int_t^T \sigma(u)du + s_3\sigma_T^2\right)\right], \tag{15.80}$$

[3]In deriving this result we are relying essentially on the result (8.13), appropriately modified to take account of the different interval of integration.

where

$$s_1 = -\frac{\zeta^2(1-\rho^2)}{2} - \zeta\left(\frac{\rho(k+\lambda b)}{b} - \frac{1}{2}\right),$$

$$s_2 = \zeta\frac{\rho k\overline{\sigma}}{b},$$

$$s_3 = \frac{\zeta\rho}{2b},$$

and

$$\mathscr{K}^* = \exp\left[\zeta\ln(S_t e^{(r-q)(T-t)}) - r(T-t) - \frac{\zeta\rho}{2b}\sigma_t^2 - \frac{\zeta\rho b}{2}(T-t)\right].$$

For convenience we set

$$y(\sigma,t) = \tilde{\mathbb{E}}_t\left[\exp\left(-s_1\int_t^T \sigma^2(u)du - s_2\int_t^T \sigma(u)du + s_3\sigma_T^2\right)\right], \tag{15.81}$$

and note that $y(\sigma,t)$ is of a functional of the form to which the Feynman–Kac formula (as in Proposition 8.3) may be applied. Thus $y(\sigma,t)$ is the solution of the partial differential equation

$$\frac{1}{2}b^2\frac{\partial^2 y}{\partial\sigma^2} + [k\overline{\sigma} - (k+\lambda b)\sigma]\frac{\partial y}{\partial\sigma} - (s_1\sigma^2 + s_2\sigma)y + \frac{\partial y}{\partial t} = 0. \tag{15.82}$$

Equation (15.82) must be solved subject to the final time condition

$$y(\sigma,T) = e^{s_3\sigma^2}. \tag{15.83}$$

It is known that (15.82) with terminal condition (15.83) has a solution of the form

$$\begin{aligned} y(\sigma,t) &= \exp\left(\frac{1}{2}A(t)\sigma^2 + B(t)\sigma + C(t) + s_3\sigma^2\right) \\ &= \exp\left(\frac{1}{2}D(t)\sigma^2 + B(t)\sigma + C(t)\right), \end{aligned} \tag{15.84}$$

where $D(t) = A(t) + 2s_3$. Substituting (15.84) into (15.82), and gathering coefficients of $\sigma^0$, $\sigma^1$ and $\sigma^2$, we obtain ordinary differential equations for $D$, $B$ and $C$, namely

$$\begin{aligned} \frac{1}{2}\dot{D} &= -s_1 - (k + \lambda b)D + \frac{1}{2}b^2 D^2, \\ \dot{B} &= -s_2 + k\bar{\sigma} D - (k + \lambda b)B + b^2 BD, \\ \dot{C} &= k\bar{\sigma} B + \frac{1}{2}b^2 B^2 + \frac{1}{2}b^2 D, \end{aligned} \tag{15.85}$$

with

$$D(T) = 2s_3, \quad B(T) = 0 \quad \text{and} \quad C(T) = 0.$$

The solution to the set of ordinary differential equations (15.85) turns out to be

$$\begin{aligned} D(t) &= \frac{(k + \lambda b)}{b^2} - \frac{\gamma_1 \sin h(\gamma_1 \tau) + \gamma_2 \cos h(\gamma_1 \tau)}{b^2 m}, \\ B(t) &= \frac{(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)(1 - \cos h(\gamma_1 \tau)) - (k\bar{\sigma}\gamma_1\gamma_2 - \gamma_3)\sin h(\gamma_1 \tau)}{\gamma_1 b^2 m}, \end{aligned}$$

and

$$\begin{aligned} C(t) = -\frac{1}{2}\ln m &+ \frac{[(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)^2 - \gamma_3^2(1 - \gamma_2^2)]\sin h(\gamma_1 \tau)}{2\,\gamma_1^3\, b^2\, m} \\ &+ \frac{(k\bar{\sigma}\gamma_1 - \gamma_2\gamma_3)\gamma_3(m-1)}{\gamma_1^3\, b^2\, m} \\ &+ \frac{\tau}{2\gamma_1^2 b^2}[(k + \lambda b)\gamma_1^2(b^2 - (k + \lambda b)k\bar{\sigma}) + \gamma_3^2], \end{aligned}$$

where

$$\begin{aligned} \gamma_1 &= \sqrt{2b^2 s_1 + (k + \lambda b)^2}, \\ \gamma_2 &= \frac{k + \lambda b - 2b^2 s_3}{\gamma_1}, \\ \gamma_3 &= (k + \lambda b)k\bar{\sigma} - s_2 b^2, \\ m &= \cos h(\gamma_1 \tau) + \gamma_2 \sin h(\gamma_1 \tau), \\ \tau &= T - t. \end{aligned}$$

The solution $y(\sigma, t)$ then allows us to form (from (15.80))

$$\psi(\varsigma) = \mathscr{K}^* y(\sigma, t).$$

Then we use this to obtain

$$\phi_1(u) = \frac{\psi(1+iu)}{Se^{-q(T-t)}},$$
$$\phi(u) = \frac{\psi(iu)}{e^{-r(T-t)}}.$$

We can then use the inversion formula to obtain the expectations in (15.73). Thus

$$\int_{\ln E}^{\infty} \tilde{\pi}_1(x_T, \sigma_T, T|x, \sigma, t)dx_T = \frac{1}{2} - \frac{1}{2\pi}\int_0^{\infty} \frac{\phi_1(-u)e^{iux} - \phi_1(u)e^{-iux}}{iu}du,$$

and

$$\int_{\ln E}^{\infty} \tilde{\pi}(x_T, \sigma_T, T|x, \sigma, t)dx_T = \frac{1}{2} - \frac{1}{2\pi}\int_0^{\infty} \frac{\phi(-u)e^{iux} - \phi(u)e^{-iux}}{iu}du.$$

## Appendix 15.3 The Characteristic Function for the Heston Model

We follow the traditional approach (following from Heston) and use the bivariate characteristic function. This approach works directly from the stochastic differential equations, the dynamics of which generate the transition density function $G$. We note that in terms of the log stock price $x = lnS$ the stochastic differential equation system (15.50), (15.51) becomes

$$dx = (r - q - \frac{1}{2}v)dt + \sqrt{1-\rho^2}\sqrt{v}d\tilde{w}_s + \rho\sqrt{v}d\tilde{w}_v, \tag{15.86}$$

$$dv = (k_v\bar{v} - (\lambda + k_v)v)dt + \sigma_v\sqrt{v}d\tilde{w}_v. \tag{15.87}$$

Integrating (15.86) from 0 to $T$ we obtain

$$\begin{aligned} x_T &= x(0) + (r-q)T - \frac{1}{2}\int_0^T v_t dt + \rho\int_0^T \sqrt{v_t}d\tilde{w}_v + \sqrt{1-\rho^2}\int_0^T \sqrt{v_t}d\tilde{w}_s \\ &= x(0) + (r-q)T - \frac{1}{2}\int_0^T v_t dt + \frac{\rho}{\sigma_v}\left(v_T - v(0) - k_v\bar{v}T + k_v\int_0^T v_t dt\right) \\ &\quad + \sqrt{1-\rho^2}\int_0^T \sqrt{v_t}d\tilde{w}_s. \end{aligned} \tag{15.88}$$

The last equality follows by integrating (15.87) from 0 to $T$ and solving for the term $\int_0^T \sqrt{v_t} d\tilde{w}_v$. Under the risk-neutral measure $\tilde{\mathbb{P}}$ the bivariate characteristic function of the two random variables $v$ and $x$ at time $T$ is defined as[4]

$$\varphi(u,w) = \tilde{\mathbb{E}}\left[\exp(iux_T + iwv_T)\right].$$

Using (15.52) we find that

$$\begin{aligned}\varphi(u,w) &= \tilde{\mathbb{E}}\Bigg[\exp(iu\Bigg(x(0) + (r-q)T - \frac{1}{2}\int_0^T v_t dt + \frac{\rho}{\sigma_v}\Bigg(v_T - v(0) - k_v\bar{v}T \\ &\quad + k_v\int_0^T v_t dt\Bigg) + \sqrt{1-\rho^2}\int_0^T \sqrt{v_t} d\tilde{w}_s\Bigg) + iwv_T)\Bigg] \\ &= \exp\left(iu(x(0) + (r-q)T - \frac{\rho}{\sigma_v}v(0) - \frac{\rho}{\sigma_v}k_v\bar{v}T)\right)\tilde{\mathbb{E}}\Bigg[\exp\Bigg(iu\Bigg(\frac{k_v\rho}{\sigma_v} \\ &\quad -\frac{1}{2}\Bigg)\int_0^T v_t dt + \left(iu\frac{\rho}{\sigma_v} + iw\right)v_T + iu\sqrt{1-\rho^2}\int_0^T \sqrt{v_t} d\tilde{w}_s\Bigg)\Bigg].\end{aligned}$$

In the following step, we take the conditional expectation value with respect to $\tilde{w}_v(s)$ for $0 \le s \le T$, which represents the filtration generated by the Brownian motion of the volatility process. Since all other terms in the above expectation are $\tilde{w}_v$-measurable, application of the law of iterated expectations leads the calculation of the expectation of $\exp\left(iu\sqrt{1-\rho^2}\int_0^T \sqrt{v_t} d\tilde{w}_s\right)$. Thus we have

$$\begin{aligned}\varphi(u,w) &= \exp\left(iu(x(0) + (r-q)T - \frac{\rho}{\sigma_v}v(0) - \frac{\rho}{\sigma_v}k_v\bar{v}T)\right) \\ &\quad \times\tilde{\mathbb{E}}\Bigg[\exp\left(iu\left(\frac{k_v\rho}{\sigma_v} - \frac{1}{2}\right)\int_0^T v_t dt + \left(iu\frac{\rho}{\sigma_v} + iw\right)v_T\right) \\ &\quad \times\tilde{\mathbb{E}}\left[\exp\left(iu\sqrt{1-\rho^2}\int_0^T \sqrt{v_t} d\tilde{W}_2\right) \mid \sigma(\tilde{w}_v(s) : 0 \le s \le T)\right]\Bigg].\end{aligned}$$

Given the knowledge of $\tilde{w}_v$, the path of $v$ is known from time 0 until $T$ and therefore deterministic. Since $\tilde{w}_v$ and $\tilde{w}_s$ are independent it follows that the integral $\int_0^T \sqrt{v_t} d\tilde{w}_s$ is normally distributed with zero mean. The variance can be calculated via the Ito isometry (see (5.24)) and is equal to $\int_0^T v_t dt$. With the

[4]This amounts to taking the double Fourier transform.

use of the characteristic function for a normally distributed variable $X$,$\mathbb{E}[e^{iaX}] = e^{ia\mathbb{E}X - \frac{1}{2}a^2 \operatorname{var} X}$ (see Sect. 6.3.4), the above becomes

$$\begin{aligned}
\varphi(u,w) &= \exp\left(iu(x(0) + (r-q)T - \frac{\rho}{\sigma_v}v(0) - \frac{\rho}{\sigma_v}k_v\bar{v}T)\right) \\
&\quad \times\tilde{\mathbb{E}}\Bigg[\exp\left(iu\left(\frac{k_v\rho}{\sigma_v} - \frac{1}{2}\right)\int_0^T v_t dt + \left(iu\frac{\rho}{\sigma_v} + iw\right)v_T\right. \\
&\qquad \left. -\frac{1}{2}u^2(1-\rho^2)\int_0^T v_t dt\right)\Bigg] \\
&= \exp\left(iu(x(0) + (r-q)T - \frac{\rho}{\sigma_v}v(0) - \frac{\rho}{\sigma_v}k_v\bar{v}T)\right) \\
&\quad \times\tilde{\mathbb{E}}\left[\exp\left(\left(\frac{k_v\rho}{\sigma_v}iu - \frac{1}{2}iu - \frac{1}{2}u^2(1-\rho^2)\right)\int_0^T v_t dt + \left(iu\frac{\rho}{\sigma_v} + iw\right)v_T\right)\right].
\end{aligned}$$

Introducing the abbreviations

$$\begin{aligned}
a(u,w) &= iw + iu\frac{\rho}{\sigma_v}, \\
b(u) &= \frac{k_v\rho}{\sigma_v}iu - \frac{1}{2}iu - \frac{1}{2}u^2(1-\rho^2),
\end{aligned}$$

the characteristic function takes this form

$$\begin{aligned}
\varphi(u,w) = \exp&\left(iu(x(0) + (r-q)T - \frac{\rho}{\sigma_v}v(0) - \frac{\rho}{\sigma_v}k_v\bar{v}T)\right) \\
&\times\tilde{\mathbb{E}}\left[\exp\left(b(u)\int_0^T v_t dt + a(u,w)v_T\right)\right].
\end{aligned}$$

This expectation contains two random variables, $\int_0^T v_t dt$ and $v_T$, which both depend on the volatility process up to time $T$. We apply the Feynman–Kac formula (see Sect. 8.6) to calculate this expectation.

Define the function $y$ as the value of the above expectation, so that

$$y(T, v(0)) = \tilde{\mathbb{E}}\left[\exp\left(a(u,w)v_T + b(u)\int_0^T v_t dt\right)\right].$$

Since $v$ follows the mean-reverting square-root process (15.87) by the Feynman–Kac formula, we have that $y$ must satisfy the partial differential equation

$$\frac{\partial y}{\partial \tau} = b(u)vy + k_v(\bar{v} - v)\frac{\partial y}{\partial v} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 y}{\partial v^2}, \tag{15.89}$$

with the boundary condition

$$y(0, v(0)) = \exp(a(u, w)v(0)).$$

If we assume that $y$ is log-linear and given by

$$y(T, v(0)) = \exp\left[A(\tau)v(0) + B(\tau)\right],$$

then the derivatives of $y$ are

$$\begin{aligned}
\frac{\partial y}{\partial \tau} &= y(A'v + B'), \\
\frac{\partial y}{\partial v} &= yA, \\
\frac{\partial^2 y}{\partial v^2} &= yA^2.
\end{aligned}$$

Therefore $A$ and $B$ satisfy the ODEs

$$\begin{aligned}
A'(\tau) - \frac{1}{2}\sigma_v^2 A^2(\tau) + k_v A(\tau) + b(u) &= 0, \\
A(0) &= a(u, w),
\end{aligned}$$

and

$$\begin{aligned}
B'(\tau) - k_v \bar{v} A(\tau) &= 0, \\
B(0) &= 0.
\end{aligned}$$

Solving these ODEs will give the unique solution to the above PDE. Making the usual transformation for Riccati-type ODEs we have

$$U(\tau) = \exp\left(-\frac{\sigma_v^2}{2}\int A(s)ds\right),$$

or equivalently

$$A(\tau) = -\frac{2}{\sigma_v^2}\frac{U'(\tau)}{U(\tau)} \text{ and } B(\tau) = \frac{2k_v\bar{v}}{\sigma_v^2}\ln U(\tau).$$

We obtain the linear homogeneous second-order ODE

$$U''(\tau) + k_v U'(\tau) - \frac{1}{2}\sigma_v^2 U(\tau) b(u) = 0. \tag{15.90}$$

This second order ordinary differential equation has the general solution with parameters $V$ and $R$ of the form

$$
\begin{aligned}
U(\tau) &= Ve^{x_1\tau} + Re^{x_2\tau} \\
x_1 &= -\frac{k_v}{2} + \sqrt{\frac{k_v^2}{4} + \frac{1}{2}\sigma_v^2 b(u)} \\
x_2 &= -\frac{k_v}{2} - \sqrt{\frac{k_v^2}{4} + \frac{1}{2}\sigma_v^2 b(u)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
U'(\tau) &= Vx_1e^{x_1\tau} + Rx_2e^{x_2\tau} \\
&= \frac{1}{k_v}\left(\frac{1}{2}\sigma_v^2 b(u)U(\tau) - U''(\tau)\right) \\
&= \frac{V}{k_v}e^{x_1\tau}\left(\frac{1}{2}\sigma_v^2 b(u) - x_1^2\right) + \frac{R}{k_v}e^{x_2\tau}\left(\frac{1}{2}\sigma_v^2 b(u) - x_2^2\right),
\end{aligned}
$$

where the second equality follows from use of (15.90). The boundary conditions $U(0) = V + R$ and $U'(0) = -\frac{1}{2}\sigma_v^2 a(u,w)(V + R)$ then yield

$$
\begin{aligned}
V &= \frac{U(0)\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right)}{d}, \\
R &= \frac{U(0)\left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right)}{d},
\end{aligned}
$$

with

$$
d = x_1 - x_2 = \sqrt{k_v^2 + 2\sigma_v^2 b(u)}.
$$

Thus

$$
\begin{aligned}
A(\tau) &= -\frac{2}{\sigma_v^2}\frac{U'(\tau)}{U(\tau)} \\
&= -\frac{2}{\sigma_v^2}\frac{U(0)\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right)x_1e^{x_1\tau} + U(0)\left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right)x_2e^{x_2\tau}}{U(0)\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right)e^{x_1\tau} + U(0)\left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right)e^{x_2\tau}} \\
&= -\frac{2}{\sigma_v^2}\frac{\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right)x_1e^{x_1\tau} + \left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right)x_2e^{x_2\tau}}{\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right)e^{x_1\tau} + \left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right)e^{x_2\tau}} \\
&= -\frac{2}{\sigma_v^2}\frac{\left(-\sigma_v^2 a(u,w) + (k_v + d)\right)x_1e^{x_1\tau} + \left(2d + \sigma_v^2 a(u,w) - (k_v + d)\right)x_2e^{x_2\tau}}{2de^{-d\tau} + \left(\sigma_v^2 a(u,w) - k_v - d\right)\left(e^{-d\tau} - 1\right)}
\end{aligned}
$$

$$= \frac{1}{\sigma_v^2} \frac{\left(-\sigma_v^2 a(u,w) + k_v + d\right)(k_v - d) + \left(\sigma_v^2 a(u,w) - k_v + d\right)(k_v + d)e^{-d\tau}}{2de^{-d\tau} + \left(\sigma_v^2 a(u,w) - k_v - d\right)\left(e^{-d\tau} - 1\right)}$$

$$= \frac{da\left(1 + e^{-d\tau}\right) - \left(1 - e^{-d\tau}\right)(k_v a(u,w) + 2b(u))}{2de^{-d\tau} + \left(\sigma_v^2 a(u,w) - k_v - d\right)\left(e^{-d\tau} - 1\right)}$$

Finally, we can solve for $B(\tau)$ by integration

$$\begin{aligned} B(\tau) &= -\frac{2k_v \bar{v}}{\sigma_v^2} \int_0^\tau \frac{U'(s)}{U(s)} ds = -\frac{2k_v \bar{v}}{\sigma_v^2} \ln \frac{U(\tau)}{U(0)} \\ &= -\frac{2k_v \bar{v}}{\sigma_v^2} \ln \frac{\left(-\frac{1}{2}\sigma_v^2 a(u,w) - x_2\right) + \left(d + \frac{1}{2}\sigma_v^2 a(u,w) + x_2\right) e^{-d\tau}}{de^{-x_1 \tau}} \\ &= \frac{2k_v \bar{v}}{\sigma_v^2} \ln \frac{2de^{\frac{1}{2}(k_v - d)\tau}}{2de^{-d\tau} + \left(\sigma_v^2 a(u,w) - k_v - d\right)\left(e^{-d\tau} - 1\right)}. \end{aligned}$$

Therefore the one-dimensional PDE (15.89) has been solved and the solution is

$$\begin{aligned} \tilde{\mathbb{E}}\left[\exp\left(a(u,w)v_T + b(u)\int_0^T v_t dt\right)\right] &= \exp[A(T, a(u,w), b(u))v(0) \\ &\quad + B(T, a(u,w), b(u))], \end{aligned}$$

with

$$\begin{aligned} A(\tau, a(u,w), b(u)) &= \frac{-(1 - \exp(-d\tau)(2b(u) + k_v a(u,w)) + da(1 + \exp(-d\tau))}{\gamma}, \\ B(\tau, a(u,w), b(u)) &= \frac{k_v \bar{v}}{\sigma_v^2}(k_v - d)\tau + \frac{2k_v \bar{v}}{\sigma_v^2} \ln \frac{2d}{\gamma}, \\ d &= \sqrt{k_v^2 + 2\sigma_v^2 b(u)}, \\ \gamma &= 2d \exp(-d\tau) + (k_v + d - \sigma_v^2 a(u,w))(1 - \exp(-d\tau)). \end{aligned}$$

The characteristic function is therefore given by

$$\begin{aligned} \varphi(u,w) = \exp\Big( & iux(0) + iurT - iu\frac{\rho}{\sigma_v} v(0) - iu\frac{\rho}{\sigma_v} k_v \bar{v} T + A(T, a(u,w), b(u))v(0) \\ & + B(T, a(u,w), b(u))\Big). \end{aligned}$$

Note that the marginal characteristic functions $\varphi(u, 0)$ and $\varphi(0, v)$ are equal to the univariate characteristic function of the logarithmic spot value and the volatility value at time $T$.

## 15.7 Problems

**Problem 15.1** a) Consider the stock price process

$$\frac{dS}{S} = \mu dt + \sigma dZ.$$

Using simulation obtain and graph the distribution of $S(t)$ when $\mu = 0.20$, $\sigma = 0.18$.

b) Consider the stock price process

$$\frac{dS}{S} = \mu dt + \sigma dZ,$$

where $\sigma$ follows the diffusion process

$$d\sigma = k(\bar{\sigma} - \sigma)dt + \delta dW,$$

and

$$\mathbb{E}(dZdW) = \rho dt.$$

Using simulation obtain and graph the distribution of $S(t)$ when $\mu = 0.20$, $k = 1, \delta = 0.25, \bar{\sigma} = 0.18$ and $\rho = -0.5, 0, 0.5$.

c) Compare the distributions at part a) and b) and discuss the differences; particularly comments on any differences caused by the correlation between $dZ$ and $dW$.

**Problem 15.2** Consider the mean reverting stochastic volatility model given by Eqs. (15.1) and (15.2). The partial differential equation for the option price in this case is given by the partial differential equation (15.18) with appropriately defined coefficients. Consider this partial differential equation in the case when the volatility between the noise terms is zero i.e. $\rho = 0$.

a) Make the change of variables

$$y = \ln S, \qquad \phi = (k + \lambda b)\sigma - k\bar{\sigma}$$

and obtain the partial differential equation for the transformed option price

$$F(y, \phi, t) \equiv f\left(e^y, \frac{\phi + k\bar{\sigma}}{k + \lambda b}, t\right).$$

b) Consider solving this partial differential equation by the Fourier transform technique of Chap. 9. To do so you will need to define the two-dimensional Fourier transform

$$\bar{F}(\omega, \eta, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(y, \phi, t) e^{-i\omega y - i\eta\phi} dy d\phi.$$

Use the same formal manipulations as in Chap. 9 and show that $\bar{F}(\omega, \eta, t)$ satisfies a first order partial differential equation of the form

$$\frac{\partial \bar{F}}{\partial t} + [\ldots]\bar{F} + (k + \lambda b)\eta \frac{\partial \bar{F}}{\partial \eta} + \frac{1}{2} b^2 (k + \lambda b)^2 \eta^2 \bar{F} = 0.$$

**Problem 15.3** Consider the stochastic volatility model

$$\frac{dS}{S} = \mu dt + \sigma dz_s,$$
$$d(\ln \sigma) = k_v(\ln \bar{\sigma} - \ln \sigma)dt + \delta dz_\sigma$$

where

$$E(dz_s dz_\sigma) = \rho dt.$$

What stochastic process does $\sigma$ follow in terms of independent Wiener processes?

**Problem 15.4** Johnson and Shanno (1987) consider the stochastic volatility model

$$\frac{dS}{S} = \mu dt + \sigma S^{\alpha_1} dz_s,$$

$$d\sigma = \kappa \sigma dt + \delta \sigma^{\alpha_2} dz_\sigma,$$

where

$$\mathbb{E}(dz_s dz_\sigma) = \rho dt$$

and $\alpha_1, \alpha_2 \geq 0$.

This model allows the instantaneous stock return to depend on both $S$ (the $S^{\alpha_1}$ term—for which there is some empirical evidence) and a stochastic $\sigma$, which follows a so-called CEV (= constant elasticity of variance process) process.

a) Obtain the PDE that determines derivative prices under these dynamics.
b) Write out the dynamics for $S$ and $\sigma$ under the risk neutral dynamics.
c) Give the expression for a European call option in terms of an expectation operator.

**Problem 15.5** A slight variant of the model considered in Problem 15.4 is the so-called stochastic $\alpha\beta\rho$ model, more popularly known as the SABR model. It has become widely adopted in recent years because it enjoys convenient properties for calibrating to the smiles and skews observed in markets because it is able to capture the correct dynamics of the smile. The model is of the form

$$dS = \mu S dt + \alpha S^{\beta} dz_s,$$
$$d\alpha = \sigma_\alpha \alpha dz_\alpha,$$

with

$$\mathbb{E}[dz_s dz_\alpha] = \rho dt.$$

a) Obtain the PDE that determines derivative prices under these dynamics.
b) Write out the dynamics for $S$ and $\alpha$ under the risk-neutral dynamics.
c) Give the expression for a European call option in terms of an expectation operator under the risk-neutral measure.

For more on the SABR model the interested reader should consult Hagan et al. (2002) and West (2005).

**Problem 15.6** A model which generalises several of the models considered in this chapter is the so-called double CEV (constant elasticity of variance) model. This model has the general form

$$\frac{dS}{S} = \mu dt + \sqrt{v} dZ_s,$$
$$dv = k_v(u - v)dt + \sigma_v v^{\alpha} dZ_v,$$
$$du = k_u(\bar{u} - u)dt + \sigma_u u^{\beta} dZ_u,$$

with

$$\mathbb{E}[dZ_s dZ_v] = \rho_{sv} dt, \quad \mathbb{E}[dZ_s dZ_u] = \rho_{su} dt, \quad \mathbb{E}[dZ_v dZ_u] = \rho_{vu} dt.$$

and typically $\alpha, \beta \in [\frac{1}{2}, 1]$. An advantage of this model is that it introduces an additional stochastic factor into the volatility dynamics (the Wiener process $Z_u$) and this caters for the empirically observed fact that three factors seem to be driving the dynamics of observed implied volatility surfaces. Note that when $\alpha = \beta = \frac{1}{2}$ we have the double Heston model and when $\alpha = \beta = 1$ the double log-normal model.

a) Obtain the PDE that determines the derivative prices under these dynamics. Make whatever assumptions about market price of risk factors that are necessary to ensure affine coefficients of the first derivative terms.
b) Write out the dynamics for $S, u$ and $v$ under the risk-neutral dynamics.
c) Give the expression for a European call option in terms of an expectation operator under the risk-neutral measure.

**Problem 15.7 Computational Problem**—Consider the stochastic volatility process

$$\frac{dS}{S} = \mu dt + \sqrt{v} dZ_S$$
$$dv = \kappa_v(\bar{v} - v)dt + \delta_v \sqrt{v} dZ_v$$

where $Z_S$ and $Z_v$ are correlated Wiener processes so that

$$\mathbb{E}(dz_S dz_v) = \rho dt.$$

Simulate the dynamics for $S$ and $v$ and obtain the distribution for $S$ at $t = 1$ conditional on the initial stock price $S_o$.

Take $\mu = 0.15$, $\bar{v} = 0.04$, $\delta_v = 0.2$, $\kappa_v = 1$, and $S_o = 0.5$. Simulate from $t = 0$ to $t = 1$, using 10,000 simulations. Experiment with the step size until you get a "good" distribution. Take $\rho = -0.5$, 0 and 0.5. Compare the distribution for $S$ you obtain with the log normal distributions obtained by using the mean and variance of the simulated time series.

In order to see clearly the differences in the tails of the distribution also plot the distributions on a log scale. As a further experiment to gauge the impact of $\delta_v$, simulate the $\rho = 0$ case with $\delta_v = 0.5$, 1 and 1.5 and compare the distributions (on both standard scale and log scale).

# Chapter 16
# Pricing the American Feature

**Abstract** To understand the problems and techniques of pricing the American feature of an option, this chapter introduces the American option pricing problem from the conventional approach based on compound options and the free boundary value problem which can be solved by using either the Fourier transform technique or a simple approximation procedure. The framework developed is readily extended to other option pricing problems.

## 16.1 Introduction

In this chapter we consider the problem of pricing an American put option on non-dividend paying stock. It may be optimal to exercise the American put early in this case, unlike the case of the American call. The American put is the simplest framework in which to understand the problems and techniques of pricing the American feature of an option. The framework developed is readily extended to other option pricing situations.

This chapter considers the American option pricing problem from three perspectives.

First, in Sect. 16.2, the conventional approach to American option evaluation that consists of dividing the time to maturity into a number of subintervals and breaking up the valuation problem into a sequence of compound option valuations over those subintervals; this approach was developed by Geske and Johnson (1984).

Second, Sect. 16.3 describes how the problem of valuing an American option can be posed as a free boundary value problem and describes a general formulation for the solution, the Fourier transform technique used in Chap. 9 to evaluate the European option is extended to solve this free boundary value problem; this approach is based on the Mckean (1965) discussion and its further elaboration by Kucera and Ziogas (2001).

Third, Sect. 16.4 discusses a simple approximation procedure for the solution of the free boundary value problem, due to Macmillan (1986) and Barone-Adesi and Whaley (1987).



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_16

## 16.2 The Conventional Approach Based on Compound Options

Let $P(S, \tau)$ denote the value of an American put when the underlying stock price is $S$ and time to maturity is $\tau(= T - t)$. The put value satisfies the partial differential equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = \frac{\partial P}{\partial \tau}, \tag{16.1}$$

subject to the maturity condition

$$\lim_{\tau \to 0} P(S, \tau) = \max[E - S, 0], \tag{16.2}$$

and boundary condition

$$\lim_{S \to \infty} P(S, \tau) = 0. \tag{16.3}$$

The put option value will also satisfy the early exercise condition

$$P(S, \tau) \geq \max[E - S, 0]. \tag{16.4}$$

The time to maturity is divided into a finite number of subintervals $t_0, t_1, \ldots, t_n$ of length $\Delta t \left(= \dfrac{T - t}{n}\right)$ and define $t_0 = T$. We consider put values $P^{(1)}, P^{(2)}, \ldots, P^{(n)}$ defined on the domains $D^{(1)}, D^{(2)}, \ldots, D^{(n)}$ respectively, see Fig. 16.1, where

$$\begin{aligned}
D^{(1)} &= \{(S, \tau); \quad 0 < S \leq \infty, \quad 0 < \tau \leq \Delta t\} \\
D^{(2)} &= \{(S, \tau); \quad 0 < S \leq \infty, \quad \Delta t < \tau \leq 2\Delta t\} \\
&\vdots \quad \vdots \\
D^{(n)} &= \{(S, \tau); \quad 0 < S \leq \infty, \quad (n-1)\Delta t < \tau \leq n\Delta t\}.
\end{aligned} \tag{16.5}$$

In the domain $D^{(1)}$, with one period left to maturity the American put is equivalent to a European put, i.e. the American put value $P^{(1)}$, for $0 < \tau \leq \Delta t$, is the solution to

$$\frac{1}{2}\sigma^2 S^2 P_{SS}^{(1)} + rSP_S^{(1)} - rP^{(1)} = P_\tau^{(1)},$$

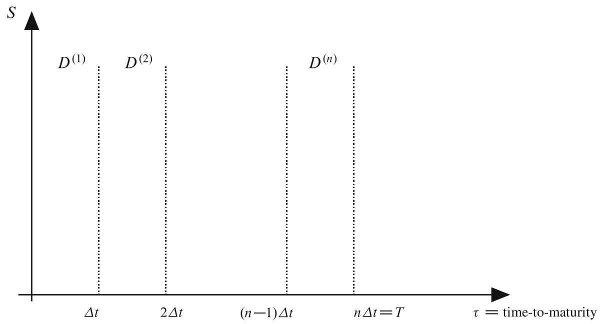


**Fig. 16.1** Domains for American put option valuation

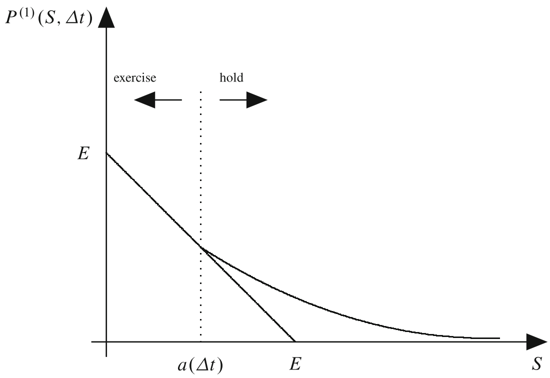


**Fig. 16.2** Value of the American put option unexercised at $\tau = \Delta t$

subject to

$$\lim_{\tau \to 0} P^{(1)}(S, \tau) = \max[E - S, 0],$$

$$\lim_{S \to \infty} P^{(1)}(S, \tau) = 0.$$

Here we denote $f_x = \frac{\partial f(x,y)}{\partial x}$, $f_{xx} = \frac{\partial^2 f(x,y)}{\partial x^2}$ and so on. The solution for $P^{(1)}(S, \tau)$ is given by the Black–Scholes European put formula.

At $\tau = \Delta t$, $P^{(1)}(S, \Delta t)$ represents the value of the American put *unexercised* (see Fig. 16.2). The American put will be exercised if its value is less than the immediate exercise value $(E - S)$. This will occur if the stock price falls below a critical value $a(\Delta t)$ defined by

$$P^{(1)}(a(\Delta t), \Delta t) = E - a(\Delta t). \tag{16.6}$$

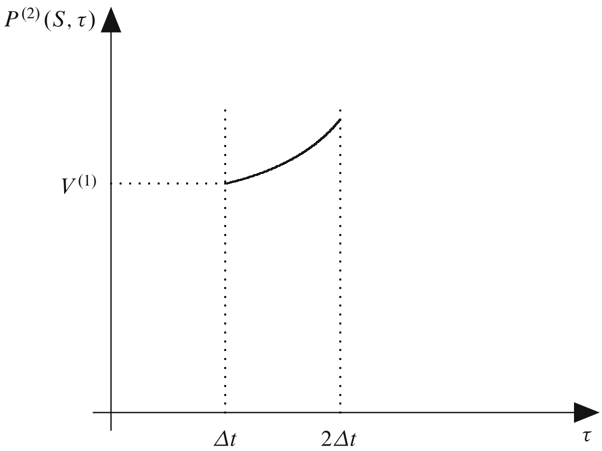


**Fig. 16.3** Value of American put option over the interval $\Delta t < \tau \le 2\Delta t$

Thus the value of the American put, at $\tau = \Delta t$, is

$$V^{(1)} = \max[P^{(1)}(S, \Delta t), E - S].$$

The value, $V^{(1)}$, will be the maturity value for the American put over the interval $\Delta t < \tau \le 2\Delta t$ (Fig. 16.3). That is over this interval $P^{(2)}(S, \tau)$ must satisfy

$$\begin{aligned} &\frac{1}{2}\sigma^2 S^2 P^{(2)}_{SS} + rSP^{(2)}_S - rP^{(2)} = P^{(2)}_\tau, \\ &\lim_{\tau \to \Delta t} P^{(2)}(S,\tau) = V^{(1)} = \max[E - S, P^{(1)}(S, \Delta t)], \\ &\lim_{S \to \infty} P^{(2)}(S,\tau) = 0. \end{aligned} \tag{16.7}$$

The problem of solving (16.7) subject to a maturity condition which is itself an option is known as a compound option problem and has been solved by Geske (1979) and involves both univariate and trivariate normal distribution functions. As with $P^{(1)}(S,\tau)$, the value $P^{(2)}(S,\tau)$ is the value of the American put *unexercised*, over the interval $\Delta t < \tau \le 2\Delta t$, if its value is less than the immediate exercise value $(E - S)$. Early exercise will occur if the stock price falls below the critical value $a(2\Delta t)$ given by

$$P^{(2)}(a(2\Delta t), 2\Delta t) = E - a(2\Delta t).$$

The value of the American put, at $\tau = 2\Delta t$, is then

$$V^{(2)} = \max[P^{(2)}(S, 2\Delta t), E - S],$$

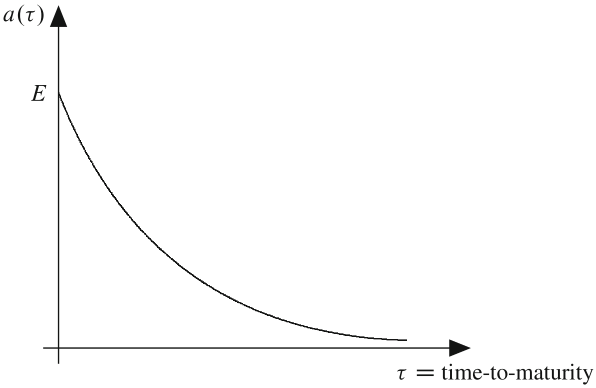


**Fig. 16.4** Time profile of the critical stock price $a$

which becomes the maturity condition at $\tau = 2\Delta t$ for the problem of finding $P^{(3)}(S, \tau)$ over the subinterval $2\Delta t < \tau \leq 3\Delta t$. This procedure is repeated recursively until current time ($\tau = T$) is reached. Over the subinterval $(n-1)\Delta t < \tau \leq n\Delta t$, the value of $P^{(n)}$ is given by

$$\frac{1}{2}\sigma^2 S^2 P_{SS}^{(n)} + rsP_S^{(n)} - rP^{(n)} = P_\tau^{(n)},$$

subject to

$$\lim_{\tau \to (n-1)\Delta t} P^{(n)}(S, \tau) = \max[E - S, P^{n-1}(S, (n-1)\Delta t)],$$

$$\lim_{S \to \infty} P^{(n)}(S, \tau) = 0.$$

The solution to this compound option problem involves the $n$-dimensional cumulative normal distribution function. As a product of this computational procedure, we obtain the time profile of the critical stock price $a$, below which the option would be exercised early (Fig. 16.4).

## 16.3 A General Formulation

### *16.3.1 The Free Boundary Value Problem*

Early work on the problem of valuing the American feature of an option was carried out in apparent ignorance of the fact that this problem is essentially a free boundary value problem as pointed out by Mckean (1965). Crucial to formulating the American put valuation problem as a free boundary value problem is an understanding of the behaviour of the critical stock price $a(\tau)$.

It can be shown using no riskless arbitrage arguments that

(i) $a(\tau)$ is non-increasing in $\tau$,
(ii) $\lim_{\tau \to 0} a(\tau) = a(0) = E$,
(iii) $a(\tau)$ is continuous.

Furthermore, an American put held to expiration will have a zero value at $\tau = 0$. This follows because from properties (i) and (ii) it is not possible for the stock price to end up below $E$ at $\tau = 0$ without crossing $a(\tau)$ for some larger value of $\tau$. This would mean $S \geq a(0) = E$ at maturity and so the unexercised American put would have zero value at maturity.

If we let $U(S, \tau; a(\tau))$ be the *unexercised* value of the American put when the critical stock price is $a(\tau)$ then $U$ is the solution of

$$\frac{1}{2}\sigma^2 S^2 U_{SS} + rSU_S - rU = U_\tau, \tag{16.8}$$

subject to

$$\lim_{\tau \to 0} U(S, \tau; a(\tau)) = 0, \tag{16.9a}$$

$$\lim_{S \to a(\tau)} U(S, \tau; a(\tau)) = E - a(\tau), \tag{16.9b}$$

$$\lim_{S \to \infty} U(S, \tau; a(\tau)) = 0. \tag{16.9c}$$

Thus at any time $\tau$ the value of the American put is given by

$$P(S, \tau) = \begin{cases} U[S, \tau; a(\tau)], & \text{for } a(\tau) < S < \infty, \\ E - S, & \text{for } 0 < S \leq a(\tau). \end{cases}$$

We would like to choose $a(\tau)$ so as to maximise $P(S, \tau)$ and Merton (1973) has shown that this is equivalent to imposing the condition

$$\lim_{S \to a(\tau)} U_S = -1. \tag{16.9d}$$

The solution to the partial differential equation (16.8) subject to the boundary conditions (16.9a)–(16.9d) is a classical free boundary value problem whose solution is discussed by Mckean (1965) and Kolodner (1956). It may be solved by use of the Fourier transform technique. Application of their techniques leads to the solution

$$U(S,\tau) = EV\left(\ln\left(\frac{S}{E}\right),\tau\right)$$

where

$$V(x,\tau) = \int_0^\tau \frac{e^{-g(x,s)}}{\sigma\sqrt{2\pi(\tau-s)}}\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s) + k_1 - \left[\frac{x-b(s)+k_1(\tau-s)}{2(\tau-s)}\right]\right)(c(s)-1)\right]ds, \tag{16.10}$$

with

$$g(x,s) = \frac{(x-b(s)+k_1(\tau-s))^2}{2\sigma^2(\tau-s)} + k_2(\tau-s), \tag{16.11}$$

$$k_1 = (r - \frac{1}{2}\sigma^2). \tag{16.12}$$

Furthermore $c(\tau)(=e^{b(\tau)})$ is given as the solution of the integral equation

$$\frac{1-c(\tau)}{2} = \int_0^\tau \frac{e^{-g(b(\tau),s)}}{\sigma\sqrt{2\pi(\tau-s)}}\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s) + k_1 - \left[\frac{b(\tau)-b(s)+k_1(\tau-s)}{2(\tau-s)}\right]\right)(c(s)-1)\right]ds. \tag{16.13}$$

Since it is nonlinear, the integral (16.13) needs to be solved numerically. Once $c(\tau)$ is obtained, (16.10) becomes a simple exercise in numerical integration to obtain $U(S,\tau)$.

In the following subsections we derive the result (16.10)–(16.13) by use of the Fourier transform technique.[1]

### *16.3.2 Transforming the Partial Differential Equation*

Firstly, transform the partial differential equation (16.8) to an equation with constant coefficients and a "standardised" strike of 1. Let

$$U(S,t) = EV(x,\tau), \tag{16.14}$$

[1]This discussion draws heavily on Kucera and Ziogas (2001).

where

$$S = Ee^x. \tag{16.15}$$

The transformed PDE is then

$$\frac{\partial V}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + k_1 \frac{\partial V}{\partial x} - k_2 V, \quad 0 \le \tau \le T, \tag{16.16}$$

in the region $b(\tau) < x < \infty$, where

$$k_1 = \left(r - \frac{1}{2}\sigma^2\right),$$
$$k_2 = r,$$

and the transformed free boundary is given by

$$b(\tau) = \ln\left(\frac{a(T-\tau)}{E}\right) = \ln\left(\frac{a(t)}{E}\right). \tag{16.17}$$

The new initial and boundary conditions are

$$V(x, 0) = \max(1 - e^x, 0), \quad -\infty < x < \infty \tag{16.18}$$

$$\lim_{x \to -\infty} V(x, \tau) = 0, \quad \tau \ge 0 \tag{16.19}$$

$$V(b(\tau), \tau) = 1 - c(\tau), \quad \tau \ge 0 \text{ with } c(\tau) = e^{b(\tau)} \tag{16.20}$$

$$\lim_{x \to b(\tau)} \frac{\partial V}{\partial x} = -c(\tau). \tag{16.21}$$

Henceforth, $b \equiv b(\tau)$ and $c \equiv c(\tau)$, for simplicity. In order to solve this PDE for $V(x, \tau)$, the $x$ domain shall be extended to $-\infty < x < \infty$ by expressing the PDE as

$$\mathscr{H}(x-b)\left(\frac{\partial V}{\partial \tau} - \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} - k_1 \frac{\partial V}{\partial x} + k_2 V\right) = 0, \quad -\infty < x < \infty, \quad 0 \le \tau \le T, \tag{16.22}$$

where $\mathscr{H}(x-b)$ is the Heaviside step function, defined as

$$\mathscr{H}(x) = \begin{cases} 1, & x > 0 \\ \frac{1}{2}, & x = 0 \\ 0, & x < 0. \end{cases} \tag{16.23}$$

The reason for this choice, in particular the $1/2$ at $x = b$, will become clear later on. The initial and boundary conditions remain as before (Fig. 16.5).

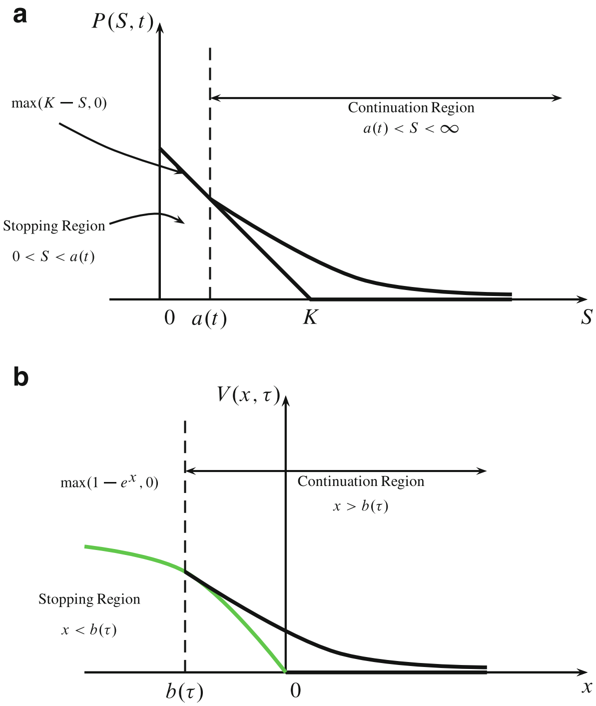


**Fig. 16.5** The continuation region. (**a**) Continuation region in $S$-space. (**b**) Continuation region in $x$-space

### *16.3.3 Applying the Fourier Transform*

Since the $x$ domain is now $-\infty < x < \infty$, the Fourier transform of the PDE can be found. Define $\mathscr{F}\{V(x,\tau)\}$ as

$$\mathscr{F}\{V(x,\tau)\} = \int_{-\infty}^{\infty} e^{i\eta x} V(x,\tau) dx.$$

Applying the Fourier transform to (16.22), we obtain

$$\mathscr{F}\left\{\mathscr{H}(x-b)\frac{\partial V}{\partial \tau}\right\} = \frac{1}{2}\sigma^2 \mathscr{F}\left\{\mathscr{H}(x-b)\frac{\partial^2 V}{\partial \tau^2}\right\} + k_1 \mathscr{F}\left\{\mathscr{H}(x-b)\frac{\partial V}{\partial x}\right\} - k_2 \mathscr{F}\{\mathscr{H}(x-b)V\}. \tag{16.24}$$

By definition

$$\mathscr{F}\{\mathscr{H}(x-b)V(x,\tau)\} = \int_{-\infty}^{\infty} e^{i\eta x}\mathscr{H}(x-b)V(x,\tau)dx$$
$$= \int_{b}^{\infty} e^{i\eta x}V(x,\tau)dx \equiv \mathscr{F}^b\{V(x,\tau)\}. \quad (16.25)$$

Additionally, denote

$$\hat{V}(\eta,\tau) \equiv \mathscr{F}^b\{V(x,\tau)\}.$$

We call $\mathscr{F}^b$ an incomplete Fourier transform (applied to $V(x,\tau)$ for $b < x < \infty$). In Appendix 16.1 we show how the incomplete Fourier transform may be derived as a consequence of the standard Fourier transform and there derive the corresponding inversion theorem.

**Proposition 16.1** *The incomplete Fourier transform of (15.25) enjoys the properties*

$$\mathscr{F}^b\left\{\frac{\partial V}{\partial x}\right\} = (c-1)e^{i\eta b} - i\eta\hat{V}(\eta,\tau),$$
$$\mathscr{F}^b\left\{\frac{\partial^2 V}{\partial x^2}\right\} = ce^{i\eta b} - i\eta\left[(c-1)e^{i\eta b} - i\eta\hat{V}\right],$$
$$\mathscr{F}^b\left\{\frac{\partial V}{\partial \tau}\right\} = \frac{\partial \hat{V}}{\partial \tau} - b'e^{i\eta b}(1-c).$$

*Proof* First consider

$$\mathscr{F}^b\left\{\frac{\partial V}{\partial x}\right\} = \int_b^{\infty} e^{i\eta x}\frac{\partial V}{\partial x}dx = [Ve^{i\eta x}]_b^{\infty} - i\eta\int_b^{\infty} e^{i\eta x}Vdx$$
$$= -V(b,\tau)e^{i\eta b} - i\eta\hat{V}(\eta,\tau). \qquad \text{(by definition of } \hat{V}\text{)}$$

Note that as in the corresponding derivation in Sect. 9.3 we tentatively assume that $lim_{x\to\infty}V(x,\tau)e^{i\eta x} = 0$. Finally by use of boundary condition (16.20),

$$\mathscr{F}^b\left\{\frac{\partial V}{\partial x}\right\} = (c-1)e^{i\eta b} - i\eta\hat{V}. \quad (16.26)$$

Next consider

$$\mathscr{F}^b\left\{\frac{\partial^2 V}{\partial x^2}\right\} = \int_b^\infty e^{i\eta x}\frac{\partial^2 V}{\partial x^2}dx = \left[\frac{\partial V}{\partial x}e^{i\eta x}\right]_b^\infty - i\eta\int_b^\infty \frac{\partial V}{\partial x}e^{i\eta x}dx$$
$$= -\frac{\partial V(x,\tau)}{\partial x}|_{x=b}\cdot e^{i\eta b} - i\eta\mathscr{F}^b\left\{\frac{\partial V}{\partial x}\right\},$$

where the last equality follows by use of the boundary condition (16.21) and use of (16.26). As in Sect. 9.3 we tentatively assume that $lim_{x\to\infty}V(x,\tau)e^{i\eta x} = 0$ so that the last equation simplifies to

$$\mathscr{F}^b\left\{\frac{\partial^2 V}{\partial x^2}\right\} = e^{i\eta b}(c - i\eta(c-1)) - \eta^2\hat{V}. \tag{16.27}$$

Finally consider

$$\mathscr{F}^b\left\{\frac{\partial V}{\partial \tau}\right\} = \int_b^\infty e^{i\eta x}\frac{\partial V(x,\tau)}{\partial \tau}dx = \frac{\partial}{\partial \tau}\left[\int_b^\infty e^{i\eta x}V(x,\tau)dx\right] + b'e^{i\eta b}V(b,\tau)$$
$$= \frac{\partial}{\partial \tau}[\mathscr{F}^b\{V\}] + b'e^{i\eta b}V(b,\tau),$$

where $b' \equiv \frac{db(\tau)}{d\tau}$. Applying the boundary condition (16.20) we have

$$\frac{\partial}{\partial \tau}[\mathscr{F}^b\{V\}] = b'e^{i\eta b}(c-1) + \mathscr{F}^b\left\{\frac{\partial V}{\partial \tau}\right\}.$$

Hence, finally

$$\mathscr{F}^b\left\{\frac{\partial V}{\partial \tau}\right\} = \frac{\partial \hat{V}}{\partial \tau} - b'e^{i\eta b}(c-1). \tag{16.28}$$

■

By use of Proposition 16.1, (16.26) becomes

$$\frac{\partial \hat{V}}{\partial \tau} - b'e^{i\eta b}(c-1) = \frac{1}{2}\sigma^2[e^{i\eta b}(c-i\eta(c-1)) - \eta^2\hat{V}] + k_1[e^{i\eta b}(c-1) - i\eta\hat{V}] - k_2\hat{V},$$

which may be re-arranged to yield

$$\frac{\partial \hat{V}}{\partial \tau} + \left(\frac{1}{2}\sigma^2\eta^2 + k_1 i\eta + k_2\right)\hat{V} = e^{i\eta b}\left[b'(c-1) + \frac{1}{2}\sigma^2(c - i\eta(c-1)) + k_1(c-1)\right]$$
$$= e^{i\eta b}\left[\frac{1}{2}\sigma^2 c + \left(b' - \frac{1}{2}\sigma^2 i\eta + k_1\right)(c-1)\right]$$
$$= F(\eta, \tau), \tag{16.29}$$

where we set

$$F(\eta, \tau) = e^{i\eta b}\left[\frac{\sigma^2 c}{2} + \left(b' - \frac{\sigma^2 i\eta}{2} + k_1\right)(c-1)\right]. \tag{16.30}$$

Equation (16.29) is an ordinary differential equation for $\hat{V}(\eta, \tau)$.

**Proposition 16.2** *The solution to the ordinary differential equation* (16.29) *is*

$$\hat{V}(\eta, \tau) = \int_0^\tau e^{-\left(\frac{\sigma^2\eta^2}{2} + k_1 i\eta + k_2\right)(\tau - S)} F(\eta, S)dS \tag{16.31}$$

*Proof* The initial condition becomes

$$\mathscr{F}\{V(x, 0)\} \equiv \hat{V}(\eta, 0).$$

The ordinary differential equation is of the form

$$\frac{d\hat{V}}{d\tau} + a_1(\eta)\hat{V} = F(\eta, \tau). \tag{16.32}$$

Using the integrating factor $e^{a_1(\eta)\tau}$, (16.32) becomes

$$\frac{d}{d\tau}(\hat{V}e^{a_1(\eta)\tau}) = F(\eta, \tau)e^{a_1(\eta)\tau}$$

whose solution may be expressed as

$$\hat{V}(\eta, \tau)e^{a_1(\eta)\tau} - \hat{V}(\eta, 0) = \int_0^\tau F(\eta, s)e^{a_1(\eta)s}ds. \tag{16.33}$$

Rearranging and substituting the explosion for $a_1(\eta)$, the solution for $\hat{V}(\eta, \tau)$ is given by

$$\hat{V}(\eta, \tau) = \hat{V}(\eta, 0)e^{-\left(\frac{1}{2}\sigma^2\eta^2 + k_1 i\eta + k_2\right)\tau} + \int_0^\tau e^{-\left(\frac{\sigma^2\eta^2}{2} + k_1 i\eta + k_2\right)(\tau - s)} F(\eta, s)ds. \tag{16.34}$$

From the definition of $\hat{V}(\eta, \tau)$, we can see that

$$\hat{V}(\eta, 0) = \int_{b(0)}^{\infty} V(x, 0)e^{i\eta x}dx.$$

At $\tau = 0$, $b(0) = 0$. Hence

$$\hat{V}(\eta, 0) = \int_0^{\infty} V(x, 0)e^{i\eta x}dx = \int_0^{\infty} H(-x)(1 - e^x)e^{i\eta x}dx = \int_{-\infty}^{0} 0 \cdot e^{i\eta x}dx = 0.$$

■

### 16.3.4 Inverting the Fourier Transform

Taking the inverse (complete) Fourier transform of (16.31) gives

$$\begin{aligned} V(x, \tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\eta x}\left[\int_0^{\tau} e^{-\left(\frac{\sigma^2\eta^2}{2}+k_1 i\eta+k_2\right)(\tau-s)}F(\eta, s)ds\right]d\eta \qquad (16.35)\\ &= \frac{1}{2\pi}\int_0^{\tau}\left[\int_{-\infty}^{\infty} e^{-\frac{\sigma^2\eta^2}{2}(\tau-s)-i\eta x-k_1 i\eta(\tau-s)-k_2(\tau-s)}F(\eta, s)d\eta\right]ds \end{aligned}$$

where $b(\tau) < x < \infty$.

**Proposition 16.3** *The inverse Fourier transform* (16.35) *may be expressed as*

$$\begin{aligned} V(x, \tau) = \int_0^{\tau} \frac{e^{-g(x,s)}}{\sigma\sqrt{2\pi(\tau - s)}}\Bigg[\frac{\sigma^2 c(s)}{2} + \Bigg(b'(s) + k_1 \\ - \left[\frac{x - b(s) + k_1(\tau - s)}{2(\tau - s)}\right]\Bigg)(c(s) - 1)\Bigg]ds, \qquad (16.36) \end{aligned}$$

*where*

$$g(x, s) = \frac{(x - b(s) + k_1(\tau - s))^2}{2\sigma^2(\tau - s)} + k_2(\tau - s). \qquad (16.37)$$

*Proof* Consider

$$\begin{aligned} F(\eta, s) &= e^{i\eta b(s)}\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s) - \frac{\sigma^2 i\eta}{2} + k_1\right)(c(s) - 1)\right]\\ &= e^{i\eta b(s)}\left[\frac{\sigma^2 c(s)}{2} + (b'(s) + k_1)(c(s) - 1) - \frac{\sigma^2 i}{2}(c(s) - 1)\eta\right]\\ &= e^{i\eta b(s)}\{f_1(s) - \eta f_2(s)\}, \qquad (16.38) \end{aligned}$$

where

$$f_1(s) = \frac{\sigma^2 c(s)}{2} + (b'(s) + k_1)(c(s) - 1),$$

and

$$f_2(s) = \frac{\sigma^2 i}{2}(c(s) - 1).$$

Thus we can rewrite (16.33) as

$$V(x,\tau) = \frac{1}{2\pi}\int_0^{\tau} e^{-k_2(\tau-s)}\left[\int_{-\infty}^{\infty} e^{-p\eta^2-q\eta}\{f_1(s) - \eta f_2(s)\}d\eta\right]ds, \tag{16.39}$$

where

$$p = \frac{\sigma^2}{2}(\tau - s), \text{ and } q = i(x + k_1(\tau - s) - b).$$

Using the result that

$$\int_{-\infty}^{\infty} e^{-p\eta^2-q\eta}\eta^k d\eta = (-1)^k\sqrt{\frac{\pi}{p}}\frac{\partial^k}{\partial q^k}e^{\frac{q^2}{4p}},$$

subject to $Re(p) \geq 0$, (which is true since $p = \frac{\sigma^2}{2}(\tau - s)$, $0 < s < \tau$), we have

$$\int_{-\infty}^{\infty} e^{-p\eta^2-q\eta^2}d\eta = \sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}}, \tag{16.40}$$

and

$$\int_{-\infty}^{\infty} e^{-p\eta^2-q\eta}\eta d\eta = -\sqrt{\frac{\pi}{p}}\frac{\partial}{\partial q}e^{\frac{q^2}{4p}} = -\sqrt{\frac{\pi}{p}}\frac{q}{2p}e^{\frac{q^2}{4p}}. \tag{16.41}$$

By use of (16.40) and (16.41), (16.39) becomes

$$\begin{aligned}
V(x,\tau) &= \frac{1}{2\pi}\int_0^{\tau} e^{-k_2(\tau-s)}\left[f_1(s)\sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}} + f_2(s)\sqrt{\frac{\pi}{p}}e^{\frac{q^2}{4p}}\frac{q}{2p}\right]ds \\
&= \int_0^{\tau} e^{-k_2(\tau-s)}\frac{e^{\frac{q^2}{4p}}\sqrt{\pi}}{2\pi\sqrt{p}}\left[f_1(s) + \frac{qf_2(s)}{2p}\right]ds \qquad (16.42) \\
&= \int_0^{\tau}\frac{e^{-k_2(\tau-s)+\frac{q^2}{4p}}}{2\sqrt{\pi p}}\left[\frac{\sigma^2 c(s)}{2} + (b'(s) + k_1)(c(s) - 1) + \frac{\sigma^2 i(c(s) - 1)q}{2\sigma^2(\tau - s)}\right]ds.
\end{aligned}$$

To further simplify this expression and eliminate $p$ and $q$, consider the following manipulations:

(i)

$$2\sqrt{\pi p} = 2\sqrt{\frac{\pi\sigma^2}{2}(\tau - s)} = \sigma\sqrt{2\pi(\tau - s)},$$

(ii)

$$\frac{q^2}{4p} = \frac{i^2(x + k_1(\tau - s) - b)^2}{2\sigma^2(\tau - s)} = -\frac{(x + k_1(\tau - s) - b)^2}{2\sigma^2(\tau - s)},$$

(iii)

$$\begin{aligned}&(b'(s) + k_1)(c(s) - 1) + \frac{\sigma^2 i(c(s) - 1)q}{2\sigma^2(\tau - s)}\\ &= (c(s) - 1)\left(b'(s) + k_1 + \frac{\sigma^2 i^2(x + k_1(\tau - s) - b(s))}{2\sigma^2(\tau - s)}\right)\\ &= (c(s) - 1)\left(b'(s) + k_1 - \left[\frac{x + k_1(\tau - s) - b(s)}{2(\tau - s)}\right]\right).\end{aligned}$$

By use of the foregoing results (16.42) can be written

$$\begin{aligned}V(x, \tau) = \int_0^\tau &\left(\frac{\exp\{-k_2(\tau - s) - (x + k_1(\tau - s) - b(s))^2/2\sigma^2(\tau - s)\}}{\sigma\sqrt{2\pi(\tau - s)}}\right)\\ &\times\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s) + k_1 - \left[\frac{x - b(s) + k_1(\tau - s)}{2(\tau - s)}\right]\right)(c(s) - 1)\right]ds.\end{aligned}$$

■

Equation (16.36) expresses the value of the American call option in terms of the early exercise boundary $b(\tau)$. At this point it remains unknown, but we are able to obtain an integral equation that determines it by requiring the expression for $V(x, \tau)$ in (16.36) to satisfy the early exercise boundary condition (16.20), application of which yields

$$\begin{aligned}\frac{1 - c(\tau)}{2} = \int_0^\tau \frac{e^{-g(b(\tau),s)}}{\sigma\sqrt{2\pi(\tau - s)}}&\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s)\right.\right.\\ &\left.\left.+ k_1 - \left[\frac{b(\tau) - b(s) + k_1(\tau - s)}{2(\tau - s)}\right]\right)(c(s) - 1)\right]ds, \qquad (16.43)\end{aligned}$$

where

$$g(b(\tau), s) = \frac{(b(\tau) - b(s) + k_1(\tau - s))^2}{2\sigma^2(\tau - s)} + k_2(\tau - s). \tag{16.44}$$

The factor of $\frac{1}{2}$ on the left in (16.43) arises because we originally found the Fourier transform of the function $\mathscr{H}(x - b)V(x, \tau)$, which is piecewise continuous in the region $-\infty < x < \infty$. From the theory of Fourier transforms, it can be shown that if $f(x)$ is piecewise continuous, then

$$\frac{f(x^+) + f(x^-)}{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\eta x} \int_{-\infty}^{\infty} f(x) e^{i\eta x} dx d\eta \tag{16.45}$$

for all $x \in \mathbb{R}$.

In (16.45) we set $f(x) = \mathscr{H}(x - b)V(x, \tau)$ and consider the point of discontinuity of this function at $x = b$. Hence, at the point of discontinuity, $x = b(\tau)$, we have

$$\begin{aligned} \mathscr{F}^{-1}\{\hat{V}(\eta, \tau)\} &= \frac{\mathscr{H}(b(\tau)^- - b)V(b(\tau)^-, \tau) + \mathscr{H}(b(\tau)^+ - b)V(b(\tau)^+, \tau)}{2} \\ &= \frac{0.V(b(\tau)^-, \tau) + 1.V(b(\tau)^+, \tau)}{2} \\ &= \frac{0 + (1 - e^{b(\tau)})}{2} = \frac{1 - c(\tau)}{2}. \end{aligned} \tag{16.46}$$

The convergence of the inverse of the incomplete Fourier transform to the mid-point at $x = b(\tau)$ is illustrated in Fig. 16.6.

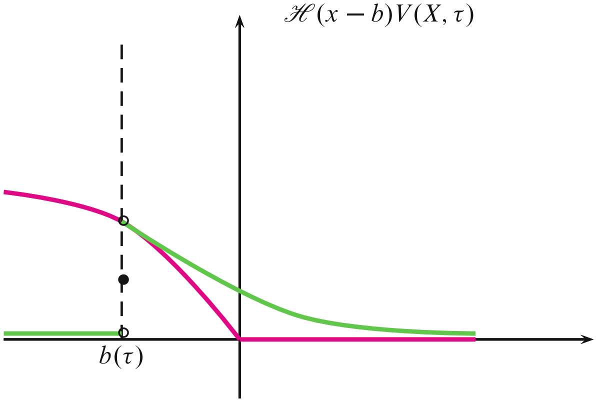


**Fig. 16.6** The convergence of the inverse of the incomplete Fourier transform to the mid-point at $x = b(\tau)$

Now the reasoning behind our initial choice for the form of $\mathscr{H}(x - b(\tau))$ in (16.23) is clear. To more properly account for the discontinuity, the solution could also be expressed as

$$\begin{aligned}
&\mathscr{H}(x - b(\tau))V(x,\tau) \\
&= \int_0^\tau \frac{e^{-g(x,s)}}{\sigma\sqrt{2\pi(\tau - s)}}\left[\frac{\sigma^2 c(s)}{2} + \left(b'(s) + k_1 - \left[\frac{x - b(s) + k_1(\tau - s)}{2(\tau - s)}\right]\right)\right. \\
&\quad \left. \times (c(s) - 1)\right] ds,
\end{aligned}$$

where $g(x,s)$ is given by Eq. (16.37), and $-\infty < x < \infty$. This will "neatly" account for the behaviour of the solution at the discontinuity $x = b(\tau)$ when taking the inverse Fourier transform of $\hat{V}(\eta,\tau)$.

Using the fact that $b(\tau) = \ln(c(\tau))$ the integral (16.43) can be written as

$$\begin{aligned}
\frac{1 - c(\tau)}{2} = \int_0^\tau \frac{e^{-g(c(\tau),s)}}{\sigma\sqrt{2\pi(\tau - s)}}&\left[\frac{\sigma^2 c(s)}{2} + \left(\frac{c'(s)}{c(s)}\right.\right. \\
&\left.\left. + k_1 - \left[\frac{\ln(\frac{c(\tau)}{c(s)}) + k_1(\tau - s)}{2(\tau - s)}\right]\right)(c(s) - 1)\right] ds
\end{aligned} \tag{16.47}$$

where

$$g(c(\tau),s) = \frac{\left(\ln\left(\frac{c(\tau)}{c(s)}\right) + k_1(\tau - s)\right)^2}{2\sigma^2(\tau - s)} + k_2(\tau - s). \tag{16.48}$$

## 16.4 An Approximate Solution

The formulation in the previous subsection solves the American valuation problem in a fairly general way. However often it is useful to obtain a simpler approximate solution. Here we outline one such solution obtained by Macmillan (1986), this solution has been generalised to the general commodity option situation by Barone-Adesi and Whaley (1987).

Let $e(S,\tau)$ be the amount by which the American put option is worth more than the European put option i.e.

$$P(S,\tau) = p(S,\tau) + e(S,\tau), \tag{16.49}$$

where $p$ is the European put option value. Since $P$ and $p$ both satisfy the partial differential (16.1) it follows that $e$ satisfies a partial differential equation of the same form

$$\frac{1}{2}\sigma^2 S^2 e_{SS} + rSe_S - re = e_\tau, \tag{16.50}$$

but with the boundary condition $e(S,0) = 0$. Defining $M = \frac{2r}{\sigma^2}$, (16.50) can be written

$$S^2 e_{SS} + MSe_S - Me - \frac{M}{r} e_\tau = 0.$$

Introduce the change of time variable from $\tau$ to $\theta$ according to

$$\theta = K(\tau) \tag{16.51}$$

where the functional form of $K$ will be specified later. The partial differential equation for $e$ then becomes

$$S^2 e_{SS} + MSe_S - Me - \frac{M}{r} K_\tau e_\theta = 0.$$

We then write the early exercise premium in the form

$$e(S,\tau) = e(S, K^{-1}(\theta)) \equiv \theta f(S,\theta)$$

so that $f$ satisfies

$$S^2 f_{SS} + MSf_S - Mf\left[1 + \frac{K_\tau}{r\theta}\left(1 + \theta\frac{f_\theta}{f}\right)\right] = 0.$$

A useful choice of $K(\tau)$ would be such that most of the time dependence of the early exercise premium is contained in this factor. One such choice turns out to be

$$K(\tau) = 1 - e^{-r\tau}, \tag{16.52}$$

for which $K_\tau = r(1-\theta)$. With this choice the last partial differential equation becomes

$$S^2 f_{SS} + MSf_S - \frac{Mf}{\theta}\left[1 + (1-\theta)\theta\frac{f_\theta}{f}\right] = 0. \tag{16.53}$$

The term $\theta(1-\theta)$ is zero at $\tau = 0$ and $\tau = \infty$ and has a maximum value of $\frac{1}{4}$ at $\theta = \frac{1}{2}$. Thus ignoring this term it should produce a reasonable approximation for small $\tau$ and large $\tau$ without too much error for intermediate $\tau$. Setting this term to zero reduces (16.53) to the ordinary differential equation

$$S^2 f_{SS} + MSf_S - \left(\frac{M}{\theta}\right) f = 0$$

which has the solution

$$f(S) = a_1 S^{q_1} + a_2 S^{q_2},$$

where $q_1$ and $q_2$ are the negative and positive roots of the quadratic equation

$$q^2 + (M-1)q - \frac{M}{\theta} = 0. \tag{16.54}$$

In order that $\lim_{S\to\infty} P(S,\tau) = 0$ be satisfied we ignore the positive root. Hence

$$f(S) = a_1 S^{q_1}. \tag{16.55}$$

To complete the solution we need to determine the constant $a_1$ and the critical stock price $B$. Those are determined by forcing the solution derived from (16.55) to satisfy the early exercise condition (16.9b) and the tangency condition (16.9d).

Details are given in Macmillan (1986) and the solution turns out to be

$$P(S,\tau) = p(S,\tau) + A\left(\frac{S}{B}\right)^q, \tag{16.56}$$

where $q$ is the negative root of (16.54), $B$ is given by

$$B = -\frac{q(E-p(B))}{\mathcal{N}(d_1(B)) - q}, \tag{16.57}$$

and $A$ is given by

$$A = \frac{B\mathcal{N}(d_1(B))}{-q}. \tag{16.58}$$

## 16.5 Appendix

### Appendix 16.1 The Incomplete Fourier Transform

To prove that if

$$f(x,\tau) = \mathscr{H}(a-x)g(x,\tau), \quad a \equiv a(\tau), \quad \mathscr{H}(a-x) \equiv \text{Heaviside Function},$$

then the standard Fourier [inversion theorem]

$$f(x,\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta, \quad -\infty < x < \infty,$$

yields

$$g(x,\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a} g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta, \quad -\infty < x \le a,$$

On the left,

$$\mathscr{H}(a-x)g(x,\tau) = \begin{cases} g(x,\tau), & -\infty < x \le a \\ 0, & a < x < \infty \end{cases}$$

On the right,

$$\begin{aligned}
&\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} \mathscr{H}(a-x)g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a} \mathscr{H}(a-x)g(x,\tau)e^{i\eta x}dx\right. \\
&\quad \left. + \int_{a}^{\infty} H(a-x)g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a} \mathscr{H}(a-x)g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta \\
&\quad + \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{a}^{\infty} \mathscr{H}(a-x)g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a} g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta.
\end{aligned}$$

Hence

$$\mathscr{H}(a-x)g(x,\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a}g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta,\quad -\infty<x<\infty$$

or alternatively,

$$g(x,\tau)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a}g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta,\quad -\infty<x<a$$

and

$$\frac{g(x,\tau)}{2}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{a}g(x,\tau)e^{i\eta x}dx\right]e^{-i\eta x}d\eta,\quad x=a.$$

## 16.6 Problems

**Problem 16.1** Verify the approximate solution to the American put problem given in (16.56) by using the boundary conditions for the American option to determine the coefficients for $A$ and $B$ in Eqs. (16.57) and (16.58).

# Chapter 17
# Pricing Options Using Binomial Trees

**Abstract** This chapter presents the binomial tree approach to the option pricing problem. We first illustrate the basic ideas of option pricing by considering the one-period binomial tree model and then extend to a multi-period binomial tree model. We then show that, by taking limits in an appropriate way, the binomial expression for the option price converges to the Black–Scholes option price and pricing equation. Alternatively, the continuous time model can be discretised in a way that yields the same expressions as obtained by the binomial tree approach.

## 17.1 Introduction

What distinguishes the approach of this chapter from the finite differences approach is that the point in the modelling process where discretisation takes place. In the finite-difference case we start with a continuous time model of the price process for the underlying asset, derive a continuous time equation for the option price and then discretise this. In the binomial tree case we start with a discrete time model of the price process for the underlying asset and then derive a discrete time expression for the option price.

Of course these two approaches can be shown to be equivalent. The continuous time model can be discretised in a way that yields the same expressions as obtained by the binomial tree approach. Similarly, by taking limits in an appropriate way the binomial expression for the option price can be shown to converge to the Black–Scholes option pricing equation. The distinction between the discrete time and continuous time approaches is illustrated in Fig. 17.1.



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_17

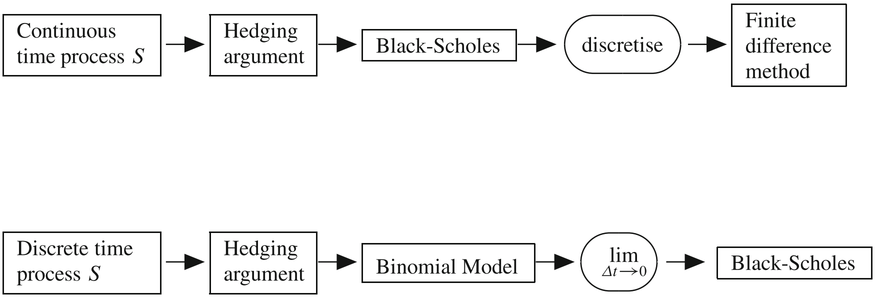


**Fig. 17.1** The different ways to view discretisation of the option pricing problem

## 17.2 The Binomial Model

### *17.2.1 The Binomial Stock Price Process*

Instead of the continuous time processes that we have employed hitherto we now assume time evolves discretely in intervals of length $h$ ($h = 1$ day, 1 week etc.) as shown in Fig. 17.2.

In this discrete time setting we allow the stock price to have only two possible movements from one time period to the next, namely an up or a down movement as illustrated in Fig. 17.3. Mathematically, this may be described as

$$S_{i+1} = \begin{cases} uS_i, & \text{with prob } p, \\ dS_i, & \text{with prob } (1-p), \end{cases} \tag{17.1}$$

where the probability $p$ relates to the underlying binomial distribution and the parameters $u, d$ relate to the variance of this distribution.

To see the connection with the continuous time framework discretise the continuous time process

$$\frac{dS}{S} = \mu dt + \sigma dz, \tag{17.2}$$

using the Euler–Maruyama method to obtain

$$S_{i+1} = (1 + \mu h)S_i + \sigma S_i \sqrt{h}\, \tilde{\xi}_i, \tag{17.3}$$

where $\tilde{\xi}_i \sim N(0, 1)$.

Suppose we replace the normal variate $\tilde{\xi}_i$ with a Bernoulli variate $\tilde{\beta}_i$ where

$$\tilde{\beta}_i = \begin{cases} +1, & \text{with prob } p \\ -1, & \text{with prob } (1-p) \end{cases} \tag{17.4}$$

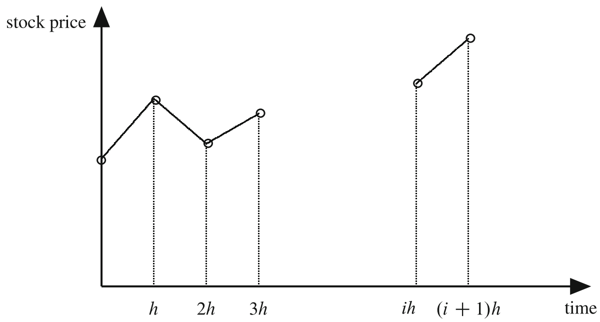


**Fig. 17.2** Discrete time evolution of the stock price

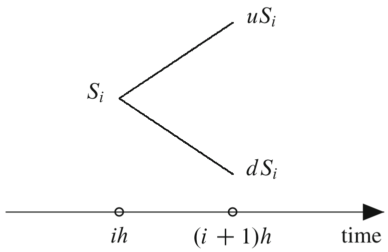


**Fig. 17.3** The one-period binomial tree for the stock price

then the discrete time process (17.3) reads

$$S_{i+1} = (1 + \mu h)S_i + \sigma S_i \sqrt{h}\ \tilde{\beta}_i. \tag{17.5}$$

In other words

$$S_{i+1} = \begin{cases} (1 + \mu h + \sigma\sqrt{h})S_i, & \text{with prob } p \\ (1 + \mu h - \sigma\sqrt{h})S_i, & \text{with prob } (1-p). \end{cases} \tag{17.6}$$

Hence we can relate the $u, d$ of (17.1) to the $\mu, \sigma$ of the continuous time approach, namely

$$\begin{aligned} u &= 1 + \mu h + \sigma\sqrt{h}, \\ d &= 1 + \mu h - \sigma\sqrt{h}. \end{aligned} \tag{17.7}$$

From the last equations we are able to derive the important result that

$$\sigma = \frac{u-d}{2\sqrt{h}}. \tag{17.8}$$

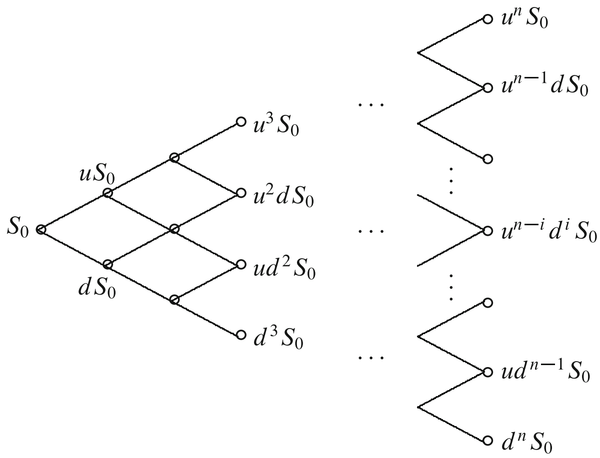


**Fig. 17.4** The $n$-period binomial tree for the stock price

We also note from (17.6) that

$$\mathbb{E}\left[\frac{S_{i+1} - S_i}{S_i}\right] = \mu h + (2p - 1)\sigma\sqrt{h}, \tag{17.9}$$

and

$$\text{var}\left[\frac{S_{i+1} - S_i}{S_i}\right] = 4p(1 - p)\sigma^2 h. \tag{17.10}$$

From (17.9) and (17.10) we observe that the choice $p = 1/2$ will yield the same mean and variance for expected stock returns as in continuous time.

By extending the binomial stock price movements over $n$ periods we generate a binomial tree for the stock price process as illustrated in Fig. 17.4. Note that the general period $n$ stock price, $u^{n-i} d^i S_0$, can be arrived at by

$$\frac{n!}{i!(n-i)!} = \binom{n}{i}$$

paths.

### 17.2.2 Option Pricing in the One-Period Model

Consider a one-period binomial stock price process and consider a European call option written on this stock, as illustrated in Fig. 17.5. The option has exercise price $E$ and matures at the end of the period.

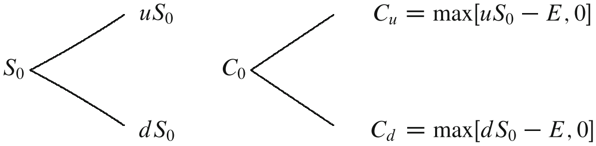


**Fig. 17.5** The one-period European call option

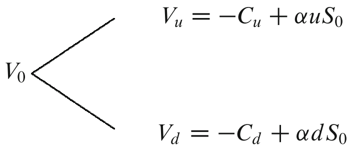


**Fig. 17.6** Evolution of the hedging portfolio

We follow the now standard procedure of forming a hedging portfolio of short 1 call and long $\alpha$ stock. The value of the portfolio at 0 is

$$V_0 = C_0 - \alpha S_0, \tag{17.11}$$

and its evolution over the period is as shown in Fig. 17.6. If $\alpha$ is chosen so that at $t = 1$

$$V_u = V_d, \tag{17.12}$$

then the hedging portfolio is riskless. Equation (17.12) implies

$$-C_u + \alpha u S_0 = -C_d + \alpha d S_0,$$

from which

$$\alpha = \frac{C_u - C_d}{(u-d)S_0}. \tag{17.13}$$

Since the hedging portfolio is riskless, the initial investment of $\alpha S_0 - C_0$ must earn the risk free rate. That is

$$(\alpha S_0 - C_0)(1 + rh) = \alpha u S_0 - C_u, \tag{17.14}$$

where $r$ is the continuous time risk free interest rate. Since $\alpha$ is given by (17.13), (17.14) may be solved for $C_0$ to yield

$$C_0 = \frac{p^* C_u + (1 - p^*) C_d}{R}, \tag{17.15}$$

where

$$p^* = \frac{R - d}{u - d}, \tag{17.16}$$

and

$$R = 1 + rh. \tag{17.17}$$

We note how the expression (17.15) for $C_0$ may be given a discounted cash flow interpretation if we interpret $p^*$ as the probability of the stock price moving up over the period. Of course, in general, $p^*$ will not equal the actual (or historical) probability of an up-movement which we have denoted by $p$ in (17.4). In fact $p^*$ is a discretised version of the equivalent or risk neutral measure $p^*(S_T, T|S, t)$ which we encountered in Sect. 8.4. We recall from (8.107) that

$$C(S,t) = e^{-r(T-t)} \int_0^\infty \max[S_T^* - E, 0] p^*(S_T^*, T \mid S, t) dS_T^*. \tag{17.18}$$

Here $t = 0, T = h$ and under the measure $p^*$ there are only two possible outcomes for $S_h$, i.e. $p^*(S_h, h|S, 0)$ is approximated by the binomial distribution shown in Fig. 17.7. Noting further that

$$e^{-rh} \simeq 1 - rh \simeq \frac{1}{1 + rh}, \tag{17.19}$$

we see that indeed the continuous time expression (17.18) reduces to the one-period binomial expression (17.15).

To prove this last assertion formally we note that the binomial approximation to $p^*(S_h, h|S, 0)$ may by written

$$p^*(S_h, h|S, 0) = p^* \delta(uS - S_h) + (1 - p^*)\delta(dS - S_h) \tag{17.20}$$

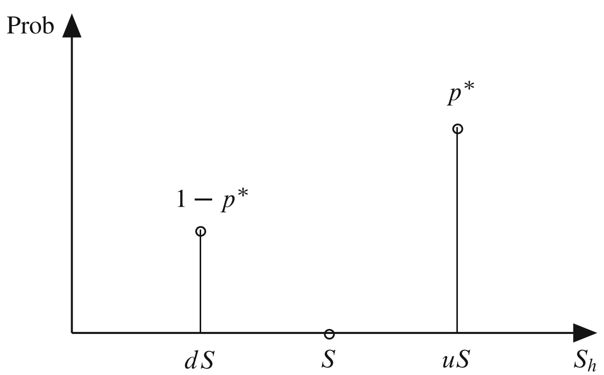


**Fig. 17.7** The one-period binomial distribution

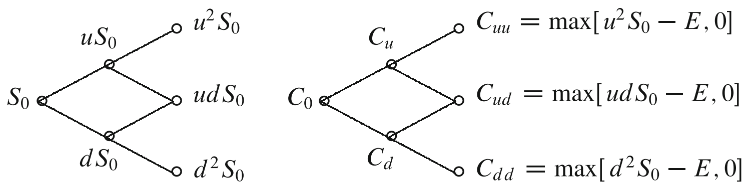


**Fig. 17.8** Two-period stock price and option price trees

and hence

$$\int_0^\infty \max[S_h - E, 0] p^*(S_h, h|S, 0) dS_h = \\ p^* \max[uS - E, 0] + (1 - p^*) \max[dS - E, 0]. \tag{17.21}$$

Substituting (17.19) and (17.21) into (17.18) yields the one-period binomial option price (17.15).

### *17.2.3 Two Period Binomial Option Pricing*

Now we follow the evolution of the stock price over two periods $(0, h)$, $(h, 2h)$. The evolution of the option price also follows a two period binomial process as shown in Fig. 17.8.

There are a number of approaches to pricing the European call option in the two-period binomial tree. These include:

(i) Breaking the two-period problem into a sequence of one-period problems. Starting from maturity, use the one-period argument to compute $C_u$, $C_d$. Then again use the one period model to compute $C_0$. This approach provides the basis for most computational schemes and extends easily to handle American options.
(ii) Apply over the two-periods the discounted cash flow argument using $p^*$ as the probability of an up-movement over one-period.
(iii) In the integral expression (17.18) approximate $p^*(S_{2h}, 2h|S, 0)$ by a two-period binomial distribution.

Of course, all of these are equivalent, as we shall show by considering each one in turn.

Applying the one-period argument over $(h, 2h)$ we readily obtain

$$C_u = \frac{p^* C_{uu} + (1 - p^*) C_{ud}}{R}, \\ C_d = \frac{p^* C_{ud} + (1 - p^*) C_{dd}}{R}. \tag{17.22}$$

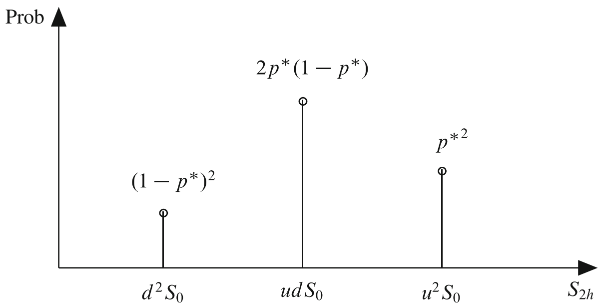


**Fig. 17.9** The two-period binomial distribution

Next applying the one-period argument over $(0, h)$ we obtain the option value at current time

$$C_0 = \frac{p^* C_u + (1 - p^*) C_d}{R}. \tag{17.23}$$

By means of this approach we are able to fill in the option values at all nodes of the tree by stepping back from maturity to initial time. As we shall see, this approach is easily adapted to evaluate options with special features such those as on dividend paying stock and American options. By substituting (17.22) into (17.23) and performing some algebraic manipulations we find that $C_0$ may be expressed as

$$C_0 = \frac{p^{*2} C_{uu} + 2p^*(1 - p^*) C_{ud} + (1 - p^*)^2 C_{dd}}{R^2}. \tag{17.24}$$

This expression has an obvious discounted cash flow interpretation once we recognize that

$$p^{*2}, \qquad 2p^*(1 - p^*), \qquad (1 - p^*)^2 \tag{17.25}$$

are the risk neutral probabilities associated with the final time stock prices

$$u^2 S_0, \qquad udS_0, \qquad d^2 S_0.$$

In fact the probabilities (17.25) represent the two period binomial distribution approximation to $p^*(S_{2h}, 2h|S, 0)$, as shown in Fig. 17.9. If we substitute the binomial distribution into the continuous time integral expression (17.18) we would again obtain the two period binomial option price (17.24).

### 17.2.4 n-Period Binomial Option Pricing

Now we follow the evolution of the stock and option prices over $n$ periods. We work through the option pricing tree by backward induction. If the option prices at

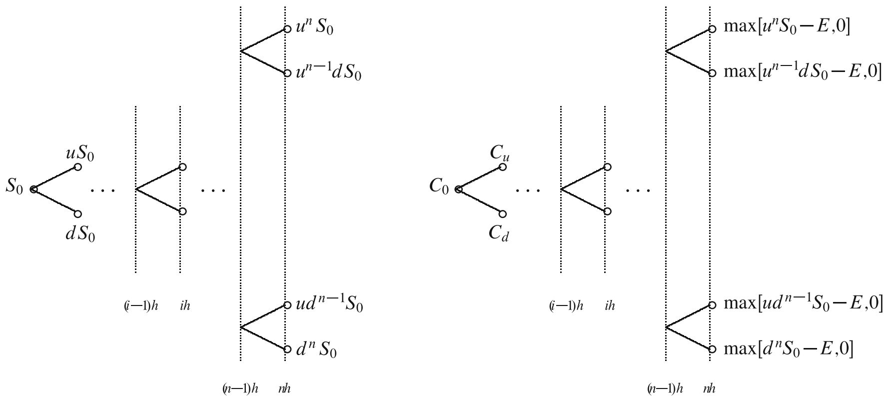


**Fig. 17.10** The $n$-period binomial tree

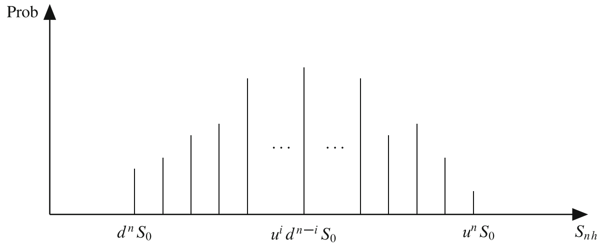


**Fig. 17.11** The $n$-period binomial distribution

all nodes at time $ih$ are known then those at the nodes at time $(i-1)h$ are obtained by considering the sequence of one-period problems linking $(i-1)h$ and $ih$. We proceed backwards from $i = n$ to $i = 0$. This procedure is easily programmed (Fig. 17.10).

To obtain the alternative view-point note first of all that the risk-neutral probability associated with the stock price $u^i d^{n-i} S_0$ is

$$\binom{n}{i} p^{*i} (1-p^*)^{n-i}. \tag{17.26}$$

Hence we have an $n$-period binomial distribution as shown in Fig. 17.11. Substituting the n-period binomial distribution into the continuous time integral (17.18) we obtain

$$C_0 = \frac{1}{R^n} \sum_{i=0}^{n} \binom{n}{i} p^{*i} (1-p^*)^{n-i} \max[u^i d^{n-i} S_0 - E, 0], \tag{17.27}$$

which has the, by now familiar, discounted cash flow interpretation.

Suppose we let

$$a = \begin{array}{c}\text{the minimum number of upward moves that the stock must make}\\ \text{over the } n\text{-periods for the call option to finish in the money,}\end{array}$$

then $a$ is the smallest non-negative integer such that

$$S_0 u^a d^{n-a} > E. \tag{17.28}$$

The expression (17.27) for $C_0$ may then be written

$$C_0 = \frac{1}{R^n} \sum_{i=a}^{n} \binom{n}{i} p^{*i} (1-p^*)^{n-i} \, (u^i d^{n-i} S_0 - E). \tag{17.29}$$

The last equation may be re-expressed as

$$\begin{aligned} C_0 &= S_0 \sum_{i=a}^{n} \binom{n}{i} p^{*i} (1-p^*)^{n-i} \left(\frac{u}{R}\right)^i \left(\frac{d}{R}\right)^{n-i} - ER^{-n} \sum_{i=a}^{n} \binom{n}{i} p^{*i} (1-p^*)^{n-i} \\ &= S_0 \sum_{i=a}^{n} \binom{n}{i} \tilde{p^*}^i (1-\tilde{p^*})^{n-i} - ER^{-n} \sum_{i=a}^{n} \binom{n}{i} p^{*i} (1-p^*)^{n-i}, \end{aligned} \tag{17.30}$$

where $\tilde{p^*} = \dfrac{u}{R} p^*$. Before proceeding we note that,

$$\begin{aligned} \Phi(a; n, p) &= \text{the complementary binomial distribution} \\ &= \sum_{i=a}^{n} \binom{n}{i} p^i (1-p)^{n-i} \qquad\qquad (17.31) \\ &= \text{the probability of at least } a \text{ upmoves} \\ &\quad \text{of the stock where } p \text{ is the probability of an upmove.} \end{aligned}$$

Hence we can write (17.30) as,

$$C_0 = S_0 \Phi(a; n, \tilde{p^*}) - ER^{-n} \Phi(a; n, p^*). \tag{17.32}$$

## 17.3 The Continuous Limit

We shall defer to a later point the discussion on the choice of $u$ and $d$. However suppose we choose them according to

$$u = e^{\sigma\sqrt{h}} \qquad \text{and} \qquad d = e^{-\sigma\sqrt{h}}, \tag{17.33}$$

where $\sigma$ is the volatility of the corresponding continuous time stock price process. Then[1]

$$\frac{u-d}{2\sqrt{h}} = \frac{e^{\sigma\sqrt{h}} - e^{-\sigma\sqrt{h}}}{2\sqrt{h}} \simeq \frac{(1+\sigma\sqrt{h}) - (1-\sigma\sqrt{h})}{2\sqrt{h}} = \sigma. \tag{17.34}$$

So that the above choice of $u$ and $d$ satisfies reasonably well the relation (17.8) between $u, d$ and $\sigma$.

### 17.3.1 The Limiting Binomial Distribution

With this choice of $u$ and $d$ it is possible to show that

$$\lim_{h\to 0} \Phi(a; n, \tilde{p^*}) = \mathcal{N}(d_1) \tag{17.35}$$

and

$$\lim_{h\to 0} \Phi(a; n, p^*) = \mathcal{N}(d_2), \tag{17.36}$$

where as usual

$$d_1 = \frac{\ln(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}.$$

Noting also that

$$\lim_{h\to 0} R^{-n} = \lim_{h\to 0}(1+rh)^{-n} = \lim_{n\to\infty}\left(1 + r\frac{(T-t)}{n}\right)^{-n} = e^{-r(T-t)},$$

we see in the limit $h \to 0, n \to \infty$ that (17.32) reduces to the Black–Scholes formula

$$C_0 = S_0\mathcal{N}(d_1) - Ee^{-r(T-t)}\mathcal{N}(d_2).$$

[1]Note that

$$e^{\sigma\sqrt{h}} \simeq 1 + \sigma\sqrt{h} + \frac{1}{2}\sigma^2 h, \qquad e^{-\sigma\sqrt{h}} \simeq 1 - \sigma\sqrt{h} + \frac{1}{2}\sigma^2 h.$$

### 17.3.2 The Black–Scholes Partial Differential Equation as the Limit of the Binomial

Now we show how the binomial expression (17.15) reduces to the Black–Scholes partial differential equation in the limit $h \to 0$ and $n \to \infty$.

We consider (17.15) in the form

$$\left(\frac{R-d}{u-d}\right) C_u + \left(\frac{u-R}{u-d}\right) C_d - RC_0 = 0. \tag{17.37}$$

If we denote current time and stock price by $t, S$, respectively, then

$$\begin{aligned} C_0 &= C(S,t), \\ C_u &= C(e^{\sigma\sqrt{h}}S, t+h), \\ C_d &= C(e^{-\sigma\sqrt{h}}S, t+h), \end{aligned} \tag{17.38}$$

where again we use the values $u, d$ given in (17.33). Next we note that

$$\begin{aligned} C_u &= C(e^{\sigma\sqrt{h}}S, t+h) \simeq C((1+\sigma\sqrt{h}+\frac{1}{2}\sigma^2 h)S, t+h) \\ &= C(S+(\sigma\sqrt{h}+\frac{1}{2}\sigma^2 h)S, t+h) \\ &= C(S,t) + \frac{\partial C}{\partial S}\sigma S\sqrt{h} + \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\frac{\partial C}{\partial S}\sigma^2 S + \frac{\partial C}{\partial t}\right) h + o(h), \end{aligned}$$

whilst

$$\begin{aligned} C_d &= C(e^{-\sigma\sqrt{h}}S, t+h) \simeq C((1-\sigma\sqrt{h}+\frac{1}{2}\sigma^2 h)S, t+h) \\ &= C(S+(-\sigma\sqrt{h}+\frac{1}{2}\sigma^2 h)S, t+h) \\ &= C(S,t) - \frac{\partial C}{\partial S}\sigma S\sqrt{h} + \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\frac{\partial C}{\partial S}\sigma^2 S + \frac{\partial C}{\partial t}\right) h + o(h). \end{aligned}$$

Using these expressions as well as (17.34) the expression (17.37) becomes

$$\begin{aligned} &\frac{(R-e^{-\sigma\sqrt{h}})}{2\sigma\sqrt{h}} \left[ C(S,t) + \frac{\partial C}{\partial S}\sigma S\sqrt{h} + \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\frac{\partial C}{\partial S}\sigma^2 S + \frac{\partial C}{\partial t}\right) h \right] \\ &+ \frac{(e^{\sigma\sqrt{h}}-R)}{2\sigma\sqrt{h}} \left[ C(S,t) - \frac{\partial C}{\partial S}\sigma S\sqrt{h} + \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2}\frac{\partial C}{\partial S}\sigma^2 S + \frac{\partial C}{\partial t}\right) h \right] \\ &- RC(S,t) + o(h) = 0. \end{aligned}$$

Combining these terms and using $R = 1 + rh$ reduces to

$$\frac{1}{2}\sigma^2 S^2 h \frac{\partial^2 C}{\partial S^2} + rhS \frac{\partial C}{\partial S} - rhC(S,t) + h\frac{\partial C}{\partial t} + o(h) = 0.$$

Dividing by $h$ and letting $h \to 0$, we obtain the Black–Scholes partial differential equation.

### 17.3.3 The Binomial as a Discretisation of the Black–Scholes Partial Differential Equation

It is also instructive to see how the Black–Scholes partial differential equation can be discretised to yield the binomial model, however first we need to express the Black–Scholes partial differential equation in terms of the log of the stock price.

Consider the Black–Scholes partial differential equation

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC + \frac{\partial C}{\partial t} = 0, \tag{17.39}$$

and the change of variable

$$y = \ln S. \tag{17.40}$$

We readily calculate that

$$\frac{\partial C}{\partial S} = \frac{1}{S}\frac{\partial C}{\partial y} \text{ and } \frac{\partial^2 C}{\partial S^2} = \frac{1}{S^2}\frac{\partial^2 C}{\partial y^2}, \tag{17.41}$$

so that in the transformed variable the Black–Scholes equation becomes

$$\frac{1}{2}\sigma^2 \frac{\partial^2 C}{\partial y^2} + r\frac{\partial C}{\partial y} - rC + \frac{\partial C}{\partial t} = 0. \tag{17.42}$$

If we wish to discretise the Black–Scholes partial differential equation, so as to make a comparison with the binomial model then the changes in $S$ have to be proportional. This means the grid for $S$ would be as shown in Fig. 17.12.

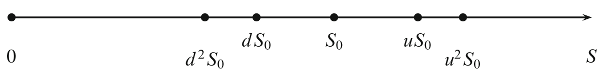


**Fig. 17.12** The unequally spaced grid in $S$ space

It should be stressed that the grid points in $S$ space are *not* equally spaced, unlike the standard finite difference schemes described earlier that had a fixed step length. To see that the above scheme for $S$ leads to fixed a step length for $y$, consider the change in $S$ to $uS$. At $uS$ the value of $y$ is

$$y_u = \ln(uS) = \ln S + \ln u = y + h,$$

where

$$h = \ln u.$$

Thus the difference scheme in $y$ has equally spaced intervals, as shown in Fig. 17.13.

Referring to Fig. 17.13 we have

$$\begin{aligned} C_u &= C(y+h, t+k) = C(y,t) + h\frac{\partial C}{\partial y} + \frac{1}{2}h^2\frac{\partial^2 C}{\partial y^2} + k\frac{\partial C}{\partial t}, \\ C_d &= C(y-h, t+k) = C(y,t) - h\frac{\partial C}{\partial y} + \frac{1}{2}h^2\frac{\partial^2 C}{\partial y^2} + k\frac{\partial C}{\partial t}. \end{aligned} \tag{17.43}$$

By subtraction of Eq. (17.43) we obtain

$$\frac{\partial C}{\partial y} = \frac{C_u - C_d}{2h}, \tag{17.44}$$

whilst by addition we find that

$$\frac{\partial^2 C}{\partial y^2} = \frac{1}{h^2}(C_u + C_d - 2C - 2k\frac{\partial C}{\partial t}). \tag{17.45}$$

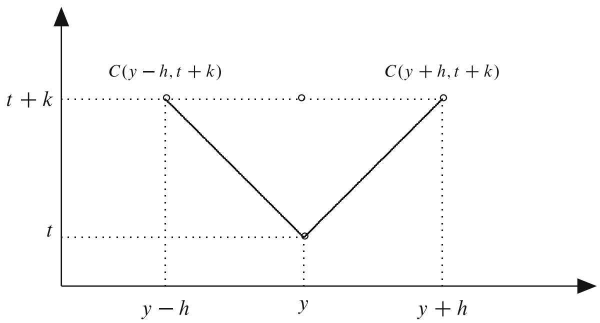


**Fig. 17.13** The equally spaced grid in $y = \ln S$ space

Substituting the derivatives (17.44), (17.45) into the transformed Black–Scholes partial differential (17.42) we obtain

$$\frac{\sigma^2}{2h^2}\left(C_u + C_d - 2C - 2k\frac{\partial C}{\partial t}\right) + r\left(\frac{C_u - C_d}{2h}\right) - rC + \frac{\partial C}{\partial t} = 0. \tag{17.46}$$

If we set

$$h = \sigma\sqrt{k}$$

then the $\dfrac{\partial C}{\partial t}$ terms are eliminated. Also $k$ is the step-size in $t$, i.e.,

$$k = t/n$$

and so

$$h = \sigma\sqrt{\frac{t}{n}}.$$

With this specification of $h$ and $k$, (17.46) reduces to

$$\left(1 + r\frac{t}{n}\right)C = C_u\left(\frac{1}{2} + \frac{1}{2}\frac{r}{\sigma}\sqrt{\frac{t}{n}}\right) + C_d\left(\frac{1}{2} - \frac{1}{2}\frac{r}{\sigma}\sqrt{\frac{t}{n}}\right). \tag{17.47}$$

This last equation describes the binomial option pricing formula with[2]

$$u = 1 + \sigma\sqrt{h}, \qquad d = 1 - \sigma\sqrt{h}, \qquad p^* = \frac{1}{2} + \frac{1}{2}\frac{r}{\sigma}\sqrt{h}, \tag{17.48}$$

which gives us some insight into the choice of parameters $u$ and $d$.

## 17.4 Choice of the Parameters $u, d$

Consider again the one-period binomial framework. Thus far we have not discussed how to choose the parameters $u, d$. We saw from (17.8) that we would require as a first property that

$$u - d = 2\sigma\sqrt{h}, \tag{17.49}$$

[2] In obtaining (17.48) we continue to impose $u - d = 2\sigma\sqrt{h}$. Thus $d = e^{rh} - 2\sigma\sqrt{h}p^* \simeq 1 - \sigma\sqrt{h}$, from which $u = 1 + \sigma\sqrt{h}$.

at least in some approximate sense, in order that the binomial process be seen as a first order approximation to the continuous time process. We note that the choice of $u, d$ in (17.33) in the discussion of the continuous limit viz

$$u = e^{\sigma\sqrt{h}} \qquad \text{and} \qquad d = e^{-\sigma\sqrt{h}}, \tag{17.50}$$

satisfies the condition (17.49) to a first order of approximation.

The choice of $u, d$ just discussed focussed on a comparison of the sample paths generated under the continuous time and discrete time processes. An alternative approach is to compare the distributions generated by both processes. We know that under the continuous time process over the interval $(0, h)$, the quantity $S_h$ is lognormally distributed. In fact, the first two moments of the continuous distribution are given by[3]

$$\mathbb{E}^*\left[\ln\left(\frac{S_h}{S_0}\right)\right] = (r - \frac{1}{2}\sigma^2)h, \tag{17.51}$$

and

$$\text{var}^*\left[\ln\left(\frac{S_h}{S_0}\right)\right] = \sigma^2 h, \tag{17.52}$$

where the $*$ indicates the risk-neutral distribution $p^*$. One set of conditions we could impose on $u$ and $d$ is that they be chosen so that the first two moments of the discrete binomial distribution match the first two moments of the continuous distribution.

We note that under the binomial distribution

$$\mathbb{E}_{bin}[\ln S_h] = p^* \ln(uS_0) + (1 - p^*)\ln(dS_0) = p^* \ln u + (1 - p^*)\ln d + \ln S_0$$

i.e.

$$\mathbb{E}_{bin}\left[\ln\frac{S_h}{S_0}\right] = p^* \ln u + (1 - p^*)\ln d, \tag{17.53}$$

whilst

$$\text{var}_{bin}\left[\ln\frac{S_h}{S_0}\right] = p^*(1 - p^*)\left[\ln(\frac{u}{d})\right]^2. \tag{17.54}$$

[3]These results follow from (3.14) with $T - t = h$.

Equating (17.51) with (17.53) and (17.52) with (17.54) we obtain

$$p^* \ln u + (1 - p^*) \ln d = (r - \frac{1}{2}\sigma^2)h, \tag{17.55}$$

$$\sqrt{p^*(1 - p^*)} \ln(\frac{u}{d}) = \sigma\sqrt{h}. \tag{17.56}$$

Solving (17.55), (17.56) we obtain

$$u = e^{\sigma\sqrt{h}}, \qquad d = e^{-\sigma\sqrt{h}} \tag{17.57}$$

which is the parameter set (17.33) used in the discussion of the continuous time limit. This is the set originally proposed by Cox et al. (1979).

There exist many other ways to choose the parameters $u, d$. For instance, by matching the moments of relative price changes $(S_h - S_0)/S_0$, by considering trinomial models etc. This issue has been thoroughly investigated by de Jager (1995).

# Chapter 18
# Volatility Smiles

**Abstract** It is commonly observed across many underlying assets that the implied volatility of the Black Scholes model varies across exercise price and time-to-maturity and has a pattern known as the volatility smile. In this chapter, we first address the volatility smile using the stochastic volatility models which may underestimate the size of the smile. We then develop an approach to calibrate the smile by choosing the volatility function as a deterministic function of the underlying asset price and time so as to fit the model option price to the observed volatility smile.

## 18.1 Introduction

In this chapter we continue to consider options written on assets which pay a continuous dividend such as stock, foreign exchange and futures.

Under the assumption that the underlying asset price, $x$, follows the geometric Brownian motion

$$\frac{dx}{x} = \mu dt + \sigma dz, \tag{18.1}$$

we know that the European call option price, $f$, is given by the Black–Scholes formula

$$f(x,t) = x\mathcal{N}(d_1) - e^{-r(T-t)}E\mathcal{N}(d_2), \tag{18.2}$$

with

$$d_1 = \frac{\ln(\frac{x}{E}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}. \tag{18.3}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_18

We know also from Chap. 9 that the Black–Scholes model holds when $\sigma$ becomes a function of time $\sigma(t)$. In this case in Eq. (18.3) we merely replace $\sigma$ by

$$\bar{\sigma} = \left[ \frac{1}{(T-t)} \int_t^T \sigma^2(s) ds \right]^{1/2}. \tag{18.4}$$

It is common practice amongst finance practitioners to calibrate the Black–Scholes model by finding the value of $\sigma$ that makes the theoretical model match option prices observed in the market, i.e. to solve

$$f(x, t; \sigma) = f_{market}, \tag{18.5}$$

to obtain the implied volatility $\hat{\sigma}$. A common observation across many underlying assets is that $\hat{\sigma}$ varies across exercise price and time-to-maturity. A pattern something like that in Fig. 18.1 is usually observed, this pattern is known as the volatility smile.

Whilst (18.4) provides some theoretical basis for the dependency of $\hat{\sigma}$ on $\tau = T - t$, the dependency of $\hat{\sigma}$ on the exercise price $E$ is incompatible with the theory underlying (18.2).

In the next section we show that the volatility smile suggests that the "true" model for option pricing requires that we employ a stochastic process for the volatility. However as we have seen in Chap. 15 models involving stochastic volatility introduce a parameter which measures the market price of volatility risk. Since this risk is not traded this parameter is difficult to estimate. An alternative approach has developed which consists of replacing $\sigma$ in (18.1) by $\sigma(x, t)$ and

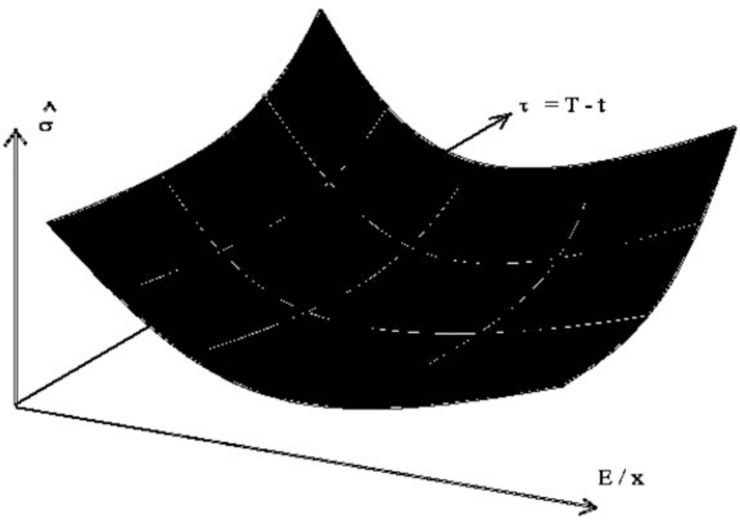


**Fig. 18.1** The volatility smile

choosing this function so as to fit the model option price to the observed volatility smile. In the final sections of this chapter we describe these methods.

## 18.2 Stochastic Volatility as the Origin of the Smile

Hull and White (1987) assume that volatility follows the stochastic process

$$d\sigma^2 = \alpha\sigma^2 dt + \xi\sigma^2 dz_\sigma, \tag{18.6}$$

with $z$ and $z_\sigma$ being uncorrelated and volatility has zero systematic risk.

For convenience we set

$$v_t = \sigma_t^2,$$

so that (18.6) may be written

$$dv = \alpha v dt + \xi v dz_\sigma. \tag{18.7}$$

For any given evolution of the process $v$ over the internal $[t, T]$ we can define the integrated variance (see Fig. 18.2)

$$\bar{v}_{T,t} = \frac{1}{T-t}\int_t^T v(u)du. \tag{18.8}$$

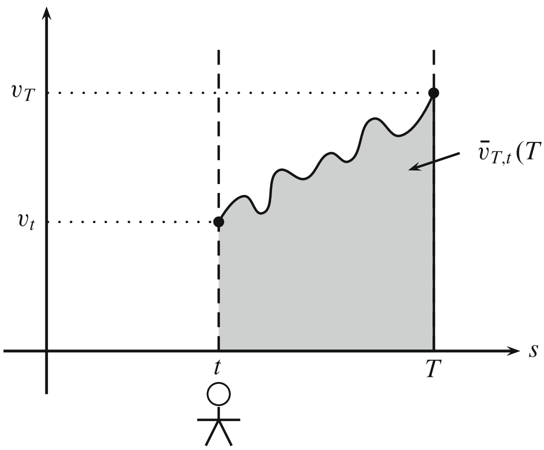


**Fig. 18.2** The evolution of the variance and integrated variance

We note that

$$\bar{v}_{t,t} = \lim_{T \to t} \bar{v}_{T,t} = v_t, \tag{18.9}$$

and that by taking the differential of (18.8) the integrated variance satisfies

$$d\bar{v}_{T,t} = \frac{(\bar{v}_{T,t} - v_t)}{T - t} dt. \tag{18.10}$$

Equations (18.7) and (18.10) should be viewed as a linked system of stochastic equations. The Kolmogorov equation for the transitional probability density function $p(\bar{v}_{T,s}, v_T, T | \bar{v}_{s,s}, v_s, s)$ $(t \leq s \leq T)$ satisfies (see Sect. 5.4)

$$\frac{\partial p}{\partial s} + \frac{(\bar{v}_{T,s} - v_s)}{T - s} \frac{\partial p}{\partial \bar{v}} + \alpha v \frac{\partial p}{\partial v} + \frac{1}{2} \xi^2 v^2 \frac{\partial^2 p}{\partial v^2} = 0, \tag{18.11}$$

subject to the initial condition

$$p(\bar{v}_{T,T}, v_T, T | \bar{v}'_{T,T}, v'_T, T) = \delta(\bar{v}_{T,T} - \bar{v}'_{T,T}) \delta(v_T - v'_T).$$

In this notation Hull and White show that the price of a European call option is given by

$$f(x, v_t, t) = \int f_{BS}(x, \bar{v}_{T,t}, t) p(\bar{v}_{T,t}, v_T, T | \bar{v}_{t,t}, v_t, t) d\bar{v}_{T,t}, \tag{18.12}$$

where

$$f_{BS}(x, v, t) = \begin{cases} \text{the Black–Scholes value} \\ \text{with volatility given by } v. \end{cases}$$

To emphasise the dependence on the exercise price $E$ we shall write $f(E)$ to denote $f(x, v_t, t)$ and $f_{BS}(E, v)$ to denote $f_{BS}(x, v, t)$, so that (18.12) is rewritten

$$f(E) = \int f_{BS}(E, \bar{v}_{T,t}) p(\bar{v}_{T,t}, v_T, T | \bar{v}_{t,t}, v_t, t) d\bar{v}_{T,t}. \tag{18.13}$$

If we assume that the price given by the stochastic volatility model is the "true" price then the implied volatility is given by the solution of

$$f(E) = f_{BS}(E, \hat{\sigma}^2).$$

Using $\mu_v, \sigma_v$ to denote the mean and standard deviation of the distribution $p(v|x, \sigma, t)$ we expand $f_{BS}(E, v)$ as

$$f_{BS}(E, v) \simeq f_{BS}(E, \mu_v) + (v - \mu_v) \frac{\partial f_{BS}}{\partial v} + \frac{1}{2} (v - \mu_v)^2 \frac{\partial^2 f_{BS}}{\partial v^2}, \tag{18.14}$$

where the partial derivatives are evaluated at $v = \mu_v$. It follows that

$$\int f_{BS}(E, \bar{v}_{T,t}) p(\bar{v}_{T,t}, v_T, T | \bar{v}_{t,t}, v_t, t) d\bar{v}_{T,t} \simeq f_{BS}(E, \mu_v) + \frac{1}{2}\sigma_v^2 \frac{\partial^2 f_{BS}}{\partial \bar{v}^2}. \tag{18.15}$$

Next we linearise $f_{BS}(E, \hat{\sigma}^2)$ about $\mu_v$, i.e.

$$f_{BS}(E, \hat{\sigma}^2) \simeq f_{BS}(E, \mu_v) + (\hat{\sigma}^2 - \mu_v)\frac{\partial f_{BS}}{\partial \bar{v}}. \tag{18.16}$$

Equating (18.15) and (18.16) we obtain

$$\hat{\sigma}^2 \simeq \mu_v + \frac{1}{2}\sigma_v^2 \frac{\partial^2 f_{BS}}{\partial \bar{v}^2}\left[\frac{\partial f_{BS}}{\partial \bar{v}}\right]^{-1}. \tag{18.17}$$

We note that the partial derivatives on the right-hand side of (18.17) depend on $E$ (as well as on other variables). Hence this equation relates $\hat{\sigma}^2$ to the exercise price $E$. We now seek to determine the nature of this relationship.

In terms of volatility $v$ the partial derivatives $\frac{\partial f_{BS}}{\partial \bar{v}}$ and $\frac{\partial^2 f_{BS}}{\partial \bar{v}^2}$ are given by

$$\frac{\partial f_{BS}}{\partial \bar{v}} = \frac{1}{2}\sqrt{\frac{T-t}{\bar{v}}} x e^{-r(T-t)} n(d_1),$$

and

$$\frac{\partial^2 f_{BS}}{\partial \bar{v}^2} = \frac{1}{4}\frac{\sqrt{T-t}}{\bar{v}^{3/2}} x e^{-r(T-t)} n(d_1)(d_1 d_2 - 1),$$

where $n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$. Hence

$$\frac{\partial^2 f_{BS}}{\partial \bar{v}^2}\left[\frac{\partial f_{BS}}{\partial \bar{v}}\right]^{-1} = \frac{d_1 d_2 - 1}{2\bar{v}}. \tag{18.18}$$

Recalling that the partial derivatives are evaluated at $v = \mu_v$ we find upon substituting (18.18) into (18.17) that

$$\hat{\sigma}^2(E) \simeq \mu_v + \frac{\sigma_v^2}{4\mu_v}\left[\frac{\{\ln(xe^{r\tau}/E)\}^2 - \mu_v\tau - \frac{1}{4}\mu_v^2\tau^2}{\mu_v\tau}\right]. \tag{18.19}$$

Noting that $\{\ln(xe^{r\tau}/E)\}^2$ has a quadratic shape as a function of $E$, we see that (18.19) predicts a smile of the type observed. The above derivation comes from Xu and Taylor (1994) who also undertake an empirical analysis of (18.19). They find that it underestimates the size of the smile. This is probably due to the assumptions about correlation of the noise terms and not pricing volatility risk.

## 18.3 Calibrating Deterministic Models to the Smile

The analysis of Sect. 18.2 suggests that a proper option pricing model should allow for stochastic volatility in the process for the underlying asset. Furthermore such a model should probably allow for correlation between the noises driving the price process and the volatility process, as well as the pricing of volatility risk (i.e. the market price of volatility risk $\neq 0$). However these further parameters, particularly the market price of volatility risk, may be difficult to estimate in practice.

Therefore a technique has evolved amongst finance industry practitioners of modelling the volatility as a deterministic function of the underlying asset price $x$ and time $t$, i.e. to replace (18.1) with

$$\frac{dx}{x} = \mu dt + \sigma(x,t)dz,$$

and then at each point in time choose the function $\sigma(x,t)$ so as to be compatible with the currently observed volatility smile.

In this section we discuss how this procedure is implemented in the binomial lattice framework. Our discussion follows closely that of Derman and Kani (1994), who build on the earlier contribution of Dupire (1994).

Firstly we assume that at each point in time an interpolating polynomial has been fitted to the observed volatility smile in Fig. 18.1. From this we are able to calculate a market option price for any strike price and time-to-maturity which we denote $f_{smile}(\frac{E}{x},\tau)$.

Next we observe that since the volatility is not constant the regular binomial tree of Chap. 17 will become distorted as shown in Fig. 18.3.

The idea of the Dupire and Derman/Kani approaches is to determine risk neutral probabilities and stock prices (i.e. nodes) at the next time level so that the tree does

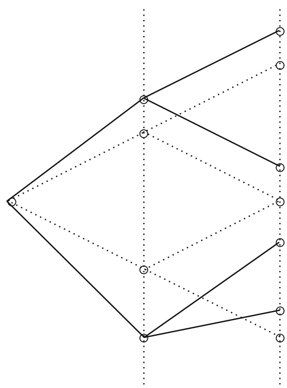

**Fig. 18.3** The regular (*dotted line*) and distorted (*solid line*) binomial tree

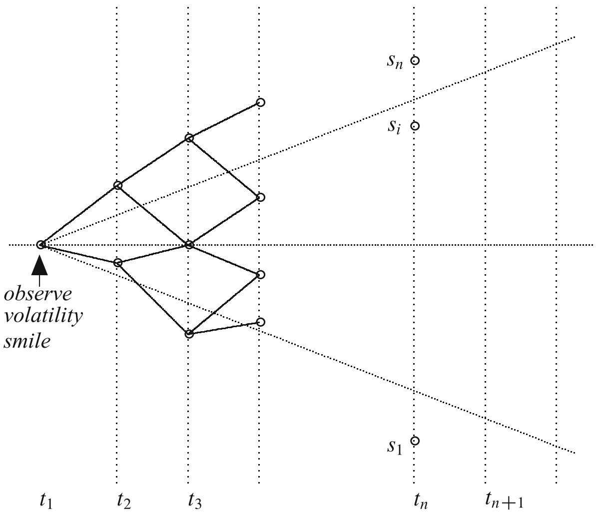


**Fig. 18.4** The implied binomial tree

recombine. This will result in the so called implied tree (which is also distorted). A similar idea is also implemented by Rubinstein (1994).

Thus we seek to determine the implied tree displayed in Fig. 18.4. In particular to determine the nodes and risk neutral transition probabilities so as to match the volatility smile currently observed at the root of the tree.

In order to develop the required pricing relationships we need to recall the concept of the *Arrow–Debreu* price. Consider the node $(n, i)$, at time level $t_n$ and spot price $s_i$, in Fig. 18.4. The associated *Arrow–Debreu* price is denoted $\lambda_i$ and it is the value at the root of the tree (i.e. at $t_1$) of a claim that pays \$1 if the node $(n, i)$ is reached at time $t_n$ and \$0 if any other node is reached at time $t_n$. Thus

$$\lambda_i = \begin{cases} \textit{the sum over all paths, from the root of the tree} \\ \textit{to node (n,i) of the product of the riskless-discounted} \\ \textit{transition probabilities.} \end{cases}$$

As the implied binomial tree is developed by forward induction the risk-neutral transition probabilities to the next time level are calculated and these are then used to calculate the *Arrow–Debreu* prices at the next time level.

Consider the first step of the forward induction displayed in Fig. 18.5. We seek to determine three quantities. The risk-neutral transition probability $p_1$ and the spot price node values $S_1, S_2$ at the next time step. To calculate these we develop three relationships.

At time $t_1$ we are able to calculate from the volatility smile (i.e. the interpolated function $f_{smile}(\frac{E}{x}, \tau)$) the market value of an option with strike price $s_1$ and maturing at time $t_2$, we denote this value by $C(s_1, t_2)$. By the principal of risk-neutral

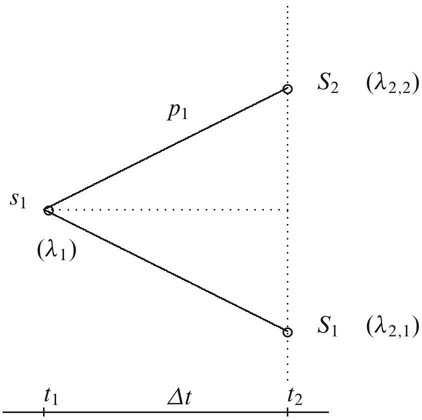


**Fig. 18.5** The first step of the forward induction

valuation we have the first relationship

$$C(s_1, t_2) = \lambda_1 e^{-r\Delta t} p_1 (S_2 - s_1). \tag{18.20}$$

The $\lambda_1$ on the right-hand side discounts all quantities to the root of the tree. Of course this discounting is trivial at the first time step since we are at the root of the tree and $\lambda_1 = 1$.

The second condition is that under the risk-neutral measure asset prices are expected to grow at the risk-free rate i.e.

$$e^{r\Delta t} s_1 = p_1 S_2 + (1 - p_1) S_1. \tag{18.21}$$

The final condition is to make the centre of the implied tree coincide with the centre of the standard binomial tree that is obtained when volatility is assumed to be constant i.e.[1]

$$s_1^2 = S_1 S_2. \tag{18.22}$$

Equations (18.20)–(18.22) can be solved for $p_1, S_1$ and $S_2$. Finally we use forward induction to calculate the *Arrow–Debreu* prices at $t_2$ (which at this point we label $\lambda_{2,1}, \lambda_{2,2}$) from the knowledge of $p_1$, thus

$$\lambda_{2,2} = \lambda_1 p_1 e^{-r\Delta t}, \quad \lambda_{2,1} = \lambda_1 (1 - p_1) e^{-r\Delta t}. \tag{18.23}$$

[1] In the standard binomial model the up and down probabilities are respectively $e^{\sigma\sqrt{\Delta t}}$ and $e^{-\sigma\sqrt{\Delta t}}$ where $\sigma$ is the constant volatility (see Sect. 17.4). Thus $S_1 = s_1 e^{-\sigma\sqrt{\Delta t}}$ and $S_2 = s_1 e^{\sigma\sqrt{\Delta t}}$ from which Eq. (18.22) follows.

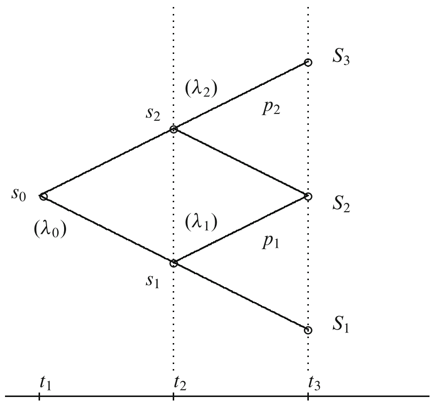


**Fig. 18.6** The second step of the forward induction

Next consider the second forward induction step from $t_2$ to $t_3$ which is illustrated in Fig. 18.6.

Note the notation of using $s_i$ to denote the stock price node values which have just been calculated and $\lambda_i$ the associated *Arrow–Debreu* prices (which have been relabelled from $\lambda_{n,i}$; we always reserve this latter notation for the *Arrow–Debreu* prices at the next step); $S_i$ denotes the unknown stock price nodes that we seek to calculate and $p_i$ the associated up-movement risk-neutral transition probabilities. From the volatility smile we can calculate $C(s_1, t_3)$ and $C(s_2, t_3)$, the value of call options maturing at time $t_3$ and having exercise prices $s_1, s_2$ respectively. It is important to stress that these values are in terms of dollars at time $t_1$.

Now we seek the five unknowns $S_1, S_2, S_3, p_1$ and $p_2$ which are obtained from the following relationships. By the principal of risk-neutral valuation of the options

$$\begin{aligned} C(s_1, t_3) &= \lambda_1 e^{-r\Delta t} p_1(S_2 - s_1) + \lambda_2 e^{-r\Delta t}(1 - p_2)(S_2 - s_1) + \lambda_2 e^{-r\Delta t} p_2(S_3 - s_1) \\ &= e^{-r\Delta t}[\{\lambda_1 p_1 + \lambda_2(1 - p_2)\}(S_2 - s_1) + \lambda_2 p_2(S_3 - s_1)], \qquad (18.24) \end{aligned}$$

$$C(s_2, t_3) = \lambda_2 e^{-r\Delta t} p_2(S_3 - s_2). \qquad (18.25)$$

From the fact that under the risk-neutral distribution the stock price grows at the risk free rate we have the further relationships

$$e^{r\Delta t} s_2 = p_2 S_3 + (1 - p_2) S_2, \qquad (18.26)$$

$$e^{r\Delta t} s_1 = p_1 S_2 + (1 - p_1) S_1. \qquad (18.27)$$

Finally the centering condition in this case of an even number of steps becomes

$$s_0 = S_2, \qquad (18.28)$$

where $s_0$ is the current spot price at time $t_1$. Equations (18.24)–(18.28) may be solved simultaneously for the five unknowns $S_1$, $S_2$, $S_3$, $p_1$ and $p_2$.

To facilitate the development of the general algorithm note that (18.24) may be rewritten

$$\begin{aligned} e^{r\Delta t}C(s_1, t_3) &= \lambda_1 p_1(S_2 - s_1) + \lambda_2\{(1 - p_2)(S_2 - s_1) + p_2(S_3 - s_1)\} \\ &= \lambda_1 p_1(S_2 - s_1) + \lambda_2\{(1 - p_2)S_2 + p_2 S_3 - s_1\}. \end{aligned}$$

The second bracket is further simplified by use of (18.26), so that

$$e^{r\Delta t}C(s_1, t_3) = \lambda_1 p_1(S_2 - s_1) + \lambda_2\{e^{r\Delta t}s_2 - s_1\}. \tag{18.29}$$

Using the notation $F_i$ to denote the forward price at level $(n+1)$ of the known stock price $s_i$ at level $n$ i.e.

$$F_i = s_i e^{r\Delta t},$$

then (18.29) can be written

$$e^{r\Delta t}C(s_1, t_3) = \lambda_1 p_1(S_2 - s_1) + \lambda_2(F_2 - s_1). \tag{18.30}$$

Note also that for consistency of notation (18.25) may be written

$$e^{r\Delta t}C(s_2, t_3) = \lambda_2 p_2(S_3 - s_2), \tag{18.31}$$

and (18.26), (18.27) may be written

$$p_1 = \frac{F_1 - S_1}{S_2 - S_1}, \tag{18.32}$$

$$p_2 = \frac{F_2 - S_2}{S_3 - S_2}. \tag{18.33}$$

The five unknowns $S_1$, $S_2$, $S_3$, $p_1$ and $p_2$ may be calculated recursively as follows: from Eqs. (18.30) and (18.32) we have

$$S_2 = \frac{S_1[e^{r\Delta t}C(s_1, t_3) - \lambda_2(F_2 - s_1)] - \lambda_1 s_1(F_1 - S_1)}{[e^{r\Delta t}C(s_1, t_3) - \lambda_2(F_2 - s_1)] - \lambda_1(F_1 - S_1)}, \tag{18.34}$$

and from (18.31) and (18.33)

$$S_3 = \frac{S_2 e^{r\Delta t}C(s_2, t_3) - \lambda_2 s_2(F_2 - S_2)}{e^{r\Delta t}C(s_2, t_3) - \lambda_2(F_2 - S_2)}. \tag{18.35}$$

Recalling the centering condition (18.28) which yields the value of $S_2$, we use (18.34) to calculate $S_1$ (moving down the tree) and (18.35) to calculate $S_3$

(moving up the tree). We can then use (18.32) and (18.33) to calculate $p_1$ and $p_2$. We shall see shortly that this recursive procedure generalises to the $n$-node situation.

Using $\lambda_{3,i}(i = 1, 2, 3)$ to denote the *Arrow–Debreu* prices at $t_3$ we calculate these from

$$\lambda_{3,1} = \lambda_1(1 - p_1)e^{-r\Delta t},$$

$$\lambda_{3,2} = \lambda_1 p_1 e^{-r\Delta t} + \lambda_2(1 - p_2)e^{-r\Delta t},$$

$$\lambda_{3,3} = \lambda_2 p_2 e^{-r\Delta t}.$$

The procedure for going from $t_n$ to $t_{n+1}$ should now be clear.

From the smile we know

$$C(s_1, t_{n+1}), C(s_2, t_{n+1}), \cdots, C(s_n, t_{n+1}).$$

By the principal of risk neutral valuation these option values should satisfy

$$C(s_i, t_{n+1}) = e^{-r\Delta t} \sum_{j=i}^{n} \{\lambda_j p_j + \lambda_{j+1}(1 - p_{j+1})\}(S_{j+1} - s_i), \tag{18.36}$$

for $i = 1, 2, \cdots, n$, where we adopt the convention $\lambda_{n+1} = 0$ to handle the last term.

From the condition that the spot price grows at the risk-free rate through the implied tree

$$F_i = p_i S_{i+1} + (1 - p_i)S_i, \quad (i = 1, 2, \cdots, n). \tag{18.37}$$

Finally to write the centering condition we define

$$n^* = \text{ integer part of } \left(\frac{n+2}{2}\right),$$

then (see Fig. 18.7)

$$\begin{cases} s_0 = S_{n^*} & \text{if } n \text{ is even,} \\ s_{n^*}^2 = S_{(n^*+1)} S_{n^*} & \text{if } n \text{ is odd.} \end{cases} \tag{18.38}$$

Equations (18.36)–(18.38) provide $(2n + 1)$ equations for the $(2n + 1)$ unknowns $S_1, S_2, \cdots, S_{n+1}, p_1, p_2, \cdots, p_n$. Using manipulations completely analogous to those leading to (18.34), (18.35) it is possible to show that

$$S_{i+1} = \frac{S_i \left[e^{r\Delta t} C(s_i, t_{n+1}) - \sum_{j=i+1}^{n} \lambda_j (F_j - s_i)\right] - \lambda_i s_i (F_i - S_i)}{\left[e^{r\Delta t} C(s_i, t_{n+1}) - \sum_{j=i+1}^{n} \lambda_j (F_j - s_i)\right] - \lambda_i (F_i - S_i)}. \tag{18.39}$$

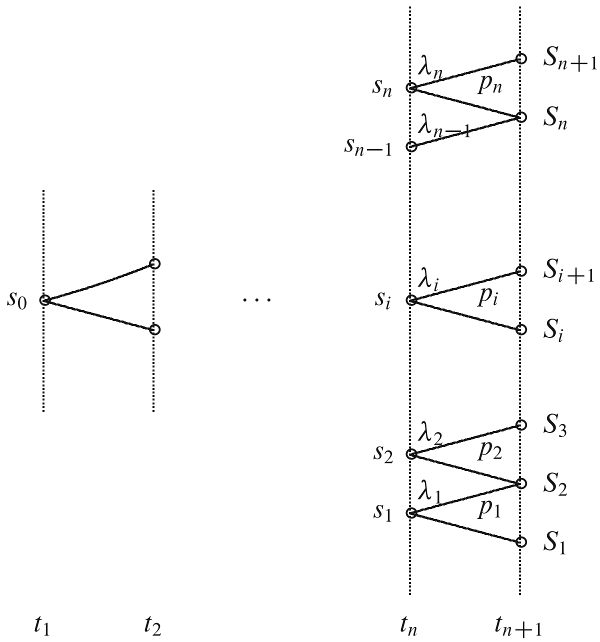


**Fig. 18.7** The *n*th step of the forward induction

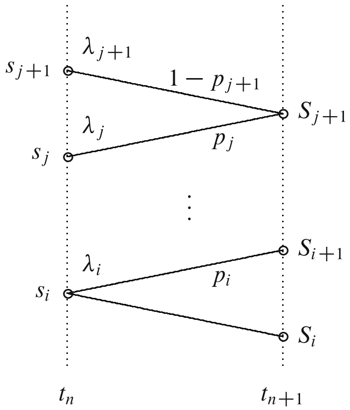


**Fig. 18.8** Risk-neutral option valuation

Once the central value has been fixed using (18.38), the above equation may be used recursively, moving up and down from the centre to calculate all remaining $S_i$. We then use (18.37) in the form

$$p_i = \frac{F_i - S_i}{S_{i+1} - S_i}, \tag{18.40}$$

to calculate the $p_i$ (see Fig. 18.8). Finally we calculate the *Arrow–Debreu* prices at $t_{n+1}$ from

$$\lambda_{n+1,i+1} = \lambda_i p_i e^{-r\Delta t} + \lambda_{i+1}(1 - p_{i+1})e^{-r\Delta t}, \tag{18.41}$$

which we apply for $i = 0, 1, \cdots, n$ with the notational convention $\lambda_0 = \lambda_{n+1} = 0$.

## 18.4 Problems

**Problem 18.1 Computational Problem**—Use the data below for BHP Billiton call options quoted in the Australian Financial Review on Tuesday 18 March 2014 to calculate the implied binomial tree by using the algorithm in Sect. 18.3 of the notes. Use the maturities only up to and including June 14.

Take a time step of 1 week. Simply use linear interpolation to find option values at strikes and maturities that do not correspond to traded ones.

Interest rate data is also given. For the purposes of this exercise apply the 90 day rate for all maturities.

# Part II
# Interest Rate Modelling

# Chapter 19
# Allowing for Stochastic Interest Rates in the Black–Scholes Model

**Abstract** The discussion in Chaps. 12 and 15 considered a relaxation of one of the key assumptions of the Black–Scholes framework, namely that the asset price changes follow a geometric Brownian motion. Another crucial assumption is the assumption of a constant interest rate over the life of the option. In this chapter we consider the specific case of stock options and retain all the assumptions of the original Black–Scholes model, except that we now allow interest rates to vary stochastically. Along the lines of Merton (Bell J Econ Manag Sci 4:141–183, 1973b), we develop the appropriate hedging argument to derive the stock option pricing partial differential equation and provide the technical details of its solution.

## 19.1 Introduction

The issue raised by the interest rate being stochastic is that we can no longer discount the expected future option pay-off using the deterministic discount factor $\exp[-r(T-t)]$. Now we must find a corresponding stochastic discount factor and adjust the pricing relationship accordingly. Our discussion follows closely that of Merton (1973) who uses the price of a bond having the same maturity as the option to capture the effect of stochastic interest rates. Problems 19.1 and 19.2 lead to a formulation of this problem in the framework of Chap. 10.

Let $T$ denote option maturity and $P(t,T)$ the price at time $t$ of a riskless discount bond which pays \$1 at time $T(>t)$. We shall use $\tau(=T-t)$ to denote time to maturity. We allow for stochastic interest rates by allowing the bond price to vary stochastically, in particular by assuming

$$\frac{dP}{P} = \alpha(P,t)dt + \delta(P,t)dv, \tag{19.1}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_19

where $dv$ are the increments of a Wiener process which are the source of the uncertainty in the evolution of the bond price.[1] The mean return $\alpha(\cdot)$ could depend on the level of bond prices as well as the time-to-maturity. To ensure that $P(T, T) = 1$ (i.e. the bond pays \$1 at maturity) we could use a Brownian bridge process as discussed in Sect. 6.3.6. In fact, the choice

$$\alpha(P, t) = \frac{1}{2}\delta^2 - \frac{\ln P(t, T)}{T - t},$$
$$\delta(P, t) = \delta,$$

when $\delta$ is constant will provide the appropriate drift and diffusion coefficients. However for the purposes of this chapter we do not need to be too precise about the nature of the drift coefficient $\alpha(P, t)$.

The traditional assumption of constant interest rate is recovered by setting $\delta(P, t) = 0$ and $\alpha(P, t) = r$, in which case

$$P(t, T) = e^{-r(T-t)}. \tag{19.3}$$

We retain the assumption that the stock price follows a diffusion process given by

$$\frac{dS}{S} = \mu dt + \sigma dz, \tag{19.4}$$

and allow for correlation between the Wiener increments $dv$ and $dz$, i.e.,

$$\mathbb{E}[dvdz] = \rho dt. \tag{19.5}$$

Here we do not adopt our standard procedure of re-expressing the stochastic dynamics for the stock and bond prices in terms of independent Wiener processes. This is so as to enable us to follow Merton's original derivation. However the expression in terms of independent Wiener increment probably provides a "cleaner" approach and this is the viewpoint adopted in Problems 19.1 and 19.2. Fig. 19.1 depicts possible paths of stock prices and bond prices under the Black-Scholes model with a stochastic interest rate.

---

[1] Since the sources of uncertainty may be different for bonds of differing maturities we should write $dv(t, \tau)$ to denote this dependency on $\tau$, and since the $dv$ at differing maturities would not be perfectly correlated we would have

$$\mathbb{E}[dv(t, \tau_1)dv(t, \tau_2)] = \rho_{12}dt. \tag{19.2}$$

However since we only consider a bond having the same maturity as the option we do not need such a notation here. Furthermore we shall see more clearly how to capture the correlation between bonds at different maturities when we come to study the Heath–Jarrow–Morton interest rate model.

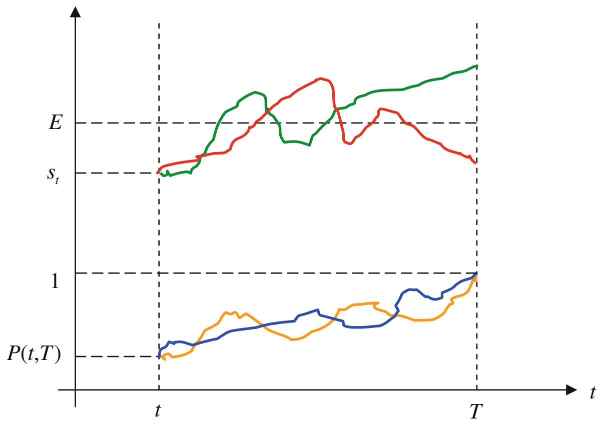


**Fig. 19.1** Illustrating stock prices and bond prices (of same maturity as the option) in the B-S world with a stochastic interest rate

## 19.2 The Hedging Portfolio

The option value will now be a function of both $S$ and $P$ as well as time, i.e.,

$$f = f(S, P, t). \tag{19.6}$$

The pricing of the option in the present situation would therefore seem to fall into the case of pricing a derivative security dependent on several underlying state variables which we encountered in Chap. 10. However the risk free rate in the discussion of that chapter was constant, or at most a deterministic function of time. Hence we need to reconsider the hedging argument for the situation at hand. Since $S$ and $P$ are driven by stochastic differential equations we may apply Ito's lemma to determine the stochastic differential equation followed by $f$. Thus, the option price $f$ is found to follow the stochastic differential equation

$$\frac{df}{f} = \mu_f dt + \sigma_{f_z} dz + \sigma_{f_v} dv, \tag{19.7}$$

where

$$\mu_f = \frac{1}{f}\left[\frac{\partial f}{\partial t} + \mu S \frac{\partial f}{\partial S} + \alpha P \frac{\partial f}{\partial P} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP \frac{\partial^2 f}{\partial S \partial P} + \frac{1}{2}\delta^2 P^2 \frac{\partial^2 f}{\partial P^2}\right],$$

$$\sigma_{f_z} = \frac{\sigma S}{f}\frac{\partial f}{\partial S}, \quad \text{and} \quad \sigma_{f_v} = \frac{\delta P}{f}\frac{\partial f}{\partial P}. \tag{19.8}$$

Extending the continuous hedging approach to the current problem, we form a portfolio of the stock, the option and the riskless bond having time to maturity equal to the expiration of the option. Letting $Q_1$, $Q_2$ and $Q_3$ denote respectively the number of dollars of the portfolio invested in the stock, the option and the bonds then the condition of zero aggregate investment can be written

$$Q_1 + Q_2 + Q_3 = 0. \tag{19.9}$$

The instantaneous dollar return on the portfolio is given by[2]

$$\begin{aligned} &Q_1\frac{dS}{S} + Q_2\frac{df}{f} + Q_3\frac{dP}{P} \\ &= [Q_1(\mu - \alpha) + Q_2(\mu_f - \alpha)]dt + [Q_1\sigma + Q_2\sigma_{f_z}]dz \\ &\quad + [Q_2\sigma_{f_v} - (Q_1 + Q_2)\delta]dv. \end{aligned} \tag{19.10}$$

Following a now familiar argument we choose the proportions $Q_1$, $Q_2$ so that the stochastic $dz$ and $dv$ terms vanish i.e.

$$Q_1\sigma + Q_2\sigma_{f_z} = 0, \tag{19.11}$$

$$Q_2\sigma_{f_v} - (Q_1 + Q_2)\delta = 0. \tag{19.12}$$

From Eq. (19.11) we have that

$$\frac{Q_1}{Q_2} = -\frac{\sigma_{f_z}}{\sigma},$$

and from Eq. (19.12)

$$\frac{Q_1}{Q_2} = \frac{\sigma_{f_v}}{\delta} - 1.$$

From these two equations we find that

$$\frac{\sigma_{f_z}}{\sigma} = 1 - \frac{\sigma_{f_v}}{\delta},$$

which becomes [from the definitions of $\sigma_{f_z}$ and $\sigma_{f_v}$ in Eq. (19.8)]

$$f = S\frac{\partial f}{\partial S} + P\frac{\partial f}{\partial P}. \tag{19.13}$$

[2]Note that the $Q_i\,(i = 1, 2, 3)$ are in monetary units and since the rates of return $dS/S$, $df/f$ and $dP/P$ are dimensionless, the units in Eq. (19.10) must be in monetary units.

With this choice of $Q_1$, $Q_2$ the hedging portfolio is riskless and its instantaneous dollar return is given by

$$\left[Q_1(\mu-\alpha) - \frac{Q_1\sigma}{\sigma_{f_z}}(\mu_f-\alpha)\right]dt = Q_1\sigma\left[\frac{\mu-\alpha}{\sigma} - \frac{\mu_f-\alpha}{\sigma_{f_z}}\right]dt.$$

This instantaneous dollar return should be zero, given that it involves zero net investment. Thus

$$\frac{\mu-\alpha}{\sigma} = \frac{\mu_f-\alpha}{\sigma_{f_z}}, \tag{19.14}$$

which is a modified form of equality of risk adjusted excess return of risky assets in the portfolio. Here, the constant risk free rate $r$ is replaced by $\alpha$, the instantaneous return on a riskless bond having the same maturity as the option. Applying the definitions of $\mu_f$ and $\sigma_{f_z}$ from (19.8) we may reduce (19.14) to the partial differential equation

$$\begin{aligned}\frac{\partial f}{\partial t} + \mu S\frac{\partial f}{\partial S} + \alpha P\frac{\partial f}{\partial P} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP\frac{\partial^2 f}{\partial S\partial P} + \frac{1}{2}\delta^2 P^2\frac{\partial^2 f}{\partial^2 P} - \alpha f \\ = S\frac{\partial f}{\partial S}(\mu-\alpha),\end{aligned}$$

which simplifies to

$$\frac{\partial f}{\partial t} + \alpha(S\frac{\partial f}{\partial S} + P\frac{\partial f}{\partial P}) + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP\frac{\partial^2 f}{\partial S\partial P} + \frac{1}{2}\delta^2 P^2\frac{\partial^2 f}{\partial P^2} - \alpha f = 0.$$

Note that by use of (19.13) the second term on the left hand side of the last equation becomes $\alpha f$ and so (19.14) has reduced to the partial differential equation

$$\frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP\frac{\partial^2 f}{\partial S\partial P} + \frac{1}{2}\delta^2 P^2\frac{\partial^2 f}{\partial P^2} + \frac{\partial f}{\partial t} = 0. \tag{19.15}$$

Equation (19.15) needs to be solved subject to the boundary conditions appropriate for the option of interest, for instance a European call or put option.

For future reference note that it is often convenient to consider (19.15) in terms of time to maturity $\tau = T - t$, so that [3]

$$\frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP\frac{\partial^2 f}{\partial S\partial P} + \frac{1}{2}\delta^2 P^2\frac{\partial^2 f}{\partial P^2} = \frac{\partial f}{\partial \tau}. \tag{19.16}$$

[3]Note that $\frac{\partial f}{\partial \tau} = \frac{\partial f}{\partial t}\cdot\frac{\partial t}{\partial \tau} = -\frac{\partial f}{\partial t}$.

## 19.3 Solving for the Option Price

Equation (19.15) is a linear partial differential equation of the parabolic type. Its particular feature (and difficulty) lies in the fact that it involves two spatial variables ($S$ and $P$). The solution of such equations is normally a difficult procedure and we would expect to have to resort to numerical procedures. However, the relationship (19.13) suggests that $f$ is first degree homogeneous in $(S, P)$.[4] If this were so, then the type of transformation used in Appendix 9.1 to solve the basic Black–Scholes equation could be used to reduce the dimensionality of the pricing equation (19.15) to the first order heat equation, whose solution in our context is the Black–Scholes equation. The transformations used in Appendix 9.1 would suggest introducing the new state variable $X$ defined by

$$X = \frac{S}{EP}. \tag{19.17}$$

Here $P$ is used to discount between maturity and current time, so that $X$ may be interpreted as the price per share of stock in units of the present value of the exercise price. In the case of a constant interest rate, $P(t, T) = e^{-r(T-t)}$, and we would recover the change of variable given in (9.41). Note that since $P$ and $S$ are driven by the stochastic differential equations (19.1) and (19.4) respectively, then by an application of the Ito's lemma we find that $X$ is driven by the stochastic differential equation

$$\frac{dX}{X} = [\mu - \alpha + \delta^2 - \rho\sigma\delta]dt + \sigma dz - \delta dv. \tag{19.18}$$

We note that the instantaneous variance of the return on $X$ is given by

$$\begin{aligned} \text{var}\left[\frac{dX}{X}\right] &= \mathbb{E}[(\sigma dz - \delta dv)^2] \\ &= \mathbb{E}[\sigma^2(dz)^2 + \delta^2(dv)^2 - 2\sigma\delta dz dv] \\ &= (\sigma^2 + \delta^2 - 2\rho\sigma\delta)dt \\ &\equiv V^2(\tau)dt, \end{aligned} \tag{19.19}$$

[4] The function $g(u, v)$ is homogeneous of degree $n$ in $(u, v)$ if $g(\lambda u, \lambda v) = \lambda^n g(u, v)$. Differentiation with respect to $\lambda$ yields

$$u g_u(\lambda u, \lambda v) + v g_v(\lambda u, \lambda v) = n\lambda^{n-1} g(u, v).$$

Setting $n = 1$, then $\lambda = 1$ yields an expression of the form (19.13).

where we define

$$V^2(\tau) = \sigma^2 + \delta^2 - 2\rho\sigma\delta. \tag{19.20}$$

At this point it is convenient to switch the time unit to $\tau = T - t$. Following further the transformation of the option price used in (9.41) and also motivated by the homogeneity property[5] of $f$ we define the price function $h$ by

$$h(X, \tau) = \frac{f(S, P, \tau)}{EP}. \tag{19.21}$$

In transforming from the variables $(f, S, P, \tau)$ to $(h, X, \tau)$ we make use of the following partial derivative relationships[6]

$$\frac{\partial f}{\partial S} = \frac{\partial h}{\partial X}, \quad \frac{\partial f}{\partial P} = Eh - \frac{S}{P}\frac{\partial h}{\partial X},$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{1}{EP}\frac{\partial^2 h}{\partial X^2}, \quad \frac{\partial^2 f}{\partial S \partial P} = \frac{-S}{EP^2}\frac{\partial^2 h}{\partial X^2}, \quad \frac{\partial^2 f}{\partial P^2} = \frac{S^2}{EP^3}\frac{\partial^2 h}{\partial X^2},$$

$$\frac{\partial f}{\partial \tau} = EP\frac{\partial h}{\partial \tau}.$$

From the above we find that

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \rho\sigma\delta SP\frac{\partial^2 f}{\partial S \partial P} + \frac{1}{2}\delta^2 P^2 \frac{\partial^2 f}{\partial P^2} = \frac{1}{2}V^2(\tau)X^2\frac{\partial^2 h}{\partial X^2}EP,$$

---

[5]If $f$ is first order homogenous in $S, P$, then $f(\lambda S, \lambda P, \tau) = \lambda f(S, P, \tau)$. Choose $\lambda = \frac{1}{EP} \Rightarrow f(\frac{S}{EP}, \frac{1}{E}, \tau) = \frac{f(S,P,\tau)}{PE} \Rightarrow h(X, \tau) = \frac{f(S,P,\tau)}{EP}$, since $X = \frac{S}{EP}$.

[6]These have been calculated as follows:

$$\frac{\partial f}{\partial S} = EP\frac{\partial h}{\partial X}\cdot\frac{\partial X}{\partial S} = EP\cdot\frac{1}{EP}\cdot\frac{\partial h}{\partial X} = \frac{\partial h}{\partial X},$$

$$\frac{\partial f}{\partial P} = Eh + EP\frac{\partial h}{\partial X}\frac{\partial X}{\partial P} = Eh + EP\left(-\frac{S}{EP^2}\right)\frac{\partial h}{\partial X} = Eh - \frac{S}{P}\frac{\partial h}{\partial X},$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{\partial}{\partial X}\left(\frac{\partial h}{\partial X}\right)\cdot\frac{\partial X}{\partial S} = \frac{1}{EP}\frac{\partial^2 h}{\partial X^2},$$

$$\frac{\partial^2 f}{\partial S \partial P} = \frac{\partial}{\partial X}\left(\frac{\partial h}{\partial X}\right)\frac{\partial X}{\partial P} = -\frac{S}{EP^2}\frac{\partial^2 h}{\partial X^2},$$

$$\frac{\partial^2 f}{\partial P^2} = E\frac{\partial h}{\partial X}\frac{\partial X}{\partial P} + \frac{S}{P^2}\frac{\partial h}{\partial X} - \frac{S}{P}\frac{\partial^2 h}{\partial X^2}\cdot\frac{\partial X}{\partial P} = \frac{S^2}{EP^3}\frac{\partial^2 h}{\partial X^2},$$

$$\frac{\partial f}{\partial \tau} = EP\frac{\partial h}{\partial \tau}.$$

and so the second-order partial differential equation (19.15) for $f$ in terms of $S$, $P$ and $\tau$ has been reduced to the first-order partial differential equation

$$\frac{1}{2}V^2(\tau)X^2\frac{\partial^2 h}{\partial X^2} - \frac{\partial h}{\partial \tau} = 0, \tag{19.22}$$

for $h$ in terms of $X$ and $\tau$. In the case of a European call option, Eq. (19.22) must be solved subject to the boundary condition

$$h(0,\tau) = 0,$$

and initial condition[7]

$$h(X,0) = \max[X-1,0].$$

We can solve Eq. (19.22) using the solution framework set up in Chap. 9. In Eq. (9.7), if we set $q(t) = r(t) = 0$, then

$$\frac{1}{2}\sigma^2(t)S^2\frac{\partial^2 f}{\partial S^2} + \frac{\partial f}{\partial t} = 0.$$

This is basically Eq. (19.22) with $\sigma(t) \rightarrow V(t), f \rightarrow h, S \rightarrow X$. Apply the solution (9.30) with

$$\begin{aligned}
E &= 1,\\
\bar{r} &= 0,\\
\bar{\sigma} &= \bar{V}^2 = \frac{1}{T-t}\int_t^T V^2(s)ds\\
&= \frac{1}{T-t}\int_0^{T-t} V^2(u+t)du\\
&= \frac{1}{\tau}\int_0^{\tau} V^2(u+t)du, \qquad \text{(change of variable } u = s-t).
\end{aligned}$$

[7]Note that $P(0) = 1$ and $X(0) = S(0)/E$. So

$$h(X,0) = \frac{f(S,1,0)}{E} = \frac{1}{E}\max[S-E,0] = \max\left[\frac{S}{E}-1,0\right] = \max[X-1,0].$$

The solution then becomes

$$h(X,\tau) = X\mathcal{N}(d_1) - 1\mathcal{N}(d_2),$$

with

$$d_1 = \frac{\ln(\frac{X}{1}) + \frac{\bar{V}^2}{2}(T-t)}{\bar{V}\sqrt{T-t}}, \qquad d_2 = d_1 - \bar{V}\sqrt{T-t}.$$

Changing back to the original variables we obtain

$$\frac{f(S,P,\tau)}{EP} = \frac{S}{EP}\mathcal{N}(d_1) - \mathcal{N}(d_2)$$

i.e.

$$f(S,P,\tau) = S\mathcal{N}(d_1) - EP\mathcal{N}(d_2) \tag{19.23}$$

with

$$d_1 = \frac{\ln(X/1) + \frac{\bar{V}^2}{2}(T-t)}{\bar{V}\sqrt{T-t}} = \frac{\ln(S/EP) + \frac{\bar{V}^2}{2}\tau}{\bar{V}\sqrt{\tau}}$$
$$= \frac{\ln(S/E) - \ln P(\tau) + \frac{\tau}{2}\bar{V}^2}{\bar{V}\sqrt{\tau}}$$

and

$$d_2 = d_1 - \bar{V}\sqrt{\tau}.$$

The solution technique originally used by Merton involved various changes of variables, and we reproduce his analysis in Appendix 19.1.

The pricing formula (19.23) generalises the Black–Scholes formula in a very natural way. The bond with same maturity as the option is used to do stochastic discounting and the "average" volatility $\bar{V}$ replaces the $\sigma$ of the standard case. If we adopt the common practice of calculating an implied $\bar{V}$ from market data then there is no need to estimate separately the $\rho, \delta(\tau)$ and $\sigma$. This observation also helps to explain the robustness of the Black–Scholes model, (when used with implied volatility) as the volatility so calculated is compatible with a wide class of deterministic time functions of $\sigma$, $\rho$ and $\delta$, and not just with a constant $\sigma$.

## 19.4 Appendix

### Appendix 19.1 Solving the P.D.E. by Change of Variable

If we introduce the new time variable[8]

$$\theta = \int_0^\tau V^2(s)ds, \tag{19.24}$$

and define

$$g(X, \theta) = h(X, \tau(\theta)),$$

then $g$ satisfies

$$\frac{1}{2}X^2\frac{\partial^2 g}{\partial X^2} - \frac{\partial g}{\partial \theta} = 0, \tag{19.25}$$

subject to

$$\begin{aligned} g(0, \theta) &= 0, \\ g(X, 0) &= \max[0, X - 1]. \end{aligned} \tag{19.26}$$

We can interpret (19.25) as the Black–Scholes option pricing equation for an option with time $\theta$ to maturity, exercise price of one dollar, when the underlying stock has variance of unity and the market interest rate is zero. So we can use the known solution to write

$$g(X, \theta) = X\mathcal{N}(d_1) - \mathcal{N}(d_2), \tag{19.27}$$

where

$$\begin{aligned} d_1 &= \frac{\ln X + \frac{1}{2}\theta}{\sqrt{\theta}}, \\ d_2 &= d_1 - \sqrt{\theta}. \end{aligned}$$

Working back through the transformations we obtain

$$f(S, P, \tau) = EP(\tau)g\left(\frac{S}{EP(\tau)}, \int_0^\tau V^2(s)ds\right). \tag{19.28}$$

[8]Note that Eq. (19.24) defines a functional relationship between $\tau$ and $\theta$.

In addition to the usual inputs of the Black–Scholes model, (19.28) requires $P(\tau)$ as well as $\rho$ and $\delta(\tau)$. If we define

$$\bar{V}^2 = \frac{1}{\tau}\int_0^\tau V^2(s)ds. \tag{19.29}$$

and make use of (19.27) then the expression for the option price in (19.28) can be written

$$f(S, P, \tau) = S\mathcal{N}(d_1) - EP(\tau)\mathcal{N}(d_2), \tag{19.30}$$

where

$$d_1 = \frac{\ln(\frac{S}{EP(\tau)}) + \frac{\tau}{2}\bar{V}^2}{\bar{V}\sqrt{\tau}} = \frac{\ln(\frac{S}{E}) - \ln P(\tau) + \frac{1}{2}\tau\bar{V}^2}{\bar{V}\sqrt{\tau}},$$

and

$$d_2 = d_1 - \bar{V}\sqrt{\tau}.$$

## 19.5 Problems

**Problem 19.1**

(a) Redo the analysis of Sect. 19.2 when the underlying asset price and bond price dynamics are specified in the following way:

$$\frac{dx_1}{x_1} = m_1dt + s_{11}dw_1 + s_{12}dw_2,$$
$$\frac{dx_P}{x_P} = m_Pdt + s_{P1}dw_1 + s_{P2}dw_2,$$

respectively. Here $dw_1$, $dw_2$ are independent Wiener increments (recall that by appropriate transformations we can always reduce to this situation).

(b) Allow for the situation in which the underlying asset pays a continuously compounded dividend at the rate $q_1$. In this case show that the pricing partial differential equation is given by

$$\frac{\partial f}{\partial t} - q_1x_1\frac{\partial f}{\partial x_1} + \mathcal{D}f = 0$$

Hence show that the pricing formula (19.21) becomes

$$f(x_1, x_p, \tau) = x_1 e^{-\bar{q}_1(T-t)} \mathscr{N}(d_1) - E x_p \mathscr{N}(d_2). \tag{19.31}$$

$$d_1 = \frac{\ln\left(\frac{x_1}{E}\right) - \ln x_p + \left(\frac{1}{2}\bar{V}^2 - \bar{q}_1\right)(T-t)}{\bar{V}\sqrt{T-t}}$$

and

$$\bar{q}_1 = \frac{1}{T-t}\int_0^{T-t} \bar{q}_1(s)ds.$$

*Hint*: In part (a) you will need to recall that in order that the system

$$a_{11}Q_1 + a_{12}Q_2 = 0$$
$$a_{21}Q_1 + a_{22}Q_2 = 0$$

have non-zero solutions then

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 0. \tag{19.32}$$

A consequence of (19.32) is that the columns of the matrix are linearly dependent, this means that there exists a constant $\lambda$ such that

$$a_{12} = \lambda a_{11}, \tag{19.33}$$
$$a_{22} = \lambda a_{21}.$$

If you use (19.32) you will need to persevere with a lot of algebra but you will eventually obtain the key result that

$$f = x_1 \frac{\partial f}{\partial x_1} + x_P \frac{\partial f}{\partial x_P}$$

and hence the pricing equation

$$\frac{\partial f}{\partial t} + \mathscr{D} f = 0, \qquad \mathscr{D} f = \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} S_{ij} x_i x_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

This may be called the “brute force approach”.

If you use (19.33), and use matrix notation at the appropriate point, you will obtain the same results with much less algebraic manipulations. This may be called the “elegant approach”.

**Problem 19.2** Consider again the situation in Sect. 19.2 where we have a stochastic interest rate in a Black–Scholes world.

Suppose that the option is written such that its payoff is a function of two underlying assets $x_1$ and $x_2$, and that the dynamics of the asset prices and bond price have been put into the form

$$\frac{dx_1}{x_1} = m_1 dt + s_{11} dw_1 + s_{12} dw_2 + s_{1p} dw_p,$$
$$\frac{dx_2}{x_2} = m_2 dt + s_{21} dw_1 + s_{22} dw_2 + s_{2p} dw_p,$$
$$\frac{dx_P}{x_P} = m_P dt + s_{P1} dw_1 + s_{P2} dw_2 + s_{PP} dw_P,$$

where the $dw_i \, (i = 1, 2, P)$ are independent. Note that the option price will now be a function $f(x_1, x_2, x_P, t)$.

Allow the underlying assets to pay continuously compounded dividends at the rates $q_1$ and $q_2$ respectively. Show that

$$f = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_P \frac{\partial f}{\partial x_P}$$

and hence derive the pricing partial differential equation.

Determine the transformation that is the equivalent of the transformation used in Sect. 19.3. What does the pricing equation look like after this transformation?

*Hint*: You will need to use what is called the "elegant approach" in the hint to Problem 19.1.

# Chapter 20
# Change of Numeraire

**Abstract** Many computational applications of derivative pricing models such as the determination of derivative prices by simulation or the estimation of derivative pricing models can be significantly simplified by a change of numeraire. In this chapter we discuss the main idea behind the change of numeraire technique and the formation of equivalent probability measures under which options can be priced. In addition, the connection of the associated numeraires via the Radon–Nikodym derivative are presented. We also consider an application of the technique for the option pricing models with stochastic interest rate discussed in Chap. 19 and an extension of the technique to accommodate multiple sources of risk in the dynamics of the underlying assets is also considered.

## 20.1 A Change of Numeraire Theorem

Consider some traded asset $S$ (e.g. stock) whose price follows the diffusion process

$$dS = \mu_s S dt + \sigma_s S dW,$$

where the Wiener process $W$ is generated by the measure $\mathbb{P}$. Let $U$, $V$ be the prices of derivative instruments dependent on $S$, so that $U = U(S,t)$ and $V = V(S,t)$. For instance $U$ and $V$ could be options of different maturity or $U$ could be an option and $V$ could be a bond.

From the discussion of Sect. 8.4, we know that if the economy is arbitrage free then under $\mathbb{P}$ the dynamics of the stock price and the derivative securities are given by

$$dS = (r + \lambda\sigma_s)S dt + \sigma_s S dW,$$

$$dU = (r + \lambda\sigma_u)U dt + \sigma_u U dW,$$

$$dV = (r + \lambda\sigma_v)V dt + \sigma_v V dW,$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_20

where $\lambda$ is the market price of risk associated with the uncertainty $W$ and the expressions for $\sigma_u \left(= \sigma_s \frac{S}{U}\frac{\partial U}{\partial S}\right)$ and $\sigma_v \left(= \sigma_s \frac{S}{V}\frac{\partial V}{\partial S}\right)$ are obtained from application of Ito's lemma. Applying Girsanov's theorem we can re-express the above dynamics as

$$dS = rSdt + \sigma_s S d\tilde{W}, \tag{20.1}$$

$$dU = rUdt + \sigma_u U d\tilde{W}, \tag{20.2}$$

$$dV = rVdt + \sigma_v V d\tilde{W}, \tag{20.3}$$

where

$$\tilde{W}(t) = W(t) + \int_0^t \lambda(\tau)d\tau, \tag{20.4}$$

is a Wiener process under the equivalent risk-neutral measure $\tilde{\mathbb{P}}$. If we let $r(t)$ denote the (possibly stochastic) risk-free rate of interest then we can form the money market account

$$A(t) = \exp\left(\int_0^t r(s)ds\right),$$

satisfying the dynamics

$$dA = rAdt.$$

Applying the results of Sect. 6.6 concerning the stochastic differential equation followed by the quotient of two diffusions, we show that

$$d\left(\frac{U}{A}\right) = \sigma_u \left(\frac{U}{A}\right) d\tilde{W}, \tag{20.5}$$

$$d\left(\frac{V}{A}\right) = \sigma_v \left(\frac{V}{A}\right) d\tilde{W}. \tag{20.6}$$

Since the drift terms in Eqs. (20.5) and (20.6) are zero it follows that the relative prices $U/A$, $V/A$ are martingales under $\tilde{\mathbb{P}}$. That is,

$$\frac{U_t}{A_t} = \tilde{\mathbb{E}}_t\left[\frac{U_T}{A_T}\right], \tag{20.7}$$

and

$$\frac{V_t}{A_t} = \tilde{\mathbb{E}}_t\left[\frac{V_T}{A_T}\right], \tag{20.8}$$

where $\tilde{\mathbb{E}}_t$ denotes the expectation under $\tilde{\mathbb{P}}$ conditional on information at time $t$. Equations (20.7) and (20.8) can easily be manipulated to yield

$$U_t = \tilde{\mathbb{E}}_t \left[ \exp\left( - \int_t^T r(s) ds \right) U_T \right], \tag{20.9}$$

and

$$V_t = \tilde{\mathbb{E}}_t \left[ \exp\left( - \int_t^T r(s) P ds \right) V_T \right]. \tag{20.10}$$

Note that since we are now considering situations in which $r$ is possibly stochastic, we leave (stochastic) integrals involving it under the expectation operator. In Eqs. (20.9) and (20.10) the term $\exp(-\int_t^T r(s))$ acts as a stochastic discount factor. For a particular realisation of the interest rate process (under $\tilde{\mathbb{P}}$) it discounts back to time $t$ the realised values of $U_T$ and $V_T$. The operation $\tilde{\mathbb{E}}_t$ then basically averages over all such expected payoffs as is illustrated in Fig. 20.1. In deriving (20.9) and (20.10) we have used the money market account as the numeraire. However it is possible to use other instruments as numeraire. For example, in pricing $U$ we may use $V$ as the numeraire (e.g. $V$ could be a bond price with the same maturity as $U$), then we would consider the relative price

$$Y = \frac{U}{V}.$$

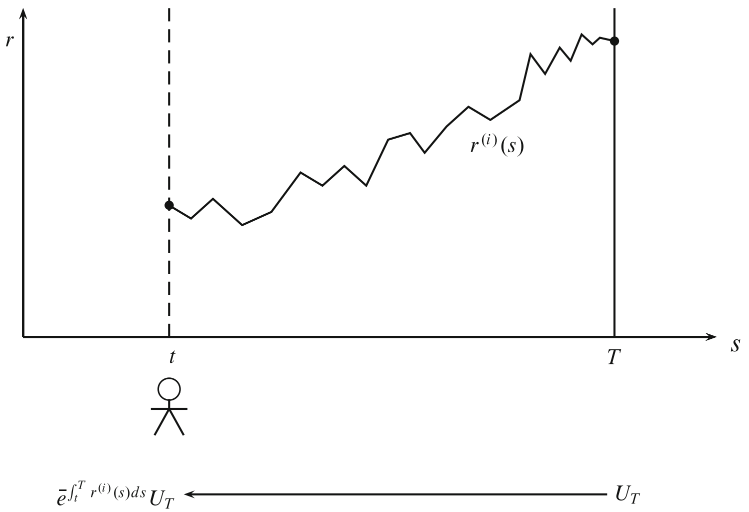


**Fig. 20.1** Discounting the value of $U_T$ from $T$ back to $t$ for the particular $i$th realisation $r^{(i)}(s)$ of the interest rate process. The expectation $\tilde{\mathbb{E}}_t$ in (20.9) averages over many such paths

We continue to use the arbitrage free dynamics for $U$ and $V$, viz. Eqs. (20.2) and (20.3). Using again the results of Sect. 6.6, we find that

$$dY = (\sigma_v - \sigma_u)\sigma_v Y\,dt + (\sigma_u - \sigma_v)Y d\tilde{W}. \tag{20.11}$$

If we introduce the new process

$$W^*(t) = \tilde{W}(t) - \int_0^t \sigma_v(s)ds, \tag{20.12}$$

then by Girsanov's theorem, we can obtain an equivalent probability measure $\mathbb{P}^*$ under which $W^*(t)$ will be a Wiener process. It would then follow that under $\mathbb{P}^*$, $Y$ is a martingale since

$$dY = (\sigma_u - \sigma_v)Y\,dW^*. \tag{20.13}$$

Thus

$$Y_t = \mathbb{E}_t^*(Y_T),$$

or, on using the definition of $Y$

$$U_t = V_t \mathbb{E}_t^*\left[\frac{U_T}{V_T}\right]. \tag{20.14}$$

Equation (20.14) may provide a more convenient pricing relationship, especially if $V$ is the price of a bond that matures at the same time as the instrument $U$. From (20.4) and (20.12) the relation between the processes $W(t)$ and $W^*(t)$ becomes

$$W^*(t) = W(t) + \int_0^t \lambda(s)ds - \int_0^t \sigma_v(s)ds. \tag{20.15}$$

In Table 20.1 we represent the relation between the three measures $\mathbb{P}$, $\tilde{\mathbb{P}}$ and $\mathbb{P}^*$ under which the financial market dynamics may be considered. We also show the relation between the related processes $W(t)$, $\tilde{W}(t)$ and $W^*(t)$, in particular under which measures they are and are not Wiener processes.

**Table 20.1** Summarising the relation between the processes $W(t)$, $\tilde{W}(t)$ and $W^*(t)$ and the measures $\mathbb{P}$, $\tilde{\mathbb{P}}$ and $\mathbb{P}^*$

| | $\mathbb{P}$ | $\tilde{\mathbb{P}}$ | $\mathbb{P}^*$ |
|---|---|---|---|
| $W$ | Wiener | Not Wiener | Not Wiener |
| | $\mathbb{E}(dW) = 0$ | $\tilde{\mathbb{E}}(dW) = -\lambda dt \neq 0$ | $\mathbb{E}^*(dW) = (\sigma_v - \lambda)dt \neq 0$ |
| $\tilde{W}$ | Not Wiener | Wiener | Not Wiener |
| | $\mathbb{E}(d\tilde{W}) = \lambda dt \neq 0$ | $\tilde{\mathbb{E}}(d\tilde{W}) = 0$ | $\mathbb{E}^*(d\tilde{W}) = \sigma_v dt \neq 0$ |
| $W^*$ | Not Wiener | Not Wiener | Wiener |
| | $\mathbb{E}(dW^*) = -(\sigma_v - \lambda)dt \neq 0$ | $\tilde{\mathbb{E}}(dW^*) = -\sigma_v dt \neq 0$ | $\mathbb{E}^*(dW^*) = 0$ |

Here $\tilde{W}(t) = W(t) + \int_0^t \lambda(s)ds$, $W^*(t) = \tilde{W}(t) - \int_0^t \sigma_v(s)ds$ and so $W^*(t) = W(t) + \int_0^t \lambda(s)ds - \int_0^t \sigma_v(s)ds$

## 20.2 The Radon–Nikodym Derivative

Comparing (20.7) and (20.14) we obtain a general result which allows us to change numeraire, viz.

$$U_t = V_t \mathbb{E}_t^* \left[ \frac{U_T}{V_T} \right] = A_t \tilde{\mathbb{E}}_t \left[ \frac{U_T}{A_T} \right].$$

Rearranging we obtain

$$\mathbb{E}_t^* \left[ \frac{V_t}{V_T} \cdot U_T \right] = \tilde{\mathbb{E}}_t \left[ \frac{A_t}{A_T} \cdot U_T \right]. \tag{20.16}$$

This last result allows us to obtain the Radon–Nikodym derivative which underlies the change of numeraire result in Eq. (20.16). (Refer also to Sect. 8.2 for the Radon–Nikodym derivative.) Some formal manipulations allow us to re-express (20.16) as

$$\mathbb{E}_t^* \left[ \frac{U_T}{V_T} \right] = \tilde{\mathbb{E}}_t \left[ \frac{A_t V_T}{A_T V_t} \cdot \frac{U_T}{V_T} \right]. \tag{20.17}$$

Introducing a notation similar to one employed in Sect. 8.2 we set

$$\xi(t, T) = \frac{A_t V_T}{A_T V_t}, \tag{20.18}$$

so that (20.17) may be written

$$\mathbb{E}_t^* \left[ \frac{U_T}{V_T} \right] = \tilde{\mathbb{E}}_t \left[ \xi(t, T) \cdot \frac{U_T}{V_T} \right]. \tag{20.19}$$

The quantity $\xi(t, T)$ is thus the Radon–Nikodym derivative which allows us to switch from calculating expectations under $\tilde{\mathbb{P}}$ to calculating expectations under $\mathbb{P}^*$, i.e.

$$d\mathbb{P}^* = \xi(t, T)d\tilde{\mathbb{P}}.$$

Considering the interval $(0, T)$ we note from (20.18) that

$$\xi(0, T) = \frac{A_0 V_T}{A_T V_0} = \frac{V_T}{A_T V_0}, \tag{20.20}$$

since $A_0 = 1$.

## 20.3 Option Pricing Under Stochastic Interest Rates

An important application of the change of numeraire results is the Black–Scholes model under stochastic interest rates that we have considered in Chap. 19. Consider the no-arbitrage condition (19.14)

$$\frac{\mu - \alpha}{\sigma} = \frac{\mu_f - \alpha}{\sigma_{f_z}} = \phi, \tag{20.21}$$

where $\phi$ is the market price of risk associated with the stock price noise, $z$. Substituting (20.21) into the stochastic differential equations (19.4) for $S$ and (19.7) for $f$ we obtain

$$\frac{dS}{S} = (\alpha + \phi\sigma)dt + \sigma dz, \tag{20.22}$$

and

$$\frac{df}{f} = (\alpha + \phi\sigma_{f_z})dt + \sigma_{f_z}dz + \sigma_{f_v}dv, \tag{20.23}$$

which together with the bond pricing equation

$$\frac{dP}{P} = \alpha dt + \delta dv, \tag{20.24}$$

give the arbitrage free financial market dynamics under $\mathbb{P}$. By defining new processes $\hat{z}$ and $\hat{v}$ Eqs. (20.22) and (20.23) may be written

$$\frac{dS}{S} = \alpha dt + \sigma d\hat{z}, \tag{20.25}$$

$$\frac{df}{f} = \alpha dt + \sigma_{f_z} d\hat{z} + \sigma_{f_v} d\tilde{v}, \tag{20.26}$$

where

$$\hat{z}(t) = z(t) + \int_0^t \phi(s)ds,$$

and[1]

$$\hat{v}(t) = v(t).$$

By Girsanov's theorem we can find an equivalent measure $\hat{\mathbb{P}}$ under which $\hat{z}(t)$ and $\hat{v}(t)$ are Wiener processes.[2] Clearly by introducing the quantity

$$B(t) = \exp\left(\int_o^t \alpha(s)ds\right),$$

then $S/B$ and $f/B$ become martingales under the measure $\hat{\mathbb{P}}$ i.e.

$$f_t = \hat{\mathbb{E}}_t\left[B(t)\frac{f_T}{B(T)}\right]. \tag{20.27}$$

However following the discussion of Sect. 20.1, we could also allow the bond price $P$ to be the numeraire. Using $\mathbb{P}^*$ to denote the probability measure when $P$ is the numeraire, by the foregoing discussion we have the result that

$$f_t = \hat{\mathbb{E}}_t\left[B(t)\frac{f_T}{B(T)}\right] = \mathbb{E}_t^*\left[P(t,T)\frac{f_T}{P(T,T)}\right],$$

i.e.,

$$\frac{f_t}{P(t,T)} = \mathbb{E}_t^*\left[f_T\right], \tag{20.28}$$

since $P(T,T) = 1$ (recall the bond matures when the option does). In order to operationalise (20.28), we need to make explicit the stochastic price dynamics from which $\mathbb{P}^*$ and hence $\mathbb{E}_t^*$ can be calculated. The dynamics are those for the stochastic

[1] Since $v$ and $\hat{v}$ are the same process this seems a redundant transformation. However it is 'cleaner' to think of two new processes for $z$ and $v$.

[2] In fact Girsanov's theorem in a multidimensional setting is stated in terms of independent Wiener processes. So the correct approach would be to follow the approach to stochastic interest rates in the Black–Scholes model as suggested in Problem 19.1. This is pursued in Problem 20.1. Then we find that we avoid the redundant transformation discussed in footnote 1.

differential equation for $S/P$ derived from (20.24) and (20.25). By an application of Ito's lemma we derive the dynamics for $S/P$ as

$$\frac{d(S/P)}{S/P} = (\delta^2 - \rho\sigma\delta)dt + \sigma d\hat{z} - \delta d\hat{v}. \tag{20.29}$$

Under $\hat{\mathbb{P}}$ this is a geometric Brownian motion with instantaneous variance

$$V^2 dt = (\sigma^2 + \delta^2 - 2\rho\sigma\delta),$$

since under $\hat{\mathbb{P}}$

$$\hat{\mathbb{E}}[\sigma d\hat{z} - \delta d\hat{v}] = 0,$$

and

$$\text{var}[\sigma d\hat{z} - \delta d\hat{v}] = \hat{\mathbb{E}}[(\sigma d\hat{z} - \delta d\hat{v})^2] = V^2 dt.$$

Equation (20.29) corresponds precisely to Eq. (19.18). Since a sum of a linear combination of normal random variables is also normal we can define under $\hat{\mathbb{P}}$ a new Wiener process $\hat{w}$ such that

$$\sigma d\hat{z} - \delta d\hat{v} = V d\hat{w}.$$

Equation (20.29) can then be written

$$\frac{d(S/P)}{S/P} = (\delta^2 - \rho\sigma\delta)dt + V d\hat{w}. \tag{20.30}$$

Introducing the new Wiener process

$$w^*(t) = \hat{w}(t) + \int_0^t \frac{\delta^2 - \rho\sigma\delta}{V} ds$$

we can again apply Girsanov's theorem to write (20.30) as

$$\frac{d(S/P)}{S/P} = V dw^*, \tag{20.31}$$

under $\mathbb{P}^*$. It is the stochastic equation (20.31) which generates paths under the measure $\mathbb{P}^*$. The associated Kolmogorov partial differential equation for the transition probability density function under $\mathbb{P}^*$, denoted $p^*$, is

$$\frac{1}{2} V^2 X^2 \frac{\partial^2 p^*}{\partial X^2} + \frac{\partial p^*}{\partial t} = 0, \tag{20.32}$$

where $X = S/P$. Hence applying the Feyman-Kac formula to the expectation on the right-hand side of (20.28), we see that $f/P$ is given by the solution of

$$\frac{1}{2}V^2X^2\frac{\partial^2(f/P)}{\partial X^2} + \frac{\partial(f/P)}{\partial t} = 0,$$

subject to[3]

$$f(X,T) = \max[0, X - E].$$

We are in fact dealing with the Black–Scholes model with $r = 0$, and will thus again recover the solution given by Eq. (19.28).

## 20.4 Change of Numeraire with Multiple Sources of Risk

The previous sections introduced the change of numeraire idea in an economy with just one source of risk. Suppose we have $(n+1)$-risky assets whose dynamics are driven by

$$\frac{dS_i}{S_i} = \mu_i dt + \sum_{j=0}^{n} s_{ij} dW_j(t) \ (i = 0, 1, \ldots, n), \tag{20.33}$$

where the $W_j(t)$ are independent Wiener processes. We consider here the case when all assets are traded, though the framework we develop is easily extended to allow the case in which some of the $S_i$ may not be traded. We use $\boldsymbol{S}$ to denote the vector of risky asset prices $(S_0, S_1, \ldots, S_n)$. Let $f_k(\boldsymbol{S}, t)$ $(k = 1, 2, \ldots, m)$ denote a derivative instrument written on the vector of processes $\boldsymbol{S}$.

From the discussion in Chap. 10 (see Sect. 10.4) we know that there exists a unique risk-neutral measure $\tilde{\mathbb{P}}$ under which the dynamic for the $S_i$ and $f_k$ can be written

$$\frac{dS_i}{S_i} = (r - q_i)dt + \sum_{j=0}^{n} s_{ij} d\tilde{W}_j(t), \ (i = 0, 1, \ldots, n), \tag{20.34}$$

$$\frac{df_k}{f_k} = rdt + \sum_{j=0}^{n} \sigma_{kj} d\tilde{W}_j(t) \ (k = 1, 2, \ldots, m), \tag{20.35}$$

where the $\tilde{W}_j(t)$ are the Wiener processes under $\tilde{\mathbb{P}}$. In Eq. (20.34) the $q_i$ is the dividend yield on $S_i$, and in (20.35) the $\sigma_{kj}$ are the volatility factors whose

[3]Recall that $P(T,T) = 1$.

calculation is outlined in Sect. 10.4, and are given by

$$f_k \sigma_{kj} = \sum_{l=1}^{n} s_{lj}\, S_l\, \Delta_{kl},$$

where

$$\Delta_{kl} = \frac{\partial f_k}{\partial S_l}.$$

Defining prices in terms of the money market account $A_t = \exp\left(\int_0^t r(s)ds\right)$ we would of course obtain

$$f_k(S,t) = \tilde{\mathbb{E}}_t\left[e^{-\int_t^T r(s)ds} f_k(\boldsymbol{S}(T),T)\right], \tag{20.36}$$

where the distribution $\tilde{\mathbb{P}}$ under which $\tilde{\mathbb{E}}_t$ is calculated is generated by the $(n+1)$ processes (20.34). Suppose instead we decide that it is convenient to use $S_0(t)$ as the numeraire (it is always possible to relabel the assets so that $S_0$ is the new numeraire). Thus we are interested in the processes

$$Z_i(t) = \frac{S_i(t)}{S_0(t)},\ (i = 1,\ldots,n), \tag{20.37}$$

and

$$Y_k(t) = \frac{f_k(t)}{S_0(t)}\ (k = 1,\ldots,m). \tag{20.38}$$

Applying the result (6.83) for the stochastic differential of the quotient of two diffusions the dynamics for the $Z_i$ and $Y_k$ processes become

$$\frac{dZ_i}{Z_i} = \left[(q_0 - q_i) - \sum_{j=0}^{n} s_{0j}(s_{ij} - s_{0j})\right] dt + \sum_{j=0}^{n}(s_{ij} - s_{0j})d\tilde{W}_j, \tag{20.39}$$

and

$$\frac{dY_k}{Y_k} = \left[q_0 - \sum_{j=1}^{m} s_{0j}(\sigma_{kj} - s_{0j})\right] dt + \sum_{j=1}^{m}(\sigma_{kj} - s_{0j})d\tilde{W}_j. \tag{20.40}$$

The dividend yields prevent us from obtaining martingales under the new measure, but this problem is easily overcome by defining the new processes

$$Z_i^*(t) = Z_i(t)e^{-(q_0-q_i)t}, \tag{20.41}$$

$$Y_k^*(t) = Y_k(t)e^{-q_0 t}, \tag{20.42}$$

whose dynamics are easily calculated as

$$\frac{dZ_i^*}{Z_i^*} = -\sum_{j=1}^{n} s_{0j}(s_{ij} - s_{0j})dt + \sum_{j=1}^{n}(s_{ij} - s_{0j})d\tilde{W}_j, \tag{20.43}$$

and

$$\frac{dY_k^*}{Y_k^*} = -\sum_{j=1}^{m} s_{0j}(\sigma_{kj} - s_{0j})dt + \sum_{j=1}^{m}(\sigma_{kj} - s_{0j})d\tilde{W}_j. \tag{20.44}$$

We know by Girsanov's theorem that we can find a new measure $\mathbb{P}^*$ under which the processes

$$W_j^*(t) = \tilde{W}_j(t) - \int_0^t s_{0j}(\tau)d\tau \tag{20.45}$$

are Wiener processes. Thus under this new measure the dynamics for the $Z_i^*$ and $Y_k^*$ become

$$\frac{dZ_i^*}{Z_i^*} = \sum_{j=1}^{n}(s_{ij} - s_{0j})dW_j^*. \tag{20.46}$$

and

$$\frac{dY_k^*}{Y_k^*} = \sum_{j=1}^{m}(\sigma_{kj} - s_{0j})dW_j^*. \tag{20.47}$$

The last equation indicates that the $Y_k^*$ are martingales under $\mathbb{P}^*$, so that

$$Y_k^*(t) = \mathbb{E}_t^*[Y_k^*(T)]. \tag{20.48}$$

By use of (20.42) this expression becomes

$$Y_k(t) = \mathbb{E}_t^*[e^{-q_0(T-t)}Y_k(T)].$$

In terms of the original variable $f_k(t)$ we have

$$f_k(t) = S_0(t)\mathbb{E}_t^*\left[e^{-q_0(T-t)}\frac{f_k(T)}{S_0(T)}.\right] \tag{20.49}$$

Suppose for example that

$$f_k(T) = \max\left[0, \sum_{j=0}^{n}\alpha_j S_j(T)\right]$$

when the $\alpha_j$ are constants. Thus

$$\frac{f_k(T)}{S_0(T)} = \max\left[0, \alpha_0 + \sum_{j=1}^{n}\alpha_j Z_j(T)\right] = \max\left[0, \alpha_0 + \sum_{j=1}^{n}\alpha_j e^{(q_0-q_i)T} Z_j^*(T)\right].$$

Equation (20.49) then becomes

$$f_k(t) = S_0(t)\mathbb{E}_t^*\left[e^{q_t-q_i T}\max\left[0, \alpha_0 e^{(q_i-q_0)T} + \sum_{j=1}^{n}\alpha_j Z_j^*(T)\right]\right].$$

The process under which $\mathbb{E}_t^*$ is calculated is given by the system (20.46).

## 20.5 Problems

**Problem 20.1** Consider the problem of stochastic interest rates in the Black–Scholes model, but formulated as in Problem 19.1. Rework the change of numeraire argument of Sect. 20.3 in this formulation (see the comments in footnote 2).

**Problem 20.2 Computational Problem**—Consider the valuation relationship (20.9). Suppose the stochastic interest rate follows the stochastic differential equation

$$dr = \kappa(\bar{r} - r)dt + \sigma dW.$$

Suppose that $U(0, T)$ is a zero-coupon bond that pays \$1 at time $T$.

Use simulation to evaluate $U(0, 0.5)$. Take the parameter values

$$\kappa = 0.6, \bar{r} = 0.07, \sigma = 0.024.$$

Experiment with the step size and number of paths so as to ensure two decimal accuracy.

# Chapter 21
# The Paradigm Interest Rate Option Problem

**Abstract** There are a number of instruments in interest rate markets that are equivalent to an option on an interest rate or an option on a bond. In this chapter we focus on the interest rate caps, which are call options on an interest rate. We show that they can be interpreted as a put option on a bond. The problem of pricing such bonds, and hence the interest rate cap, shall motivate much of the discussion in subsequent chapters. In the last section we briefly discuss the issues associated with the interest rate option problem that distinguish it from the option pricing problem in a world of deterministic interest rates.

## 21.1 Interest Rate Caps, Floors and Collars

### *21.1.1 Interest Rate Caps*

An interest rate cap is an agreement written on some reference rate $R$ (e.g. 6-month LIBOR) that sets the borrowing rate at the market rate $R$ if $R < R_{Cap}$ and limits the rate to $R_{Cap}$ if the market rate $R > R_{Cap}$, see Fig. 21.1. The date at which the comparison between $R$ and $R_{Cap}$ is made is known as the reset date.

An interest rate cap is a call option on the interest rate, and is an insurance against the interest rate on an underlying floating rate asset rising above a certain level (i.e., the strike price or cap rate), Fig. 21.2.

If the interest rate rises above the cap rate, the buyer of the cap effectively receives a payoff which is the difference between the current market rate and the cap rate. There can be a series of rate resets over the life of the cap. Hence, the cap is a portfolio of call options. Each component of the cap is known as a caplet. For the caplet over the period between $t_i$ and $t_{i+1}$, the cap rate $R_{Cap}$ is compared with the reference rate at $t_i$ (i.e., $R_i$). However, the payoff for this period is settled at $t_{i+1}$ (i.e., payment in arrears).



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_21

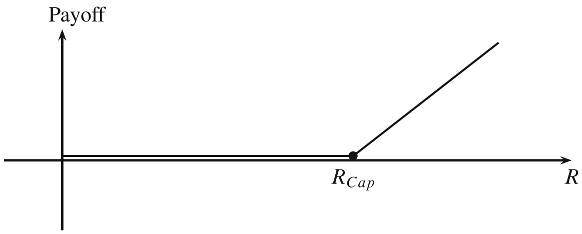


**Fig. 21.1** Payoff on a long position in an interest rate cap

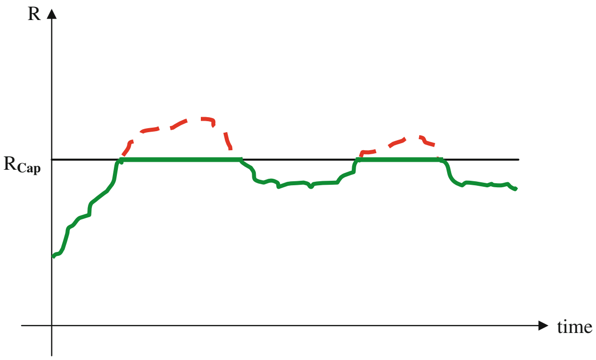


**Fig. 21.2** Illustrating the actual (*red*) and capped (*green*) interest rates

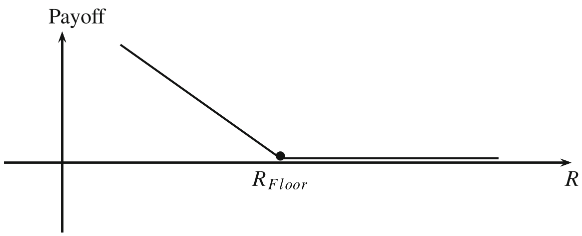


**Fig. 21.3** Payoff on a long position in an interest rate floor

### *21.1.2 Interest Rate Floors*

An interest rate floor is the reverse of an interest rate cap. A long position would be of interest to a lender who wants to guarantee that a lending rate will not fall below a certain pre-specified rate. If the market rate $R > R_{Floor}$ the lender receives the market rate $R$. If the market rate $R < R_{Floor}$ then the lender receives the floor rate $R_{Floor}$, see Fig. 21.3.

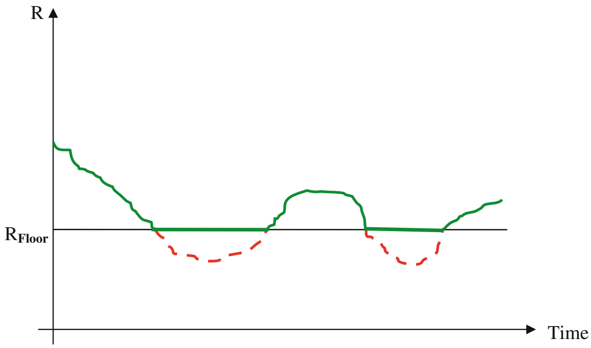


**Fig. 21.4** Illustrating the actual (*red*) and floor (*green*) rates

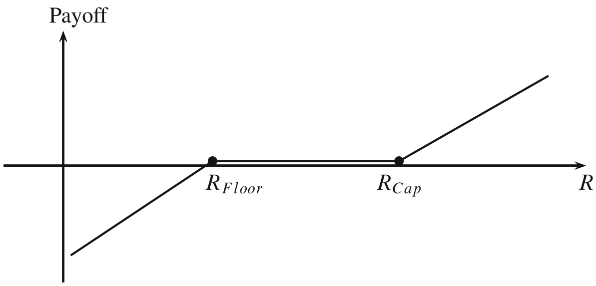


**Fig. 21.5** Long a cap and short a floor

The interest rate floor is a put option on the interest rate and is an insurance against the interest rate on an underlying floating rate asset falling below the strike price or floor rate. The buyer will only receive a payoff if the current interest rate is below the floor rate, see Fig. 21.4.

### *21.1.3 Interest Rate Collars*

A collar is a combination of a long position in an interest rate cap and a short position in an interest rate floor, see Fig. 21.5. The price for a collar will be lower than that of just a cap or a floor. Investors often enter into such arrangements to offset the cost of a cap with the premium from the floor. Effectively, they are selling off some of the cap's downside protection.

## 21.2 Payoff Structure of Interest Rate Caps and Floors

A cap is a series of caplets whose payoff is the difference between the spot rate at the reset date and the agreed cap rate. The present value at time $t_i$ of a caplet payoff received at time $t_{i+1}$ for a \$1 principal amount is:

$$PV(\text{caplet payoff}_i) = \frac{\tau \max[R_i - R_{Cap}, 0]}{1 + R_i\tau} \tag{21.1}$$

$$= \tau \max[R_i - R_{Cap}, 0] P(t_i, t_{i+1}) \tag{21.2}$$

where $\tau = t_{i+1} - t_i$, and $P(t_i, t_{i+1})$ is the price at $t_i$ of a pure discount bond maturing at time $t_{i+1}$. It is assumed in this framework that the underlying reference interest rate $R_i$ is quoted so that

$$(1 + R_i\tau)^{-1} = P(t_i, t_{i+1}), \tag{21.3}$$

which allows us to use the prices of pure discount bonds as discount factors in Eq. (21.2).

Conversely for an interest rate floor, the present value at time $t_i$ of the floorlet payoff for period $(t_i, t_{i+1})$ is:

$$PV(\text{floorlet payoff}_i) = \frac{\tau \max[R_{Floor} - R_i, 0]}{1 + R_i\tau} = \tau \max[R_{Floor} - R_i, 0] P(t_i, t_{i+1}).$$

Figure 21.6 illustrates the payoff situation for an $n$-period interest rate cap.

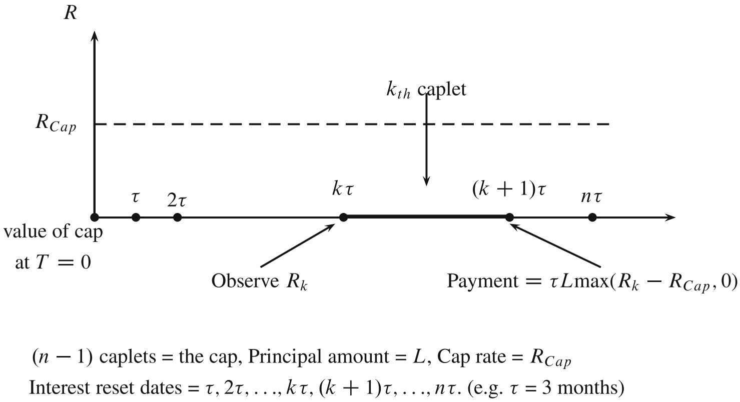


**Fig. 21.6** The payoff structure of an interest rate cap

## 21.3 Relationship to Bond Options

The purpose of this section is to demonstrate that the problem of pricing interest rate caplets and floorlets reduces to the problem of pricing put and call options on bonds. Hence interest rate caps and floors can be reduced to a portfolio of options on bonds. The significance of this observation lies in the fact that bonds are traded instruments and so we may use them to form the hedging portfolios that are the basis of derivative security pricing methodology. This reinterpretation of interest rate caps and floors is necessary since we cannot form hedging portfolios directly with the underlying reference interest rates as these are not traded instruments. From Eq. (21.1), the caplet payoff can be written as:

$$PV(\text{caplet payoff}_i) = \frac{\tau}{1 + R_i\tau} \max\left[\frac{R_i\tau - R_{Cap}\tau}{\tau}, 0\right] \tag{21.4}$$

$$= \frac{1}{1 + R_i\tau} \max\left[1 + R_i\tau - (1 + R_{Cap}\tau), 0\right] \tag{21.5}$$

$$= 1 \max\left[1 - \frac{1 + R_{Cap}\tau}{1 + R_i\tau}, 0\right] \tag{21.6}$$

$$= \left(1 + R_{Cap}\tau\right) \max\left[\frac{1}{1 + R_{Cap}\tau} - \frac{1}{1 + R_i\tau}, 0\right]. \tag{21.7}$$

Let $X_c = 1/(1 + R_{Cap}\tau)$, Eq. (21.7) becomes:

$$\begin{aligned} PV(\text{caplet payoff}_i) &= (1 + R_{Cap}\tau) \max\left[X_c - \frac{1}{1 + R_i\tau}, 0\right] \\ &= (1 + R_{Cap}\tau) \underbrace{\max\left[X_c - P(t_i, t_{i+1}), 0\right]}_{\text{payoff of a bond put option}} . \end{aligned}$$

Hence, the caplet payoff is equivalent (within the proportionality factor $1+R_{Cap}\tau$) to the payoff of a bond put option maturing at time $t_i$ on an underlying bond maturing at time $t_{i+1}$ with exercise price being $X_c$.

## 21.4 The Inherent Difficulty of the Interest Rate Option Problem

In this section we briefly highlight why it is that the interest rate option problem is a so much more difficult problem than the option pricing problem in a world of deterministic interest rates. There are essentially two main reasons. The first is that the number of underlying "assets" is infinite. The second is that it is not obvious which underlying "asset" to use.

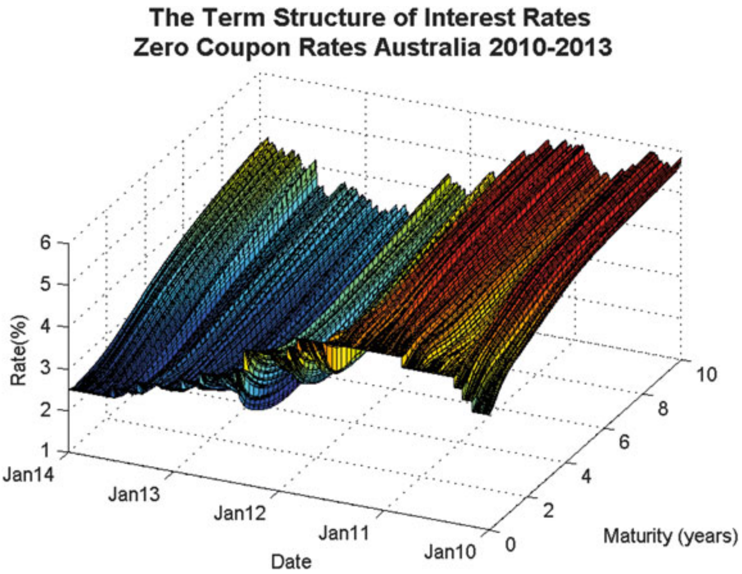


**Fig. 21.7** The term structure of interest rates in Australia 2010–2013

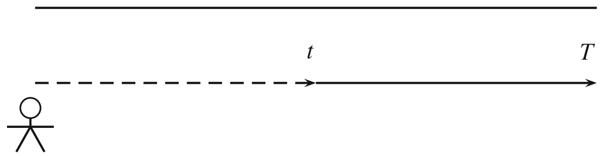


**Fig. 21.8** The time line for interest rate processes

Consider the first issue, namely the infinite dimensional nature of the problem. The basic object whose dynamics we seek to model is the yield curve out to a range of maturities. Figure 21.7 graphs the yield curve in Australia during 2010–2013 with maturities up to 10 years. Interest rate derivatives derive their value from the evolution of this surface. In principle this surface is an infinite dimensional object, though in later chapters we shall see that its dynamic evolution can be captured reasonably well by a finite number of fixed maturity forward rates.

On the second issue we note that one may use at least three different quantities to describe the par rates that are measured on the vertical axis in Fig. 21.7, these are (see the time line in Fig. 21.8);

$r(t)$ = instantaneous interest rate agreed at $t$ for borrowing starting at $t$,
$P(t,T)$ = price at time $t$ of pure discount bond maturing at time $T$,
$f(t,T)$ = instantaneous interest rate agreed at $t$ for borrowing starting at $T$.

The relationship between these quantities is illustrated in Fig. 21.9, the details of which will become clearer in subsequent chapters. We have used a ? in the

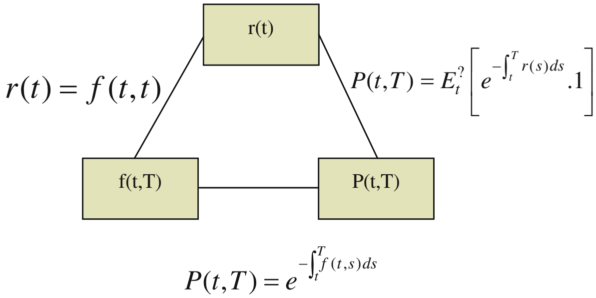


**Fig. 21.9** Illustrating the relationship between spot rate, forward rate and bond price

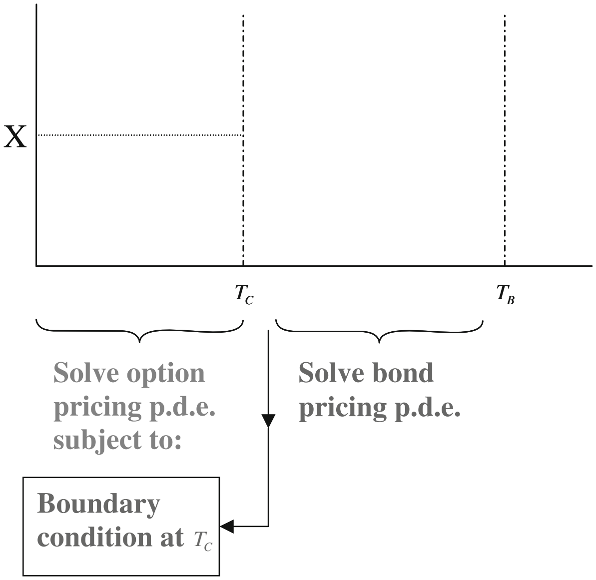


**Fig. 21.10** Illustrating the two pass nature of bond option pricing

expectation operator linking $r(t)$ and $P(t,T)$ to indicate that at this point the probability measure with respect to which this expectation is calculated is not known. Other quantities such as yields and the discretely compounded rates that are quoted in markets can also be determined in terms of $r(t)$, $f(t,T)$ or $P(t,T)$.

We have pointed out that the paradigm problem of pricing an interest cap (or floor) reduces to the problem of pricing an option on a bond. A further difficulty is that this process is in fact a two pass process. Suppose $T_B$ is the date of bond maturity (at which point the bond has a payoff of \$1) and $T_C$ is the bond option maturity date. We first need to solve for the bond price over the interval $T_B$ to $T_C$ to determine the possible option payoffs at $T_C$. We then need to solve the bond option pricing problem over the interval $T_C$ to 0. The nature of this two pass process is illustrated in Fig. 21.10.

# Chapter 22
# Modelling Interest Rate Dynamics

**Abstract** In this chapter, we establish the fundamental relationships between interest rates, bond prices and forward rates. We further discuss the modelling of interest rates and analyse typical models for the spot interest rate and the forward rates. As we desire interest rates to be non-negative, we seek stochastic processes with this feature such as the Feller process. Thus we present the motivation of the Feller process and its relevance to the interest rate modelling. We also summarise the main results of Fubini's theorem, that are very useful for modelling forward rates.

## 22.1 The Relationship Between Interest Rates, Bond Prices and Forward Rates

In this section we clarify the relationship between interest rates, bond prices, yield to maturity and forward rates. Let $P(t, T)$ denote the price at time $t$ of a pure discount bond paying \$1 at time $T$. The yield to maturity $\rho(t, T)$ is the continuously compounded rate of return causing the bond price to satisfy the maturity condition

$$P(T, T) = 1, \tag{22.1}$$

that is, $\rho(t, T)$ satisfies (see Fig. 22.1)

$$P(t, T)e^{\rho(t,T)(T-t)} = 1. \tag{22.2}$$

The yield may also be expressed as

$$\rho(t, T) = -\frac{\ln P(t, T)}{T - t}. \tag{22.3}$$

The instantaneous spot interest rate, $r(t)$, is the yield on the currently maturing bond, i.e.,

$$r(t) = \rho(t, t). \tag{22.4}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_22

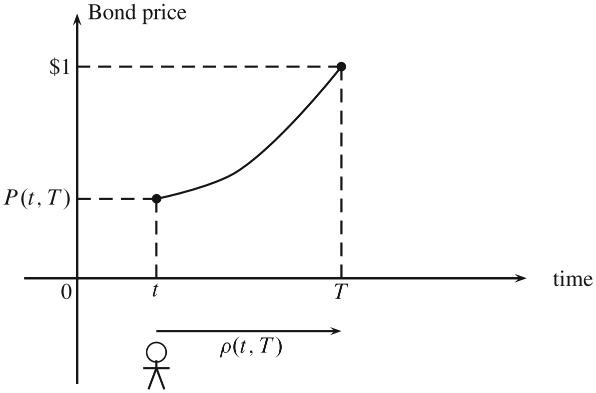


**Fig. 22.1** The yield to maturity $\rho(t,T)$; satisfies $P(t,T)e^{\rho(T-t)} = 1$

By allowing $t \to T$ in (22.3) and applying L'Hôpital's rule we find that[1]

$$r(T) = P_t(T,T), \tag{22.5}$$

where $P_t(T,T)$ denotes the partial derivative of $P(t,T)$ with respect to its first argument (running time $t$), evaluated at the time $t = T$.

The forward rate arises when we consider an investor who holds a bond maturing at $T_1$ and asking what return he or she would earn between $T_1$ and $T_2(> T_1)$, if he or she contracted now at time $t$. Figure 22.2 displays a time line that is useful when thinking about the relation between forward rates and bond prices. The required rate of return is the forward rate $f(t,T_1,T_2)$ defined by

$$P(t,T_1) = P(t,T_2)e^{f(t,T_1,T_2)(T_2-T_1)},$$

i.e.

$$f(t,T_1,T_2) = \frac{1}{T_2-T_1}\ln\left[\frac{P(t,T_1)}{P(t,T_2)}\right]. \tag{22.6}$$

To see how $f(t,T_1,T_2)$ represents the implicit rate of interest currently available (at time $t$) on riskless loans from $T_1$ to $T_2$ consider the set of transactions illustrated

[1]We need to apply L'Hôpital's rule since setting $t = T$ in Eq. (22.3) yields the meaningless ratio $0/0$. Thus by L'Hôpital's rule

$$\lim_{t\to T}\rho(t,T) = \lim_{t\to T} -\frac{\frac{\partial}{\partial t}\ln P(t,T)}{\frac{\partial}{\partial t}(T-t)} = \lim_{t\to T}\frac{\frac{-1}{P(t,T)}P_t(t,T)}{-1} = \lim_{t\to T}\frac{P_t(t,T)}{P(t,T)} = P_t(T,T).$$

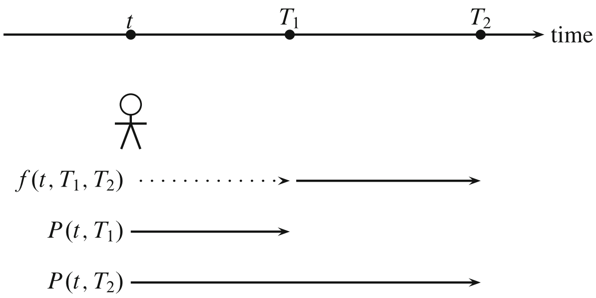


**Fig. 22.2** The forward rate $f(t,T_1,T_2)$ and bond prices $P(t,T_1)$, $P(t,T_2)$

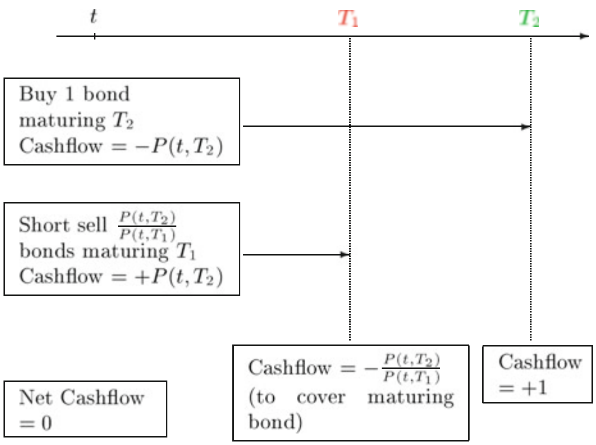


**Fig. 22.3** The net zero transaction that relates the forward $f(t,T_1,T_2)$ and bond prices $P(t,T_1)$ and $P(t,T_2)$

in Fig. 22.3. A set of transactions at time $t$ involving zero net cashflow has the net effect of investing $P(t,T_2)/P(t,T_1)$ dollars at time $T_1$ to yield a dollar for sure at time $T_2$. The forward rate defined above is simply the continuously compounded interest rate earned on this investment. Note that calculation of $f(t,T_1,T_2)$ involves only bond prices observable at time $t$.

The Heath–Jarrow–Morton model to be discussed in a later chapter focuses on the instantaneous forward rate defined by[2]

$$f(t,T) \equiv f(t,T,T), \tag{22.7}$$

[2]There is a slight abuse of notation in that we use the symbol $f$ to denote both the three variable function $f(t,T_1,T_2)$ and the two variable function $f(t,T)$.

and is the instantaneous rate of return the bond holder can earn by extending the investment an instant beyond $T$. Letting $T_2 \to T_1$, in (22.6) and applying L'Hôpital's rule we see that[3]

$$f(t,T) = -\frac{P_T(t,T)}{P(t,T)}. \tag{22.8}$$

It follows from (22.8) that the price of the pure discount bond may be written in terms of the forward rate as

$$P(t,T) = e^{-\int_t^T f(t,s)ds}. \tag{22.9}$$

In a world of certainty all securities, in equilibrium, must earn the same instantaneous rate of return so as to exclude the possibility of riskless arbitrage opportunities. The application of this equilibrium condition to discount bonds implies

$$\frac{P_t(t,T)}{P(t,T)} = r(t), \tag{22.10}$$

from which[4] we obtain the relationship between pure discount bond prices and the instantaneous spot rate in a world of certainty,

$$P(t,T) = e^{-\int_t^T r(s)ds}. \tag{22.11}$$

A comparison of (22.11) and (22.2) shows the relationship between the yield to maturity and the spot rate in a world of certainty,

$$\rho(t,T) = \frac{1}{T-t}\int_t^T r(s)ds. \tag{22.12}$$

---

[3]It is best to set $T_1 = T$, $T_2 = T + x$ in Eq. (22.6) and calculate as follows:

$$\begin{aligned} f(t,T) &= \lim_{x\to 0} \frac{1}{x}[\ln P(t,T) - \ln P(t,T+x)] = \lim_{x\to 0} -\frac{\frac{\partial}{\partial x}\ln P(t,T+x)}{\frac{\partial}{\partial x}x} \\ &= \lim_{x\to 0} \frac{-P_T(t,T+x)}{P(t,T+x)} = \frac{-P_T(t,T)}{P(t,T)}. \end{aligned}$$

[4]From Eq. (22.10)

$$\frac{d}{ds}\ln P(s,T) = r(s).$$

Integrating $(t,T)$ we obtain

$$\ln[P(s,T)]_t^T = \ln P(T,T) - \ln P(t,T) = \ln 1 - \ln P(t,T) = -\ln P(t,T) = \int_t^T r(s)ds.$$

Hence (22.11) follows.

Next, substituting (22.11) into (22.8) reveals the relationship between the spot rate and the forward rate in a world of certainty viz.

$$f(t,T) = r(T), \qquad \text{for all} \quad T \geq t. \tag{22.13}$$

Note that this last equation is a degenerate version of the expectations hypothesis, i.e., the expected instantaneous spot rate for time $T$ is equal to the instantaneous forward rate for time $T$, calculated at time $t$.

In subsequent chapters we shall see how the relationships (22.11)–(22.13) generalise quite naturally in a world of uncertainty to be the corresponding relationships under suitable probability measures.

Another important set of interest rates are LIBOR rates. They are akin to the forward rates $f(t,T_1,T_2)$ in that they are rates that one can contract at time $t$ for borrowing over a fixed period in the future. However the time difference between current time and that fixed period in the future remains constant. Whereas with $f(t,T_1,T_2)$ the time difference between $t$ and $T_1$ decreases as time evolves, in fact it would make no sense to consider $t$ beyond $T_1$. We shall discuss LIBOR rates in Chap. 26 when we discuss the Brace–Gatarek–Musiela (BGM) LIBOR market model.

## 22.2 Modelling the Spot Interest Rate

First we consider models for the dynamics of the spot interest rate. A number of such models have been proposed and most of these are of the general form

$$dr = \mu(r,t)dt + \sigma(r,t)dW. \tag{22.14}$$

Typically the drift term is of the form

$$\mu(r,t) = \kappa(\overline{r} - r(t)), \tag{22.15}$$

where $\overline{r}$ is the long run level of the spot rate of interest. This form implies mean reverting behaviour of the spot interest rate which is confirmed in the empirical studies discussed below. It is less clear what form the volatility function should take. Forms usually used in empirical studies and in development of term structure and interest rate derivative models to be discussed later are of the general form

$$\sigma(r,t) = \sigma r(t)^{\gamma}, \tag{22.16}$$

where usually $\gamma \geq 0$ is assumed.

With these drift and diffusion terms the interest rate process (22.14) assumes the form

$$dr = \kappa(\bar{r} - r(t))dt + \sigma r(t)^{\gamma} dW. \tag{22.17}$$

Chan et al. (1992) have estimated a discretised version of this model on US Treasury bill data for the period 1964–1989 and found (in terms of our notation) $\kappa = 0.5921$, $\bar{r} = 0.0689$, $\sigma^2 = 1.6704$ and $\gamma = 1.4999$. The value of $\sigma^2$ may seem high but it should be borne in mind that the average volatility is measured by $\sigma\bar{r}^{\gamma} = 0.0234 (= 2.34\,\%$p.a.) which is a reasonable value for interest rate markets. Noting the interpretation that $1/\kappa$ is the average time for reversion back to the mean we see that the estimated value of $\kappa = 0.5921$ implies that this average time is about 1.69 years. This value also looks reasonable.

In Fig. 22.4 we illustrate simulations of $r(t)$ given by (22.17) for varying values of $\gamma$ which use the same sequence of random numbers. In these simulations we have used the values $\kappa = 0.6, \bar{r} = 0.07$ [which are close to those found by Chan et al. (1992)] and the values of $\sigma$ displayed in Table 22.1. The values have been chosen so that the overall volatility term at $r = \bar{r}$, viz. $\sigma\ \bar{r}^{\gamma}$ remains constant at the value

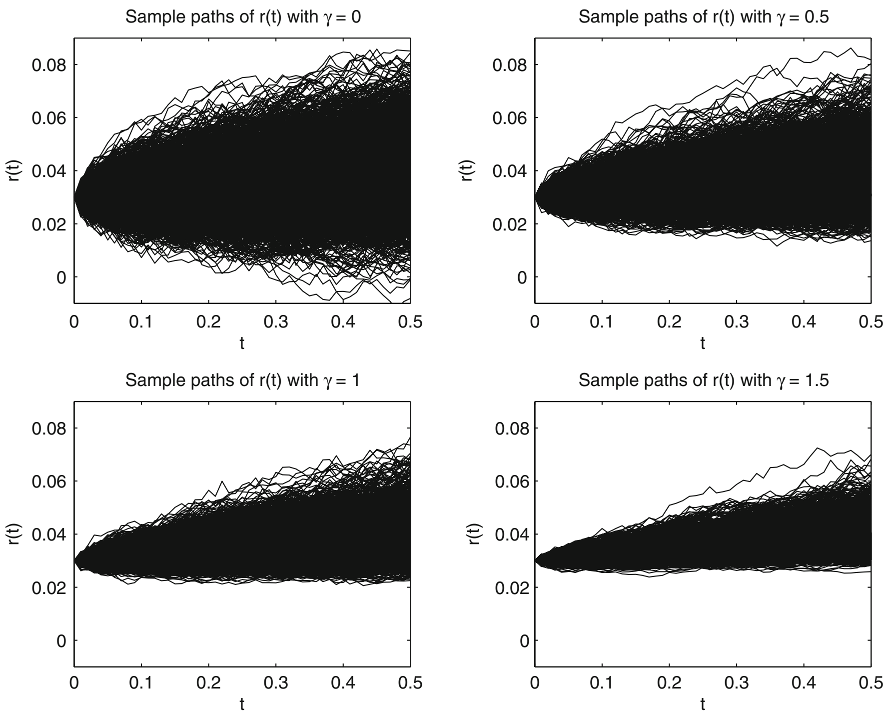


**Fig. 22.4** Bundles of simulated paths for the interest rate process (22.17) for various values of $\gamma$. The parameters used are calibrated to the estimates of Chan et al. (1992)

**Table 22.1** Values of $\sigma$ and $\gamma$ used in simulations of the interest rate process; here $\sigma \bar{r}^{\gamma} = 2.34\,\%$

| $\gamma$ | 0 | 0.5 | 1 | 1.5 |
|---|---|---|---|---|
| $\sigma$ | 0.0234 | 0.08844 | 0.33429 | 1.26348 |

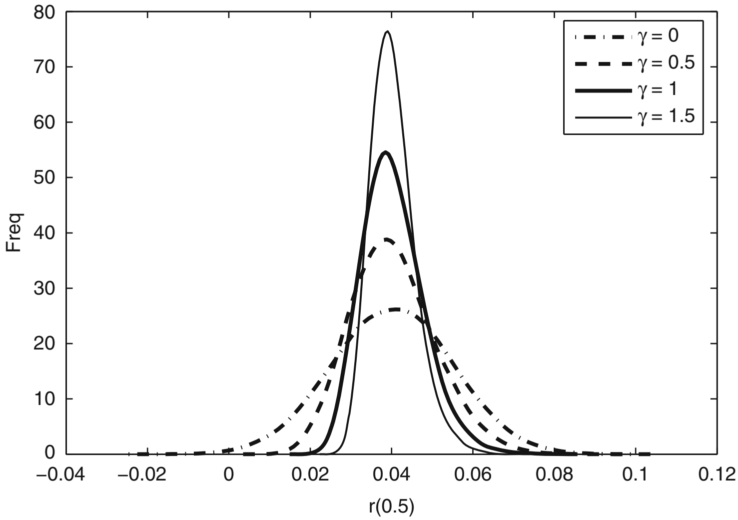


**Fig. 22.5** The density functions for $r$ at $t = 0.05$ corresponding to the simulations in Fig. 22.4

0.0234 as $\gamma$ varies. We have used the initial value, $r_0 = 0.03$. It is of interest to note that when $\gamma = 0$, interest rates can become negative though the probability of this event is very low. As the value of $\gamma$ increases we can see that the bundle of interest rate paths moves away from zero and the distribution becomes more peaked, as can also be seen from Figs. 22.4 and 22.5.

It can be shown that for $\gamma \geq 1/2$ the probability of interest rates becoming negative tends to zero. Hence a commonly used form of the volatility function is (22.16) with $\gamma = 1/2$ and this is the basis of the Cox–Ingersoll–Ross model to be discussed in a later chapter. Stochastic processes having a linear drift term such as (22.15) and a square root volatility have been extensively investigated by Feller (1951). He showed that the conditional transition density function for Eq. (22.14) with

$$\mu(r,t) = \kappa(\bar{r} - r(t)), \qquad \sigma(r,t) = \sigma\sqrt{r(t)}, \tag{22.18}$$

where $\kappa, \bar{r}$ and $\sigma$ are positive constants, is given by

$$p(r,t|r_0,0) = g_t \sum_{n=0}^{\infty} \frac{e^{-\lambda_t/2}(\lambda_t/2)^n(1/2)^{n+d/2}}{n!\Gamma(n+d/2)} (g_t r)^{n-1+d/2} e^{-(g_t r)/2}$$
$$\equiv g_t P_{\chi^2(d,\lambda_t)}(g_t r), \tag{22.19}$$

where $P_{\chi^2(d,\alpha)}(x)$ is the density function of a non-central Chi-squared distribution with $d$ degrees of freedom and non-centrality parameter $\alpha$. Alternatively, the density function may be expressed as (which is really the form given by Feller)

$$p(r,t|r_0,0) = g_t \frac{e^{-(g_t r+\lambda_t)}(g_t r)^{(d-1)/2}\sqrt{\lambda_t}}{2(\lambda_t g_t r)^{d/4}} I_{d/2-1}(\sqrt{\lambda_t g_t r}), \tag{22.20}$$

where the modified Bessel function of the first kind $I_\alpha(x)$ is defined as

$$I_\alpha(x) = \sum_{n=0}^{\infty} \frac{(x/2)^{2n+\alpha}}{n!\Gamma(n+\alpha+1)}.$$

In the above formula the parameters $d, g_t$ and $\lambda_t$ are defined as

$$d = 4\kappa\bar{r}/\sigma^2,$$
$$g_t = \frac{4\kappa}{\sigma^2(1-e^{-\kappa t})},$$
$$\lambda_t = \frac{4\kappa r_0}{\sigma^2(e^{\kappa t}-1)}.$$

We need the condition $2\kappa\bar{r} > \sigma^2$ to ensure that $r(t)$ is positive. A plot of $p(r,t|r_0,0)$ for a range of values of $\sigma$ and the values of $\kappa, \bar{r}$ indicated earlier is displayed in Fig. 22.6. The process having drift and diffusion terms as specified in (22.18) has become known variously as the square root process, the Cox–Ingersoll–Ross process and the Feller process. We shall use the latter term as it is the results obtained by Feller in the cited paper that are used in the quantitative finance literature. In Sect. 22.3 we shall give some motivation for the Feller process that will give some insight into why the underlying distribution is non-central chi squared.

An important criticism of modelling the volatility by use of functional forms such as (22.16) is that volatility of interest rates depend on the level of interest rates. This seems to run against what is often observed as it is possible to identify historical periods when rates were high but relatively stable or rates were low but fairly volatile. It is also often argued that volatility of interest rates should also be a function of the news arrival process (i.e. the $dW$'s). One way to capture this effect would be to allow $\sigma$ to itself follow a stochastic process—in this regard see Problem 22.8.

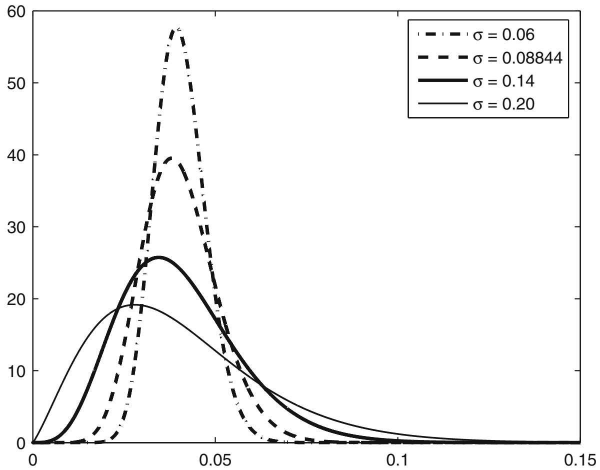


**Fig. 22.6** The transition density function of the Feller process for different values of $\sigma$. The remaining parameters of the process correspond to those used for the simulations in Fig. 22.4

## 22.3 Motivating the Feller (or Square Root) Process

A very widely used stochastic process in finance is the Feller or square root process, frequently referred to as the Cox, Ingersoll and Ross (or just CIR) process. As we have stated earlier this process was originally studied by Feller (1951), but introduced into the finance literature by Cox et al. (1985b). Its interest for finance applications lies in the fact that (for certain parameter constellations) the process is guaranteed to generate positive values and its density function is known. This is clearly a desirable feature for a model of interest rate processes, but also in other areas of finance e.g. models of credit spreads or risk-premia. In this section we seek to give some insight into why this process generates positive (or at least non-negative outcomes), and also why its conditional transitional probability density function is a chi-squared distribution.

Consider the process (22.14) when $\mu(r,t) = (\alpha - \beta r(t))$ and $\sigma(r,t) = \sigma(t)$, i.e.

$$dr = (\alpha - \beta r(t))dt + \sigma(t)dW(t). \tag{22.21}$$

This is essentially an Ornstein–Uhlenbeck process (see Sect. 6.3.5) whose solution can be written

$$r(t) = r(0)e^{-\beta t} + \int_0^t \alpha e^{-\beta(t-s)}ds + \int_0^t e^{-\beta(t-s)}\sigma(s)dW(s). \tag{22.22}$$

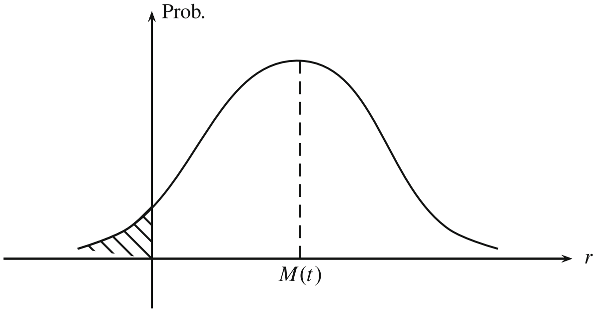


**Fig. 22.7** The possibility of negative interest rates from the distribution in Eq. (22.23)

It follows from Eq. (22.22) that

$$r(t) \sim N(M(t), V^2(t)), \tag{22.23}$$

where

$$M(t) = r_0 e^{-\beta t} + \int_0^t \alpha e^{-\beta(t-s)} ds,$$

and

$$\begin{aligned} V^2(t) &= E_0\left[\left(\int_0^t e^{-\beta(t-s)}\sigma(s)dW(s)\right)^2\right] \\ &= \int_0^t e^{-2\beta(t-s)}\sigma^2(s)ds. \end{aligned}$$

The fact that $r(t)$ is normally distributed means that there is a positive probability that $r(t)$ can become negative as illustrated in Fig. 22.7. This is not a desirable feature of a process for interest rate dynamics, though in practice this probability may be quite low as can be seen from the simulated distribution in Fig. 22.5 for the value $\gamma = 0$.

We are therefore motivated to find a process that has zero probability of $r(t)$ becoming zero or negative. Consider the set of $u$ (mean reverting to zero) Ornstein–Uhlenbeck processes

$$dX_j = -\frac{1}{2}\beta X_j(t)dt + \frac{1}{2}\sigma dW_j, \quad (j = 1, \cdots, u). \tag{22.24}$$

By solving (22.24) we obtain

$$X_j(t) = X_j(0)e^{-\frac{1}{2}\beta t} + \frac{1}{2}\sigma \int_0^t e^{-\frac{\beta}{2}(t-v)} dW_j(v), \tag{22.25}$$

implying that

$$X_j \sim N(M_{X_j}(t), V_X(t)),$$

where

$$M_{X_j}(t) = X_j(0)e^{-\frac{1}{2}\beta t},$$

and

$$V_X(t) = \frac{\sigma^2}{4}\int_0^t e^{-\beta(t-v)}dv.$$

Consider the process

$$r(t) = X_1^2(t) + X_2^2(t) + \cdots + X_u^2(t).$$

Ito's lemma for a multi-variable function and some algebraic manipulations imply that

$$dr = \left(\frac{u\sigma^2}{4} - \beta r(t)\right)dt + \sigma\sqrt{r(t)}\left(\sum_{i=1}^{u}\frac{X_i(t)}{\sqrt{r(t)}}dW_i\right). \tag{22.26}$$

Consider the new process

$$W(t) = \sum_{i=1}^{u}\int_0^t \frac{X_i(v)}{\sqrt{r(v)}}dW_i(v), \tag{22.27}$$

so that

$$dW(t) = \sum_{i=1}^{u}\frac{X_i(t)}{\sqrt{r(t)}}dW_i(t).$$

We calculate[5]

$$E[dW(t)] = 0,$$

and

$$var[dW(t)] = \sum_{i=1}^{u}\frac{X_i^2(t)}{r(t)}E[dW_i^2] = \left(\sum_{i=1}^{u}X_i^2(t)\right)\frac{dt}{r(t)} = dt.$$

[5] Recall that $E(dW_i dW_j) = 0 \quad (i \neq j) \quad (j = 1, 2, \cdots, d)$.

Furthermore we easily calculate that

$$\mathbb{E}[(dW(t))^m] = 0 \text{ for } m \geq 3.$$

So $dW(t)$ has all the statistical properties of a Wiener increment. It thus follows that $W(t)$ is a Wiener process. So we can write (22.26) as

$$dr = \left(\frac{u\sigma^2}{4} - \beta r(t)\right) dt + \sigma\sqrt{r(t)}dW, \tag{22.28}$$

which enjoys the property $r(t) \geq 0$ by construction. The process (22.28) is simply the Feller process already encountered in Sect. 22.2. It is more simply written here as

$$dr = (\alpha - \beta r(t))dt + \sigma\sqrt{r(t)}dW, \tag{22.29}$$

where we choose $u$ as

$$u = \frac{4\alpha}{\sigma^2} > 0. \tag{22.30}$$

Now, for arbitrary $\alpha$ and $\sigma$, $u$ is *not* an integer. It can be shown that results for the CIR process with $u$ equal to an integer are also true for $u$ *not* an integer. The value of $u$ for the process affects very much its behaviour. We distinguish two cases:

(i) $u < 2$

We note that $u < 2$ implies $\alpha < \sigma^2/2$. In this case it can be shown that

$$\Pr\left\{\begin{array}{c} r(t) = 0 \text{ at infinite number} \\ \text{of values of } t \end{array}\right\} = 1.$$

This implies behaviour shown in Fig. 22.8. The excursions at $r(t) = 0$ are also not a desirable feature of a process for interest rate dynamics since it does not accord with what we observed empirically.

(ii) $u > 2$

In this case $\alpha \geq \sigma^2/2$ and it can be shown that

$$\Pr\{\exists \text{ at least one } t > 0 \text{ s.t. } r(t) = 0) = 0$$

which is the desirable feature of the process for $r$ that we seek. Figure 22.9 illustrates what a typical sample path would look like.

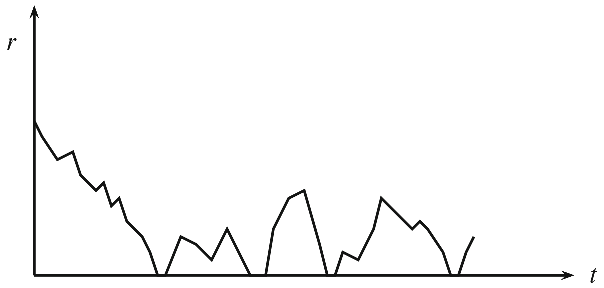


**Fig. 22.8** The Feller process for $r(t)$ with $\alpha < \sigma^2/2$

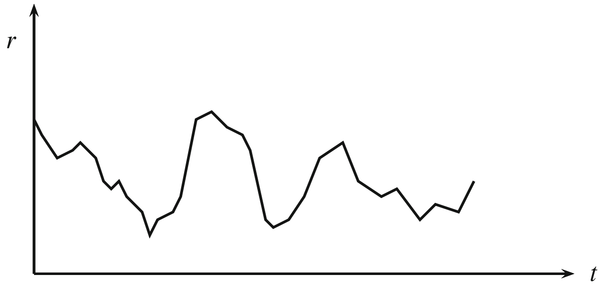


**Fig. 22.9** The Feller process for $r(t)$ with $\alpha > \sigma^2/2$

In order to give an intuitive feel for the result that the distribution for $r(t)$ is a non-central Chi-squared distribution, we set

$$\begin{aligned} X_1(0) &= X_2(0) = \cdots X_{u-1}(0) = 0 \\ X_u(0) &= \sqrt{r(0)}. \end{aligned}$$

Then from (22.25) it follows that

$$X_i(t) \sim N(0, V_X^2(t)) \quad i = 1, \cdots, u-1,$$

and so

$$X_u(t) \sim N(e^{-\frac{1}{2}\beta t}\sqrt{r(0)}, V_X^2(t)).$$

Thus

$$r(t) = V_X^2(t) \sum_{i=1}^{u-1} \left( \frac{X_i(t)}{V_X(t)} \right)^2 + X_u^2(t).$$

Noting that $X_u^2(t)$ is the square of a normally distributed variable with non-zero mean and that each of the $(u-1)$ terms in the sum on the RHS is the square of a (zero-mean) normally distributed variable, it follows that $r(t)$ has a non-central Chi-squared distribution with $u-1 = \frac{4\alpha-\sigma^2}{\sigma^2}$ degrees of freedom. The actual density function is derived at Eq. (22.20).

## 22.4 Fubini's Theorem

In the next section we shall consider models of the forward rate and shall use these to derive corresponding processes for the bond price and spot interest rate. In this analysis we shall frequently encounter stochastic double integrals of the form

$$\int \left( \int g(u, v, \omega) dW(u) \right) dv, \tag{22.31}$$

with a variety of different limits in the integrals. Here $\omega$ can be any stochastic variable such as the forward rate, the bond price or the spot interest rate. We will very often need to calculate the stochastic differential of terms like Eq. (22.31) and this is more simply done if we can express it in the form

$$\int \left( \int g(u, v, \omega) dv \right) dW(u). \tag{22.32}$$

In other words we need to change the order of integration in stochastic double integrals. The result that allows us to do this under fairly mild conditions on $g$ is Fubini's theorem. Essentially this result allows us to manipulate stochastic double integrals in the same way that we manipulate ordinary double integrals. Three main results may be derived from Fubini's theorem, these are:

(I)

$$\int_0^\tau \left( \int_0^t g(u, v, \omega) dW(u) \right) dv = \int_0^t \left( \int_0^\tau g(u, v, \omega) dv \right) dW(u), \tag{22.33}$$

for all $0 \leq t \leq \tau$.

(II)

$$\int_t^T \left( \int_0^\tau g(u, v, \omega) dW(u) \right) dv = \int_0^\tau \left( \int_t^T g(u, v, \omega) dv \right) dW(u), \tag{22.34}$$

for all $0 \leq \tau \leq t \leq T$.

In these two results the integration limits do not involve the integration variables $u$ or $v$. These results state that in this case we may change the order of integration as we do for the same situation with ordinary integrals.

(III)

$$\int_0^t \left( \int_0^v g(u,v,\omega) dW(u) \right) dv = \int_0^t \left( \int_u^t g(u,v,\omega) dv \right) dW(u). \tag{22.35}$$

Now the integration limits depend on the integration variables, however the result again shows that we can handle such stochastic double integrals in the same way that we handle ordinary double integration. Figures 22.10, 22.11 and 22.12 illustrate the relevant regions of integration for cases I, II and III respectively. In each case the region of integration (the shaded area) can be swept out by left to right motion of

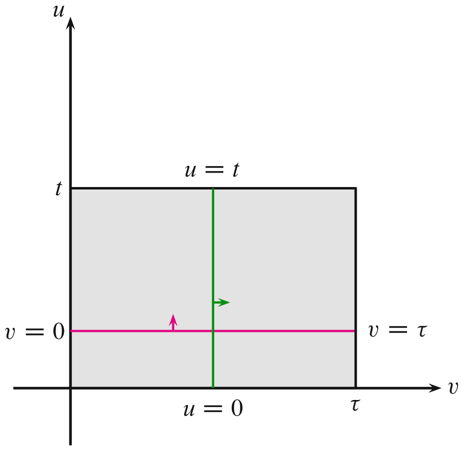


**Fig. 22.10** Change of limits for version I of Fubini's theorem

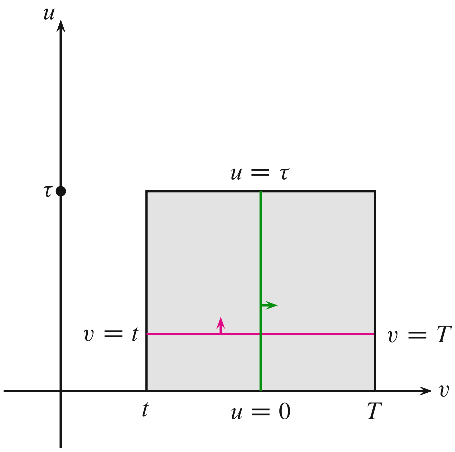


**Fig. 22.11** Change of limits for version II of Fubini's theorem

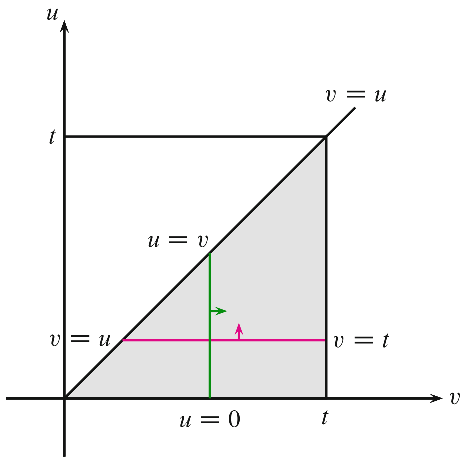


**Fig. 22.12** Change of limits for version III of Fubini's theorem

the vertical line (the LHS double integrals) or, from top to bottom motion of the horizontal line (the RHS double integrals).

## 22.5 Modelling Forward Rates

We recall that $f(t,T)$ denotes the forward rate at time $t$ for instantaneous borrowing at time $T$. The time line for $f(t,T)$ is illustrated in Fig. 22.13, where we display $f(t,T)$ and $f(t+dt,T)$. The important point to keep in mind is that the maturity date $T$ is fixed, as so the time to maturity decreases as $t$ evolves. This is in contrast to the Brace–Musiela notation that considers fixed period ahead forward rates, as we shall see in Chap. 26 when we come to consider the LIBOR market model.

Heath et al. (1992a) propose to model the forward rate as the driving stochastic process. They write the process for $f(t,T)$ in the form of a stochastic integral equation as

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T,\omega(u))du + \int_0^t \sigma(u,T,\omega(u))dW(u), \qquad (22.36)$$

for $0 \leq t \leq T$.

Figure 22.14 illustrates how the forward curve at time $(t+dt)$ evolves from that at $t$ by the addition of two terms. The first, $\alpha(t,T,\cdot)dt$, represents the average change across the maturities. The second, $\sigma(t,T,\cdot)dW$, represents the impact at different maturities of the shock $dW$ that occurs during $(t, t+dt)$.

More specifically, the quantities $\alpha(u,T,\omega(u))$ and $\sigma(u,T,\omega(u))$ are the drift and volatility of the forward rate process and $f(0,T)$ is the initial forward rate curve which can be obtained from the currently observed yield curve. We allow for possible dependence of $\alpha$ and $\sigma$ on $\omega(t)$, a vector of path dependent variables such

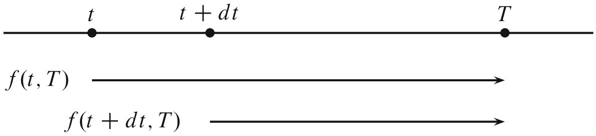


**Fig. 22.13** The time line for the fixed terminal data instantaneous forward rate $f(t, T)$

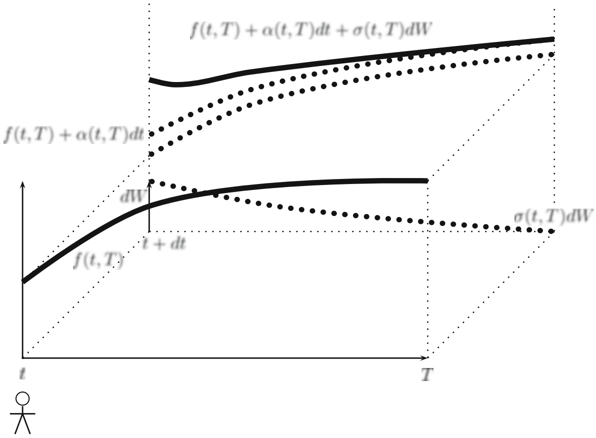


**Fig. 22.14** The evolution of the forward rate curve

as the instantaneous spot interest rate $r(t)$, discrete maturity forward rates or even the entire forward curve itself. This would allow the model to capture the effect of the level of interest rates on volatility. This general notation was indeed employed in the original Heath–Jarrow–Morton paper. In the discussion of this section we only allow one shock term $dW$ in order to alleviate the mathematical notation. Heath–Jarrow–Morton allow for multiple shock terms and we consider this case in Chap. 25. In differential form we may write Eq. (22.36) as the stochastic differential equation

$$df(t, T) = \alpha(t, T, \omega(t))dt + \sigma(t, T, \omega(t))dW(t). \tag{22.37}$$

We shall defer for the moment discussion of appropriate functional forms for $\alpha(t, T, \omega(t))$ and $\sigma(t, T, \omega(t))$. The aim of the calculations in the rest of this section is to determine the stochastic processes for the spot rate $r(t)$ and bond price $P(t, T)$ implied by Eq. (22.36).

The reader needs to be wary that the correct path from the forward rate dynamics to the spot rate dynamics has to occur at the level of the stochastic integral equations. Thus we must set $T = t$ in the stochastic integral equation (22.36) for $f(t, T)$ to obtain the stochastic integral equation for $r(t)$. For this latter we can obtain the stochastic differential equation for $r(t)$. It is not correct to obtain the stochastic differential equation for $r(t)$ simply by setting $T = t$ in the stochastic differential

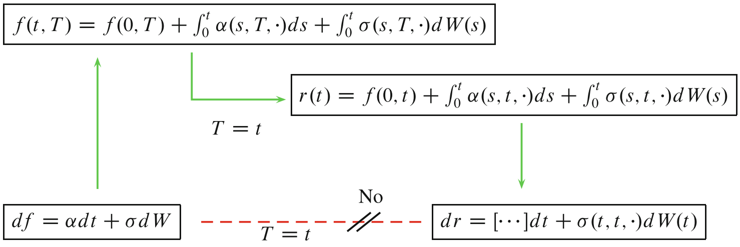


**Fig. 22.15** The path from forward rate dynamics to spot rate dynamics

equation (22.37) for $f(t,T)$. This has to do with the fact that the path history matters here and we need to work with the stochastic integral equations, which are properly defined mathematical objects. The stochastic differential equation is merely a (very convenient) short-hand notation as we stressed in Chap. 4. We illustrate the correct path from forward rate dynamics to spot rate dynamics in Fig. 22.15.

Recalling that the spot interest rate and forward rate are related by $r(t) = f(t,t)$ we see from Eq. (22.36) that $r(t)$ satisfies the stochastic integral equation

$$r(t) = f(0,t) + \int_0^t \alpha(u,t,\omega(u))du + \int_0^t \sigma(u,t,\omega(u))dW(u). \tag{22.38}$$

This may alternatively be expressed as the stochastic differential equation[6]

$$dr = \left[ f_2(0,t) + \alpha(t,t,\omega(t)) + \int_0^t \alpha_2(u,t,\omega(u))du + \int_0^t \sigma_2(u,t,\omega(u))dW(u) \right] dt \\ + \sigma(t,t,\omega(t))dW(t). \tag{22.39}$$

Here $f_2, \alpha_2, \sigma_2$ denote partial differentiation of $f, \alpha, \sigma$ with respect to their second arguments. Close inspection of (22.39) reveals that this process for $r(t)$ is more general than the types of processes considered in Sect. 22.2. The difference stems from the third component of the drift term, viz.,

$$\int_0^t \sigma_2(u,t,\omega(u))dW(u). \tag{22.40}$$

This term is a weighted sum of the shock terms $dW(u)$ from initial time 0 to current time $t$, thus it is a path-dependent term. Hence the stochastic differential equation

[6] Here we make use of the result for the differential of a stochastic integral in Sect. 5.5, a result that is used frequently in this and subsequent chapters.

for $r(t)$, Eq. (22.39) is non-Markovian.[7] Thus the spot interest rate process implied by the forward rate process (22.36) is far more general than those discussed in Sect. 22.2.

### 22.5.1 *From Forward Rate to Bond Price Dynamics*

In this section we determine the bond price dynamics implied by the forward rate dynamics (22.36). Before proceeding we compute a quantity that will be useful in simplifying the expression for the bond price that we shall calculate below. We note from Eq. (22.38) that

$$\int_0^t r(s)ds = \int_0^t f(0,s)ds + \int_0^t \int_0^s \alpha(u,s,\omega(u))duds + \int_0^t \int_0^s \sigma(u,s,\omega(u))dW(u)ds,$$

which by use of result (III) of Fubini' theorem in Sect. 22.4 can be written

$$\begin{aligned}\int_0^t r(s)ds = \int_0^t f(0,s)ds + \int_0^t \left(\int_u^t \alpha(u,s,\omega(u))ds\right) du \\ + \int_0^t \left(\int_u^t \sigma(u,s,\omega(u))ds\right) dW(u).\end{aligned} \tag{22.41}$$

In order to determine the stochastic differential equation for the bond price that is implied by the forward rate process (22.36), we recall first the definitional relationship between bond prices and forward rates, viz.,

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right),$$

or

$$\ln P(t,T) = -\int_t^T f(t,s)ds. \tag{22.42}$$

Substituting Eq. (22.36) into this last equation we have that

$$\begin{aligned}&\ln P(t,T)\\ &= -\int_t^T f(0,s)ds - \int_t^T \int_0^t \alpha(u,s,\omega(u))duds - \int_t^T \int_0^t \sigma(u,s,\omega(u))dW(u)ds.\end{aligned}$$

[7] In fact even without the term Eq. (22.40) we have non-Markovian dynamics because of the term $\int_0^t \alpha_2(u,t,\omega(u))du$, which depends on the path followed by the path dependent variable $\omega$ between 0 and $t$.

Applying result (II) of Fubini' theorem (see Sect. 22.4) we can re-express $\ln P(t,T)$ as

$$\ln P(t,T) = -\int_t^T f(0,s)ds - \int_0^t \left(\int_t^T \alpha(u,s,\omega(u))ds\right)du - \int_0^t \left(\int_t^T \sigma(u,s,\omega(u))ds\right)dW(u),$$

which may further be re-arranged to[8]

$$\begin{aligned}\ln P(t,T) = &-\int_0^T f(0,s)ds - \int_0^t \left(\int_u^T \alpha(u,s,\omega(u))ds\right)du \\ &- \int_0^t \left(\int_u^T \sigma(u,s,\omega(u))ds\right)dW(u) + \int_0^t f(0,s)ds \\ &+ \int_0^t \left(\int_u^t \alpha(u,s,\omega(u))ds\right)du + \int_0^t \left(\int_u^t \sigma(u,s,\omega(u))ds\right)dW(u).\end{aligned} \tag{22.43}$$

From Eq. (22.41) we see that we can represent the last three terms of (22.43) as $\int_0^t r(s)ds$, and from Eq. (22.42) we see that

$$-\int_0^T f(0,s)ds = \ln P(0,T).$$

Hence Eq. (22.43) simplifies to

$$\begin{aligned}\ln P(t,T) = \ln P(0,T) + \int_0^t r(s)ds - \int_0^t \left(\int_u^T \alpha(u,s,\omega(u))ds\right)du \\ - \int_0^t \left(\int_u^T \sigma(u,s,\omega(u))ds\right)dW(u).\end{aligned} \tag{22.44}$$

Taking differentials it follows immediately that the log bond price $B(t,T) \equiv \ln P(t,T)$ satisfies the stochastic differential equation

$$dB(t,T) = [r(t) - \alpha_B(t,T)]\,dt + \sigma_B(t,T)dW(t), \tag{22.45}$$

where we define

$$\alpha_B(t,T) = \int_t^T \alpha(t,s,\omega(t))ds \quad \text{and} \quad \sigma_B(t,T) = -\int_t^T \sigma(t,s,\omega(t))ds. \tag{22.46}$$

[8]Since $\int_t^T f(0,s)ds = \int_0^T f(0,s)ds - \int_0^t f(0,s)ds$ and for any function $g$ we have $\int_0^t \left(\int_t^T g(u,s)ds\right)du = \int_0^t \left(\int_t^u g(u,s)ds + \int_u^T g(u,s)ds\right)du = \int_0^t \left(-\int_u^t g(u,s)ds + \int_u^T g(u,s)\, ds\right)du$.

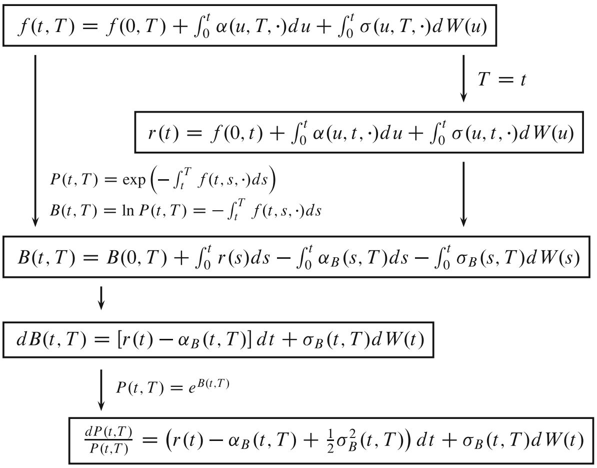


**Fig. 22.16** The steps from forward rate to bond price dynamics

By a straightforward application of Ito's Lemma (since $P(t,T) = e^{B(t,T)}$) we find the corresponding stochastic differential equation for the bond price to be

$$dP(t,T) = \left[r(t) - \alpha_B(t,T) + \frac{1}{2}\sigma_B^2(t,T)\right]P(t,T)dt + \sigma_B(t,T)P(t,T)dW(t). \tag{22.47}$$

The stochastic differential equations (22.39) and (22.47) form a linked system for the instantaneous spot rate and the bond price. This system will typically be non-Markovian and many applications in the Heath–Jarrow–Morton framework seek simplifications which reduce this system to Markovian form. We give one example of how this may be done in the next section.

Since the steps from the forward price dynamics to the bond price dynamics are rather round-about we feel it may be useful to summarise these in Fig. 22.16.

## *22.5.2 A Specific Example*

For the forward rate volatility in Eq. (22.37) we assume that the path dependent variable $\omega$ is associated to the instantaneous spot interest rate $r(t)$ in the functional form

$$\sigma(u,t,\omega(u)) = \sigma(u,t,r(u)) = \sigma e^{-\lambda(t-u)}r(u)^{\gamma}. \tag{22.48}$$

Note that in this case

$$\sigma_2(u,t,r(u)) = -\lambda\sigma e^{-\lambda(t-u)}r(u)^\gamma = -\lambda\sigma(u,t,r(u)), \tag{22.49}$$

so that the non-Markovian term in Eq. (22.39) may be written

$$\begin{aligned}\int_0^t \sigma_2(u,t,r(u))dW(u) &= -\lambda\int_0^t \sigma(u,t,r(u))dW(u)\\ &= -\lambda\int_0^t \sigma e^{-\lambda(t-u)}r(u)^\gamma dW(u).\end{aligned} \tag{22.50}$$

Define the subsidiary stochastic variable

$$\phi(t) = \int_0^t \sigma e^{-\lambda(t-u)}r(u)^\gamma dW(u), \tag{22.51}$$

and note that it satisfies the stochastic differential equation

$$d\phi = -\lambda\phi(t)dt + \sigma r(t)^\gamma dW(t). \tag{22.52}$$

Using $\phi(t)$ to substitute the non-Markovian term in Eq. (22.39) we see that $r(t)$ and $\phi(t)$ form a linked *Markovian* system[9] viz.

$$dr = \left[f_2(0,t) + \alpha(t,t,r(t)) + \int_0^t \alpha_2(u,t,r(u))du - \lambda\phi(t)\right]dt + \sigma r(t)^\gamma dW(t), \tag{22.53}$$

$$d\phi = -\lambda\phi(t)dt + \sigma r(t)^\gamma dW(t). \tag{22.54}$$

In this system the past levels of volatility and the news arrival process affect the dynamics of $r(t)$ via the variable $\phi(t)$. This captures the type of processes estimated by Brenner et al. (1996) though in a different way to their specification. Here the drift rather than the volatility is affected directly by the $\phi(t)$ process.

In Fig. 22.17 we illustrate some simulations of $r(t)$ and $\phi(t)$ in the system (22.53)–(22.54) for $\gamma = 0.5$. For the drift term we have taken

$$\alpha(t,T,r(t)) = \sigma(t,T,r(t))\int_t^T \sigma(t,u,r(t))du, \tag{22.55}$$

[9] We are assuming here that the term $\int_0^t \alpha_2(u,t,r(u))du$ can be expressed in terms of the state variable, $r(t)$ and $\phi(t)$. This will be the case if we assume the form Eq. (22.55) below that ensures arbitrage free dynamics. Otherwise it may be necessary to introduce further state variables to obtain a Markovian system.

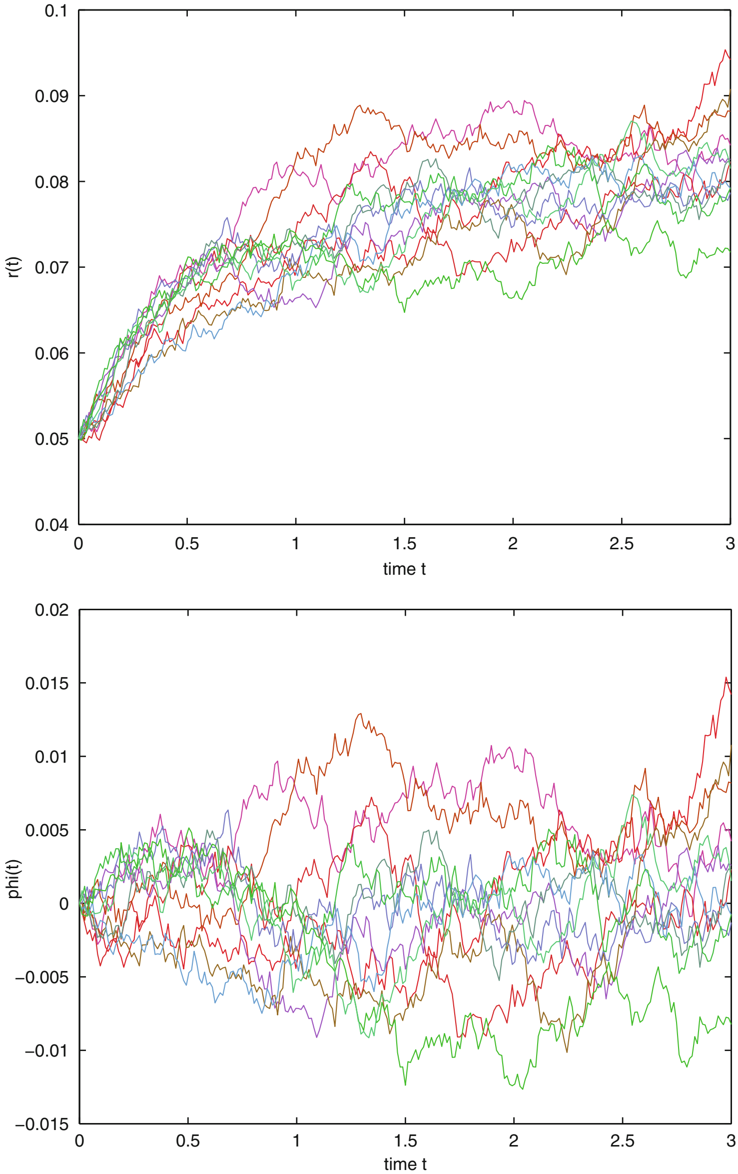


**Fig. 22.17** Simulations of $r(t)$ and $\phi(t)$ for (22.53)–(22.54) with drift specified by (22.55)

which ensures that the economy is arbitrage free as we shall show in Chap. 25. The values of $\sigma, \lambda$ were taken from Bhar and Chiarella (1997b). The initial forward rate curve $f(0,t)$ is relevant to 90-day bank bill futures data traded on the SFE in 1991, and was calculated using the polynomial fitting procedure described in Bhar and Hunt (1993).

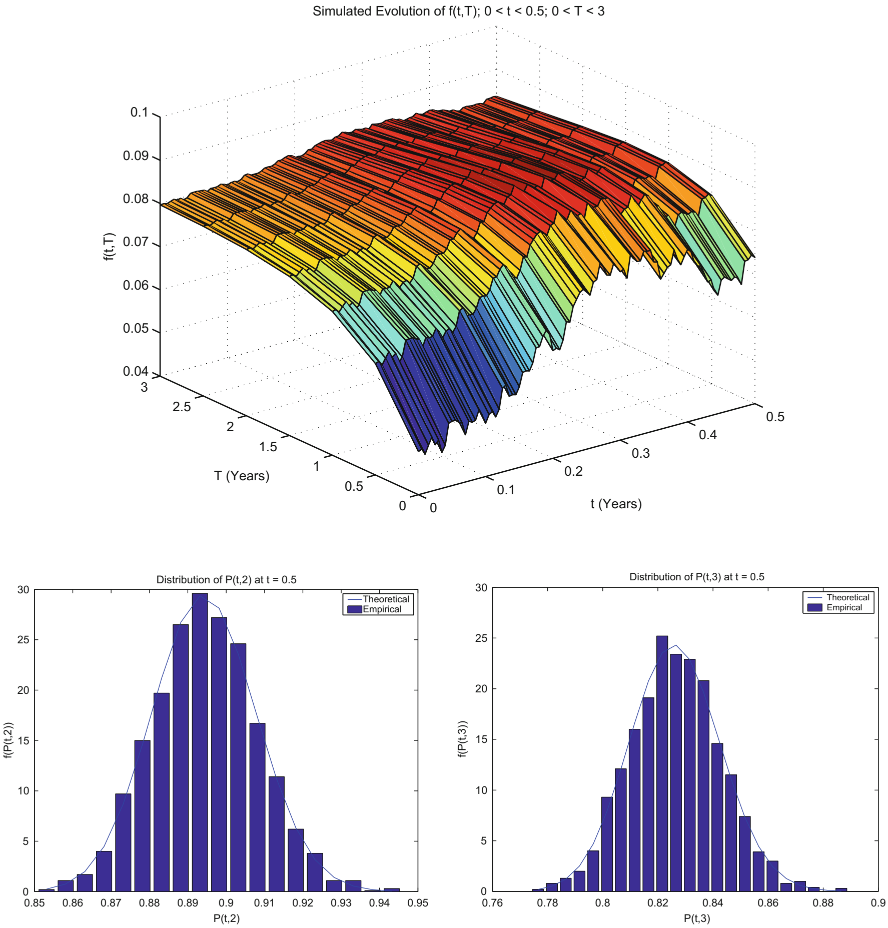


**Fig. 22.18** Simulations of the process for forward rates and simulated distribution density of bond prices

In order to give some intuitive feel for the type of dynamics implied by the forward rate process, we also display in Fig. 22.18 simulations of the forward rate process and the simulated distribution density of the corresponding bond prices $P(t, 2)$ and $P(t, 3)$ when $t = 0.5$, for the same values of $\gamma$.

## 22.6 Appendix

### Appendix 22.1 Calculation of Covariance

Consider the interest rate process equation (22.38) in the case of a deterministic drift and volatility so that

$$r(t) = f(0,t) + \int_0^t \alpha(u,t)du + \int_0^t \sigma(u,t)dW(u).$$

We can readily calculate that

$$m_r(t) = \mathbb{E}[r(t)] = f(0,t) + \int_0^t \alpha(u,t)du.$$

We can then calculate the covariance between interest rates at two maturities $s$ and $t$ as

$$\begin{aligned} cov(r(s), r(t)) &= E_0[\{r(s) - m_r(s)\}\{r(t) - m_r(t)\}] \\ &= E_0\left[\int_0^s \sigma(u,s)dW(u) \int_0^t \sigma(u,t)dW(u)\right]. \end{aligned}$$

For definiteness suppose $s < t$

$$\begin{aligned} & E_0\left[\int_0^s \sigma(u,s)dW(u) \int_0^t \sigma(u,t)dW(u)\right] \\ &= E_0\left[\int_0^s \sigma(u,s)dW(u) \int_0^s \sigma(u,t)dW(u) + \underbrace{\int_0^s \sigma(u,s)dW(u) \int_s^t \sigma(u,t)dW(u)}_{\text{these } dW(u) \text{ are non overlapping } \therefore\ E_0 = 0}\right] \\ &= E_0\left[\int_0^s \sigma(u,s)dW(u) \int_0^s \sigma(u,t)dW(u)\right] \\ &= \int_0^s \sigma^2(u,t)du. \end{aligned}$$

In general (i.e. $s < t$ or $s > t$) we have

$$\text{cov}(r(s), r(t)) = \int_0^{\min(s,t)} \sigma^2(u,t)du = \int_0^{s\wedge t} \sigma^2(u,t)du$$

where $s \wedge t \equiv \min(s,t)$.

## 22.7 Problems

**Problem 22.1** Consider the spot rate process

$$dr = (\mu_0 + \mu_1 r(t))dt + \sigma_r dW,$$

where $\mu_0$, $\mu_1$ and $\sigma_r$ are constants. Calculate the expected average value of the spot rate i.e.

$$\mathbb{E}_t\left[\frac{1}{T-t}\int_t^T r(u)du\right].$$

**Problem 22.2** Consider the mean-reverting process

$$dr = \kappa(\theta - r(t))dt + \sigma dW,$$

where $\kappa, \theta$ and $\sigma$ are constants. Show that for $s \geq t$

$$r(s) = e^{-\kappa(s-t)}r(t) + \theta(1 - e^{-\kappa(s-t)}) + \sigma\int_t^s e^{-\kappa(s-u)}dW(u).$$

Then by using the appropriate version of Fubini's theorem, show that

$$\int_t^T r(s)ds = (r(t) - \theta)B(t,T) + \theta(T-t) + \sigma\int_t^T B(u,T)dW(u),$$

where

$$B(t,T) \equiv \frac{1}{\kappa}(1 - e^{-\kappa(T-t)}).$$

**Problem 22.3** In Sect. 22.5 we have discussed how, starting with the dynamics for the forward rate, we can obtain the dynamics for the spot rate and bond price.

Suppose that instead we start with the dynamics for the bond price, i.e. we assume

$$\frac{dP(t,T)}{P(t,T)} = \alpha_P(t,T)dt + \sigma_P(t,T)dz(t).$$

What process does this imply for $f(t,T)$ and $r(t)$?

Suppose we take

$$\sigma_P(t,T) = \bar{\sigma}e^{-\lambda(T-t)}.$$

What will be the corresponding volatility for the forward rate process?

For this volatility function write out the stochastic differential equation for the spot rate $r(t)$.

Is it still possible to write it in Markovian form in this case?

*Hint*: Determine the stochastic integral equation satisfied by $B(t, T) = \ln P(t, T)$. Then use this to determine the stochastic dynamics for $G(t, T) = \frac{\partial}{\partial T} B(t, T)$.

**Problem 22.4** Following on from Problem 22.3, we could also take as our starting point the dynamics of the yield to maturity. That is we assume

$$d\rho(t, T) = \alpha_\rho(t, T)dt + \sigma_\rho(t, T)dz(t).$$

Now determine the stochastic integral equations and stochastic differential equations satisfied by $P(t, T)$, $f(t, T)$ and $r(t)$.

Suppose that we take

$$\sigma_\rho(t, T) = \bar{\sigma} e^{-\lambda(T-t)}$$

what will be the corresponding volatilities for the bond price $P(t, T)$ and forward rate $f(t, T)$?

Write out the stochastic differential equation satisfied by $r(t)$ in this case. Can you express it in Markovian form?

*Hint*: Recall the definition of the yield to maturity

$$\rho(t, T) = -\frac{\ln P(t, T)}{T - t}.$$

Use Ito's Lemma to find the stochastic differential equation satisfied by $B(t, T) = \ln P(t, T)$. A further application of Ito's Lemma gives the stochastic differential equation for $P(t, T)$.

**Problem 22.5** In Sect. 22.5.1 consider the case in which $\sigma$ and $\alpha$ depend on the forward rate itself, i.e. $\sigma(t, T, \omega(t)) = \sigma(t, T, f(t, T))$. Show that then the forward rate, the derivative of the forward rate with respect to maturity time, the spot interest rate and the bond price form the linked stochastic differential system

$$df(t, T) = \alpha(t, T, f(t, T))dt + \sigma(t, T, f(t, T))dW(t), \tag{22.56}$$

$$\begin{aligned} df_2(t, T) = {} & [\alpha_2(t, T, f(t, T)) + \alpha_3(t, T, f(t, T)) f_2(t, T)]dt \\ & + [\sigma_2(t, T, f(t, T)) + \sigma_3(t, T, f(t, T)) f_2(t, T)]dW(t), \end{aligned} \tag{22.57}$$

$$\begin{aligned} dr = {} & \left\{ f_2(0, t) + \alpha(t, t, r(t)) + \int_0^t [\alpha_2(u, t, f(u, t)) + \alpha_3(u, t, f(u, t)) f_2(u, t)]du \right\} dt \\ & + \int_0^t [\sigma_2(u, t, f(u, t)) + \sigma_3(u, t, f(u, t)) f_2(u, t)]\, dW(u)dt + \sigma(t, t, r(t))dW(t), \end{aligned} \tag{22.58}$$

$$dP(t,T) = \left[ r(t) - \alpha_B(t,T,\cdot) + \frac{1}{2}\sigma_B^2(t,T,\cdot) \right] P(t,T)dt + \sigma_B(t,T,\cdot)P(t,T)dz(t), \tag{22.59}$$

where

$$\alpha_B(t,T,\cdot) = \int_t^T \alpha(t,s,f(t,s))ds \quad \text{and} \quad \sigma_B(t,T,\cdot) \equiv -\int_t^T \sigma(t,s,f(t,s))ds. \tag{22.60}$$

Comment on whether this system is Markovian or non-Markovian.

**Problem 22.6** Work through the calculations leading to the stochastic differential equation system (22.56)–(22.59). Obtain the form taken by this stochastic differential equation system when it is assumed that

$$\sigma(t,T,f(t,T)) = \sigma_0 e^{-\lambda(T-t)} f(t,T)^{\gamma}.$$

**Problem 22.7** **Computational Problem**—Simulate the interest rate process

$$dr = \kappa(\bar{r} - r(t))dt + \sigma\, r^{\gamma}\, dW.$$

Use the values that were used to generate Fig. 22.4.

Reproduce Figs. 22.4 and 22.5.

**Problem 22.8** **Computational Problem**—Consider again the interest rate process of Problem 22.7, but now suppose the volatility is stochastic. That is we have

$$dr = \kappa_r(\bar{r} - r(t))dt + \sqrt{v(t)}\, r(t)^{\gamma} dZ_r,$$
$$dv = \kappa_v(\bar{v} - v(t))dt + \sigma_v \sqrt{v(t)}\, dZ_v,$$

where

$$\mathbb{E}[dZ_r\, dZ_v] = \rho\, dt.$$

For $\kappa_r$ and $\bar{r}$ use the same values as in Problem 22.7. For $\bar{v}$ use the values of $\sigma^2$ in Table 22.1. Set the initial value of the $v$ process to $\bar{v}$.

Simulate this model to determine the impact of stochastic volatility on the distributions in Fig. 22.5. In particular see how the skewness and kurtosis of the distribution is affected. Experiment with values of $\kappa_v$ between 1 and 3, and values of $\sigma_v$ equal to $\sqrt{\kappa_v \bar{v}}/2$. Consider $\rho = 0.5, 0.0$ and $-0.5$.

**Problem 22.9** **Computational Problem**—Consider the forward rate process

$$f(t,T) = f(0,T) + \int_0^t \bar{\sigma}(s,T)ds + \int_0^t \sigma(s,T)dW(s),$$

where

$$\bar{\sigma}(t, T) = \sigma(t, T) \int_t^T \sigma(t, v) dv,$$

so that the dynamics are arbitrage free.

Take the volatility function

$$\sigma(t, T) = \beta_0 \left[1 + \alpha_0 r(t) + a_1 f(t, T_1)\right]^{\gamma} e^{-\lambda(T-t)},$$

and the initial forward curve having the functional form

$$f(0, T) = 0.08 - 0.03 \, e^{-1.5T}.$$

Using simulation obtain and graph the distribution for $r(t)$, $f(t, T)$ and $P(t, T)$ when $T = 3$, $T_1 = 5$, at $t = 0.5$ and $t = 1.0$ and for the following sets of parameter specifications

$$(a) \quad \gamma = 0, \quad \beta_0 = 0.02, \quad \lambda = 0.6;$$

$$(b) \quad \gamma = \frac{1}{2}, \quad \beta_0 = 0.015, \quad \alpha_0 = 4.41, \quad \alpha_1 = 6.97, \quad \lambda = 0.6.$$

**Problem 22.10 Computational Problem**—Consider the stochastic differential equation system (22.56)–(22.59) when it is assumed that $\sigma(t, T, f(t, T)) = \sigma_0 e^{-\lambda(T-t)} f(t, T)^{\gamma}$ (see Problems 22.6 and 22.5). Using simulation obtain and graph the distribution for $r(t)$, $f(t, T)$ and $P(t, T)$ when $T = 3$ at $t = 0.5$ and $t = 1.0$ for the parameter specification

(a) $\gamma = 0, \sigma_0 = 0.02, \lambda = 0.6;$
(b) $\gamma = \frac{1}{2}, \sigma_0 = 0.02236, \lambda = 0.6.$

Take $f(0, T)$ as in Problem 22.9 and $\alpha(t, T)$ is given by Eq. (22.55) with $r$ replaced by $f(t, T)$.

# Chapter 23
# Interest Rate Derivatives: One Factor Spot Rate Models

**Abstract** In this chapter we survey models of interest rate derivatives which take the instantaneous spot interest rate as the underlying factor. The continuous hedging argument is extended so as to model the term structure of interest rates and other interest rate derivative securities. This basic approach is due to Vasicek (J Financ Econ 5:177–188, 1977) and hence we shall often refer to it as the Vasicek approach. By specifying different functional forms for the drift, the diffusion and the market price of risk, we develop three well known spot rate models, namely the Vasicek model, the Hull–White model and the Cox–Ingersoll–Ross model. Then we present a general framework for pricing bond options and we apply this framework to obtain closed form solutions for bond options under the specifications of the Hull–White and the Cox–Ingersoll–Ross model. Finally we discuss the calibration of the Hull–White model to the currently observed yield curve.

## 23.1 Introduction

The essential feature of pricing options on interest rate derivative securities is that we need to take account of the stochastic nature of interest rates. Chapter 19 illustrated one approach to this problem, namely modelling the price of pure discount bonds as a stochastic process and making this one of the stochastic factors upon which the value of the option depends. The general approach is due to Merton (1973). There are however a number of practical difficulties in attempting to implement this approach. In particular it requires specification of the average expected return variance over the time interval to maturity, together with the covariance between return and the instantaneous short term rate. It is not clear in practice how best to estimate these variances and covariances. Nevertheless, Merton's approach has guided the development of many of the subsequent interest rate option models.

A characteristic of the stock option model is that there is one basic approach to which can be added embellishments to account for different stochastic processes for the underlying asset (e.g. a jump-diffusion process) or to account for different boundary conditions (e.g. European or American options). For interest rate contingent claims however there does not seem to be one basic approach but rather



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_23

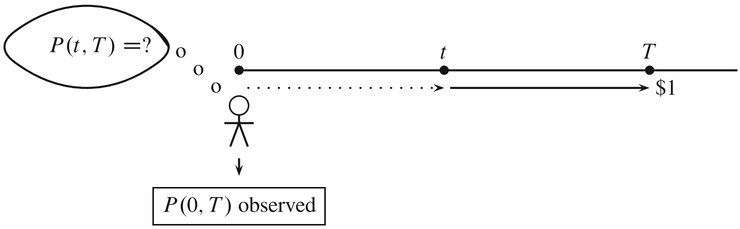


**Fig. 23.1** The time line for the bond pricing problem

a range of alternative approaches. These differ according to what is taken as the underlying factor, which is usually one of the instantaneous spot interest rate, the bond price or the forward rate. Further, some models are presented in a discrete time framework and some in a continuous time framework. An important distinction between alternative approaches is whether the initial term structure (i.e. the currently observed yield curve) is itself to be modelled or to be taken as given. This modelling choice will determine whether the resulting models involve the market price of interest rate risk.

## 23.2 Arbitrage Models of the Term Structure

In this section we consider the perspective of an investor who is standing at time 0 and observes various market rates that enable him/her to compute the initial forward $f(0, T)$ (or equivalently the initial bond price curve $P(0, T)$) for any maturity $T$ out to some maximum maturity (e.g. 30 years in US markets, 10 years in Australian markets). This investor wishes to price at any time $t$ ($< T$) a pure default-free discount bond that pays \$1 at time $T$. The investor in particular seeks the arbitrage-free bond price, i.e. one that does not allow the possibility of riskless arbitrage opportunities between bonds of differing maturities. Furthermore the investor wishes the bond price so obtained to be consistent with currently observed initial bond price curve. Figure 23.1 illustrates the time line for the bond pricing problem.

Initially, we assume that the price of a default-free bond is a function of only the current short term interest rate, time and maturity. Thus we write $P(r(t), t, T)$ to denote the price at time $t$ of a discount bond maturing at time $T$, having maturity value of \$1, when the current instantaneous spot interest rate is $r(t)$, (which is assumed to be riskless in the sense that money invested at this rate will always be paid back) i.e.,

$$P(r(T), T, T) = 1. \tag{23.1}$$

We assume the instantaneous spot rate follows the diffusion process

$$dr = \mu_r(r,t)dt + \sigma_r(r,t)dz. \tag{23.2}$$

By Ito's lemma the bond price therefore satisfies the stochastic differential equation

$$\frac{dP}{P} = \mu_P(r,t,T)dt + \sigma_P(r,t,T)dz, \tag{23.3}$$

where

$$\mu_P(r,t,T) = \frac{1}{P}\left(\frac{\partial P}{\partial t} + \mu_r\frac{\partial P}{\partial r} + \frac{1}{2}\sigma_r^2\frac{\partial^2 P}{\partial r^2}\right), \tag{23.4}$$

$$\sigma_P(r,t,T) = \frac{\sigma_r}{P}\frac{\partial P}{\partial r}. \tag{23.5}$$

Consider an investor who at time $t$ invests \$1 in a hedge portfolio containing two default free bonds maturing at times $T_1$ and $T_2$ respectively and held in the dollar amounts $Q_1$ and $Q_2$. Using $P_i$ to denote the price of the bond maturing at time $T_i$, we can write

$$\left.\begin{array}{r}\text{dollar return on the}\\ \text{hedge portfolio}\end{array}\right\} = Q_1\frac{dP_1}{P_1} + Q_2\frac{dP_2}{P_2}$$

$$= (Q_1\mu_{P_1} + Q_2\mu_{P_2})dt + (Q_1\sigma_{P_1} + Q_2\sigma_{P_2})dz, \tag{23.6}$$

where $\mu_{P_i}$, $\sigma_{P_i}$ denote respectively the expected return and standard deviation of the bond of maturity $T_i$ $(i = 1,2)$. This return can be made certain by choosing the amounts $Q_1$, $Q_2$, so that

$$\frac{Q_1}{Q_2} = -\frac{\sigma_{P_2}}{\sigma_{P_1}}. \tag{23.7}$$

Thus from (23.6) the dollar return on the now riskless hedge portfolio is

$$(Q_1\mu_{P_1} + Q_2\mu_{P_2})dt.$$

Absence of riskless arbitrage then implies that this return must be the same return as a loan at the instantaneous spot interest rate $r$. Given that the original investment is \$1 (i.e. $Q_1 + Q_2 = 1$) then this last condition states that

$$(Q_1\mu_{P_1} + Q_2\mu_{P_2})dt = 1 \cdot r(t)dt.$$

Rearranging we obtain

$$Q_1(\mu_{P_1} - r(t)) + Q_2(\mu_{P_2} - r(t)) = 0,$$

which when combined with (23.7) yields the condition for no-riskless arbitrage between bonds of any two maturities, namely

$$\frac{\mu_{P_1} - r(t)}{\sigma_{P_1}} = \frac{\mu_{P_2} - r(t)}{\sigma_{P_2}}. \tag{23.8}$$

Since the maturity dates $T_1$, $T_2$ were arbitrary, it must be the case that the ratio

$$\frac{\mu_P(r,t,T) - r(t)}{\sigma_P(r,t,T)}$$

is independent of maturity $T$. Let $\lambda(r,t)$ denote the common value of this ratio for bonds of an arbitrary maturity $T$. Thus

$$\frac{\mu_P(r,t,T) - r(t)}{\sigma_P(r,t,T)} = \lambda(r,t). \tag{23.9}$$

The quantity $\lambda$ can be interpreted as the market price of interest rate risk per unit of bond return volatility. Thus Eq. (23.9) asserts that in equilibrium bonds are priced so that instantaneous bond returns equal the instantaneous risk free interest rate plus a risk premium equal to the market price of interest rate risk times instantaneous bond return volatility. Substitution from (23.4) and (23.5) of the expressions for $\mu_P(r,t,T)$ and $\sigma_P(r,t,T)$ respectively, after some simplification, in the partial differential equation for the bond price,

$$\frac{\partial P}{\partial t} + (\mu_r - \lambda(r,t)\sigma_r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma_r^2\frac{\partial^2 P}{\partial r^2} - r(t)P = 0, \tag{23.10}$$

which must be solved subject to the boundary condition

$$P(r(T),T,T) = 1. \tag{23.11}$$

In order to solve (23.10), either analytically or numerically, we need to specify the drift $\mu_r$ and diffusion $\sigma_r$ as well as form of the market price of risk term $\lambda(r,t)$. One common assumption is that this latter term is constant. However, to formally derive this result, involves some very particular assumptions about how the capital market operates. These conditions are discussed briefly in the next section, however to give a proper theoretical basis to the choice of $\lambda(r,t)$ it would be necessary to construct a dynamic general equilibrium model and relate $\lambda(r,t)$ to investor preferences. This is the approach adopted by Cox et al. (1985b).

## 23.3 The Martingale Representation

Just as in the case of the stock option model we were able to obtain a martingale representation of the pricing relationships, so can we do the same thing in the present context. We note from the no riskless arbitrage condition (23.9) that (we use $\lambda(t)$ instead of $\lambda(r,t)$ for notational convenience here)

$$\mu_P(r,t,T) = r(t) + \lambda(t)\sigma_P(r,t,T). \tag{23.12}$$

Substitution of (23.12) into (23.3) yields the stochastic bond price dynamics under the condition of no-riskless arbitrage viz.

$$\frac{dP}{P} = (r(t) + \lambda(t)\sigma_P(r,t,T))dt + \sigma_P(r,t,T)dz. \tag{23.13}$$

Following a line of reasoning identical to that used in Chap. 10 for the stock option model we define a modified Wiener process $\tilde{z}(t)$ by

$$\tilde{z}(t) = z(t) + \int_0^t \lambda(s)ds. \tag{23.14}$$

Under the historical measure $\mathbb{P}$, $\tilde{z}(t)$ is not a standard Wiener process (i.e. $\mathbb{E}(\tilde{z}(t) \neq 0$ where $\mathbb{E}$ is the expectation operation under $\mathbb{P}$) but by an application of Girsanov's theorem we can obtain an equivalent measure $\tilde{\mathbb{P}}$ under which $\tilde{z}(t)$ is a standard Wiener process (i.e. $\tilde{\mathbb{E}}(\tilde{z}(t)) = 0$ where $\tilde{\mathbb{E}}$ is the expectation operation under $\tilde{\mathbb{P}}$). Thus in terms of $\tilde{z}(t)$ the stochastic differential equation (23.13) for $P$ under the measure $\tilde{\mathbb{P}}$ becomes

$$\frac{dP}{P} = r(t)dt + \sigma_P(r,t,T)d\tilde{z}. \tag{23.15}$$

However unlike in the stock option situation, the spot rate $r$ is here stochastic, so we need to define the money market account (the accumulated value by time $t$ of \$1 continuously compounded at $r$ since time 0) as

$$A(t) = e^{\int_0^t r(s)ds}. \tag{23.16}$$

It is a simple matter to demonstrate that

$$dA = r(t)A(t)dt. \tag{23.17}$$

We then define the bond price in units of the money market account,[1]

$$Z(r,t,T) = \frac{P(r,t,T)}{A(t)}, \tag{23.18}$$

and a simple application of Ito's lemma reveals that $Z$ satisfies

$$\frac{dZ}{Z} = \sigma_P(r,t,T)d\tilde{z}. \tag{23.19}$$

Thus Eq. (23.19) implies $Z(r,t,T)$ is a martingale under $\tilde{\mathbb{P}}$, i.e.

$$Z(r,t,T) = \tilde{\mathbb{E}}_t[Z(r(T),T,T)], \tag{23.20}$$

which in terms of the original bond price can be expressed (after slight re-arrangement) as

$$P(r,t,T) = \tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T)}P(r(T),T,T)\right],$$

or since $P(r(T),T,T) = 1$, more simply as

$$P(r,t,T) = \tilde{\mathbb{E}}_t\left[e^{-\int_t^T r(s)ds}\right]. \tag{23.21}$$

In order to derive the interest rate dynamics under $\tilde{\mathbb{P}}$ we use (23.14) to replace $dz$ by $(d\tilde{z} - \lambda(t)dt)$ in Eq. (23.2) to obtain

$$dr = (\mu_r(r,t) - \lambda(r,t)\sigma_r(r,t))dt + \sigma_r(r,t)d\tilde{z}. \tag{23.22}$$

An application of the Feynman–Kac formula[2] (in particular Proposition 8.2) to (23.21) and (23.22) would take us back to the partial differential equation (23.10). Thus, just as in the stock option situation, we have two representations of the bond price, the partial differential equation (23.10) and the expectation operator (23.21) under the interest rate dynamics (23.22). To use these representations we need to specify the function $\mu_r$,$\sigma_r$ and also the functional form for the market price of interest rate risk $\lambda(r,t)$. This we do in the following section for specific term structure models.

---

[1]Recall that by the rules of stochastic calculus

$$\frac{dZ}{Z} = \frac{dP}{P} - \frac{dA}{A} - \frac{dP}{P}\cdot\frac{dA}{A} + \left(\frac{dA}{A}\right)^2.$$

[2]Make the identifications $x \to r$, $v(t,r) \to P(r,t,T)$, $\lambda = -1$, $f[s,x(s)] \to r(s)$.

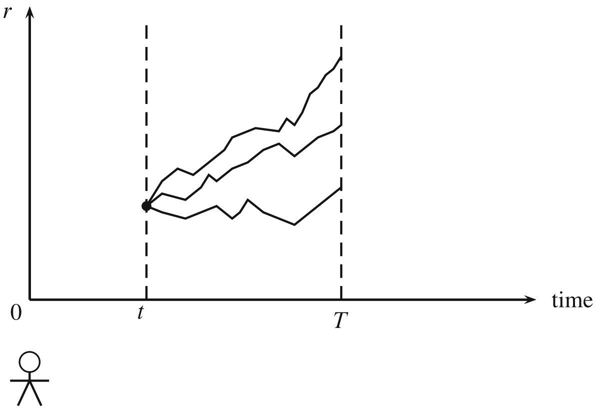


**Fig. 23.2** Typical paths for the $r$ process over $[t, T]$. Equation (23.21) averages the quantity $e^{-\int_t^T r(s)ds}$ over many such paths under the $\tilde{\mathbb{P}}$ measure

Before leaving this section we wish to emphasize the discounted cash flow interpretation of the representation (23.21). The factor $\exp(-\int_t^T r(s)ds)$ discounts back to $t$ the dollar received at $T$, for one particular path followed by $r(s)$. Since $r(s)$ is stochastic this quantity is in fact a stochastic discount factor. To obtain the discounted value at $t$ of the \$1 received at $T$ we need to average over the range of possible paths followed by $r(s)$ under the measure $\tilde{\mathbb{P}}$. This is effectively what the $\tilde{\mathbb{E}}_t$ does; Fig. 23.2 illustrates this idea. It is also of interest to contrast the bond price expression (23.21) with the corresponding expression in Eq. (22.11) for a world of certainty, and we see how this is generalised in a natural way to the world of uncertainty.

We thus have a complete analogy with the stock option price derivation of Chaps. 6 and 7 with the exception that the pricing relationships here involve the market price of interest rate risk $\lambda$. But from our discussion in Chap. 10 this is to be expected since the underlying factor, the spot interest rate $r(t)$, is not a traded factor.

## 23.4 Some Specific Term Structure Models

A variety of term structure models are obtained by specifying different forms for $\mu_r(r,t)$ and $\sigma_r(r,t)$ in the interest rate process, Eq. (23.2), and/or different forms for the market price of risk term $\lambda(r,t)$.

### *23.4.1 The Vasicek Model*

The Vasicek (1977) model holds a special place in the interest rate term structure literature as it was the earliest model. Its basic assumptions are to take

$$\mu_r(r,t) = \kappa(\gamma - r(t)) \text{ and } \sigma_r(r,t) = \sigma, \tag{23.23}$$

where $\kappa > 0$ and $\sigma > 0$ are constant. We also assume a constant market price of interest rate risk, i.e., $\lambda(r,t) = \lambda$. The bond pricing partial differential equation (23.10) in this case becomes

$$\frac{\partial P}{\partial t} + (\theta - \kappa r(t))\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} - r(t)P = 0, \tag{23.24}$$

where we set $\theta = \kappa\gamma - \lambda\sigma$. One way to solve this partial differential equation is to see whether it is possible to find functions $a(t,T)$ and $b(t,T)$ such that the solution can be written in the form[3]

$$P(t,T) = e^{-a(t,T)-b(t,T)r(t)}. \tag{23.25}$$

In order that the boundary condition (23.11) be satisfied for all possible $r(T)$ it must be the case that $a(t,T)$ and $b(t,T)$ satisfy

$$a(T,T) = 0 \text{ and } b(T,T) = 0. \tag{23.26}$$

From (23.25) we note that[4]

$$\frac{\partial P}{\partial r} = -bP, \quad \frac{\partial^2 P}{\partial r^2} = b^2 P \text{ and } \frac{\partial P}{\partial t} = (-a_t - b_t r(t))P. \tag{23.27}$$

Substituting these relations into (23.24) and gathering terms in powers of $r(t)(r(t)^0 = 1)$ we obtain

$$\left[-a_t - b\theta + \frac{1}{2}\sigma^2 b^2\right] + [-b_t + \kappa b - 1]r(t) = 0. \tag{23.28}$$

If (23.28) is to hold for all $t$ and all $r$ it must be the case that each bracket is separately equal to zero. Thus we obtain for $a$ and $b$ the ordinary differential equations

$$b_t = \kappa b - 1, \tag{23.29}$$

and

$$a_t = -b\theta + \frac{1}{2}\sigma^2 b^2, \tag{23.30}$$

[3]One motivation for the functional form Eq. (23.25) is that when $r$ is constant, we have $P(t,T) = e^{-r(T-t)}$, and (23.25) is an obvious generalisation of this relation.

[4]We employ the notation $a_t = \frac{\partial}{\partial t}a(t,T)$, $b_t = \frac{\partial}{\partial t}b(t,T)$.

which must be solved subject to the boundary conditions (23.26). From (23.29) we obtain[5]

$$b(t,T) = \frac{\left[1 - e^{-\kappa(T-t)}\right]}{\kappa}. \tag{23.31}$$

Substituting (23.31) into (23.30), integrating from $t$ to $T$ and, using the boundary condition $a(T,T) = 0$ we find that

$$a(t,T) = \int_t^T \left[\theta b(s,T) - \frac{1}{2}\sigma^2 b^2(s,T)\right] ds, \tag{23.32}$$

which after some elementary integrations reduces to[6]

$$a(t,T) = \left(\frac{\theta}{\kappa} - \frac{\sigma^2}{2\kappa^2}\right)(T-t) + \left(\frac{\theta}{\kappa^2} - \frac{\sigma^2}{2\kappa^3}\right)(e^{-\kappa(T-t)} - 1) + \frac{\sigma^2}{4\kappa^3}(e^{-\kappa(T-t)} - 1)^2. \tag{23.33}$$

[5] Equation (23.29) can be re-arranged into

$$\frac{d}{dt}\left(b(t,T)e^{-\kappa t}\right) = -e^{-\kappa t}$$

and integrating $t$ to $T$ we obtain

$$b(T,T)e^{-\kappa T} - b(t,T)e^{-\kappa t} = -\int_t^T e^{-\kappa s} ds = -\frac{1}{\kappa}(e^{-\kappa t} - e^{-\kappa T}).$$

Use of the boundary condition $b(T,T) = 0$ and some re-arrangement yields (23.31).

[6] Note that

$$\int_t^T b(s,T)ds = \int_t^T \frac{(1 - e^{-\kappa(T-s)})}{\kappa} ds = \int_0^{T-t} \frac{(1 - e^{-\kappa u})}{\kappa} du = \frac{(T-t)}{\kappa} + \frac{1}{\kappa^2}(e^{-\kappa(T-t)} - 1)$$

and

$$\begin{aligned}\int_t^T b^2(s,T)ds &= \int_0^{T-t} \frac{(1 - e^{-\kappa u})^2}{\kappa^2} du \\ &= \frac{1}{\kappa^2}\left[(T-t) + \frac{2}{\kappa}(e^{-\kappa(T-t)} - 1) - \frac{1}{2\kappa}(e^{-2\kappa(T-t)} - 1)\right] \\ &= \frac{(T-t)}{\kappa^2} - \frac{1}{2\kappa^3}\left\{(e^{-\kappa(T-t)} - 1)^2 - 2(e^{-\kappa(T-t)} - 1)\right\}.\end{aligned}$$

By considering the corresponding expression for the yield to maturity $\rho(t, T)$ $(= -\ln P(t,T)/(T-t) = (a(t,T) + b(t,T)r(t))/(T-t))$ we find that

$$\rho(t,T) = \left(\frac{\theta}{\kappa} - \frac{\sigma^2}{2\kappa^2}\right)\left(1 + \frac{e^{-\kappa(T-t)} - 1}{\kappa(T-t)}\right) + \frac{\sigma^2}{4\kappa^3}\frac{(e^{-\kappa(T-t)} - 1)^2}{T-t} + \frac{(1 - e^{-\kappa(T-t)})}{T-t} r(t). \tag{23.34}$$

By letting $(T - t) \to \infty$ we find that the yield at infinite maturity is given by[7]

$$\rho_\infty = \frac{\theta}{\kappa} - \frac{\sigma^2}{2\kappa^2}. \tag{23.35}$$

One may then express the bond price as

$$P(r,t,T) = \exp\left[\frac{(e^{-\kappa(T-t)} - 1)}{\kappa}(r(t) - \rho_\infty) - \rho_\infty(T-t) - \frac{\sigma^2}{4\kappa^3}(e^{-\kappa(T-t)} - 1)^2\right]. \tag{23.36}$$

Whilst the Vasicek model may now be of historical interest, it nevertheless contains all the basic ingredients needed to deal with the more sophisticated models which we consider in the following subsections. Namely a technique to solve the pricing partial differential equation and the idea of relating the parameters of the model to information that can be obtained from the currently observed yield curve. The last observation also makes evident one of the widely perceived shortcomings of the Vasicek model. By setting $t = 0$ in (23.36) we obtain

$$P(r(0), 0, T) = \exp\left[\frac{(e^{-\kappa T} - 1)}{\kappa}(r(0) - \rho_\infty) - \rho_\infty T - \frac{\sigma^2}{4\kappa^3}(e^{-\kappa T} - 1)^2\right]. \tag{23.37}$$

If $\theta$ is chosen so as to match the long term yield $\rho_\infty$ then we only have two parameters, $\kappa$ and $\sigma$, left to make expression (23.36) consistent with the entire currently observed yield curve $P(r(0), 0, T)$. Clearly this is impossible as at the most we could choose $\kappa$ and $\sigma$ to fit two points exactly, or alternatively choose them to obtain some sort of least squares fit. These observations suggest that one possibility to develop a model that fits the currently observed yield curve is to make at least one, if not more, of the quantities $\kappa$, $\gamma$ and $\sigma$ time varying. We would then have at our disposition a whole set of values of say $\kappa$ (if it were allowed to be time varying) with which to match the theoretical model to the currently observed yield

[7] One could use this insight to infer a value for the unknown factor $\theta$ from the currently observed yield curve, from which one could obtain an estimate of $\rho_\infty$.

curve. This is the essential insight of the Hull–White model to which we turn in the next subsection.

However before turning to the Hull–White model we consider what the solution (23.25) implies for the bond price dynamics. We know that under $\mathbb{P}$ the bond price dynamics are given by (23.13). However at that point in our development we did not have an explicit expression for $\frac{\partial P}{\partial r}$. The solution (23.25) now enables us to calculate this expression, in fact it is given in (23.27). Substituting this expression into Eq. (23.13) and from Eq. (23.5), we find that the bond price dynamics are given by

$$\frac{dP}{P} = (r(t) - \lambda\sigma b(t,T))dt - \sigma b(t,T)dz. \tag{23.38}$$

The last equation indicates that the standard deviation of bond return is $-\sigma b$, the minus sign simply indicates that a positive shock to the interest rate dynamics (i.e., a positive $dz$) results in a negative shock to the bond price dynamics. This is merely a reflection of the fact that interest rates and bond prices are inversely related. As we have seen in the discussion of the general case in Sect. 23.2, Eq. (23.38) can be transformed, under the equivalent measure $\tilde{\mathbb{P}}$, to

$$\frac{dP}{P} = r(t)dt - \sigma b(t,T)d\tilde{z}. \tag{23.39}$$

The last equation leads, as we saw in Sect. 23.2 to the martingale representation (23.21). Furthermore the interest rate dynamics under $\tilde{\mathbb{P}}$ becomes

$$dr = (\theta - \kappa r(t))dt + \sigma d\tilde{z}, \tag{23.40}$$

where we recall that $\theta$ has already been defined as $\theta = \kappa\gamma - \lambda\sigma$. These are the dynamics with respect to which the expectation $\tilde{\mathbb{E}}_t$ in (23.21) is to be calculated.

### *23.4.2 The Hull–White Model*

Hull and White (1990) take as the process for the short rate

$$dr = \kappa(t)(\gamma(t) - r(t))dt + \sigma(t)dz. \tag{23.41}$$

The difference from the Vasicek model being the time dependence of the coefficients $\kappa(t)$, $\gamma(t)$ and $\sigma(t)$, the motivation being the points discussed at the conclusion of the previous section. The bond pricing partial differential equation (23.10) now becomes

$$\frac{\partial P}{\partial t} + (\theta(t) - \kappa(t)r(t))\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2(t)\frac{\partial^2 P}{\partial r^2} - r(t)P = 0, \tag{23.42}$$

where we set

$$\theta(t) = \kappa(t)\gamma(t) - \lambda\sigma(t). \tag{23.43}$$

The only difference from the partial differential equation (23.24) being the time dependence of the coefficients $\theta(t)$, $\kappa(t)$ and $\sigma(t)$. It seems not unreasonable to attempt again a solution of the form (23.25). In fact precisely the same manipulations yield for the time coefficients $a(t, T)$ and $b(t, T)$ the ordinary differential equations

$$b_t = \kappa(t)b - 1, \tag{23.44}$$

and

$$a_t = -b\theta(t) + \frac{1}{2}\sigma^2(t)b^2, \tag{23.45}$$

the only difference being that the two ordinary differential equations we must solve now have time varying coefficients. If we define

$$\mathscr{K}(t) = \int_0^t \kappa(s)ds, \tag{23.46}$$

then the solution to (23.44) can be written[8]

$$b(t, T) = \int_t^T e^{\mathscr{K}(t) - \mathscr{K}(s)} ds. \tag{23.47}$$

Substituting (23.47) into (23.45) and integrating $t$ to $T$ yields

$$a(t, T) = \int_t^T b(s, T)\theta(s)ds - \frac{1}{2}\int_t^T \sigma^2(s)b^2(s, T)ds. \tag{23.48}$$

For general forms of the functions $\kappa(t)$, $\sigma(t)$ and $\theta(t)$ it may be necessary to perform numerically the integrations in (23.47) and (23.48). In fact to perform the integrations we would need to also have some functional form for $\lambda$, and this would be difficult to obtain. It turns out that we can instead find from market data the

[8]Note that $\frac{d}{dt}\mathscr{K}(t) = \kappa(t)$ so that $\frac{d}{dt}e^{\mathscr{K}(t)} = e^{\mathscr{K}(t)}\kappa(t)$. Multiplying across (23.44) by $e^{-\mathscr{K}(t)}$ and re-arranging we obtain

$$\frac{d}{dt}(e^{-\mathscr{K}(t)}b(t, T)) = -e^{-\mathscr{K}(t)}.$$

The result then follows by integrating $t$ to $T$ and following manipulations similar to those in footnote 5.

function $\theta(t)$ (which contains $\lambda$) and this is (together with $\kappa(t)$ and $\sigma(t)$) all we need to use the bond pricing formula.

Let us now consider the bond price dynamics implied by the bond pricing formula (23.13) with $a(t,T)$ and $b(t,T)$ now given by (23.47) and (23.48). We follow exactly the corresponding manipulations for the Vasicek model that led to Eq. (23.38), which are not altered by the fact that $\kappa$, $\gamma$ and $\sigma$ are now time varying. From the general expression (23.25) for the bond price and Eq. (23.5) for the volatility of bond return we have

$$\sigma_P(t,T) = -\sigma(t)b(t,T). \tag{23.49}$$

Thus for the Hull–White model we obtain for the bond price dynamics

$$\frac{dP}{P} = (r(t) - \lambda\sigma(t)b(t,T))dt - \sigma(t)b(t,T)dz, \tag{23.50}$$

where we highlight the time dependence of $\sigma(t)$ and the fact that $b(t,T)$ is given by Eq. (23.47). Under the equivalent measure $\tilde{\mathbb{P}}$ the bond price dynamics are given by

$$\frac{dP}{P} = r(t)dt - \sigma(t)b(t,T)d\tilde{z}, \tag{23.51}$$

which leads to the martingale representation (23.21) as we have shown in the general case in Sect. 23.2. The interest rate dynamics under which $\tilde{\mathbb{E}}_t$ is calculated are given by (after setting $dz = d\tilde{z} - \lambda dt$ in (23.41))

$$dr = (\theta(t) - \kappa(t)r(t))dt + \sigma(t)d\tilde{z}. \tag{23.52}$$

### *23.4.3 The Cox–Ingersoll–Ross (CIR) Model*

Cox et al. (1985a) (CIR) consider the interest rate process

$$dr = \kappa(t)(\gamma - r(t))dt + \sigma\sqrt{r(t)}dz. \tag{23.53}$$

As we discussed in Sect. 22.3 the motivation for using this process is that it guarantees non-negative (or positive if $\kappa(t)\gamma > \sigma^2/2$) spot interest rate sample paths. Using a dynamic general equilibrium framework[9] in order to obtain a tractable bond pricing equation we assume that the market price of interest rate risk is a function of $r(t)$ given by

$$\lambda(r,t) = \lambda\sqrt{r(t)}, \tag{23.54}$$

[9]In fact CIR employ a dynamic general equilibrium framework to derive the bond pricing equation and under specific assumptions about investor preferences, end up with a market price of risk given by (23.54).

where $\lambda$ is a constant. The pricing partial differential equation (23.10) with this specification of the market price of risk becomes

$$\frac{1}{2}\sigma^2 r(t)\frac{\partial^2 P}{\partial r^2} + (\kappa(t)\gamma - (\kappa(t) + \lambda\sigma)r(t))\frac{\partial P}{\partial r(t)} + \frac{\partial P}{\partial t} - r(t)P = 0. \quad (23.55)$$

Given the very similar structure to the partial differential equation encountered in the Vasicek and Hull–White models (the only difference is the $r$ in front of the second derivative) it seems not unreasonable to again try a solution of the same form viz.

$$P(t, T) = e^{-a(t,T)-b(t,T)r(t)}. \quad (23.56)$$

As already discussed when obtaining the corresponding solution for the Hull–White model, the condition $P(T, T) = 1$ can only be guaranteed if

$$b(T, T) = 0, \quad a(T, T) = 0. \quad (23.57)$$

We note also that here

$$\frac{\partial P}{\partial r} = -bP, \quad \frac{\partial^2 P}{\partial r^2} = b^2 P, \quad \frac{\partial P}{\partial t} = (-rb_t - a_t)P,$$

which upon substitution into (23.55) and re-arrangement of terms yields

$$[-\kappa(t)\gamma b - a_t] + \left[\frac{1}{2}\sigma^2 b^2 + (\kappa(t) + \lambda\sigma)b - b_t - 1\right] r(t) = 0. \quad (23.58)$$

In order that this relation hold for all $r(t)$ and all $t$ it must be the case that

$$\frac{1}{2}\sigma^2 b^2 + (\kappa(t) + \lambda\sigma)b - b_t - 1 = 0, \quad (23.59)$$

and

$$-\kappa(t)\gamma b - a_t = 0. \quad (23.60)$$

The difference compared to the solution of the Hull–White model is the $b^2$ term in the ordinary differential equation (23.59), which makes its solution more difficult. However this is in fact the well-known Ricatti ordinary differential equation whose solution is known. We show in Appendix 23.1 that the solution to (23.59) is

$$b(t, T) = \frac{2}{\sigma^2}\frac{[1 - e^{-\beta(T-t)}]}{[\phi_1 e^{-\beta(T-t)} - \phi_2]}, \quad (23.61)$$

where

$$\phi_1 = -\frac{(\kappa(t)+\lambda)}{\sigma^2} + \frac{\beta}{\sigma^2}, \quad \phi_2 = \frac{-(\kappa(t)+\lambda)}{\sigma^2} - \frac{\beta}{\sigma^2},$$

and

$$\beta = \sqrt{(\kappa(t)+\lambda)^2 + 2\sigma^2}.$$

The expression (23.61) appears in many different forms in the literature. Equation (23.60) may be written

$$\frac{da}{dt} = -\kappa(t)\gamma b(t,T),$$

which upon integration from $t, T$ yields (using $a(T,T) = 0$)

$$-a(t,T) = -\kappa(t)\gamma \int_t^T b(s,T)ds. \tag{23.62}$$

We show in Appendix 23.1 that Eq. (23.62) integrates to

$$a(t,T) = \frac{2\kappa(t)\gamma}{\beta\sigma^2}\left[-\beta\frac{(T-t)}{\phi_1} - \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\ln(\frac{\phi_1-\phi_2 e^{\beta(T-t)}}{\phi_1-\phi_2})\right]. \tag{23.63}$$

There are also many alternative representations of (23.63) in the literature. We saw in the discussion on the Hull–White model, that in order to be able to calibrate the model to market data we needed the additional flexibility required by allowing the coefficients in (23.53) to be time varying. We can adopt exactly the same procedure with the CIR model so that any or all of the coefficients $\sigma$, $\kappa(t)$, $\gamma$ and $\lambda$ in the partial differential equation (23.55) become time varying.

Again we try a solution of the form (23.56) and Eqs. (23.59) and (23.60) still emerge as the equations determining the coefficients $b$ and $a$. Only now it needs to be borne in mind that the coefficients $\sigma$, $\kappa(t)$, $\lambda$ and $\gamma$ are time-varying. We show in the appendix that the functional form (23.61) is still valid for $b(t,T)$ except that the constant $\beta$ is replaced by the time averaged function

$$\bar{\beta}(t,T) = \frac{1}{T-t}\int_t^T \beta(s)ds. \tag{23.64}$$

The expression for $a(t,T)$ can only be left as the integral

$$a(t,T) = \int_t^T \kappa(s)\gamma(s)b(s,T)ds, \tag{23.65}$$

since the integration would in general be impossible analytically because $\bar{\beta}(t, T)$ could be a quite complicated time function. From (23.56) we can readily calculate that

$$\frac{\partial P}{\partial r} = -b(t, T)P$$

and hence Eq. (23.15) for the risk neutral bond price dynamics in the CIR case become

$$\frac{dP}{P} = r(t)dt - \sigma b(t, T)\sqrt{r(t)}d\tilde{z}. \tag{23.66}$$

The interest rate dynamics under the equivalent measure $\tilde{\mathbb{P}}$ are obtained by setting $dz = d\tilde{z} - \lambda\sqrt{r(t)}dt$ in Eq. (23.53) and so are given by

$$dr = (\theta(t) - \alpha(t)r(t))dt + \sigma\sqrt{r(t)}d\tilde{z}$$

where

$$\theta(t) = \kappa(t)\gamma(t) \quad \text{and} \quad \alpha(t) = \kappa(t) + \lambda\sigma.$$

## 23.5 Calculation of the Bond Price from the Expectation Operator

We have seen in Sect. 23.4 how to obtain an explicit expression for the bond price by solving the pricing partial differential equation (23.10) under various assumptions about $\mu_r$ and $\sigma_r$. It is also of interest to see how to obtain the same result by starting from the martingale or expectation operator expression (23.21). The particular spot interest rate models with which we are working provide one of the rare instances where we can carry out analytically, both the solution of the partial differential equation and the calculation of the expectation operator.

The key to carrying out the expectation operation in (23.21) is to determine the distributional characteristics of $\int_t^T r(s)ds$ under $\tilde{\mathbb{P}}$. We shall now show that in the case of the Hull–White model this quantity is normally distributed with mean and variance that we calculate below. We know from Chap. 6 how to calculate the expectation of the exponential of a normally distributed random variable. The appropriate interest rate dynamics are given by Eq. (23.52), which using the quantity $\mathscr{K}(t)$ defined by Eq. (23.46), can be written

$$d(r(t)e^{\mathscr{K}(t)}) = e^{\mathscr{K}(t)}\theta(t)dt + e^{\mathscr{K}(t)}\sigma(t)d\tilde{z}.$$

Integrating from $t$ to $s$ $(< T)$ and re-arranging we find that

$$r(s) = r(t)e^{\mathscr{K}(t)-\mathscr{K}(s)} + \int_t^s e^{\mathscr{K}(u)-\mathscr{K}(s)}\theta(u)du + \int_t^s e^{\mathscr{K}(u)-\mathscr{K}(s)}\sigma(u)d\tilde{z}(u). \tag{23.67}$$

Next integrate Eq. (23.67) from $t$ to $T$ to obtain

$$\int_t^T r(s)ds = r(t)\int_t^T e^{\mathscr{K}(t)-\mathscr{K}(s)}ds + \int_t^T \left(\int_t^s e^{\mathscr{K}(u)-\mathscr{K}(s)}\theta(u)du\right) ds$$
$$+ \int_t^T \left(\int_t^s e^{\mathscr{K}(u)-\mathscr{K}(s)}\sigma(u)d\tilde{z}(u)\right) ds.$$

Interchanging the order of integration in the second integral and applying Fubini's theorem (Sect. 22.4 version III is being used here) to the stochastic integral the last equation becomes

$$\int_t^T r(s)ds = r(t)\int_t^T e^{\mathscr{K}(t)-\mathscr{K}(s)}ds + \int_t^T \left(\int_u^T e^{\mathscr{K}(u)-\mathscr{K}(s)}ds\right)\theta(u)du$$
$$+ \int_t^T \left(\int_u^T e^{\mathscr{K}(u)-\mathscr{K}(s)}ds\right)\sigma(u)d\tilde{z}(u). \tag{23.68}$$

By making use of the definition of $b(t, T)$ at Eq. (23.47) we can write (23.68) more compactly as

$$\int_t^T r(s)ds = b(t, T)r(t) + \int_t^T b(u, T)\theta(u)du + \int_t^T b(u, T)\sigma(u)d\tilde{z}(u). \tag{23.69}$$

First we note that Eq. (23.69) implies that $\int_t^T r(s)ds$ is normally distributed (conditional on information at time $t$) since the coefficients on the right-hand side are at most time functions (as opposed to being functions of $r(t)$). The mean, $M(t)$, and variance $V^2(t)$, are easily calculated to be

$$M(t) = b(t, T)r(t) + \int_t^T b(u, T)\theta(u)du, \tag{23.70}$$

and

$$V^2(t) = \int_t^T b^2(u, T)\sigma^2(u)du. \tag{23.71}$$

From the above discussion we can assert that (under $\tilde{\mathbb{P}}$)

$$\int_t^T r(s)ds \sim N(M(t), V^2(t)), \tag{23.72}$$

and so

$$-\int_t^T r(s)ds \sim N(-M(t), V^2(t)). \tag{23.73}$$

Finally using the results of (iv) in Sect. 6.3 we obtain the result

$$\begin{aligned} P(r,t,T) &= \tilde{E}_t[e^{-\int_t^T r(s)ds}] \\ &= e^{-M(t)+\frac{1}{2}V^2(t)} \\ &= \exp\left[-b(t,T)r(t) - \int_t^T b(u,T)\theta(u)du + \frac{1}{2}\int_t^T b(u,T)^2\sigma^2(u)du\right] \\ &= \exp[-b(t,T)r(t) - a(t,T)], \end{aligned} \tag{23.74}$$

by making use of the definition of $a(t,T)$ in Eq. (23.48). We see that in Eq. (23.74) we have recovered the bond pricing formula (23.25) obtained by solving the partial differential equation (23.42).

## 23.6 Pricing Bond Options

We continue to assume that the short-term rate follows the continuous diffusion process (23.2). We also assume that there are no riskless arbitrage opportunities in the bond market. Thus the price of the discount bond of any maturity is still given by the solution to the partial differential equation (23.10). Let $C(r,t)$[10] denote the price at time $t$ of a call option of maturity $T_C$ written on a bond having maturity $T$ ($> T_C$) (see Fig. 23.3).

By Ito's lemma

$$\frac{dC}{C} = \mu_C dt + \sigma_C dz, \tag{23.75}$$

[10]A more precise notation for the option value would be $C(r,t,T_C,T)$. However to ease the notation we shall usually just write $C(r,t)$, unless we need to highlight the dependence on option maturity ($T_C$) or underlying bond maturity ($T$), as in Sect. 23.7.

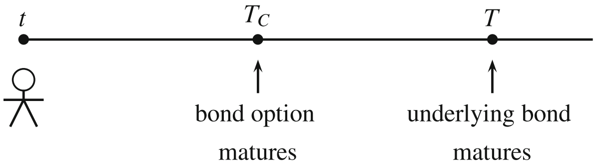


**Fig. 23.3** Time line for the bond option problem

where

$$\mu_C = \frac{1}{C}\left(\frac{\partial C}{\partial t} + \mu_r \frac{\partial C}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 C}{\partial r^2}\right), \tag{23.76}$$

$$\sigma_C = \frac{\sigma_r}{C}\frac{\partial C}{\partial r}. \tag{23.77}$$

Consider an investor who at time $t$ invests \$1 in a hedge portfolio containing the bond of maturity $T$ held in the dollar amount $Q_P$ and the option of maturity $T_C$ held in the dollar amount $Q_C$. The dollar return on this hedge portfolio over time interval $dt$ is given by

$$\left.\begin{array}{r}\text{dollar return on the}\\ \text{hedge portfolio}\end{array}\right\} = Q_P \frac{dP}{P} + Q_C \frac{dC}{C}$$
$$= (Q_P\mu_P + Q_C\mu_C)dt + (Q_P\sigma_P + Q_C\sigma_C)dz.$$

The hedge portfolio is rendered riskless by choosing $Q_P, Q_C$ such that

$$\frac{Q_P}{Q_C} = -\frac{\sigma_C}{\sigma_P}. \tag{23.78}$$

The absence of riskless arbitrage means that the hedge portfolio can only earn the same return as the original \$1 invested at the risk-free rate. In other words

$$(Q_P\mu_P + Q_C\mu_C)dt = 1 \cdot r(t)dt. \tag{23.79}$$

Recalling that $Q_P + Q_C = 1$, the conditions (23.78) and (23.79) imply

$$\frac{\mu_C - r(t)}{\sigma_C} = \frac{\mu_P - r(t)}{\sigma_P}. \tag{23.80}$$

But by Eq. (23.9) we know that in an arbitrage-free bond market $(\mu_P - r(t))/\sigma_P$ is equal to the market price of interest rate risk. Thus we arrive at the no-riskless arbitrage condition between the option and bond markets, viz.

$$\frac{\mu_C(t,s) - r(t)}{\sigma_C(t,s)} = \frac{\mu_P(t,s) - r(t)}{\sigma_P(t,s)} = \lambda(r,t). \tag{23.81}$$

Equation (23.81) has the now familiar interpretation that in the absence of riskless arbitrage the excess return risk adjusted on both the bond and the option are equal. Furthermore the common factor to which they are equal is the market price of risk of the spot interest rate, the underlying factor. Equation (23.81) yields the partial differential equation (23.10) for the bond price $P$, and for the option price $C$, the partial differential equation

$$\frac{\partial C}{\partial t} + (\mu_r - \lambda(r,t)\sigma_r)\frac{\partial C}{\partial r} + \frac{1}{2}\sigma_r^2\frac{\partial^2 C}{\partial r^2} - r(t)C = 0, \tag{23.82}$$

which in the case of a European call option (with exercise price $X$) on the bond must be solved on the time interval $0 < t < T_C$ subject to the boundary conditions

$$\begin{aligned} &C(r(T_C), T_C) = \max[0, P(r(T_C), T_C, T) - E], \\ &C(\infty, t) = 0. \end{aligned} \tag{23.83}$$

The last condition is a consequence of the result that

$$P(\infty, t, T) = 0,$$

i.e. the bond value declines to zero as the interest rate becomes large.

Note the two-pass structure of the solution process. We must first solve the partial differential equation (23.10) with boundary condition (23.1) for the bond price $P(r(s), s, T)$ on the time interval $T_C \leq s \leq T$. The value $P(r(T_C), T_C, T)$ is then used in the solution of the partial differential equation (23.82) (in fact the same partial differential equation) via the boundary condition in (23.83). This two-pass procedure is illustrated in Fig. 23.4. In order to obtain the martingale representation for the option price we follow almost identical steps to those we followed in Sect. 23.3 to obtain the martingale representation for the bond price. It follows from Eq. (23.81) that

$$\mu_C = r(t) + \lambda(r,t)\sigma_C. \tag{23.84}$$

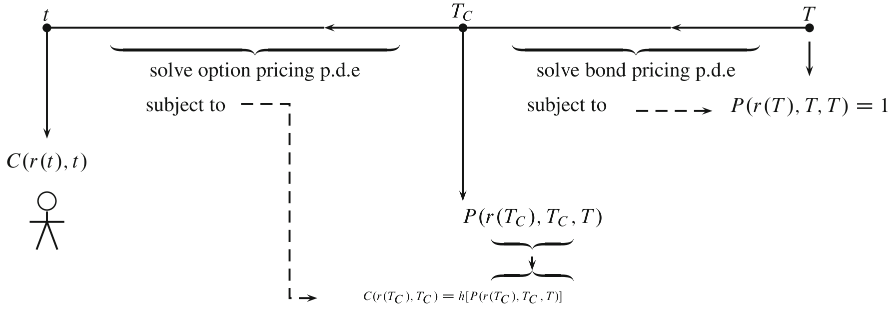


**Fig. 23.4** The two-pass procedure for solving the bond option pricing problem

Substituting (23.84) into (23.75) the arbitrage free option price dynamics are given by

$$\frac{dC}{C} = (r(t) + \lambda(r,t)\sigma_C)dt + \sigma_C dz.$$

The last equation may in turn be written in terms of $\tilde{z}(t)$ (see Eq. (23.14)) as

$$\frac{dC}{C} = r(t)dt + \sigma_C d\tilde{z}(t), \tag{23.85}$$

where we recall that under the equivalent measure $\tilde{\mathbb{P}}$, the quantity $\tilde{z}(t)$ is a standard Wiener process. If we set

$$Y(r,t) = \frac{C(r,t)}{A(t)},$$

which is the option price measured in units of the money market account, then from Ito's lemma[11]

$$\frac{dY}{Y} = \sigma_C d\tilde{z}(t).$$

The last equation implies that $Y$ is a martingale under $\tilde{\mathbb{P}}$, thus

$$Y(r,t) = \tilde{\mathbb{E}}_t[Y(r(T_C), T_C)],$$

which in terms of the option price itself can be expressed as

$$C(r,t) = \tilde{\mathbb{E}}_t[e^{-\int_t^{T_C} r(s)ds} C(r(T_C), T_C)]. \tag{23.86}$$

If for example we wish to price a European call option on a bond then the maturity condition is

$$C(r(T_C), T_C) = \max[0, P(r(T_C), T_C, T) - X].$$

The interest rate dynamics under $\tilde{\mathbb{P}}$ are still given by (23.22), viz.

$$dr = (\mu_r - \lambda(r,t)\sigma_r)dt + \sigma_r d\tilde{z}.$$

[11] Alternatively we could use the result in Sect. 6.6 for the stochastic differential equation followed by the ratio of the two diffusions (23.17) and (23.85).

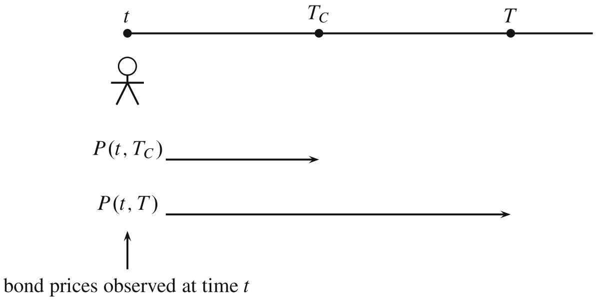


**Fig. 23.5** Using $P(t, T_C)$ as numeraire—the forward measure

Application of the Feynman–Kac formula to (23.86) (see Proposition 8.3[12]) will take us back to the option pricing partial differential equation (23.82). Recalling the discussion about stochastic discounting under $\tilde{\mathbb{P}}$ at the end of Sect. 23.3 we see that Eq. (23.86) has an obvious expected (under $\tilde{\mathbb{P}}$) discounted payoff interpretation.

One of the difficulties with evaluating the expectation in (23.86) is that one needs the joint distribution of $\exp(-\int_t^{T_C} r(s)ds)$ and $C(r(T_C), T_C)$. The calculation of this joint distribution may in practice be quite difficult. A simpler calculation may be obtained by using the so-called forward measure,[13] which consists in choosing as numeraire a bond of maturity $T_C$ (see Fig. 23.5).

That is we consider[14]

$$Y(r, t, T_C, T) = \frac{C(r, t, T_C, T)}{P(r, t, T_C)}, \tag{23.87}$$

the dynamics of which under $\tilde{\mathbb{P}}$ are given by (see Sect. 6.6 and recall that under $\tilde{\mathbb{P}}$ the dynamics for $P$ and $C$ are given respectively by Eqs. (23.15) and (23.85))

$$\frac{dY}{Y} = -\sigma_P(t, T_C)(\sigma_C - \sigma_P(t, T_C))dt + (\sigma_C - \sigma_P(t, T_C))d\tilde{z}. \tag{23.88}$$

Equation (23.88) may be rearranged to

$$\frac{dY}{Y} = (\sigma_C - \sigma_P(t, T_C))(d\tilde{z} - \sigma_P(t, T_C)dt).$$

---

[12]Make the identification $T \to T_C$, $\lambda \to -1$, $g[x(s), s] \to x(s)$, $x(s) \to r(s)$, $\lambda \to -1$, $f(x(T)) \to C(r(T_C), T_C)$.

[13]The measure $\mathbb{P}^*$ that we develop below is known as the forward measure because under this measure the instantaneous forward rate equals the expected future forward rate, as we show in Sect. 25.5.

[14]In the ensuing discussion we write the bond and option prices with their full functional dependence, e.g. $C(r, t, T_C, T)$ and $P(r, t, T_C)$ to give greater clarity.

Following the discussion in Sect. 20.1 we can define a new process

$$z^*(t) = \tilde{z}(t) - \int_0^t \sigma_P(u, T_C)du, \tag{23.89}$$

and a new measure $\mathbb{P}^*$ such that $z^*(t)$ is a Wiener process under this measure. Thus the dynamics for $Y$ become

$$\frac{dY}{Y} = (\sigma_C - \sigma_P(t, T_C))dz^*,$$

and it follows that $Y$ is a martingale under $\mathbb{P}^*$. Using $\mathbb{E}_t^*$ to denote expectations under $\mathbb{P}^*$, formed at time $t$, we can write

$$Y(r(t), t, T_C, T) = \mathbb{E}_t^*\Big[Y(r(T_C), T_C, T_C, T)\Big], \tag{23.90}$$

which upon use of the definition (23.87) becomes (keep in mind that $P(r(T_C), T_C, T_C) = 1$)

$$C(r(t), t, T_C, T) = P(r(t), t, T_C)\ \mathbb{E}_t^*\Big[C(r(T_C), T_C, T_C, T)\Big]. \tag{23.91}$$

The difference between the expressions (23.86) and (23.91) for the value of the bond option lies in the way the stochastic discounting is done. In (23.86), the stochastic discounting is done along each stochastic interest rate path from $t$ and $T_C$, and since these paths are stochastic this term must appear under the expectation operator. In (23.91) the discounting from $t$ to $T_C$ is done using the bond of maturity $T_C$, which is known to the investor at time $t$ and hence this term does not need to appear under the expectation operator. It sometimes turns out that the expectation operation in (23.91) can be calculated explicitly, as we shall see in Sect. 23.7 for the Hull–White and CIR models.

If it is necessary to evaluate the expectation in (23.91) by simulation then we will need the dynamics for $r$ under $\mathbb{P}^*$, but these are easily obtained by using (23.89) to replace $d\tilde{z}$ in (23.22) by $dz^* + \sigma_P(t, T_C)dt$ so that

$$dr = \big(\mu_r + \sigma_r(\sigma_P(t, T_C) - \lambda(r, t))\big)dt + \sigma_r dz^*. \tag{23.92}$$

It is also of interest to obtain the dynamics under $\mathbb{P}^*$ for the relative bond price

$$X(r, t, T_C, T) = \frac{P(r, t, T)}{P(r, t, T_C)}. \tag{23.93}$$

This follows by noting that under $\mathbb{P}^*$ we have

$$\frac{dP(t, T)}{P(t, T)} = \big(r(t) + \sigma_P(t, T)\,\sigma_P(t, T_C)\big)dt + \sigma_P(t, T)dz^*,$$

and using the results of Sect. 6.6, so that

$$\frac{dX}{X} = \big(\sigma_P(t,T) - \sigma_P(t,T_C)\big)dz^*. \tag{23.94}$$

## 23.7 Solving the Option Pricing Equation

In this section we apply the general spot interest rate pricing framework of Sect. 23.6 to two special models that have become well known in the literature, perhaps because they yield closed form solutions. First the Hull–White model, which assumes a Gaussian process for the spot interest rate. Second the CIR model which assumes a Feller or square root process for the spot interest rate. In both cases it is convenient to use the bond of option maturity as numeraire as just discussed in Sect. 23.6. In this way the option pricing formula is basically Black–Scholes in the Hull–White case or Black–Scholes like in the CIR case.

### *23.7.1 The Hull White Model*

We recall Eq. (23.52) that for the Hull–White model the spot interest rate dynamics under $\tilde{\mathbb{P}}$ are given by

$$dr = (\theta(t) - \kappa(t)r(t))dt + \sigma(t)d\tilde{z},$$

where $\theta(t)$ is defined at Eq. (23.43). In this case Eq. (23.82) becomes

$$\frac{\partial C}{\partial t} + (\theta(t) - \kappa(t)r(t))\frac{\partial C}{\partial r} + \frac{1}{2}\sigma^2(t)\frac{\partial^2 C}{\partial r^2} - r(t)C = 0, \tag{23.95}$$

subject to the boundary condition (23.83). In turns out that the solution to (23.95) can be very elegantly obtained by an application of the change of measure results of Chap. 20. Instead of using the money market account as the numeraire, it is more convenient to use the price of the pure discount bond $P(r,t,T_C)$ whose maturity date is $T_C$. We note from (23.49) with $T = T_C$ the volatility of the bond return for the Hull–White model is

$$\sigma_P(t,T_C) = -\sigma(t)\, b(t,T_C) \tag{23.96}$$

with $b(t,T)$ defined by (23.47). The expression for the value of the bond option under the measure $\mathbb{P}^*$ is given by (23.91) and we consider the specific case of a European call bond option so that

$$C(r(T_C),T_C,T_C,T) = \Big(P(r,T_C,T) - E\Big)^+.$$

Recalling the definition of the relative bond price, see Eq. (23.93),[15] then this payoff may be written

$$C(r(T_C), T_C, T_C, T) = \Big(X(T_C, T_C, T) - E\Big)^+.$$

Substituting (23.96) into (23.94) we find that the dynamics for $X$ become

$$\frac{dX}{X} = \sigma(t)[b(t, T_C) - b(t, T)]dz^*. \tag{23.97}$$

In terms of the relative bond price $X$ we can express (23.91) as

$$\frac{C(r, t, T_C, T)}{P(r, t, T_C)} = \mathbb{E}_t^*[(X(T_C, T_C, T) - E)^+]. \tag{23.98}$$

Since the expectation in (23.98) is with respect to outcomes for the $X$ variable, the relevant stochastic dynamics underlying the probability distribution in the calculation of $\mathbb{E}_t^*$ is the stochastic differential equation (23.97). We note from Eq. (23.97) that $dX/X$ is normally distributed under $\mathbb{P}^*$, with

$$\mathbb{E}_t^*\left[\frac{dX}{X}\right] = 0, \tag{23.99}$$

$$var^*\left[\frac{dX}{X}\right] = \sigma^2(t)[b(t, T_C) - b(t, T)]^2 dt \equiv v^2(t)dt. \tag{23.100}$$

The calculation of the expectation in (23.98) with driving dynamics (23.97) is simply the Black–Scholes European call option pricing problem with risk free rate of interest set to zero, exercise price $E$ and with time varying variance $v^2(t)$. Thus

$$\mathbb{E}_t^*[(X(T_C, T_C, T) - E)^+] = X(t, T_C, T)\mathcal{N}(d_1^*) - E\mathcal{N}(d_2^*), \tag{23.101}$$

where

$$d_1^* = \frac{\ln(X(t, T_C, T)/E) + \bar{v}^2(T_C - t)/2}{\bar{v}\sqrt{T_C - t}},$$

$$d_2^* = d_1^* - \bar{v}\sqrt{T_C - t},$$

$$\bar{v}^2 = \frac{1}{T_C - t}\int_t^{T_C} v^2(s)ds.$$

[15]Strictly speaking we should write $X(r, t, T_C, T)$. However it turns out (see Eq. (23.97) below) that the dynamics for $r$ do not enter directly into the dynamics for $X$.

Substituting (23.101) into (23.98), the final expression for the call option price is given by

$$C(r,t,T) = P(r,t,T_C)X\mathcal{N}(d_1) - EP(r,t,T_C)\mathcal{N}(d_2)$$
$$= P(r,t,T)\mathcal{N}(d_1) - EP(r,t,T_C)\mathcal{N}(d_2) \tag{23.102}$$

with $d_1$ given by

$$d_1 = \frac{ln\,(P(r,t,T)/P(r,t,T_C)E) + \bar{v}^2(T_C - t)/2}{\bar{v}\sqrt{T_C - t}}$$

and $d_2$ by

$$d_2 = d_1 - \bar{v}\sqrt{T_C - t}.$$

By the put-call parity condition, the corresponding put option price can similarly be expressed as

$$U(r,t,T_C,T) = EP(r,t,T_C)\mathcal{N}(-d_2) - P(r,t,T)\mathcal{N}(-d_1).$$

The structure of the option pricing formula (23.102) should be compared with (19.23), the one obtained for the Black–Scholes model with stochastic interest rates. One sees that they are identical in structure if one replaces the underlying traded asset (the stock $S$) of Chap. 19 with the underlying traded asset (the bond $P$) of the current situation.

### *23.7.2 The CIR Model*

In the case of the CIR model with the interest rate process given by (23.53), Eq. (23.82) becomes

$$\frac{\partial C}{\partial t} + [\kappa(t)(\gamma - r) - \lambda\sigma r(t)]\frac{\partial C}{\partial r} + \frac{1}{2}\sigma^2 r(t)\frac{\partial^2 C}{\partial r^2} - r(t)C = 0. \tag{23.103}$$

Equation (23.103) can also be solved by using the change of measure ideas of Chap. 20. The derivation follows exactly the same lines as in the previous subsection, the only difference is that now the bond price dynamics are given by Eq. (23.66). As a result the dynamics for $X$ are given by

$$\frac{dX}{X} = v(t)\sqrt{r(t)}dW^*, \tag{23.104}$$

where

$$v(t) = \sigma[b(t,T) - b(t,T_C)], \tag{23.105}$$

with $b(t,T)$ given by Eq. (23.61). The Kolmogorov equation associated with Eq. (23.104) is

$$\frac{1}{2}v^2(t)r(t)\frac{\partial^2\pi}{\partial r^2} + \frac{\partial\pi}{\partial t} = 0. \tag{23.106}$$

The probability density function arising from Eq. (23.106) is essentially given by (22.19) in the limit $\kappa \to 0$. Integration of the call option payoff with respect to this distribution yields the option price. For instance, if the boundary condition is given by (23.83) the expression for the option price turns out to be

$$C(r,t,T_C;T,K) = P(r,t,T)\chi^2\left(2r^*[\phi+\psi+B(T_C,T)];\frac{4\alpha\gamma}{\sigma^2},\frac{2\phi^2 re^{\xi(T_C-t)}}{\phi+\psi+B(T_C,T)}\right) \tag{23.107}$$

$$- EP(r,t,T_C)\chi^2\left(2r^*[\phi+\psi];\frac{4\alpha\gamma}{\sigma^2},\frac{2\phi^2 re^{\xi(T_C-t)}}{\phi+\psi}\right), \tag{23.108}$$

where

$$\xi \equiv ((\alpha+\lambda)^2 + 2\sigma^2)^{1/2},$$

$$\phi \equiv \frac{2\xi}{\sigma^2(e^{\xi(T_C-t)} - 1)},$$

$$\psi \equiv \frac{\alpha+\lambda+\xi}{\sigma^2},$$

$$r^* \equiv \frac{1}{B(T_C,T)}\left[\log\left(\frac{A(T_C,T)}{E}\right)\right],$$

$\chi^2(\cdot)$ is the noncentral chi-square distribution function and $r^*$ is the critical interest rate below which exercise will occur, namely that obtained by solving $E = P(r^*,T_C,T)$.

## 23.8 Rendering Spot Rate Models Preference Free-Calibration to the Currently Observed Yield Curve

Consider again the Hull–White model with interest rate dynamics under the historical measure $\mathbb{P}$ given by Eq. (23.41). We know either from Sect. 23.4.2 or from Sect. 23.5 that the bond price is given by

$$P(r,t,T) = e^{-a(t,T)-b(t,T)r(t)}, \tag{23.109}$$

where

$$b(t,T) = \int_t^T e^{\mathscr{K}(t)-\mathscr{K}(s)}ds, \quad \mathscr{K}(t) = \int_0^t \kappa(s)ds, \tag{23.110}$$

$$a(t,T) = \int_t^T b(s,T)\theta(s)ds - \frac{1}{2}\int_t^T \sigma(s)^2 b(s,T)^2 ds, \tag{23.111}$$

and

$$\theta(t) = \kappa(t)\gamma(t) - \lambda(t)\sigma(t). \tag{23.112}$$

In order to use this model we need estimates (from market data) for $\sigma(t)$, $\kappa(t)$ and $\theta(t)$. Note that $\theta(t)$ impounds in itself the functions $\gamma(t)$ and $\lambda(t)$, which do not therefore need to be separately estimated, at least for the purposes of pricing derivative securities.

We assume that we already have estimates of $\sigma(t)$ from the prices of interest rate caps using the corresponding option pricing formula (see Sect. 23.7), thus it only remains to determine $\kappa(t)$ and $\theta(t)$. We assume that we also have available market information on the volatility of bonds returns of all maturities at time 0. We know from Eqs. (23.13) and (23.109) that the volatility of bond returns is $-\sigma(t)b(t,T)$.[16] Thus we assume $\sigma(0)b(0,T)$ is given as a function of maturity $T$. Putting $t = 0$ in Eq. (23.83) we have

$$b(0,T) = \int_0^T e^{-\mathscr{K}(s)}ds. \tag{23.113}$$

Differentiating with respect to maturity $T$ yields

$$\mathscr{K}(T) = -\ln(\frac{\partial}{\partial T}b(0,T)),$$

[16]From Eq. (23.13) we have that var $\left(\frac{dP}{P}\right) = \left(\frac{\sigma_r}{P}\frac{\partial P}{\partial r}\right)^2 dt$ and from Eq. (23.109), $\frac{1}{P}\frac{\partial P}{\partial r} = -b(t,T)$.

and since $\mathscr{K}'(t) = \kappa(t)$, we obtain

$$\kappa(T) = -\frac{\partial}{\partial T}[\ln(\frac{\partial}{\partial T}b(0,T))]. \tag{23.114}$$

Next set $t = 0$ in the bond pricing equation so that

$$P(r(0), 0, T) = e^{-a(0,T)-b(0,T)r(0)}. \tag{23.115}$$

The function $P(r(0), 0, T)$ would be available from the currently observed yield curve. We consider (23.115) in the form

$$a(0, T) = -\ln P(r(0), 0, T) - b(0, T)r(0). \tag{23.116}$$

From the last equation $a(0, T)$ can be considered as known (from market data) as a function of $T$, we shall further assume that this function is sufficiently smooth to be at least twice differentiable. Thus our remaining task is to determine the function $\theta(t)$. We recall from Eq. (23.111) that

$$a(0, T) = \int_0^T b(s, T)\theta(s)ds - \frac{1}{2}\int_0^T \sigma^2(s)b(s, T)^2 ds. \tag{23.117}$$

The second term on the right hand side, perhaps via numerical integration, will simply be a known function of $T$. Thus Eq. (23.117) constitutes an integral equation for the unknown function $\theta$. By a process of successive differentiations we find that (see Appendix 23.2)

$$\begin{aligned}\theta(T) &= e^{-\mathscr{K}(T)}\frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}a(0,T)\right)\\ &\quad + e^{-\mathscr{K}(T)}\frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b(s,T)^2 ds\right)\right).\end{aligned} \tag{23.118}$$

Whilst Eq. (23.118) involves awkward looking algebraic expressions, its numerical evaluation would be a routine task. Consider the case where $\sigma$, $\kappa$ are constant, so that $\theta(t)$ is the only time varying parameter. Now we simply have $\mathscr{K}(t) = \kappa t$, and hence

$$b(t, T) = \int_t^T e^{\kappa t - \kappa s} ds = \frac{1}{\kappa}(1 - e^{\kappa(t-T)}), \tag{23.119}$$

from which

$$b(0, T) = \frac{1}{\kappa}(1 - e^{-\kappa T}). \tag{23.120}$$

Furthermore

$$a(t,T) = \int_t^T b(s,T)\theta(s)ds - \frac{\sigma^2}{2}\int_t^T b^2(s,T)ds, \tag{23.121}$$

and so

$$a(0,T) = \int_0^T b(s,T)\theta(s)ds - \frac{\sigma^2}{2}\int_0^T b^2(s,T)ds. \tag{23.122}$$

Differentiating (23.122) with respect to $T$ we obtain

$$\begin{aligned}\frac{\partial a(0,T)}{\partial T} &= \int_0^T \theta(s)\frac{\partial}{\partial T}\left(\frac{1}{\kappa}(1-e^{-\kappa(T-s)})\right)ds - \frac{\sigma^2}{2}\frac{\partial}{\partial T}\left(\int_0^T b^2(s,T)ds\right)\\ &= e^{-\kappa T}\int_0^T \theta(s)e^{\kappa s}ds - \frac{\sigma^2}{2}\frac{\partial}{\partial T}\left(\int_0^T b^2(s,T)ds\right).\end{aligned} \tag{23.123}$$

Now, using (23.119), Eq. (23.122) may be written

$$\begin{aligned}a(0,T) &= \int_0^T \frac{1}{\kappa}(1-e^{-\kappa(T-s)})\theta(s)ds - \frac{1}{2}\int_0^T \sigma^2 b^2(s,T)ds \\ &= \frac{1}{\kappa}\int_0^T \theta(s)ds - \frac{e^{-\kappa T}}{\kappa}\int_0^T e^{\kappa s}\theta(s)ds - \frac{1}{2}\int_0^T \sigma^2 b^2(s,T)ds.\end{aligned} \tag{23.124}$$

Using (23.124) to eliminate the $e^{-\mathscr{K}T}\int_0^T \theta(s)e^{\mathscr{K}s}ds$ term in (23.123) we obtain

$$\begin{aligned}\frac{\partial a(0,T)}{\partial T} =& -\kappa a(0,T) + \int_0^T \theta(s)ds - \frac{\kappa\sigma^2}{2}\int_0^T b^2(s,T)ds \\ & - \frac{\sigma^2}{2}\frac{\partial}{\partial T}\left(\int_0^T b^2(s,T)ds\right),\end{aligned}$$

which upon re-arrangement yields

$$\begin{aligned}\int_0^T \theta(s)ds = \frac{\partial a(0,T)}{\partial T} + \kappa a(0,T) + \frac{\sigma^2}{2}\Bigg[\kappa\int_0^T b^2(s,T)ds \\ + \frac{\partial}{\partial T}\left(\int_0^T b^2(s,T)ds\right)\Bigg].\end{aligned}$$

Differentiating the last equation with regard to $T$ yields,

$$\theta(T) = \frac{\partial}{\partial T}\left[\frac{\partial a(0,T)}{\partial T} + \kappa a(0,T)\right]$$

$$+\frac{\sigma^2}{2}\frac{\partial}{\partial T}\left[\kappa\int_0^T b^2(s,T)ds+\frac{\partial}{\partial T}\left(\int_0^T b^2(s,T)ds\right)\right].$$

With the function $\theta(T)$ now at our disposal we can compute the time function $a(t,T)$ and hence bond prices calibrated to market data. In the case of the CIR model, the steps taken to calibrate the model to the initial yield curve and cap and swaption data for example are similar to the Hull–White model. We do not provide details here.

## 23.9 Appendix

### Appendix 23.1 Solution of the Ordinary Differential Equations (23.59) and (23.60)

Consider the ordinary differential equation

$$\frac{db}{dt}=\alpha_0 b^2+\alpha_1 b-1=\alpha_0\left[b^2+\frac{\alpha_1}{\alpha_0}b-\frac{1}{\alpha_0}\right]. \tag{23.125}$$

where $\alpha_0$ and $\alpha_1$ are constants. The quadratic in the brackets on the RHS can be factorised as

$$b^2+\frac{\alpha_1}{\alpha_0}b-\frac{1}{\alpha_0}=(b-\phi_1)(b-\phi_2),$$

where

$$\begin{aligned}\phi_1&=-\frac{\alpha_1}{2\alpha_0}+\frac{\beta}{2\alpha_0},\\ \phi_2&=-\frac{\alpha_1}{2\alpha_0}-\frac{\beta}{2\alpha_0},\\ \beta&=\sqrt{\alpha_1^2+4\alpha_0}.\end{aligned} \tag{23.126}$$

Thus the ordinary differential equation (23.125) can be written

$$\frac{db}{dt}=\alpha_0(b-\phi_1)(b-\phi_2), \tag{23.127}$$

or as

$$\frac{db}{(b-\phi_1)(b-\phi_2)}=\alpha_0 dt.$$

With some slight re-arrangement the last equation can be written

$$\left[\frac{1}{b-\phi_1} - \frac{1}{b-\phi_2}\right] db = \alpha_0(\phi_1 - \phi_2)dt = \beta dt. \tag{23.128}$$

Integrating the last equation from $t$ to $T$ we obtain

$$\left[\ln\left(\frac{b-\phi_1}{b-\phi_2}\right))\right]_t^T = \beta(T-t), \tag{23.129}$$

i.e.

$$\ln\left(\frac{b(T,T)-\phi_1}{b(T,T)-\phi_2}\right) - \ln\left(\frac{b(t,T)-\phi_1}{b(t,T)-\phi_2}\right) = \beta(T-t),$$

which on making use of $b(T,T) = 0$ becomes

$$\ln\left(\frac{b(t,T)-\phi_1}{b(t,T)-\phi_2}\right) = \ln\left(\frac{\phi_1}{\phi_2}\right) - \beta(T-t),$$

i.e.

$$\frac{b(t,T)-\phi_1}{b(t,T)-\phi_2} = \exp\left[\ln(\frac{\phi_1}{\phi_2}) - \beta(T-t)\right] = \frac{\phi_1}{\phi_2}e^{-\beta(T-t)}.$$

Solving the last equation for $b(t,T)$ we obtain

$$b(t,T) = \frac{\phi_1\phi_2(1-e^{-\beta(T-t)})}{\phi_2 - \phi_1 e^{-\beta(T-t)}}. \tag{23.130}$$

Using the fact that $\phi_1\phi_2 = -1/\alpha_0$ this simplifies slightly to

$$b(t,T) = \frac{1}{\alpha_0}\frac{[1-e^{-\beta(T-t)}]}{[\phi_1 e^{-\beta(T-t)} - \phi_2]}. \tag{23.131}$$

To solve (23.59), we set $\alpha_0 = \sigma^2/2, \alpha_1 = \alpha + \lambda$, so that we obtain

$$b(t,T) = \frac{2}{\sigma^2}\frac{[1-e^{-\beta(T-t)}]}{[\phi_1 e^{-\beta(T-t)} - \phi_2]}, \tag{23.132}$$

where

$$\phi_1 = -\frac{(\alpha+\lambda)}{\sigma^2} + \frac{\beta}{\sigma^2},$$

$$\phi_2 = -\frac{(\alpha+\lambda)}{\sigma^2} - \frac{\beta}{\sigma^2}, \tag{23.133}$$
$$\beta = \sqrt{(\alpha+\lambda)^2 + 2\sigma^2}.$$

Next from Eq. (23.62) of Sect. 23.4.3

$$a(t,T) = \alpha\gamma \int_t^T b(s,T)ds.$$

Making the transformation $u = \beta(T-s)$ we see that

$$a(t,T) = \frac{+\alpha\gamma}{\beta} \int_0^{\beta(T-t)} b(T - \frac{u}{\beta}, T)du.$$

Substituting the expression (23.132) for $b(t,T)$ (and setting $\tau = T-t$) we obtain

$$a(t,T) = +\frac{\alpha\gamma}{\beta}\frac{2}{\sigma^2}\int_0^{\beta\tau} \frac{(1-e^{-u})}{(\phi_1 e^{-u} - \phi_2)}du.$$

Consider the integral

$$\begin{aligned}
I &= \int_0^{\beta\tau} \left(\frac{1-e^{-u}}{\phi_1 e^{-u} - \phi_2}\right) du \\
&= \int_0^{\beta\tau} \frac{du}{\phi_1 e^{-u} - \phi_2} - \int_0^{\beta\tau} \frac{e^{-u}du}{\phi_1 e^{-u} - \phi_2} \\
&= \int_0^{\beta\tau} \frac{e^{u}du}{\phi_1 - \phi_2 e^{u}} - \int_0^{\beta\tau} \frac{e^{-u}du}{\phi_1 e^{-u} - \phi_2} \\
&= \left[\frac{-1}{\phi_2}\ln(\phi_1 - \phi_2 e^{u})\right]_0^{\beta\tau} + \left[\frac{1}{\phi_1}\ln(\phi_1 e^{-u} - \phi_2)\right]_0^{\beta\tau} \\
&= \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\ln(\phi_1-\phi_2) - \frac{1}{\phi_2}\ln(\phi_1 - \phi_2 e^{\beta\tau}) + \frac{1}{\phi_1}\ln(\phi_1 e^{-\beta\tau} - \phi_2) \\
&= \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\ln(\phi_1-\phi_2) - \frac{1}{\phi_2}\ln(\phi_1 - \phi_2 e^{\beta\tau}) + \frac{1}{\phi_1}\ln[e^{-\beta\tau}/(\phi_1 - \phi_2 e^{\beta\tau})] \\
&= -\frac{\beta\tau}{\phi_1} + \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\left\{\ln(\phi_1-\phi_2) - \ln(\phi_1 - \phi_2 e^{\beta\tau})\right\}.
\end{aligned}$$

Finally

$$I = -\frac{\beta}{\phi_1}\tau - \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\ln\left(\frac{\phi_1 - \phi_2 e^{\beta\tau}}{\phi_1-\phi_2}\right),$$

and so

$$a(t,T) = \frac{2\alpha\gamma}{\beta\sigma^2}\left[\frac{-\beta(T-t)}{\phi_1} - \frac{(\phi_1-\phi_2)}{\phi_1\phi_2}\ln(\frac{\phi_1-\phi_2 e^{\beta(T-t)}}{\phi_1-\phi_2})\right].$$

Using the fact that $\phi_1\phi_2 = -2/\sigma^2$ and $\phi_1 - \phi_2 = 2\beta/\sigma^2$ we finally obtain

$$a(t,T) = \frac{2\alpha\gamma}{\sigma^2}\left[-\frac{(T-t)}{\phi_1} + \ln\left(\frac{\phi_1-\phi_2 e^{\beta(T-t)}}{\phi_1-\phi_2}\right)\right]. \tag{23.134}$$

When allowing for time-varying coefficients, the steps leading to (23.128) remain the same as in the constant coefficients case, only now $\phi_1$, $\phi_2$, $\sigma$ and $\beta$ become functions of time. Thus in order to use the same functional form for the solution we need to define

$$\bar{\beta}(t,T) = \frac{1}{T-t}\int_t^T \beta(s)ds.$$

Thus integration of (23.128) will yield (23.132) with $\beta$ replaced by $\bar{\beta}(t,T)$.

## Appendix 23.2 Calculating $\theta(T)$ in the Calibration of the Hull–White Model

From Eq. (23.110) we note that

$$\frac{\partial b}{\partial T} = e^{\mathscr{K}(t)-\mathscr{K}(T)}.$$

Differentiating Eq. (23.117) with respect to $T$ yields

$$\begin{aligned}\frac{\partial}{\partial T}a(0,T) &= b(T,T)\theta(T) + \int_0^T e^{\mathscr{K}(s)-\mathscr{K}(T)}\theta(s)ds - \frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b^2(s,T)ds\right) \\ &= e^{-\mathscr{K}(T)}\int_0^T e^{\mathscr{K}(s)}\theta(s)ds - \frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b^2(s,T)ds\right).\end{aligned}$$

Rearranging

$$e^{\mathscr{K}(T)}\frac{\partial}{\partial T}a(0,T) = \int_0^T e^{\mathscr{K}(s)}\theta(s)ds - e^{\mathscr{K}(T)}\frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b^2(s,T)ds\right).$$

Differentiating again with respect to $T$ we obtain

$$\frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}a(0,T)\right) = e^{\mathscr{K}(T)}\theta(T) - \frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b^2(s,T)ds\right)\right).$$

Rearranging this last equation we obtain

$$\theta(T) = e^{-\mathscr{K}(T)}\frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}a(0,T)\right) + e^{-\mathscr{K}(T)}\frac{\partial}{\partial T}\left(e^{\mathscr{K}(T)}\frac{\partial}{\partial T}\left(\frac{1}{2}\int_0^T \sigma^2(s)b^2(s,T)ds\right)\right).$$

## 23.10 Problems

**Problem 23.1** Equation (23.15) shows that under the equivalent martingale measure the bond price dynamics are given by

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + \sigma_P(r,t,T)d\tilde{z}(t)$$

with $\sigma_P(r,t,T)$ defined by (23.5). Use Ito's lemma to show that the yield to maturity

$$\rho(t,T) = -\frac{1}{(T-t)}\ln P(t,T)$$

satisfies

$$d\rho(t,T) = \frac{1}{(T-t)}\left[\left(\rho(t,T) + \frac{1}{2}\sigma_P^2(r,t,T) - r(t)\right)dt - \sigma_P(r,t,T)d\tilde{z}(t)\right].$$

**Problem 23.2** Recall the relation (22.9) between the bond price $P(t,T)$ and the instantaneous forward rate $f(t,T)$. By considering the quantity

$$B(t,T) = \ln P(t,T)$$

and assuming that

$$d\left(\frac{\partial B(t,T)}{\partial T}\right) = \frac{\partial}{\partial T}dB(t,T),$$

show that $f(t, T)$ satisfies the stochastic differential equations

$$df(t, T) = -\sigma(t, T)\sigma_P(t, T)dt + \sigma(t, T)d\tilde{z}$$

where

$$\sigma(t, T) = -\frac{\partial}{\partial T}\sigma_P(t, T).$$

**Problem 23.3** Consider the Hull–White model of Sect. 23.4.2, for which the bond price $P(t, T)$ is given by (23.25) with $a(t, T)$ and $b(t, T)$ given by Eqs. (23.47) and (23.48). Obtain the corresponding expression for the forward rate $f(t, T)$. Show what this expression reduces to in the special case of the Vasicek model.

# Chapter 24
# Interest Rate Derivatives: Multi-Factor Models

**Abstract** In this chapter we develop a framework for term structure modelling that allows factors other than the instantaneous spot rate itself to influence the evolution of the term structure of interest rates. The framework allows for multi-factor generalisations of the Hull–White model as well as of the CIR model. First we present a two-factor extension of the Hull–White model. Then we develop a general multi-factor term structure model and the corresponding bond option pricing model. Finally as a specific application, we consider the so called Duffie–Kan affine class of term structure models, which is widely applied in practice.

## 24.1 Hull–White Two-Factor Model

Hull and White (1994) introduced their two-factor model to allow their extended Vasicek model to better calibrate to market data. The basic idea of the Hull–White two-factor model is to add to the drift of the Hull–White one factor model a mean-reverting-to-zero stochastic process. Since this second factor is also Gaussian distributed it is still possible to obtain Black Scholes type option pricing formulae. Thus the instantaneous spot rate $r(t)$ is assumed to follow the process

$$dr = \kappa(t)(\gamma(t) + h(t) - r(t))dt + \sigma_1(t)dz_1, \tag{24.1}$$

where the additional underlying factor $h(t)$ satisfies

$$dh = -\beta h(t)dt + \sigma_2(t)dz_2, \tag{24.2}$$

and the Wiener increments $dz_1, dz_2$ are correlated , i.e.

$$\mathbb{E}_t[dz_1 dz_2] = \rho dt. \tag{24.3}$$



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_24

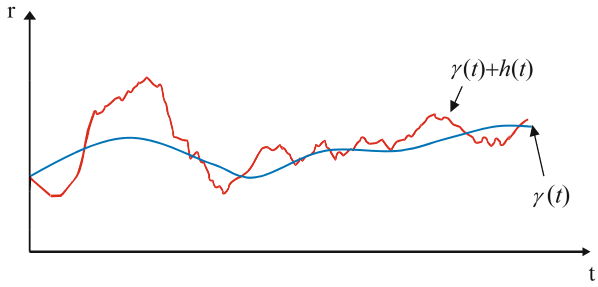


**Fig. 24.1** The noisy long run short rate of the Hull–White two factor model

The stochastic differential equation system (24.1) and (24.2) may be expressed in terms of independent Wiener processes $W_1(t)$, $W_2(t)$ as

$$dr = \kappa(\gamma + h(t) - r(t))dt + \sigma_1\sqrt{1-\rho^2}dW_1 + \sigma_1\rho dW_2, \tag{24.4}$$

$$dh = -\beta h(t)dt + \sigma_2 dW_2. \tag{24.5}$$

The factor $h(t)$ essentially introduces a noisy long run short rate, as illustrated in Fig. 24.1. Furthermore it introduces an extra source of risk, in the form of the extra Wiener process (at least for $\rho \neq 1$).

### *24.1.1 Bond Price*

The bond price is now a function of the two underlying factors $r$ and $h$ and will be denoted as $P(r,h,t,T)$. By Ito's Lemma

$$\frac{dP}{P} = \mu_P dt + \sigma_{P_1}dW_1 + \sigma_{P_2}dW_2, \tag{24.6}$$

where

$$\mu_P = \frac{1}{P}\left(\frac{\partial P}{\partial t} + \kappa(\gamma + h(t) - r(t))\frac{\partial P}{\partial r} - \beta h(t)\frac{\partial P}{\partial h} + \mathscr{D}P\right), \tag{24.7}$$

$$\sigma_{P_1} = \frac{1}{P}\sigma_1\sqrt{1-\rho^2}\frac{\partial P}{\partial r}, \tag{24.8}$$

$$\sigma_{P_2} = \frac{1}{P}\left(\sigma_1\rho\frac{\partial P}{\partial r} + \sigma_2\frac{\partial P}{\partial h}\right), \tag{24.9}$$

and

$$\mathscr{D}P = \frac{1}{2}\sigma_1^2\frac{\partial^2 P}{\partial r^2} + \rho\sigma_1\sigma_2\frac{\partial^2 P}{\partial r\partial h} + \frac{1}{2}\sigma_2^2\frac{\partial^2 P}{\partial h^2}.$$

We set up a hedging portfolio as before, except that we now need bonds of three maturities $T_1, T_2$ and $T_3$ because of the extra source of risk $W_2$. If initial $1 investment is spread in the proportions $Q_1, Q_2$ and $Q_3$ amongst the three bonds, then the return on the hedging portfolio is given by

$$\left.\begin{array}{l}\text{Return on}\\ \text{hedging portfolio}\end{array}\right\} \begin{aligned}&= Q_1\frac{dP_1}{P_1} + Q_2\frac{dP_2}{P_2} + Q_3\frac{dP_3}{P_3}\\ &= (Q_1\mu_P^{(1)} + Q_2\mu_P^{(2)} + Q_3\mu_P^{(3)})dt\\ &\quad + (Q_1\sigma_{P_1}^{(1)} + Q_2\sigma_{P_1}^{(2)} + Q_3\sigma_{P_1}^{(3)})dW_1\\ &\quad + (Q_1\sigma_{P_2}^{(1)} + Q_2\sigma_{P_2}^{(2)} + Q_3\sigma_{P_2}^{(3)})dW_2.\end{aligned}$$

Here we use superscripts to denote maturity. We choose the proportions $Q_1$, $Q_2$ and $Q_3$ to eliminate the stochastic Wiener increment terms, thus

$$Q_1\sigma_{P_1}^{(1)} + Q_2\sigma_{P_1}^{(2)} + Q_3\sigma_{P_1}^{(3)} = 0, \tag{24.10}$$

$$Q_1\sigma_{P_2}^{(1)} + Q_2\sigma_{P_2}^{(2)} + Q_3\sigma_{P_2}^{(3)} = 0. \tag{24.11}$$

The condition that the hedge portfolio can then only earn the risk free rate is expressed as

$$Q_1\mu_P^{(1)} + Q_2\mu_P^{(2)} + Q_3\mu_P^{(3)} = r(t)$$

or, using $Q_1 + Q_2 + Q_3 = 1$,

$$Q_1(\mu_P^{(1)} - r(t)) + Q_2(\mu_P^{(2)} - r(t)) + Q_3(\mu_P^{(3)} - r(t)) = 0. \tag{24.12}$$

The system (24.10)–(24.12) can only have non-zero solution for $Q_1, Q_2, Q_3$ if for each maturity there exist functions $\lambda_1(r,h,t)$ and $\lambda_2(r,h,t)$ (independent of maturity) such that

$$\mu_P^{(i)} - r(t) = \lambda_1(r,h,t)\sigma_{P_1}^{(i)} + \lambda_2(r,h,t)\sigma_{P_2}^{(i)}. \tag{24.13}$$

Since the maturities were chosen arbitrarily a relation such as (24.13) must hold for any maturity, thus we write (dropping the arguments of $\lambda_1$ and $\lambda_2$)

$$\mu_P - r(t) = \lambda_1(r,h,t)\sigma_{P_1} + \lambda_2(r,h,t)\sigma_{P_2}. \tag{24.14}$$

By use of expressions (24.7)–(24.9), Eq. (24.14) becomes the partial differential equation

$$\mathscr{D}P + [\theta_r + \kappa h(t) - \kappa r(t)]\frac{\partial P}{\partial r} + [\theta_h - \beta h(t)]\frac{\partial P}{\partial h} - r(t)P + \frac{\partial P}{\partial t} = 0 \tag{24.15}$$

where we set

$$\theta_r = \kappa\gamma - \lambda_1\sigma_1\sqrt{1-\rho^2} - \lambda_2\sigma_1\rho \quad \text{and} \quad \theta_h = -\lambda_2\sigma_2.$$

Equation (24.15) must be solved subject to

$$P(r(T), h(T), T, T) = 1. \tag{24.16}$$

Substituting (24.14) into (24.6), the bond price dynamics in the arbitrage-free economy may be written

$$\frac{dP}{P} = r(t)dt + \sigma_{P_1}d\tilde{W}_1 + \sigma_{P_2}d\tilde{W}_2, \tag{24.17}$$

where

$$\tilde{W}_i(t) = W_i(t) + \int_0^t \lambda_i(r(s), h(s), s)ds \quad (i = 1, 2). \tag{24.18}$$

Again appealing to Girsanov's theorem we can find an equivalent $\tilde{\mathbb{P}}$ under which the $\tilde{W}_i(t)$ become Wiener processes. Using the now standard argument we can derive from (24.17) that

$$P(r, h, t, T) = \tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(s)ds\right)\right], \tag{24.19}$$

where the dynamics driving the measure $\tilde{\mathbb{P}}$ are

$$dr = [\theta_r + \kappa h(t) - \kappa r(t)]dt + \sigma_1\sqrt{1-\rho^2}d\tilde{W}_1 + \sigma_1\rho d\tilde{W}_2,$$
$$dh = (\theta_h - \beta h(t))dt + \sigma_2 d\tilde{W}_2.$$

In the partial differential equation (24.15) we note the linearity of the coefficients of $\frac{\partial P}{\partial r}$ and $\frac{\partial P}{\partial h}$ in terms of $r$ and $h$, and so seek a solution of the form

$$P(r, h, t, T) = \exp(-a(t, T) - b_r(t, T)r(t) - b_h(t, T)h(t)), \tag{24.20}$$

where

$$a(T,T) = b_r(T,T) = b_h(T,T) = 0. \tag{24.21}$$

Collecting terms in $r(t)$, $h(t)$ and the constant term we obtain

$$\left[-\frac{db_r}{dt} + \kappa b_r - 1\right] r(t) + \left[-\kappa b_r + \beta b_h - \frac{db_h}{dt}\right] h(t)$$
$$+\left[-\frac{da}{dt} - \theta_r b_r - \theta_h b_h + \frac{1}{2}\sigma_1^2 b_r^2 + \rho\sigma_1\sigma_2 b_r b_h + \frac{1}{2}\sigma_2^2 b_h^2\right] = 0.$$

Thus we obtain three ordinary differential equations for $b_r$, $b_h$ and $a$, namely

$$\frac{db_r}{dt} = \kappa b_r - 1, \tag{24.22}$$

$$\frac{db_h}{dt} = \beta b_h - \kappa b_r, \tag{24.23}$$

$$\frac{da}{dt} = -\theta_r b_r - \theta_h b_h + \frac{1}{2}\sigma_1^2 b_r^2 + \rho\sigma_1\sigma_2 b_r b_h + \frac{1}{2}\sigma_2^2 b_h^2. \tag{24.24}$$

Equations (24.22)–(24.24) may be solved sequentially for $b_r(t,T)$, $b_h(t,T)$ and $a(t,T)$. Considering the special case when the coefficients $\kappa$, $\gamma$, $\beta$, $\sigma_1$ and $\sigma_2$ are all constant the solutions for $b_r$ and $b_h$ are readily obtained. Thus

$$b_r(t,T) = \frac{[1 - e^{-\kappa(T-t)}]}{\kappa}, \tag{24.25}$$

and

$$b_h(t,T) = \frac{1}{\beta}[1 - e^{-\beta(T-t)}] + \frac{1}{\kappa - \beta}[e^{-\kappa(T-t)} - e^{-\beta(T-t)}]. \tag{24.26}$$

From Eq. (24.24) the function $a(t,T)$ is obtained from the integration

$$a(t,T) = \int_t^T \left[\theta_r b_r(s,T) + \theta_h b_h(s,T) - \frac{1}{2}\sigma_1^2 b_r(s,T)^2 \right.$$
$$\left. - \rho\sigma_1\sigma_2 b_r(s,T) b_h(s,T) - \frac{1}{2}\sigma_2^2 b_h(s,T)^2\right] ds.$$

### 24.1.2 Option Prices

The bond option price in the Hull–White two-factor model will also be a function of $r$ and $h$, so we denote it as $C(r, h, t)$. By Ito's Lemma its dynamics will be

$$\frac{dC}{C} = \mu_C dt + \sigma_{C_1} dW_1 + \sigma_{C_2} dW_2, \tag{24.27}$$

where

$$\mu_C = \frac{1}{C}\left(\frac{\partial C}{\partial t} + \kappa(\gamma + h(t) - r(t))\frac{\partial C}{\partial r} - \beta h(t)\frac{\partial C}{\partial h} + \mathscr{D}C\right), \tag{24.28}$$

$$\sigma_{C_1} = \frac{1}{C}\sigma_1\sqrt{1-\rho^2}\frac{\partial C}{\partial r}, \tag{24.29}$$

$$\sigma_{C_2} = \frac{1}{C}\sigma_1\rho\frac{\partial C}{\partial r} + \frac{1}{C}\sigma_2\frac{\partial C}{\partial h}. \tag{24.30}$$

The hedging argument follows exactly as in Sect. 23.6 except that the hedging portfolio now consists of the underlying bond of maturity $T$ and two options of maturity $T_C^{(1)}$ and $T_C^{(2)}$ both written on the underlying bond. Furthermore it is assumed that bonds are priced so that there exist no riskless arbitrage opportunities in the bond market. The no-riskless arbitrage condition in the option market now becomes

$$\mu_P - r(t) = \lambda_1(r, h, t)\sigma_{P_1} + \lambda_2(r, h, t)\sigma_{P_2} \tag{24.31}$$

and

$$\mu_C^{(i)} - r(t) = \lambda_1(r, h, t)\sigma_{C_1}^{(i)} + \lambda_2(r, h, t)\sigma_{C_2}^{(i)} \quad (i = 1, 2) \tag{24.32}$$

where the superscript $i$ refers to the option of maturity $T_C^{(i)}$. Since this maturity was chosen arbitrarily a relation of the form (24.32) must hold for an option of any general maturity $T_C$, thus

$$\mu_C - r(t) = \lambda_1(r, h, t)\sigma_{C_1} + \lambda_2(r, h, t)\sigma_{C_2}. \tag{24.33}$$

Upon use of (24.28)–(24.30) Eq. (24.33) has the familiar interpretation that under the condition of no-riskless arbitrage opportunities options are priced so that their expected excess return equals the risk premium term determined by the market price of risk associated with each source of uncertainty. Equation (24.31) merely restates the corresponding condition for the underlying bond already expressed by (24.14). Equation (24.33) becomes the partial differential equation

$$\mathscr{D}C + [\theta_r + \kappa h(t) - \kappa r(t)]\frac{\partial C}{\partial r} + [\theta_h - \beta h(t)]\frac{\partial C}{\partial h} - r(t)C + \frac{\partial C}{\partial t} = 0, \tag{24.34}$$

which must be solved subject to the option payoff condition e.g. for a European call

$$C(r(T_C), h(T_C), T_C) = (P(r(T_C), h(T_C), T_C, T) - E)^+. \tag{24.35}$$

As in the case of the Hull–White one-factor model, it is easier to solve (24.34) by using the change of numeraire idea. Substituting (24.31) into (24.27) we find that in the arbitrage free economy the option dynamics are given by

$$\frac{dC}{C} = r(t)dt + \sigma_{C_1} d\tilde{W}_1 + \sigma_{C_2} d\tilde{W}_2, \tag{24.36}$$

where the $\tilde{W}_i$ are Wiener processes under the equivalent measure $\tilde{\mathbb{P}}$ introduced in Sect. 24.1.1. We use the bond whose maturity is option maturity $T_C$, i.e. $P(r, h, t, T_C)$ as the numeraire and form, respectively, the relative bond and option prices

$$X(t, T_C, T) = \frac{P(r, h, t, T)}{P(r, h, t, T_C)}, \tag{24.37}$$

and

$$Y(t) = \frac{C(r, h, t)}{P(r, h, t, T_C)}. \tag{24.38}$$

We note that by virtue of the solution (24.18) of the bond pricing equation, we have

$$\frac{\partial P}{\partial r} = -b_r P \quad \text{and} \quad \frac{\partial P}{\partial h} = -b_h P,$$

so the bond price dynamics under $\tilde{\mathbb{P}}$ may be written

$$\frac{dP}{P} = r(t)dt - \sigma_1 \sqrt{1-\rho^2} b_r d\tilde{W}_1 - (\sigma_1 \rho b_r + \sigma_2 b_h) d\tilde{W}_2, \tag{24.39}$$

where the $\tilde{W}_i$ have been defined at Eq. (24.18). By an application of Ito's Lemma we find that the dynamics for $X$ are given by

$$\begin{aligned}\frac{dX}{X} =& \sigma_1 \sqrt{1-\rho^2}[b_r(t, T_C) - b_r(t, T)]dW_1^* \\ &+ \{\sigma_1 \rho [b_r(t, T_C) - b_r(t, T)] + \sigma_2 [b_h(t, T_C) - b_h(t, T)]\} dW_2^*,\end{aligned} \tag{24.40}$$

where we set

$$W_1^*(t) = \tilde{W}_1(t) + \int_0^t \sigma_1 \sqrt{1-\rho^2} b_r(s, T_C) ds,$$

$$W_2^*(t) = \tilde{W}_2(t) + \int_0^t [\sigma_1 \rho b_r(s, T_C) + \sigma_2 b_h(s, T_C)] ds.$$

By Girsanov's theorem we can find an equivalent measure $\mathbb{P}^*$ under which $W_1^*$, $W_2^*$ are Wiener processes. Since the coefficients in (24.40) are only deterministic time functions, we can find a Wiener process $W^*$ such that

$$\frac{dX}{X} = v_X(t) dW^* \tag{24.41}$$

where

$$\begin{aligned} v_X^2(t) =& \sigma_1^2(1-\rho^2)[b_r(t, T_C) - b_r(t, T)]^2 \\ &+ \{\sigma_1 \rho [b_r(t, T_C) - b_r(t, T)] + \sigma_2 [b_h(t, T_C) - b_h(t, T)]\}^2. \end{aligned} \tag{24.42}$$

Application of Ito's Lemma (see Sect. 6.6 on the quotient of two diffusions) to (24.38) gives the dynamics for $Y$ as

$$\frac{dY}{Y} = (\sigma_{C_1} + \sigma_1 \sqrt{1-\rho^2} b_r) dW_1^* + (\sigma_{C_2} + \sigma_1 \rho b_r + \sigma_2 b_h) dW_2^*. \tag{24.43}$$

Under the $\mathbb{P}^*$ measure $Y$ is a martingale so that

$$Y(t) = \mathbb{E}_t^*[Y(T_C)]. \tag{24.44}$$

Equation (24.44) may also be expressed as

$$\frac{C(r, h, t)}{P(r, h, t, T_C)} = \mathbb{E}_t^*[(P(r, h, T_C, T) - E)^+], \tag{24.45}$$

or, in terms of the relative bond price $X$ as,

$$\frac{C(r, h, t)}{P(r, h, t, T_C)} = \mathbb{E}_t^*[(X(T_C, T_C, T) - E)^+]. \tag{24.46}$$

The calculation of (24.46) is identical to the calculation of (23.98), with the exception that the quantity $v^2(t)$ defined at (23.100) is replaced by the quantity $v_X^2(t)$ defined at Eq. (24.42). The Black–Scholes pricing formula at (23.101) still applies, with $v_X(t)$ replacing $v(t)$. We note that this change gives the model some additional flexibility, via the $\rho$, $\sigma_2$, and $\beta$, to calibrate the model to the market observed data.

## 24.2 The General Framework

We shall denote the row vector of $n$ underlying factors by $\mathbf{X} = (X_1, X_2, \cdots, X_n)$. These underlying factors may or may not be directly interpretable as economic quantities. We shall discuss some specific examples later.

Typically the instantaneous spot interest rate, $r$, itself may or may not be one of the factors. In cases that it is not, it will usually be determined as some function of the underlying factors. We shall see below how $r$ is determined as a function of the arbitrage free bond price. We assume that the dynamics of the underlying factors follow diffusion process given by

$$dX_i = \mu_i(\mathbf{X}, t)dt + \sum_{j=1}^{m} \sigma_{ij}(\mathbf{X}, t)dW_j(t) \tag{24.47}$$

for $i = 1, 2, \cdots, n$, where the $W_j(t)$ are independent Wiener process. Here we allow the vector of drift coefficients and matrix of diffusion coefficients to possibly depend on all of the underlying factors.

### *24.2.1 Bond Pricing*

Since the bond evolution may depend on all $n$ underlying factors we denote its price at time $t$ by $P(\mathbf{X}(t), t, T)$. By Ito's Lemma the bond price dynamics are given by

$$\frac{dP}{P} = \mu_P(\mathbf{X}, t)dt + \sum_{k=1}^{m} \sigma_{P_k}(\mathbf{X}, t)dW_k, \tag{24.48}$$

where

$$\mu_P(\mathbf{X}, t) = \frac{1}{P}\left(\frac{\partial P}{\partial t} + \sum_{j=1}^{n} \mu_j \frac{\partial P}{\partial X_j} + \mathscr{D}P\right), \tag{24.49}$$

$$\sigma_{P_k}(\mathbf{X}, t) = \frac{1}{P}\sum_{j=1}^{n} \sigma_{jk}(\mathbf{X}, t)\frac{\partial P}{\partial X_j}, \tag{24.50}$$

with

$$\mathscr{D}P = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{l=1}^{m} \sigma_{il}\sigma_{jl}\right)\frac{\partial^2 P}{\partial X_i \partial X_j}. \tag{24.51}$$

We assume that in this economy the only traded instruments are pure discount bonds of various maturities. Thus our investor now invests \$1 in a hedge portfolio containing $(m+1)$ bonds of maturities $T_0, T_1, \cdots, T_m$ in the dollar amounts $Q_0, Q_1, \cdots, Q_m$. As previously we use $P_i$ to denote the price of the bond maturity at time $T_i$ and $\mu_P^{(i)}, \sigma_{P_k}^{(i)}$ the corresponding vectors of drift and diffusion coefficients. The dollar return on the hedge portfolio is given by

$$\sum_{i=0}^{m} Q_i \frac{dP_i}{P_i} = \left( \sum_{i=0}^{m} Q_i \mu_P^{(i)} \right) dt + \sum_{k=1}^{m} \left( \sum_{i=0}^{m} Q_i \sigma_{P_k}^{(i)} \right) dW_k. \tag{24.52}$$

The stochastic terms are eliminated by choosing proportions $Q_0, Q_1, \cdots, Q_m$ so that

$$\sum_{i=0}^{m} Q_i \sigma_{P_k}^{(i)} = 0, \tag{24.53}$$

for $k = 1, 2, \cdots, m$. The hedge portfolio is then riskless and must thus only earn the instantaneous risk-free rate $r(t)$, i.e.

$$\sum_{i=0}^{m} Q_i \mu_P^{(i)} = r(t). \tag{24.54}$$

Since the original investment is \$1 the proportions satisfy $\sum_{i=0}^{m} Q_i = 1$, so that the last equation becomes

$$\sum_{i=0}^{m} Q_i \left( \mu_P^{(i)} - r(t) \right) = 0. \tag{24.55}$$

Equations (24.53) and (24.55) constitute a system of $(m+1)$ linear equations for the $(m+1)$ unknowns $Q_0, Q_1, \cdots, Q_m$. To clarify the exposition we write the system in matrix form as

$$\begin{bmatrix} \sigma_{P_1}^{(0)} & \sigma_{P_1}^{(1)} & \sigma_{P_1}^{(2)} & \cdots & \sigma_{P_1}^{(m)} \\ \sigma_{P_2}^{(0)} & \sigma_{P_2}^{(1)} & \sigma_{P_2}^{(2)} & \cdots & \sigma_{P_2}^{(m)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \sigma_{P_m}^{(0)} & \sigma_{P_m}^{(1)} & \sigma_{P_m}^{(2)} & \cdots & \sigma_{P_m}^{(m)} \\ (\mu_P^{(0)} - r) & (\mu_P^{(1)} - r) & (\mu_P^{(2)} - r) & \cdots & (\mu_P^{(m)} - r) \end{bmatrix} \begin{bmatrix} Q_0 \\ Q_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Q_m \end{bmatrix} = 0. \tag{24.56}$$

This system can have a non-zero solution for $Q_0, Q_1, \cdots, Q_m$ if and only if the determinant of the matrix is zero. This latter condition is satisfied if and only if there exist functions (independent of the bond maturities) $\lambda_1(\mathbf{X}, t), \lambda_2(\mathbf{X}, t), \cdots, \lambda_m(\mathbf{X}, t)$ such that

$$\mu_P^{(i)} - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{P_j}^{(i)} \tag{24.57}$$

for each maturity $i = 0, 1, \cdots, m$. Since the bond maturities were arbitrarily chosen the result (24.57) must hold for a bond of any general maturity $T$. Thus

$$\mu_P - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{P_j}. \tag{24.58}$$

If we interpret $\lambda_j(\mathbf{X}, t)$ as the market price of risk associated with the risk factor $W_j(t)$ then (24.58) can be given the interpretation, familiar from Chap. 10, that the absence of riskless arbitrage opportunities in the economy implies that bonds are priced such that the expected excess return on the bond equals a risk premium consisting of the sum of the market price of risk of each risk factor multiplied by its bond return volatility. Use of the expressions (24.49)–(24.51) turn (24.58) into the partial differential equation

$$\frac{\partial P}{\partial t} + \sum_{j=1}^{n} \left( \mu_j - \sum_{k=1}^{m} \lambda_k \sigma_{jk} \right) \frac{\partial P}{\partial X_j} + \mathscr{D}P - r(t)P = 0, \tag{24.59}$$

where we recall that $\mathscr{D}$ is the operator of second partial derivatives defined at (24.51). This partial differential equation must be solved subject to the boundary condition

$$P(\mathbf{X}(T), T, T) = 1. \tag{24.60}$$

We shall consider some specific examples shortly.

Next we obtain the martingale representation of the bond price. Substituting (24.58) into (24.48) we obtain

$$\frac{dP}{P} = r(t)dt + \sum_{k=1}^{m} \sigma_{P_k}(\mathbf{X}, t)d\tilde{W}_k, \tag{24.61}$$

here

$$\tilde{W}_k(t) = W_k(t) + \int_0^t \lambda_k(\mathbf{X}, s)ds. \tag{24.62}$$

By the now familiar procedure of appealing to Girsanov's theorem we can find a measure $\tilde{\mathbb{P}}$ under which the $\tilde{W}_k(t)$ are Wiener process.[1] The Radon–Nikodym derivative relating the two measures $\mathbb{P}$ and $\tilde{\mathbb{P}}$ is here given by

$$\xi(t) = \exp\left[-\frac{1}{2}\int_0^t (\sum_{i=1}^n \lambda_i^2(s))ds - \int_0^t (\sum_{i=1}^n \lambda_i(s)dW_i(s))\right]. \tag{24.63}$$

The money market account is still given by (23.17) and again we form the bond price measured in units of the money market account,

$$Z(\mathbf{X}, t, T) = \frac{P(\mathbf{X}, t, T)}{A(t)} \tag{24.64}$$

whose dynamics are given by

$$\frac{dZ}{Z} = \sum_{k=1}^m \sigma_{P_k}(\mathbf{X}, t)d\tilde{W}_k. \tag{24.65}$$

The last result implies that $Z(\mathbf{X}, t, T)$ is a martingale under $\tilde{\mathbb{P}}$. Manipulations completely analogous to those that led to (23.21) yield

$$P(\mathbf{X}, t, T) = \tilde{\mathbb{E}}_t\left[e^{-\int_t^T r(s)ds}\right], \tag{24.66}$$

where $\tilde{\mathbb{E}}_t$ is the expectation operator under $\tilde{\mathbb{P}}$. To derive the dynamics of the underlying factors under $\tilde{\mathbb{P}}$ we use (24.62) in Eq. (24.47) to obtain

$$dX_i = \left(\mu_i - \sum_{j=1}^m \sigma_{ij}\lambda_j\right)dt + \sum_{j=1}^m \sigma_{ij}d\tilde{W}_j. \tag{24.67}$$

The Feynman Kac formula applied to (24.66) will again take us back to the partial differential equation (24.59). It remains to determine the instantaneous spot rate,

[1] It is important to stress that the measure now is not unique because of the market incompleteness, which manifests itself through the market prices of risk $\lambda_k$.

$r(t)$, in terms on the factors $X_1(t), \ldots, X_n(t)$. Assume that either by solving the partial differential equation (24.59) or from the expectation operator expression (24.66) we have obtained a solution for the bond price $P(\mathbf{X}(t), t, T)$. We know from (22.3) that the yield to maturity is given by $\rho(t, T) = -\ln P(\mathbf{X}(t), t, T)/(T-t)$, so that from (22.4) we have

$$r(t) = -\lim_{T \to t} \frac{\ln P(\mathbf{X}(t), t, T)}{T-t} = -\left. \frac{\partial P(\mathbf{X}(t), t, T)}{\partial T} \right|_{T=t}. \tag{24.68}$$

We shall see in subsequent example how Eq. (24.68) will yield a useful expression for $r(t)$ in terms of the underlying factors $X_1(t), \ldots, X_n(t)$.

### 24.2.2 Bond Option Pricing

The bond option price will be a function of all $n$ underlying factors so we denote its value at time $t$ as $C(\mathbf{X}(t), t, T_C, T)$. The option maturity is $T_C$ and the bond on which the option is written matures at time $T \ \ (> T_C)$. By Ito's lemma the option price dynamics are given by

$$\frac{dC}{C} = \mu_C(\mathbf{X}, t)dt + \sum_{k=1}^{m} \sigma_{C_k}(\mathbf{X}, t)dW_k, \tag{24.69}$$

where

$$\mu_C(\mathbf{X}, t) = \frac{1}{C}\left( \frac{\partial C}{\partial t} + \sum_{j=1}^{n} \mu_j \frac{\partial C}{\partial X_j} + \mathscr{D}C \right), \tag{24.70}$$

$$\sigma_{C_k}(\mathbf{X}, t) = \frac{1}{C} \sum_{j=1}^{n} \sigma_{jk}(\mathbf{X}, t) \frac{\partial C}{\partial X_j}, \tag{24.71}$$

with

$$\mathscr{D}C = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \sum_{l=1}^{m} \sigma_{il}\sigma_{jl} \right) \frac{\partial^2 C}{\partial X_i \partial X_j}. \tag{24.72}$$

We assume that the traded instruments available for investment are bonds and bond options of a range of maturities. The investor invests \$1 in a hedge portfolio containing the bond of maturity $T$ in dollar amount $Q_0$ and $m$ bond options of maturities $T_C^{(1)}$, $T_C^{(2)}$, …, $T_C^{(m)}$ in the dollar amounts $Q_1$, $Q_2$, …, $Q_m$ respectively. In the subsequent discussion we shall use $C_i$ to denote the option of maturity $T_C^{(i)}$, and $\mu_C^{(i)}$ and $\sigma_{C_k}^i$ its drift and diffusion coefficients respectively. Furthermore we assume that the bonds are already priced so that there exists no riskless arbitrage opportunities in the bond market. We recall from (24.57) that this condition means that

$$\mu_P - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{P_j}.$$

The dollar return on the hedge portfolio over the time interval $(t, t + dt)$ is given by

$$\left.\begin{array}{r}\textit{dollar return on the}\\ \textit{hedge portfolio}\end{array}\right\} = Q_0\frac{dP}{P} + \sum_{i=1}^{m} Q_i\frac{dC_i}{C_i}$$

$$= (Q_0\mu_P + \sum_{i=1}^{m} Q_i\mu_C^{(i)})dt + \sum_{k=1}^{m}(Q_0\sigma_{P_k}$$

$$+ \sum_{i=1}^{m} Q_i\sigma_{C_k}^{(i)})dW_k. \tag{24.73}$$

The stochastic terms are eliminated by choosing the dollar amounts $Q_0$, $Q_1$, …, $Q_m$ so that

$$Q_0\sigma_{P_k} + \sum_{i=1}^{m} Q_i\sigma_{C_k}^{(i)}) = 0, \tag{24.74}$$

for $k = 1, 2, \ldots, m$. The hedge portfolio is then riskless and must only earn the instantaneous risk free rate $r(t)$ if riskless arbitrage opportunities are to be avoided, that is

$$Q_0\mu_P + \sum_{i=1}^{m} Q_i\mu_C^{(i)}) = r(t).$$

Since the original investment is \$1 the proportions satisfy $\sum_{i=0}^{m} Q_i = 1$ so that the last equation becomes

$$Q_0(\mu_P - r(t)) + \sum_{i=1}^{m} Q_i(\mu_C^{(i)} - r(t)) = 0. \tag{24.75}$$

Equations (24.74) and (24.75) constitute a system of $(m + 1)$ linear equations for the $(m + 1)$ unknowns $Q_0, Q_1, \ldots, Q_m$, which in matrix rotation may be written as

$$\begin{bmatrix} \sigma_{P_1} & \sigma_{C_1}^{(1)} & \sigma_{C_1}^{(2)} & \cdots & \sigma_{C_1}^{(m)} \\ \sigma_{P_2} & \sigma_{C_2}^{(1)} & \sigma_{C_2}^{(2)} & \cdots & \sigma_{C_2}^{(m)} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \sigma_{P_m} & \sigma_{C_m}^{(1)} & \sigma_{C_m}^{(2)} & \cdots & \sigma_{C_m}^{(m)} \\ (\mu_P - r) & (\mu_C^{(1)} - r) & (\mu_C^{(2)} - r) & \cdots & (\mu_C^{(m)} - r) \end{bmatrix} \begin{bmatrix} Q_0 \\ Q_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ Q_m \end{bmatrix} = 0. \tag{24.76}$$

This system can have a non-zero solution for $Q_0, Q_1, \ldots, Q_m$ if and only if the determinant of the matrix is zero. This latter condition is satisfied if and only if there exist functions (independent of the bond and bond option maturities) $\lambda_1(\mathbf{X}, t)$, $\lambda_2(\mathbf{X}, t), \cdots, \lambda_m(\mathbf{X}, t)$ such that

$$\mu_P - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{P_j}, \tag{24.77}$$

and

$$\mu_C^{(i)} - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{C_j}^{(i)}. \tag{24.78}$$

for each bond option maturity $i = 1, 2, \ldots, m$. Equation (24.77) is just (24.58) which must hold since we are assuming bonds are already priced so that no arbitrage opportunities exist. Since the bond option maturities were arbitrarily chosen the result (24.78) must hold for a bond option of any general maturity $T_C$. Thus

$$\mu_C - r(t) = \sum_{j=1}^{m} \lambda_j(\mathbf{X}, t)\sigma_{C_j}. \tag{24.79}$$

Clearly the $\lambda_j(\mathbf{X}, t)$ have the same interpretation as in Sect. 24.2.1 and (24.79) is simply stating the fact that the absence of riskless arbitrage opportunities in the economy implies that bond options are priced such that the expected excess return on the bond option equals a risk premium consisting of the sum of the market price risk of each factor multiplied by its bond option return volatility.

Substituting (24.70) and (24.71) into (24.79) we obtain the bond option pricing partial differential equation

$$\frac{\partial C}{\partial t} + \sum_{j=1}^{n}(\mu_j - \sum_{k=1}^{m}\lambda_k\sigma_{jk})\frac{\partial_C}{\partial_{X_j}} + \mathscr{D}\mathscr{C} - r(t)\mathscr{C} = 0, \tag{24.80}$$

which must be solved subject to the boundary condition relevant to the bond option of interest. For example if the bond option is a European call option then the boundary condition is

$$C(\mathbf{X}(T_C), T_C, T_C, T) = (P(\mathbf{X}(T_C), T_C, T) - E)^+. \tag{24.81}$$

The bond price $P(\mathbf{X}(T_C), T_C, T)$ would be obtained by first solving the bond pricing problem for $P(\mathbf{X}(t), t, T)$. To obtain the martingale representation for the bond option price we substitute (24.79) into (24.69) to obtain

$$\frac{dC}{C} = r(t)dt + \sum_{k=1}^{m}\sigma_{C_k}(\mathbf{X}, t)d\tilde{W}_k, \tag{24.82}$$

where the processes $\tilde{W}_k(t)$ are defined by (24.62). Just as in Sect. 24.2.1 the processes $\tilde{W}_k(t)$ will be Wiener processes under the measure $\tilde{\mathbb{P}}$ identified there. We consider the relative bond price defined in terms of the money market account [see Eq. (23.17)] as

$$V(\mathbf{X}(t), t, T_C, T) = \frac{C(\mathbf{X}(t), t, T_C, T)}{A(t)}, \tag{24.83}$$

and whose dynamics are given by

$$\frac{dV}{V} = \sum_{k=1}^{m}\sigma_{C_k}(\mathbf{X}, t)d\tilde{W}_k.$$

The last result implies that $V$ is a martingale, so that

$$V(\mathbf{X}(t), t, T_C, T) = \tilde{\mathbb{E}}_t[V(\mathbf{X}(T_C), T_C, T_C, T)],$$

which can be rearranged to

$$C(\mathbf{X}(t), t, T_C, T) = \tilde{\mathbb{E}}_t[e^{-\int_t^{T_C} r(s)ds}C(\mathbf{X}(T_C), T_C, T_C, T)]. \tag{24.84}$$

The dynamics of the underlying factors under $\tilde{\mathbb{P}}$ are given by (24.67). The Feynman–Kac formula applied to (24.84) will yield the partial differential equation (24.80). It is also possible to express the bond option price in terms of the forward measure.

For this we consider the option measured in units of the bond maturing at $T_C$, that is the quantity

$$Y(\mathbf{X}(t), t, T_C, T) = \frac{C(\mathbf{X}(t), t, T_C, T)}{P(\mathbf{X}(t), t, T_C)}. \tag{24.85}$$

Using the result concerning the stochastic differential equation followed by the quotient of two diffusions, [recall that the dynamics for $C$ under $\tilde{\mathbb{P}}$ are given by (24.82) and for $P$ by (24.61) and substitute directly into the result (6.89)] the dynamics for $Y$ are given by (we highlight the time and maturity dependence of $\sigma_{P_k}$)

$$\frac{dY}{Y} = -\sum_{k=1}^{m} \sigma_{P_k}(t, T_C)(\sigma_{C_k} - \sigma_{P_k}(t, T_C))dt + \sum_{k=1}^{m}(\sigma_{C_k} - \sigma_{P_k}(t, T_C))d\tilde{W}_k. \tag{24.86}$$

Following the same logic as in the last part of Sect. 20.4, we rearrange Eq. (24.86) as

$$\frac{dY}{Y} = \sum_{k=1}^{m}(\sigma_{C_k} - \sigma_{P_k}(t, T_C))(d\tilde{W}_k - \sigma_{P_k}(t, T_C)dt).$$

We define new processes

$$W_k^*(t) = \tilde{W}_k(t) - \int_0^t \sigma_{P_k}(u, T_C)du, \tag{24.87}$$

and we can use Girsanov's theorem to assert that there exists an equivalent measure $\mathbb{P}^*$ under which the $W_k^*(t)$ are Wiener Processes. Thus under this measure the dynamics for $Y$ become

$$\frac{dY}{Y} = \sum_{k=1}^{m}(\sigma_{C_k} - \sigma_{P_k}(t, T_C))dW_k^*,$$

which implies that $Y$ is a martingale under $\mathbb{P}^*$. Hence

$$Y(\mathbf{X}(t), t, T_C, T) = \mathbb{E}_t^*\Big[Y(\mathbf{X}(T_C), T_C, T_C, T)\Big],$$

which in terms of the original variables may be written

$$C(\mathbf{X}(t), t, T_C, T) = P(\mathbf{X}(t), t, T_C)\mathbb{E}_t^*\Big[C(\mathbf{X}(T_C), T_C, T_C, T)\Big], \tag{24.88}$$

where $\mathbb{E}_t^*$ denotes the expectation operation under the measure $\mathbb{P}^*$. The bond option of interest could then be evaluated by substituting into (24.88) the appropriate boundary condition, for instance in the case of a European bond option the condition (24.81). The Hull–White two-factor model (discussed in Sect. 24.1) provides one of the few examples where it is possible to calculate the bond option price explicitly. The dynamics of the underlying factors $X(t)$ under $\mathbb{P}^*$ are obtained by substituting (24.87) into (24.67), so that

$$dX_i = \left(\mu_i - \sum_{j=1}^{m} \sigma_{ij}\big(\lambda_j - \sigma_{P_j}(t, T_C)\big)\right) dt + \sum_{j=1}^{m} \sigma_{ij} dW_j^*. \tag{24.89}$$

If it is necessary to evaluate the expectation in (24.89) by simulation then the process (24.89) is the one that must be simulated.

## 24.3 The Affine Class of Models

In this section we consider the affine class of term structure models for which it turns out that closed form solutions to the bond price partial differential equation (24.59) are possible. The basis of this analysis is the work of Duffie and Kan (1996). For expositional purposes we will deal with the two-factor case.

### *24.3.1 The Two-Factor Case*

Consider the situation where there are two factors $X_1(t), X_2(t)$ driven by the stochastic differential equation system (in vector notation)

$$\begin{aligned} d\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} &= \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{bmatrix} \bar{X}_1 - X_1(t) \\ \bar{X}_2 - X_2(t) \end{bmatrix} dt \\ &\quad + \begin{bmatrix} \sqrt{\alpha_{10} + \alpha_{11} X_1(t)} & 0 \\ 0 & \sqrt{\alpha_{20} + \alpha_{22} X_2(t)} \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \end{bmatrix} \end{aligned} \tag{24.90}$$

In terms of the notation of Eq. (24.47) we have

$$\begin{aligned} \mu_1(X_1, X_2, t) &= \kappa_{11}(\bar{X}_1 - X_1(t)) + \kappa_{12}(\bar{X}_2 - X_2(t)), \\ \mu_2(X_1, X_2, t) &= \kappa_{21}(\bar{X}_1 - X_1(t)) + \kappa_{22}(\bar{X}_2 - X_2(t)), \\ \sigma_{11}(X_1, X_2, t) &= \sqrt{\alpha_{10} + \alpha_{11} X_1(t)}, \end{aligned}$$

$$\sigma_{21}(X_1, X_2, t) = 0,$$
$$\sigma_{12}(X_1, X_2, t) = 0,$$
$$\sigma_{22}(X_1, X_2, t) = \sqrt{\alpha_{20} + \alpha_{22} X_2(t)}.$$

The key to understanding the motivation for the choices of functional forms in the affine class of models is the second term in the bond pricing partial differential equation (24.59). The functional forms for the $\mu_j$, $\sigma_{jk}$ and $\lambda_k$ are chosen so that the coefficient $(\mu_j - \sum_{k=1}^{m} \lambda_k \sigma_{jk})$ is affine in $X_1(t)$ and $X_2(t)$. So far we have not specified a functional form for the market prices of risk $\lambda_1(X_1, X_2, t)$ and $\lambda_2(X_1, X_2, t)$. Note that with the choices of drift and diffusion coefficients as in (24.90) the coefficient of $\frac{\partial P}{\partial X_1}$ in (24.59) is

$$\kappa_{11}(\bar{X}_1 - X_1(t)) + \kappa_{12}(\bar{X}_2 - X_2(t)) - \lambda_1(X_1, X_2, t)\sqrt{\alpha_{10} + \alpha_{11} X_1(t)}. \tag{24.91}$$

Clearly if we make the choice

$$\lambda_1(X_1, X_2, t) = \lambda_1 \sqrt{\alpha_{10} + \alpha_{11} X_1(t)}, \tag{24.92}$$

with $\lambda_1$ on the R.H.S. a constant (this is a slight abuse of notation), then the coefficient (24.91) assumes the affine form

$$\beta_{10} + \beta_{11} X_1(t) + \beta_{12} X_2(t), \tag{24.93}$$

where

$$\beta_{10} = \kappa_{11}\bar{X}_1 + \kappa_{12}\bar{X}_2 - \lambda_1 \alpha_{10},$$
$$\beta_{11} = -(\kappa_{11} + \lambda_1 \alpha_{11}),$$
$$\beta_{12} = -\kappa_{12}.$$

Furthermore with the above choices for the $\mu_i$, $\sigma_{ij}$ and $\lambda_i$ the expression for $\mathscr{D}P$ in Eq. (24.51) becomes

$$\mathscr{D}P = \frac{(\alpha_{10} + \alpha_{11} X_1(t))}{2} \frac{\partial^2 P}{\partial X_1^2} + \frac{(\alpha_{20} + \alpha_{22} X_2(t))}{2} \frac{\partial^2 P}{\partial X_2^2}.$$

Similarly assume

$$\lambda_2(X_1, X_2, t) = \lambda_2 \sqrt{\alpha_{20} + \alpha_{22} X_2(t)}.$$

Thus the bond pricing partial differential equation for the two-factor model under consideration assumes the form

$$\frac{\partial P}{\partial t} + (\beta_{10} + \beta_{11} X_1(t) + \beta_{12} X_2(t)) \frac{\partial P}{\partial X_1} + (\beta_{20} + \beta_{21} X_1(t) + \beta_{22} X_2(t)) \frac{\partial P}{\partial X_2} + \left( \frac{\alpha_{10} + \alpha_{11} X_1(t)}{2} \right) \frac{\partial^2 P}{\partial X_1^2} + \left( \frac{\alpha_{20} + \alpha_{22} X_2(t)}{2} \right) \frac{\partial^2 P}{\partial X_2^2} - rP = 0 \qquad (24.94)$$

subject to the boundary condition

$$P(X_1(T), X_2(T), T, T) = 1. \qquad (24.95)$$

Equation (24.94) can be solved by exactly the same procedure used to solve the partial differential equation (24.15) for the Hull–White two-factor model. The result is summarized in the following proposition.

**Proposition 24.1** *The solution to bond pricing partial differential equation* (24.94) *is of the form*

$$P(X_1, X_2, t, T) = \exp(-A_0(t, T) - A_1(t, T) X_1(t) - A_2(t, T) X_2(t)) \qquad (24.96)$$

*where $A_0(t, T)$, $A_1(t, T)$ and $A_2(t, T)$ are solutions of the ordinary differential equation system*

$$\frac{\partial A_0}{\partial t} = -\xi_0 - \beta_{10} A_1 - \beta_{20} A_2 + \frac{\alpha_{10}}{2} A_1^2 + \frac{\alpha_{20}}{2} A_2^2, \qquad (24.97)$$

$$\frac{\partial A_1}{\partial t} = -\xi_1 - \beta_{11} A_1 - \beta_{21} A_2 + \frac{\alpha_{11}}{2} A_1^2, \qquad (24.98)$$

$$\frac{\partial A_2}{\partial t} = -\xi_2 - \beta_{12} A_1 - \beta_{22} A_2 + \frac{\alpha_{22}}{2} A_2^2, \qquad (24.99)$$

*which must be solved subject to the terminal time conditions*

$$A_0(T, T) = A_1(T, T) = A_2(T, T) = 0. \qquad (24.100)$$

*The bond price representation* (24.96) *implies that the instantaneous spot rate is given by*

$$r(t) = \xi_0 + \xi_1 X_1(t) + \xi_2 X_2(t).$$

*Proof* Given the functional form (24.96) for the bond price, from Eq. (24.68) we have for $r(t)$ the expression

$$r(t) = \xi_0 + \xi_1 X_1(t) + \xi_2 X_2(t),$$

where

$$\xi_0 = \left.\frac{\partial A_0(t,T)}{\partial T}\right|_{(T=t)}, \quad \xi_1 = \left.\frac{\partial A_1(t,T)}{\partial T}\right|_{(T=t)}, \quad \xi_2 = \left.\frac{\partial A_2(t,T)}{\partial T}\right|_{(T=t)}.$$

Substituting the expression for $r_t$ into (24.94) the partial differential equation for $P$ becomes

$$\frac{\partial P}{\partial t} + (\beta_{10} + \beta_{11}X_1(t) + \beta_{12}X_2(t))\frac{\partial P}{\partial X_1} + (\beta_{20} + \beta_{21}X_1(t) + \beta_{22}X_2(t))\frac{\partial P}{\partial X_2}$$
$$+ \left(\frac{\alpha_{10} + \alpha_{11}X_1(t)}{2}\right)\frac{\partial^2 P}{\partial X_1^2} + \left(\frac{\alpha_{20} + \alpha_{22}X_2(t)}{2}\right)\frac{\partial^2 P}{\partial X_2^2} - (\xi_0 + \xi_1 X_1(t) + \xi_2 X_2(t))P = 0. \tag{24.101}$$

The solution technique consists in substituting the functional form (24.96) into the partial differential equation (24.101) and gathering like terms. First we note that

$$\frac{\partial P}{\partial t} = -(\dot{A}_0(t,T) + \dot{A}_1(t,T)X_1(t) + \dot{A}_2(t,T)X_2(t))P$$

where $\cdot$ denotes the time derivative $\frac{\partial}{\partial t}$. Also

$$\frac{\partial P}{\partial X_1} = -A_1 P, \quad \frac{\partial^2 P}{\partial X_1^2} = A_1^2 P, \quad \frac{\partial P}{\partial X_2} = -A_2 P, \quad \frac{\partial^2 P}{\partial X_2^2} = A_2^2 P.$$

Using the foregoing results the partial differential equation (24.101) reduces to

$$\begin{aligned}
&[-\dot{A}_0 - \beta_{10}A_1 - \beta_{20}A_2 + \frac{\alpha_{10}}{2}A_1^2 + \frac{\alpha_{20}}{2}A_2^2 - \xi_0] \\
&+ [-\dot{A}_1 - \beta_{11}A_1 - \beta_{21}A_2 + \frac{\alpha_{11}}{2}A_1^2 - \xi_1]X_1(t) \\
&+ [-\dot{A}_2 - \beta_{12}A_1 - \beta_{22}A_2 + \frac{\alpha_{22}}{2}A_2^2 - \xi_2]X_2(t) = 0.
\end{aligned} \tag{24.102}$$

The only way it is possible for the left hand side of $b$ to be zero for all possible evolutions of the factors $X_1(t), X_2(t)$ is for each square bracket to be zero for all time $t$. Thus we obtain the result that $A_0$, $A_1$ and $A_2$ must satisfy the set of ordinary differential equations (24.97)–(24.99) in the proposition. To obtain the appropriate initial conditions we note that in order that the bond price satisfy the terminal payoff condition (24.95) it must be the case that

$$A_0(T,T) = A_1(T,T) = A_2(T,T) = 0$$

which are the terminal conditions (24.100) of the proposition. ■

Many authors present the ordinary differential equation system (24.97)–(24.99) in terms of time-to-maturity $\tau(\equiv T-t)$ and furthermore assume that $A_0, A_1, A_2$ are functions of $\tau$. In this case (and with slight abuse of notation) the ordinary differential equation system (24.97)–(24.99) becomes

$$\frac{dA_0}{d\tau} = \xi_0 + \beta_{10}A_1 + \beta_{20}A_2 - \frac{\alpha_{10}}{2}A_1^2 - \frac{\alpha_{20}}{2}A_2^2, \tag{24.103}$$

$$\frac{dA_1}{d\tau} = \xi_1 + \beta_{11}A_1 + \beta_{21}A_2 - \frac{\alpha_{11}}{2}A_1^2, \tag{24.104}$$

$$\frac{dA_2}{d\tau} = \xi_2 + \beta_{12}A_1 + \beta_{22}A_2 - \frac{\alpha_{22}}{2}A_2^2, \tag{24.105}$$

with initial conditions

$$A_0(0) = A_1(0) = A_2(0) = 0. \tag{24.106}$$

The ordinary differential equation system (24.103)–(24.105) can be rapidly solved using standard numerical techniques for ordinary differential equations. So far we have not specified the initial values of the processes $X_1(t), X_2(t)$ nor the parameters $\xi_0, \xi_1$ and $\xi_2$. Here the model builder has some flexibility. A reasonable choice for initial values of $X_1(t), X_2(t)$ would be their long run values, so that

$$X_1(0) = \bar{X}_1 \quad \text{and} \quad X_2(0) = \bar{X}_2. \tag{24.107}$$

For the choices of $\xi_1, \xi_2$, one could leave them to be determined through a calibration or econometric estimation procedure. Alternatively one could specify them somewhat arbitrarily, for instance

$$\xi_1 = \xi_2 = 1 \tag{24.108}$$

and that $\xi_0$ is a constant. Once $\xi_1, \xi_2$ are chosen, the value of $\xi_0$ could be obtained from

$$r(0) = \xi_0 + \xi_1 X_1(0) + \xi_2 X_2(0),$$

where $r(0)$ may be inferred from the currently observed yield curve.

## 24.4 Problems

**Problem 24.1** Consider the following stochastic volatility extension of the Hull–White model:

$$dr = \kappa_r(\gamma - r(t))dt + \sigma(t)dz_1,$$

where $\sigma(t)$ follows the diffusion process

$$d\sigma = \kappa_\sigma(\bar{\sigma} - \sigma(t))dt + \xi dz_2$$

and

$$\mathbb{E}[dz_1 dz_2] = \rho dt.$$

Now bond prices will be a function of $r(t)$, $\sigma(t)$ and $t$ and so denoted $P(r, \sigma, t, T)$, where $T$ is bond maturity. Modify appropriately the hedging argument in the bond market to obtain the arbitrage free pricing relationship for bonds in this market.

Give the bond pricing relationship in both partial differential equation form and in expectation operator form. In the latter case be sure to indicate clearly what is the s.d.e system that has to be simulated if one were to use the Monte-Carlo approach to calculate the expectation operator.

**Problem 24.2** It is assumed that the instantaneous spot interest rate depends linearly on two factors $X_1$ and $X_2$ so that

$$r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t),$$

where $X_1$ and $X_2$ are driven by the stochastic differential equation system

$$dX_1 = \kappa_{11}(\alpha_1 - X_1(t))dt + \sqrt{X_1(t)}dW_1,$$

$$dX_2 = (\kappa_{21}(\alpha_1 - X_1(t)) - \kappa_{22}X_2(t))dt + \sqrt{1 + \beta_{21}X_1(t)}dW_2,$$

where $W_1$ and $W_2$ are independent Wiener processes. Assuming that the bond price is a function of $r$ and hence a function of $X_1$ and $X_2$ follow a similar procedure to that of Sect. 24.1 to derive the bond pricing partial differential equation.

Furthermore assume that the market prices of risk, $\lambda_1$ and $\lambda_2$, associated with $W_1$ and $W_2$ respectively are given by

$$\lambda_1 = \bar{\lambda}_1\sqrt{X_1(t)}, \quad \lambda_2 = \bar{\lambda}_2\sqrt{1 + \beta_{21}X_1(t)}$$

where $\bar{\lambda}_1$, $\bar{\lambda}_2$ are constant. Try a solution of the form

$$P(X_1, X_2, t, T) = \exp(-a(t, T) - b_1(t, T)X_1(t) - b_2(t, T)X_2(t)).$$

Use a similar procedure to that outlined in Sect. 24.3 to obtain the ordinary differential equations that determine $a(t, T)$, $b_1(t, T)$ and $b_2(t, T)$. Make sure to specify the boundary conditions.

Use the fact that

$$r(t) = \lim_{T \to t} \rho(t, T),$$

where $\rho(t, T)$ is the yield to maturity, to relate $\delta_0$, $\delta_1$ and $\delta_2$ to the $a, b_1$ and $b_2$ functions.

**Problem 24.3** Repeat Problem 24.2 for the model that may be represented by

$$r(t) = \delta_0 + \delta_1 X_1(t) + \delta_2 X_2(t) + \delta_3 X_3(t),$$

where the factors $X_1(t), X_2(t)$ and $X_3(t)$ are driven by the stochastic differential equation system

$$d \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} \kappa_{11} & 0 & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{bmatrix} \begin{bmatrix} \alpha_1 - X_1(t) \\ -X_2(t) \\ -X_3(t) \end{bmatrix} dt + \begin{bmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{21} X_1(t)} & 0 \\ 0 & 0 & \sqrt{1 + \beta_{31} X_1(t)} \end{bmatrix} d \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix},$$

where $W_1, W_2$ and $W_3$ are independent Wiener processes.

Assume that the market price of risk, $\lambda_i$, associated with $W_i$ $(i = 1, 2, 3)$ are proportional to the corresponding volatility, that is $\lambda_1 = \bar{\lambda}_1 \sqrt{X_1(t)}$, $\lambda_2 = \bar{\lambda}_2 \sqrt{1 + \beta_{21} X_1(t)}$, $\lambda_3 = \sqrt{1 + \beta_{31} X_1(t)}$.

# Chapter 25
# The Heath–Jarrow–Morton Framework

**Abstract** Interest rate modelling can also be performed by starting from the dynamics of the instantaneous forward rate. As we shall see the dynamics of all other quantities of interest can then be derived from it. This approach has its origin in Ho and Lee (J Finance XLI:1011–1029, 1986) but was most clearly articulated in Heath et al. (Econometrica 60(1):77–105, 1992a), to which we shall subsequently refer as Heath–Jarrow–Morton. In this framework, the condition of no riskless arbitrage results in the drift coefficient of the forward rate dynamics being expressed in terms of the forward rate volatility function. The major weakness in implementing the Heath–Jarrow–Morton approach is that the spot rate dynamics are usually path dependent (non-Markovian). We consider a class of functional forms of the forward rate volatility that allow the model to be reduced to a finite dimensional Markovian system of stochastic differential equations. This class contains some important models considered in the literature.

## 25.1 Introduction

The interest rate derivative models developed in Chap. 23 took as their starting point the dynamics of the instantaneous spot interest rate. The models we derived there also had the characteristic that the market price of interest rate risk appears in the pricing relationships. We saw in Sect. 23.8 that at least in principle it is possible to remove this dependence of the models on preference related quantities. This can be done by expressing terms involving the market price of interest rate risk in terms of market observed quantities such as the currently observed yield curve and volatilities of traded interest rate dependent instruments. However this procedure for rendering spot rate models preference-free can be tedious and for some model specifications may be computationally intensive.

An alternative interest rate modelling approach, originated by Ho and Lee (1986), is the Heath–Jarrow–Morton approach which starts from the dynamics of the forward rate and requires the specification of the initial term structure and the volatility of the associated forward rate. The dynamics of the spot interest rate are then developed from those of the forward rate. The spot interest rate is also an important economic variable whose assessment determines the evolution of the



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_25

bond prices. Heath et al. (1992b) describe how this framework can be used to price and hedge the entire interest rate derivative book of a financial institution thus offering a consistent approach in managing interest rate exposure. However, the major shortcoming of the Heath–Jarrow–Morton approach is that the spot rate dynamics are not path independent (i.e. it is non-Markovian) and the entire history of the term structure has to be carried thus increasing the computational complexity.

The key unobserved input to this approach to term structure modelling is the aforementioned volatility of the forward rates. Many of the forms of the volatility functions reported in the literature have been chosen for analytical convenience rather than on the basis of empirical evidence. In fact apart from the study of Heath et al. (1990), Flesaker (1993), Amin and Morton (1994), Amin and Ng (1993), Ho et al. (2001), and Bhar et al. (2004), there has not been a great deal of empirical research into the appropriate form of the volatility function to be used in the arbitrage free class of models. This is due to the fact that the non-Markovian nature of the stochastic dynamical system makes difficult application of standard econometric estimation procedures.

The non-Markovian feature also makes difficult the expression for prices of term-structure contingent claims in terms of partial differential equations. In the Heath–Jarrow–Morton approach these prices are expressed as expectation operators, under the equivalent martingale measure, of appropriate payoffs. Nowhere in the existing literature is it stated how to consistently turn this expectation operator into a partial differential equation. It is important to be able to do so in order to apply to the evaluation of interest rate sensitive contingent claims many useful computational techniques, as outlined for example in Wilmott et al. (1993). These techniques are the most appropriate to value various path dependent options such as American, Asian etc., but require an expression of the contingent claim price in terms of partial differential equation operators with appropriate boundary conditions.

The notation used in the original Heath–Jarrow–Morton paper allows for a very general dependence of the forward rate volatility functions on path dependent quantities. For the sake of definiteness we shall assume in this chapter that the path dependence of the forward rate volatility functions arises from dependence on the instantaneous spot interest rate and/or a set of discrete fixed-tenor forward rates. As we shall see this specification allows us to develop a fairly broad class of interest rate derivative models.

With such a specification, the instantaneous spot rate process in the Heath–Jarrow–Morton framework can be expressed in terms of a finite dimensional Markovian system. The dimension of the resultant system of stochastic differential equations is dependent on the exact form of the volatility function and it usually includes variables that at first sight seem not to be readily observable. But we shall show how it is possible to express these in terms of forward rates or yields, which may be observable.

The transformation to the Markovian form also allows easier comparison with other approaches such as, Vasicek (1977), Cox et al. (1985a) and Hull and White (1987, 1990, 1994). This is important in the sense that it allows us to easily reconcile many of the alternative approaches to the modelling of the term structure of interest rates.

## 25.2 The Basic Structure

We denote as $f(t, T)$ the instantaneous forward rate negotiated at time $t$ for instantaneous borrowing at time $T(> t)$. The starting point of the Heath–Jarrow–Morton model of the term structure of interest rates is the stochastic integral equation for the forward rate[1]

$$f(t, T) = f(0, T) + \int_0^t \alpha(v, T, \omega(v))dv + \int_0^t \sigma(v, T, \omega(v))dW(v), \qquad (25.1)$$

for $0 \le t \le T$, where $\alpha(v, T, \omega(v))$ and $\sigma(v, T, \omega(v))$ are the instantaneous drift and the volatility function at time $v$ of the forward rate $f(v, T)$, respectively. The instantaneous drift $\alpha(v, T, \omega(v))$ and volatility function $\sigma(v, T, \omega(v))$ of the forward rate $f(v, T)$ could depend through $\omega$ on path dependent quantities, such as the instantaneous spot rate and/or a set of discrete tenor forward rates. Thus the specifications in Eq. (25.1) allow for functional forms of the type $\hat{\alpha}(v, T, r(v), f(v, \tau_1), f(v, \tau_2), \cdots, f(v, \tau_m))$ and $\hat{\sigma}(v, T, r(v), f(v, \tau_1), f(v, \tau_2), \cdots, f(v, \tau_m))$ where $f(t, \tau_i)$ is the instantaneous forward rate at time $t$ applicable at the fixed tenor $\tau_i(> t)$ with $m$ such tenors. The noise term $dW(v)$ is the increment of a standard Wiener process generated by a probability measure $\mathbb{P}$. Note that for expositional simplicity in this section we consider only one noise term affecting the evolution of the forward rate. The stochastic integral equation (25.1), may alternatively be expressed as the stochastic differential equation

$$df(t, T) = \alpha(t, T, \omega(t))dt + \sigma(t, T, \omega(t))dW(t). \qquad (25.2)$$

It follows from Eq. (25.1) that the instantaneous spot rate $r(t)(\equiv f(t, t))$ is given by the stochastic integral equation

$$r(t) = f(0, t) + \int_0^t \alpha(v, t, \omega(v))dv + \int_0^t \sigma(v, t, \omega(v))dW(v). \qquad (25.3)$$

The corresponding stochastic differential equation for $r(t)$ [see Eq. (22.39)] is

$$dr = \mu_r(t)dt + \sigma(t, t, \omega(t))dW(t), \qquad (25.4)$$

[1] We refer the reader to Sect. 22.5 for further discussion on the interpretation of (25.1).

where

$$\mu_r(t) = f_2(0,t) + \alpha(t,t,\omega(t)) + \int_0^t \alpha_2(v,t,\omega(v))dv + \int_0^t \sigma_2(v,t,\omega(v))dW(v), \tag{25.5}$$

where $f_2$, $\alpha_2$ and $\sigma_2$ denote the partial derivative of $f$, $\alpha$ and $\sigma$ respectively, with respect to their second arguments. We recall that the bond price at time $t$ is related to the forward rate by

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right), \qquad 0 \leq t \leq T. \tag{25.6}$$

By the use of Fubini's theorem for stochastic integrals and application of Ito's lemma (see Sect. 22.5.1) the bond price satisfies the stochastic differential equation

$$dP(t,T) = [r(t) + b(t,T)]P(t,T)dt + \sigma_B(t,T)P(t,T)dW(t), \tag{25.7}$$

where

$$\sigma_B(t,T) \equiv -\int_t^T \sigma(t,s,\omega(t))ds, \tag{25.8}$$

and

$$b(t,T) \equiv -\int_t^T \alpha(t,s,\omega(t))ds + \frac{1}{2}\sigma_B^2(t,T). \tag{25.9}$$

A quantity of interest is the money market account

$$A(t) = \exp\left(\int_0^t r(y)dy\right), \tag{25.10}$$

which is the value at time $t$ of a dollar continuously compounded from 0 to $t$ at the instantaneous spot rate $r$. This quantity may be used to define the relative bond price

$$Z(t,T) = \frac{P(t,T)}{A(t)}, \qquad (0 \leq t \leq T). \tag{25.11}$$

The fact that $dA = r(t)A(t)dt$ and application of the rule for the quotient of two diffusions (see Sect. 6.6) yields the result that the relative bond price satisfies the stochastic differential equation

$$dZ(t,T) = b(t,T)Z(t,T)dt + \sigma_B(t,T)Z(t,T)dW(t). \tag{25.12}$$

## 25.3 The Arbitrage Pricing of Bonds

Bonds can be priced using exactly the same hedging portfolio as was used in Chap. 23, namely we use bonds of two different maturities. We know from the arbitrage arguments of Chap. 23 that in order that there not exist riskless arbitrage opportunities between bonds of different maturities then the instantaneous excess bond return, risk adjusted by its volatility must equal the market price of interest rate risk; see Eq. (23.9). The relevant bond dynamics in the current context are given by (25.7), so that here Eq. (23.9) becomes

$$\frac{[r(t)+b(t,T)]-r(t)}{\sigma_B(t,T)} = \begin{array}{c} \textit{market price of} \\ \textit{interest rate risk} \end{array} \equiv -\phi(t), \tag{25.13}$$

which simplifies to

$$b(t,T)+\phi(t)\sigma_B(t,T)=0. \tag{25.14}$$

Using expressions (25.8) and (25.9) this last equation may be written explicitly as

$$\int_t^T \alpha(t,s,\omega(t))ds - \frac{1}{2}\left(\int_t^T \sigma(t,s,\omega(t))ds\right)^2 + \phi(t)\int_t^T \sigma(t,s,\omega(t))ds = 0.$$

Keeping $t$ fixed and differentiating with respect to maturity $T$, the above equation reduces to

$$\alpha(t,T,\omega(t)) - \left(\int_t^T \sigma(t,s,\omega(t))ds\right)\sigma(t,T,\omega(t)) + \phi(t)\sigma(t,T,\omega(t)) = 0,$$

which may be rearranged to

$$\alpha(t,T,\omega(t)) = -\sigma(t,T,\omega(t))\left[\phi(t) - \int_t^T \sigma(t,s,\omega(t))ds\right]. \tag{25.15}$$

Equation (25.15) is the forward rate drift restriction that was first reported by Heath–Jarrow–Morton (Eq. (18) of Heath et al. 1992a). It states that if the bond market is free of riskless arbitrage opportunities then the forward rate drift, the forward rate volatility and the market price of interest rate risk must be tied together as shown by this equation. Heath–Jarrow–Morton show that in fact this condition is both necessary and sufficient for the absence of riskless arbitrage opportunities.

Up to this point Heath–Jarrow–Morton have not done anything conceptually different from the standard arbitrage approach of Chap. 23. However in the Heath–Jarrow–Morton approach Eq. (25.14) is used in a different way. In the standard arbitrage approach, Eq. (25.14) becomes a partial differential equation for the bond price as a function of the assumed driving state variable (usually the instantaneous spot rate). In the Heath–Jarrow–Morton approach, the condition (25.14) becomes the forward rate drift restriction that is used, as we shall see below, to conveniently express the bond price dynamics under an equivalent probability measure. By use of (25.14), the stochastic differential equations (25.7) and (25.12) for $P(t,T)$ and $Z(t,T)$ respectively become

$$dP(t,T) = [r(t) - \phi(t)\sigma_B(t,T)]P(t,T)dt + \sigma_B(t,T)P(t,T)dW(t), \tag{25.16}$$

$$dZ(t,T) = -\phi(t)\sigma_B(t,T)Z(t,T)dt + \sigma_B(t,T)Z(t,T)dW(t). \tag{25.17}$$

At the same time, by substituting (25.15) into (25.3)[2] the stochastic integral equation for $r(t)$ becomes

$$\begin{aligned} r(t) = f(0,t) &+ \int_0^t \sigma(v,t,\omega(v)) \int_v^t \sigma(v,s,\omega(v))dsdv \\ &- \int_0^t \sigma(v,t,\omega(v))\phi(v)dv + \int_0^t \sigma(v,t,\omega(v))dW(v). \end{aligned} \tag{25.18}$$

The key advance in the Heath–Jarrow–Morton approach is that, by an application of Girsanov's theorem to (25.16), Eqs. (25.16)–(25.18), can be written in terms of a different Wiener process generated by an equivalent martingale probability measure $\tilde{\mathbb{P}}$. Thus if we define a new Wiener process $\tilde{W}(t)$ under $\tilde{\mathbb{P}}$ by

$$\tilde{W}(t) = W(t) - \int_0^t \phi(s)ds, \tag{25.19}$$

or in differential form by

$$d\tilde{W}(t) = dW(t) - \phi(t)dt, \tag{25.20}$$

then Eqs. (25.16)–(25.18) become

$$dP(t,T) = r(t)P(t,T)dt + \sigma_B(t,T)P(t,T)d\tilde{W}(t), \tag{25.21}$$

$$dZ(t,T) = \sigma_B(t,T)Z(t,T)d\tilde{W}(t), \tag{25.22}$$

[2]Note that from (25.15) we have

$$\int_0^t \alpha(v,t,\omega(v))dv = -\int_0^t \sigma(v,t,\omega(v))\phi(v)dv + \int_0^t \sigma(v,t,\omega(v)) \int_v^t \sigma(v,s,\omega(v))dsdv.$$

and

$$r(t)=f(0,t)+\int_0^t \sigma(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))dsdv+\int_0^t \sigma(v,t,\omega(v))d\tilde{W}(v). \tag{25.23}$$

Alternatively, Eq. (25.23) can be expressed as the stochastic differential equation

$$\begin{aligned} dr = \Bigg[ f_2(0,t) + \frac{\partial}{\partial t}\left(\int_0^t \sigma(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))dsdv\right) \\ + \int_0^t \frac{\partial\sigma(v,t,\omega(v))}{\partial t} d\tilde{W}(v)\Bigg] dt + \sigma(t,t,\omega(t))d\tilde{W}(t). \end{aligned} \tag{25.24}$$

It is at times convenient to deal with the ln of the bond price $B(t,T) \equiv \ln P(t,T)$. This quantity, by Ito's lemma, satisfies (under $\tilde{\mathbb{P}}$)

$$dB(t,T) = [r(t) - \frac{1}{2}\sigma_B^2(t,T)]dt + \sigma_B(t,T)d\tilde{W}(t). \tag{25.25}$$

Furthermore, the arbitrage free stochastic integral equation for the forward rate under $\tilde{\mathbb{P}}$ can be written,

$$f(t,T) = f(0,T)+\int_0^t \sigma(v,T,\omega(v))\int_v^T \sigma(v,s,\omega(v))dsdv+\int_0^t \sigma(v,T,\omega(v))d\tilde{W}(v), \tag{25.26}$$

and the corresponding stochastic differential equation as

$$df(t,T) = \sigma(t,T,\omega(t))\int_t^T \sigma(t,s,\omega(t))dsdt + \sigma(t,T,\omega(t))d\tilde{W}(t). \tag{25.27}$$

The essential characteristic of the reformulated stochastic differential and integral equations (25.21)–(25.27) expressed in terms of Brownian motion, under the equivalent probability measure $\tilde{\mathbb{P}}$, is that the empirically awkward market price of risk term, $\phi(t)$, is eliminated from explicit consideration. From the discussion of Girsanov's theorem in Sect. 8.2 [in particular Eqs. (8.38) and (8.42)] we obtain the expression for the Radon–Nikodym derivative

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^t \phi^2(s)ds + \int_0^t \phi(s)dW(s)\right). \tag{25.28}$$

If we write $\tilde{\mathbb{E}}_t$ to denote mathematical expectation with respect to the equivalent probability measure (i.e. the one associated with $d\tilde{W}(t)$) then from Eq. (25.22)

$$\tilde{\mathbb{E}}_t[dZ(t,T)] = 0.$$

This last equation implies that $Z(t, T)$ is a martingale under $\tilde{\mathbb{P}}$, i.e.

$$Z(t, T) = \tilde{\mathbb{E}}_t(Z(T, T)),$$

or, in terms of the bond price

$$P(t, T) = \tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T)}\right] = \tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(y)dy\right)\right]. \tag{25.29}$$

Equation (25.29) is the fundamental bond pricing equation of the Heath–Jarrow–Morton framework, and it has the same discounted cash flow interpretation as Eq. (23.21). Namely the quantity $\exp\left(-\int_t^T r(y)dy\right)$ should be interpreted as the stochastic discount factor under $\tilde{\mathbb{P}}$ used to discount back to time $t$ the \$1 payoff to be received at time $T$. We stress that the actual implementation of (25.29) will depend on the form chosen from the forward rate volatility function. At the simplest level, the expectation in Eq. (25.29) could be calculated by numerically simulating Eq. (25.23). Note however that if the volatility function depends on discrete tenor forward rates $f(t, \tau_1), \cdots, f(t, \tau_m)$ then these would have be simulated at the same time. Closed form analytical expressions for the bond price may be obtained with appropriate assumptions on the volatility function as we shall see below.

## 25.4 Arbitrage Pricing of Bond Options

Suppose we wish to price at time $t$ an option on the bond, for example a European call option on the bond, with the option maturing at $T_c$. As we saw in Chap. 21 this problem is relevant to the pricing of an interest rate cap. We know from the discussion at the end of the previous section that under the risk-neutral measure $\tilde{\mathbb{P}}$ we can discount the payoff at $T_c$ back to $t$ using the stochastic discount factor

$$\exp\left(-\int_t^{T_c} r(s)ds\right). \tag{25.30}$$

Multiplying the payoff by the discount factor we find that under one realisation of the spot-rate process under $\tilde{\mathbb{P}}$ the option value at $t$ is given by

$$\exp\left(-\int_t^{T_c} r(s)ds\right) \max\left[P(T_c,\ T) - X, 0\right]. \tag{25.31}$$

The value of the option, $C(t, T_c)$, is then obtained by taking the expectation of this quantity under the risk-neutral measure $\tilde{\mathbb{P}}$ (i.e. forming $\tilde{\mathbb{E}}_t$). Thus we obtain

$$C(t, T_c) = \tilde{\mathbb{E}}_t \left[ \exp\left( -\int_t^{T_c} r(s)ds \right) \max[P(T_c, T) - X, 0] \right]. \tag{25.32}$$

In general, if we have some spot interest rate contingent claim with payoff at $t = T_c$ given by $H(r(T_c),\ T_c)$,[3] then its value at $t$ is given by

$$U(t, T_c) = \tilde{\mathbb{E}}_t \left[ \exp\left( -\int_t^{T_c} r(s)ds \right) H(r(T_c),\ T_c) \right]. \tag{25.33}$$

In order to obtain a pricing partial differential equation for $U(t, T_c)$ we need to obtain the Kolmogorov backward equation associated with Eq. (25.24), the stochastic differential equation for the spot rate process $r(t)$. It is difficult to do this in general because of the non-Markovian term $\int_0^t \frac{\partial \sigma}{\partial t} d\tilde{W}(v)$ that appears in the drift in Eq. (25.24). In Sect. 25.6 we discuss assumptions on $\sigma$ which allow us to obtain Markovian representations for $r(t)$ and hence obtain pricing partial differential equations.

The reader may wonder why in this section we have not mirrored the argument of Sect. 23.6 and used the hedging argument approach to derive the bond option pricing formula in the present context. The reason is that in Sect. 23.6 there was one underlying factor, $r(t)$, driving the uncertainty of the market, so the option price could be written in the form $C(r, t)$ and Ito's Lemma applied to obtain its dynamics. Similarly in Sect. 24.1 there were two underlying factors, $r(t)$ and $h(t)$, driving the uncertainty of the market and we would write the option price as $C(r, h, t)$. Again application of Ito's Lemma gave us the dynamics for $C$. In both cases the dynamics of the hedging portfolio could then be obtained. Here we have not so far been so precise about the factors upon which the volatility function $\sigma(t, T, \cdot)$ depends, apart from stating that it could depend on a vector of discrete tenor forward rates and the instantaneous spot rate. In order to mirror the hedging argument approach used in Chap. 23 we need to be more specific about the dynamics of these underlying rates so that we could then obtain the option price dynamics by applying Ito's lemma. This we shall do in a later section, when we discuss the Markovianisation issue. At this point we stress that the expressions (25.29) for the bond price and (25.33) for the interest rate derivative hold for quite general specifications of the volatility function. Of course if we want to implement these expressions, using for example stochastic simulation, then we would need to specify the dynamics (under $\tilde{\mathbb{P}}$) of all stochastic factors entering into the specification of $\sigma(t, T, \cdot)$.

[3] Here we allow the payoff function to depend on the instantaneous spot rate. It could of course depend on various other rates as well.

## 25.5 Forward-Risk-Adjusted Measure

We saw in Eq. (25.33) that the value of a spot interest rate contingent claim at time $t$ can be written

$$U(t, T_c) = \tilde{\mathbb{E}}_t \left[ \exp\left( -\int_t^{T_c} r(s)ds \right) H(r(T_c), T_c) \right], \tag{25.34}$$

where $H(r(T_c), T_c)$ denotes the payoff on the claim at time $T_c$. Suppose $P(t, T_c)$ represents the price at time $t$ of a pure discount bond maturing at time $T_c$. Then by Eq. (25.29),

$$P(t, T_c) = \tilde{\mathbb{E}}_t \left[ \exp\left( -\int_t^{T_c} r(s)ds \right) \right]. \tag{25.35}$$

We can use the results of Chap. 20 to express the value of the interest rate contingent claim, using $P(t, T_c)$ as the numeraire. By forming the quantity $Y = U/P$ we would obtain [see Eq. (20.14)]

$$U(t, T_c) = P(t, T_c) \mathbb{E}_t^* \left[ \frac{U(T_c, T_c)}{P(T_c, T_c)} \right]. \tag{25.36}$$

But in the current notation $U(T_c, T_c) = H(r(T_c), T_c)$ and $P(T_c, T_c) = 1$, hence

$$U(t, T_c) = P(t, T_c) \mathbb{E}_t^* \left[ H(r(T_c), T_c) \right]. \tag{25.37}$$

The advantage of (25.37) over (25.34) is that the stochastic discounting term $\exp\left( -\int_t^{T_c} r(s)ds \right)$ which appears in the expectation operation of (25.34) is replaced by the non-stochastic discounting term $P(t, T_c)$ which appears outside the expectation operator of (25.37). The value of this method depends on how easy (or difficult) it is to calculate $\mathbb{E}_t^*$ in (25.37). We saw how this change of measure result was useful in obtaining Merton's bond pricing formula in Sect. 20.3. The measure associated with the $\mathbb{E}_t^*$ operation, which we shall denote as $\mathbb{P}^*$ is known as the $T$-forward measure. The reason for this nomenclature is that under $\mathbb{P}^*$ the forward rate at time $t$ is the expectation of the instantaneous spot rate at $T$ i.e.

$$f(t, T) = \mathbb{E}_t^* [r(T)]. \tag{25.38}$$

To see this result recall that

$$P(t, T) = \tilde{\mathbb{E}}_t \left[ \exp\left( -\int_t^{T} r(s)ds \right) \right].$$

Differentiating this last equation with respect to $T$ we obtain

$$\begin{aligned}\frac{\partial}{\partial T}P(t,T) &= \tilde{\mathbb{E}}_t\left[\frac{\partial}{\partial T}\exp\left(-\int_t^T r(s)ds\right)\right]\\ &= \tilde{\mathbb{E}}_t\left[-\exp\left(-\int_t^T r(s)ds\right)\cdot\frac{\partial}{\partial T}\int_t^T r(s)ds\right]\\ &= -\tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(s)ds\right)\cdot r(T)\right]\\ &= -P(t,T)\mathbb{E}_t^*[r(T)]. \qquad (25.39)\end{aligned}$$

The last line was obtained by applying (25.37) with $H(r(T),T) = r(T)$. Rearranging the last result we obtain

$$-\frac{\partial}{\partial T}\ln P(t,T) = \mathbb{E}_t^*[r(T)]. \qquad (25.40)$$

However $f(t,T) = -\frac{\partial}{\partial T}\ln P(t,T)$ hence we have established the result in Eq. (25.38).

For some later applications we need to clarify what form the Radon–Nikodym derivative assumes for the forward risk-adjusted measure. To do this we simply identify $V_T$ with $P(T,T)$ and $V_0$ with $P(0,T)$ in Eq. (20.20), so that the Radon–Nikodym derivative becomes

$$\xi(0,T) = \frac{P(T,T)}{A(T)P(0,T)} = \frac{1}{A(T)P(0,T)}. \qquad (25.41)$$

## 25.6 Reduction to Markovian Form

The principal difficulty in implementing and estimating Heath–Jarrow–Morton models arises from the non-Markovian noise term in the stochastic integral equation (25.23) for $r(t)$. This manifests itself in the third component of the drift term of the stochastic differential equation (25.24). This component depends on the history of the noise process from time 0 to current time $t$. Depending upon the specification of the volatility function the second component of the drift term could also depend on the path history up to time $t$.

Our aim in this section is to consider a class of functional forms of $\sigma(t,T,\cdot)$ that allow the non-Markovian representation of $r(t)$ and $P(t,T)$ to be reduced to a finite dimensional Markovian system of stochastic differential equations. We investigate volatility functions of the forward rate which have the general form of a deterministic function of time and maturity multiplied by a function of the path dependent variable $\omega$, i.e.

$$\sigma(t,T,\omega(t)) = Q(t,T)G(\omega(t)), \quad 0 \le t \le T, \qquad (25.42)$$

where $G$ is an appropriately well-behaved function. A useful representation for $Q(t,T)$ would be

$$Q(t,T) = P_n(T-t)e^{-\lambda(T-t)}, \tag{25.43}$$

where $P_n(u)$ is the polynomial

$$P_n(u) = a_0 + a_1 u + \ldots + a_n u^n.$$

This form would allow the term structure of the volatility to exhibit humps as observed in implied forward rate volatilities from cap prices.[4] We recall the discussion of Sect. 22.5.2 when we considered forward rate volatility functions of the form[5]

$$\sigma(t,T,\omega(t)) = \bar{\sigma}e^{-\lambda(T-t)}G(\omega(t)), \tag{25.44}$$

for $\bar{\sigma} > 0$ and $\lambda$ constant. This structure includes forward rate volatilities for a number of important cases in the literature. Some of these models include

- $\lambda > 0, G(\omega(t)) = 1$ leads to a version of the extended Vasicek model of Hull–White,
- $\lambda > 0, G(\omega(t)) = g(r(t))$ leads to the generalised spot rate model of Ritchken and Sankarasubramanian (1995).
- $\lambda > 0, G(\omega(t)) = \sqrt{r(t)}$ leads to an extended version of the CIR model,
- $\lambda > 0, G(\omega(t)) = g(r(t), f(t,\tau))$ leads to a version of the model of Chiarella and Kwon (1999).

Our aim is, under the volatility specification of (25.44), to express Eqs. (25.24) and (25.27) as a Markovian system of stochastic differential equations. By considering the drift term of the stochastic differential equation (25.27), under the volatility specification of (25.44), we obtain

$$\begin{aligned}
\sigma(t,T,\omega(t))\int_t^T \sigma(t,s,\omega(t))ds &= \bar{\sigma}^2 G^2(\omega(t))e^{-\lambda(T-t)}\int_t^T e^{-\lambda(s-t)}ds \\
&= \bar{\sigma}^2 G^2(\omega(t))e^{-\lambda(T-t)}\frac{\left(e^{-\lambda(T-t)}-1\right)}{-\lambda} \\
&= \sigma^2(t,T,\omega(t))\frac{\left(e^{\lambda(T-t)}-1\right)}{\lambda}.
\end{aligned}$$

[4] Ritchken and Chuang (1999) assume $P_n(u) = P_1(T-t) = (a_0 + a_1(T-t))$.

[5] It is possible to carry through the discussion of this subsection with Eq. (25.44) generalised to $\sigma(t,T,\omega(t)) = \bar{\sigma}e^{-\int_t^T \lambda(s)ds}G(\omega(t))$.

Thus, we express the stochastic differential equation (25.27) for the forward rate as

$$df(t,T) = \left[\sigma^2(t,T,\omega(t))\frac{e^{\lambda(T-t)}-1}{\lambda}\right]dt + \sigma(t,T,\omega(t))d\tilde{W}(t). \tag{25.45}$$

Similarly, we consider the stochastic differential equation (25.24) for r(t) and in particular the first integral term

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\int_0^t \sigma(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))ds dv\right)\\&= \int_0^t \left[\sigma_2(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))ds + \sigma^2(v,t,\omega(v))\right]dv,\end{aligned} \tag{25.46}$$

where the notation $\sigma_2$ represents the partial derivative of $\sigma$ with respect to its second argument. Given the expression for $\sigma$ in (25.44), we have that,

$$\sigma_2(v,t,\omega(v)) = -\lambda\sigma(v,t,\omega(v)). \tag{25.47}$$

Thus, the right hand side of (25.46) reduces to

$$\int_0^t \left[-\lambda\sigma(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))ds + \sigma^2(v,t,\omega(v))\right]dv. \tag{25.48}$$

By using (25.47) the other two terms of (25.24) can be expressed as

$$\begin{aligned}&\left[\int_0^t \sigma_2(v,t,\omega(v))d\tilde{W}(v)\right]dt + \sigma(t,t,\omega(t))d\tilde{W}(t)\\&= \left[-\lambda\int_0^t \sigma(v,t,\omega(v))d\tilde{W}(v)\right]dt + \sigma(t,t,\omega(t))d\tilde{W}(t).\end{aligned}$$

Thus, the stochastic differential equation (25.24) becomes

$$\begin{aligned}dr = &\left[f_2(0,t) - \lambda\int_0^t \sigma(v,t,\omega(v))\int_v^t \sigma(v,s,\omega(v))ds dv\right.\\&\left.+ \int_0^t \sigma^2(v,t,\omega(v))dv - \lambda\int_0^t \sigma(v,t,\omega(v))d\tilde{W}(v)\right]dt\\&+ \sigma(t,t,\omega(t))d\tilde{W}(t).\end{aligned} \tag{25.49}$$

We note from the stochastic integral equation (25.23) that

$$\begin{aligned} r(t) - f(0,t) = &\int_0^t \sigma(v,t,\omega(v)) \int_v^t \sigma(v,s,\omega(v)) ds dv \\ &+ \int_0^t \sigma(v,t,\omega(v)) d\tilde{W}(v). \end{aligned} \tag{25.50}$$

Then by using (25.50), the stochastic differential equation (25.49) for the spot rate is simplified to

$$dr = [f_2(0,t) + \lambda f(0,t) + \psi(t) - \lambda r(t)]\, dt + \sigma(t,t,\omega(t)) d\tilde{W}(t), \tag{25.51}$$

where we define the subsidiary variable $\psi(t)$ as

$$\psi(t) = \int_0^t \sigma^2(v,t,\omega(v)) dv. \tag{25.52}$$

The subsidiary variable $\psi(t)$ defined in Eq. (25.52) plays a central role in allowing us to transform the original non-Markovian dynamics to Markovian form. Similar subsidiary variables appear in the reduction to Markovian forms of Cheyette (1992), Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997a), Inui and Kijima (1998), and Chiarella and Kwon (1999). It is clear from (25.52) that $\psi(t)$ may be interpreted as a variable summarising the path history of the forward rate volatility.

## 25.7 Some Special Models

At this point the stochastic differential equation (25.51) is still non-Markovian because the integral in the drift term involves the history of the path dependent forward rate volatility. To proceed any further we need to consider specific functional forms for $G(\omega(t))$ in the volatility specifications (25.44).

### *25.7.1 The Hull–White Extended Vasicek Model*

If $G(\omega(t)) = 1$ then the subsidiary variable (25.52) becomes a time function, i.e.

$$\psi(t) = \int_0^t \sigma^2(v,t,\omega(v)) dv = \int_0^t \bar{\sigma}^2 e^{-2\lambda(t-v)} dv = \frac{\bar{\sigma}^2}{2\lambda}(1 - e^{-2\lambda t}).$$

Thus by setting

$$\theta(t) = f_2(0,t) + \lambda f(0,t) + \frac{\bar{\sigma}^2}{2\lambda}(1 - e^{-2\lambda t}), \tag{25.53}$$

the stochastic differential equation (25.51) for $r(t)$ finally becomes

$$dr = [\theta(t) - \lambda r(t)]dt + \bar{\sigma}\, d\tilde{W}(t), \tag{25.54}$$

which is a sought Markovian representation. Clearly this is the extended Vasicek model with the long run mean allowed to be time varying.

Furthermore, note that the expression we have obtained for $\theta(t)$ is the same as the one we obtained in Sect. 23.7 when we worked directly from the expression for the bond price obtained from the continuous arbitrage approach—which takes the spot rate process as the driving dynamics. It is also worth pointing out that by setting $\lambda = 0$ we obtain the continuous time specification of the Ho–Lee model. We have already seen how to price European options in this framework in Sect. 23.7. To price American options in this framework, the option pricing equation

$$\frac{1}{2}\bar{\sigma}^2 \frac{\partial^2 C}{\partial r^2} + [\theta(t) - \lambda r]\frac{\partial C}{\partial r} + \frac{\partial C}{\partial t} - rC = 0, \tag{25.55}$$

must be solved subject to the boundary conditions for an American option, see Chiarella and El-Hassan (1996) for details.

### 25.7.2 The General Spot Rate Model

If $G(\omega(t))$ is a function of the spot interest rate $r(t)$, i.e. $G(\omega(t)) = g(r(t))$, then we need to separately handle the non-Markovian term appearing in the drift of Eq. (25.51). The subsidiary variable (25.52) now becomes

$$\psi(t) = \int_0^t \bar{\sigma}^2 e^{-2\lambda(t-v)} g^2(r(v))dv. \tag{25.56}$$

By differentiating (25.56) we have that

$$d\psi = [\bar{\sigma}^2 g^2(r(t)) - 2\lambda\psi(t)]dt. \tag{25.57}$$

We are now dealing with the two-dimensional Markovian system

$$\begin{aligned} dr &= [f_2(0,t) + \lambda f(0,t) + \psi(t) - \lambda r(t)]dt + \bar{\sigma}\, g(r(t))d\tilde{W}(t), \\ d\psi &= [\bar{\sigma}^2 g^2(r(t)) - 2\lambda\psi(t)]dt. \end{aligned} \tag{25.58}$$

The representation (25.58) was obtained by Ritchken and Sankarasubramanian (1995). If we define the partial differential operator $\mathscr{K}$ by

$$\mathscr{K} = \frac{1}{2}\bar{\sigma}^2 g^2(r)\frac{\partial^2}{\partial r^2} + [f_2(0,t) + \lambda f(0,t) + \psi - \lambda r(t)]\frac{\partial}{\partial r} + [\bar{\sigma}^2 g^2(r) - 2\lambda\psi]\frac{\partial}{\partial \psi},$$

then the Kolmogorov equation for the transition probability density $\pi$ is

$$\mathscr{K}\pi + \frac{\partial \pi}{\partial t} = 0,$$

and derivative instruments are priced according to the partial differential equation

$$\mathscr{K}V + \frac{\partial V}{\partial t} - rV = 0,$$

(note $V = P$ for bond price, $V = C$ for option price) subject to appropriate boundary conditions, e.g. $V(r, T, T) = 1$, for bonds, $V(r, T_c, T) = \max[0, P(r, T_c, T) - K]$, for European call options, etc. To evaluate American options we need to employ numerical methods. Chiarella and El-Hassan (1998) have found the method of lines to be very effective in this context.

Note that in the special case of $g(r(t)) = \sqrt{r(t)}$, we are dealing with the extended CIR model

$$\begin{aligned} dr &= [f_2(0,t) + \lambda f(0,t) + \psi(t) - \lambda r(t)]dt + \bar{\sigma}\sqrt{r(t)}d\tilde{W}(t), \\ d\psi &= [\bar{\sigma}^2 r(t) - 2\lambda\psi(t)]dt. \end{aligned} \tag{25.59}$$

### 25.7.3 The Forward Rate Dependent Volatility Model

In this case, we further generalise the form of the volatility function to include forward interest rates. Here $G(\omega(t))$ can be a function of the spot interest rate, $r(t)$, and of the forward interest rate, $f(t,\tau)$ of a fixed maturity $\tau$, so that $G(\omega(t)) = g(r(t), f(t,\tau))$. For example, $f(t,\tau)$ could be some long-term forward rate. The intuition behind such a specification is that not only the spot interest rate but also a fixed maturity forward interest rate influence the evolution of the term structure. The particular forward rate to be used may depend on the application under consideration. This approach may be considered to be equivalent in some sense to the Brennan and Schwartz (1979) model where a short-term rate and a long-term rate are used to explain the evolution of the term structure. We need to determine the additional state variables necessary to make the system Markovian although with a higher dimension. The associated subsidiary variable (25.52) is given by

$$\psi(t) = \int_0^t \bar{\sigma}^2 e^{-2\lambda(t-v)} g^2(r(v), f(v,\tau))dv. \tag{25.60}$$

By differentiating (25.60), we obtain the stochastic differential equation for $\psi(t)$ as

$$\begin{aligned} d\psi &= [\sigma^2(t, t, \omega(t)) - 2\lambda\psi(t)]dt \\ &= [\bar{\sigma}^2 g^2(r(t), f(t, \tau)) - 2\lambda\psi(t)]dt \end{aligned} \tag{25.61}$$

We have now reduced the non-Markovian stochastic dynamics to a three dimensional Markovian stochastic dynamical system consisting of the stochastic differential equation for the spot rate $r(t)$ [recall (25.51)]

$$dr = [f_2(0, t) + \lambda f(0, t) + \psi(t) - \lambda r(t)]\, dt + \bar{\sigma} g(r(t), f(t, \tau)) d\tilde{W}(t), \tag{25.62}$$

the stochastic differential equation for the discrete forward rate $f(t, \tau)$ [recall (25.45)], namely,

$$\begin{aligned} df(t, \tau) &= \sigma^2(t, \tau, \omega(t)) \frac{(e^{\lambda(\tau - t)} - 1)}{\lambda} dt + \sigma(t, \tau, \omega(t)) d\tilde{W}(t), \\ &= \bar{\sigma}^2 g^2(r(t), f(t, \tau)) e^{-2\lambda(\tau - t)} \frac{(e^{\lambda(\tau - t)} - 1)}{\lambda} dt \\ &\quad + \bar{\sigma} e^{-\lambda(\tau - t)} g(r(t), f(t, \tau)) d\tilde{W}(t), \end{aligned} \tag{25.63}$$

and the stochastic differential equations (25.61) for $\psi(t)$. Finally we recall that the dynamics of the forward rate to any maturity $T$, $f(t, T)$ is given by (25.45) and so are determined once $r(t)$ and $f(t, \tau)$ are determined. These latter quantities are driven by the three stochastic differential equations (25.61)–(25.63) which together form the Markovian representation. The price of any derivative instrument would then have to depend on $r(t)$ and $f(t, \tau)$. Thus a bond of maturity $T$ would have a price at time t denoted by $P(t, T, r(t), f(t, \tau))$, and this price is also driven by the three-dimensional Markovian stochastic differential equation system referred to above.

#### 25.7.3.1 Interpreting the Subsidiary Variable $\psi(t)$

However, it would perhaps be more satisfying to relate $\psi(t)$ to the market rates $r(t)$ and $f(t, \tau)$. Indeed it turns out that such a relationship does exist for the forward rate volatility function assumed in Eq. (25.44) for the generalised case of $G(\omega(t)) = g(r(t), f(t, \tau))$.

**Proposition 25.1** *The subsidiary integrated square variance quantity $\psi(t)$ defined in Eq. (25.60) is related to the rates $r(t)$ and $f(t,\tau)$ via*

$$\psi(t) = \lambda\alpha\,(t,\tau)\,[r\,(t) - f\,(0,t)] - \lambda e^{-\lambda(t-\tau)}\alpha\,(t,\tau)\,[f\,(t,\tau) - f\,(0,\tau)]\,, \tag{25.64}$$

*where* $\alpha(t,\tau) \equiv \dfrac{e^{-\lambda t}}{e^{-\lambda\tau} - e^{-\lambda t}}$.

For the proof of Proposition 25.1 see Appendix 25.1.

An important consequence of Proposition 25.1 is that it allows us to reduce by one the dimension of the stochastic dynamic system (25.61)–(25.63) to the two-dimensional one consisting of the stochastic differential equations (25.62) and (25.63) with $\psi(t)$ being defined by Eq. (25.64). This reduction in dimension is quite significant if we seek to solve for derivative prices in this framework by use of partial differential equations or lattice based methods as in Bhar et al. (2000), since then we need to deal only with two rather than three spatial variables in the partial differential operator. The reduction is less significant, though still useful, when using Monte-Carlo simulation. This is so since Monte Carlo simulation requires the simulation of the one Wiener increment, $d\tilde{W}(t)$. The generation of $\psi(t)$ by Eq. (25.64) rather than discretising Eq. (25.61) should lead to some computational efficiency.

A consequence of Proposition 25.1 is that we are able to express the forward rate to any maturity $T$ in terms of the two rates $r(t)$ and $f(t,\tau)$.

**Proposition 25.2** *The forward rate $f(t,T)$ to any maturity $T$ is given by*

$$\begin{aligned} f(t,T) - f(0,T) = \\ - e^{-2\lambda(T-\tau)}\frac{\alpha(\tau,t)}{\alpha(T,t)}\,[f(t,\tau) - f(0,\tau)] + e^{-2\lambda(T-t)}\frac{\alpha(t,\tau)}{\alpha(T,\tau)}\,[r(t) - f(0,t)]\,. \end{aligned} \tag{25.65}$$

For the proof of Proposition 25.2 see Appendix 25.2.

#### 25.7.3.2 The Term Structure of Interest Rates

We recall the Heath–Jarrow–Morton approach of defining a money market account (25.10) and showing that the relative bond price

$$Z(t,T) = \frac{P(t,T)}{A(t)},$$

is a martingale, so that the bond price can be written

$$P(t,T) = \tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T)}\right] = \tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(y)dy\right)\right]. \tag{25.66}$$

Here, $\tilde{\mathbb{E}}_t$ is the expectation taken with respect to the probability distribution generated by the stochastic differential system (25.62) and (25.63). We use

$$\pi\left(r(t^*), f(t^*, \tau) \mid r(t), f(t, \tau)\right),$$

to denote the transition probability density function between $t$ and $t^*$ ($t \leq t^*$). This quantity satisfies the Kolmogorov backward partial differential equation, which for our case is given by,

$$\mathscr{K}\pi + \frac{\partial \pi}{\partial t} = 0,$$

where the operator $\mathscr{K}$ is the infinitesimal generator of the diffusion process for $f(t, \tau)$, $r(t)$ driven by the stochastic differential equations (25.62) and (25.63). It turns out that $\mathscr{K}$ is given by (see Appendix 25.3),

$$\begin{aligned}\mathscr{K}\pi \equiv\ & \sigma_1^2 \frac{\left(e^{\lambda(\tau-t)} - 1\right)}{\lambda} \frac{\partial \pi}{\partial f} + [f_2(0,t) + \lambda f(0,t) + \psi - \lambda r] \frac{\partial \pi}{\partial r} \\ & + \frac{1}{2}\sigma_1^2 \frac{\partial^2 \pi}{\partial f^2} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 \pi}{\partial r^2} + \sigma_1 \sigma_r \frac{\partial^2 \pi}{\partial f \partial r},\end{aligned} \tag{25.67}$$

where $\sigma_1(t) = \sigma(t, \tau, \omega(t))$ and $\sigma_r(t) = \sigma(t, t, \omega(t))$. By application of the Feynman–Kac formula to Eq. (25.66) we find that the bond price $P(t, T, r, f)$ satisfies the partial differential equation,

$$\frac{\partial P}{\partial t} + \mathscr{K}P - r(t)P = 0, \tag{25.68}$$

subject to the terminal condition

$$P(T, T, r, f) = 1,$$

and the boundary conditions

$$\begin{aligned} P(t, T, \infty, f) &= 0, \qquad (f \geq 0), \\ P(t, T, r, \infty) &= 0, \qquad (r \geq 0). \end{aligned}$$

The further boundary conditions $P(t, T, 0, f)$ and $P(t, T, r, 0)$ may be obtained by an extrapolation procedure to be discussed in Bhar et al. (2000). Note that in subsequent discussion we set

$$D(t) \equiv f_2(0, t) + \lambda f(0, t).$$

A consequence of Proposition 25.2 is that it turns out to be possible to obtain an analytical expression for the bond price. In fact we may state the following proposition:

**Proposition 25.3** *The price of bonds driven by the Markovian stochastic differential equation system* (25.62) *and* (25.63) *can be expressed as*

$$P(t,T,r,f) = \frac{P(0,T)}{P(0,t)} \exp\left[-\beta(t,T)\left(r(t) - f(0,t)\right) - \frac{1}{2}\beta^2(t,T)\psi(t)\right]. \tag{25.69}$$

*where* $\beta(t,T) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right)$, *and* $\psi(t)$ *is defined in Eq.* (25.64).

For the proof of Proposition 25.3 see Appendix 25.4.

The bond pricing equation (25.69) has precisely the same form as the one derived by Ritchken and Sankarasubramanian (1995) who (in current notation) assumed a form for the volatility function in Eq. (25.42) with $G(\omega(t)) = g(r(t))$ which is independent of the forward rate $f(t,\tau)$. In fact the results in Propositions 25.2 and 25.3 can be considerably generalised. Chiarella and Kwon (1999) have shown that (25.69) holds in precisely the same form even when the forward rate volatility depends on a set of discrete forward rates $f(t,\tau_1), f(t,\tau_2), \ldots, f(t,\tau_r)$ where $t \leq \tau_1 < \tau_2 < \cdots < \tau_r \leq T$. Of course, under these different specifications the history variable $\psi(t)$ will evolve differently but the functional relationship remains the same.

#### 25.7.3.3 Pricing European Bond Options

Consider an option written on the bond of maturity $T$. We suppose the option matures at time $T_c(< T)$ and denote its price by $C(t,T,r,f)$. This price satisfies the partial differential equation

$$\frac{\partial C}{\partial t} + \mathscr{K}C - rC = 0, \quad (0 \leq t \leq T_c). \tag{25.70}$$

If we are dealing with a European call option with strike price E then the terminal condition for (25.70) is

$$C(T_c,T,r,f) = [P(T_c,T,r,f) - E]^+ .$$

The boundary conditions at infinity are

$$C(t,T,\infty,f) = 0, \qquad f \geq 0,$$
$$C(t,T,r,\infty) = 0, \qquad r \geq 0.$$

We recall that the bond prices at option maturity for any given values of $r(T_c)$, $f(T_c,\tau)$ can be obtained directly from Eq. (25.69) without the need to solve the bond pricing partial differential equation (25.68). In Bhar et al. (2000), we discuss the solution of the partial differential equation (25.70) by means of the alternating directions implicit (ADI) method.

An alternative approach to pricing the European option is to use the result (also derived by Heath–Jarrow–Morton) that

$$C(t,T,r,f) = \tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^{T_c} r(y)dy\right)[P(T_c,T,r(T_c), f(T_c,\tau)) - E]^+\right]. \tag{25.71}$$

The expectation in Eq. (25.71) could be approximated by simulating an appropriate number of times the stochastic differential equation system (25.62) and (25.63) from $t$ to $T_c$.

## 25.8 Heath–Jarrow–Morton Multi-Factor Models

In our previous discussion, we focussed on the case where only one noise factor was impinging on the forward rate curve. However Heath–Jarrow–Morton framework allows for the possibility of multiple noise sources, i.e.,

$$f(t,T) = f(0,T) + \int_0^t \alpha(v,T,\cdot)dv + \sum_{i=1}^{n}\int_0^t \sigma_i(v,T,\cdot)dW_i(v), \tag{25.72}$$

where the $n$ noise terms $dW_i$ are the increments of independent Wiener processes and the $\sigma_i(t,T,\cdot)$ are the volatility functions associated with each noise term. The manipulations leading to (25.15) in Sect. 25.2 are identical in the multiple noise case and merely involve a little more algebra. Thus setting $T = t$ in (25.72) we have

$$r(t) = f(0,t) + \int_0^t \alpha(v,t,\cdot)dv + \sum_{i=1}^{n}\int_0^t \sigma_i(v,t,\cdot)dW_i(v). \tag{25.73}$$

Substituting (25.72) into (25.6) and following the same procedure that yielded (25.7) we find that the stochastic differential equation for the bond price now becomes

$$dP(t,T) = [r(t) + b(t,T,\cdot)]P(t,T)dt + \sum_{i=1}^{n} a_i(t,T,\cdot)P(t,T)dW_i(t), \tag{25.74}$$

where

$$a_i(t,T,\cdot) = -\int_t^T \sigma_i(t,v,\cdot)dv, \tag{25.75}$$

and

$$b(t,T,\cdot) = -\int_t^T \alpha(t,v,\cdot)dv + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T,\cdot). \tag{25.76}$$

The process for the relative bond price

$$Z(t,T) = \frac{P(t,T)}{A(t)},$$

is easily found to be

$$dZ(t,T) = b(t,T,\cdot)Z(t,T)dt + \sum_{i=1}^{n} a_i(t,T,\cdot)Z(t,T)dW_i(t). \tag{25.77}$$

Now by forming a portfolio of bonds of $(n+1)$ different maturities and holding these in proportions that ensure the existence of no riskless arbitrage opportunities results in the condition (for interpretation compare with (10.5) that gives the expected excess return condition in the multi-factor case)

$$[r(t) + b(t,T,\cdot)] - r(t) = -\sum_{i=1}^{n} \phi_i(t)a_i(t,T,\cdot), \tag{25.78}$$

where $\phi_i(t)$ is the market price of risk associated with the $i$th noise factor. The term on the left-hand side of Eq. (25.78) is the expected excess return on the bond, the term on the right hand side is the sum of the risk-premia $(\phi_i a_i)$ for bearing the risk associated with each source of uncertainty $(W_i(t))$. Equation (25.78) simplifies to

$$b(t,T,\cdot) + \sum_{i=1}^{n} \phi_i(t)a_i(t,T,\cdot) = 0, \tag{25.79}$$

which is the multifactor analogue of the forward rate drift restriction (25.14). By use of (25.75) Eq. (25.79) reads

$$\int_t^T \alpha(t,v,\cdot)dv = \sum_{i=1}^{n}\left[\frac{1}{2}a_i^2(t,T,\cdot) + \phi_i(t)a_i(t,T,\cdot)\right]. \tag{25.80}$$

Differentiating this last equation with respect to maturity $T$ we find that

$$\alpha(t,T,\cdot) = -\sum_{i=1}^{n}\sigma_i(t,T,\cdot)\left[\phi_i(t) - \int_t^T \sigma_i(t,v,\cdot)dv\right], \tag{25.81}$$

which is the forward rate drift restriction in the multi-factor case. Thus substituting (25.81) into (25.73), (25.74) and (25.77) the stochastic differential equations for $r(t)$, $P(t,T)$ and $Z(t,T)$ become respectively, in the arbitrage free environment,

$$\begin{aligned} dr = \Bigg[ & f_2(0,t) + \frac{\partial}{\partial t}\sum_{i=1}^{n}\int_0^t \sigma_i(v,t,\cdot)\int_v^t \sigma_i(v,y,\cdot)dydv + \sum_{i=1}^{n}\int_0^t \frac{\partial \sigma_i}{\partial t}(v,t,\cdot)dW_i(v) \\ & - \sum_{i=1}^{n}\phi_i(t)\sigma_i(t,t,\cdot) - \sum_{i=1}^{n}\int_0^t \phi_i(v)\frac{\partial \sigma_i}{\partial t}(v,t,\cdot)dv \Bigg] dt + \sum_{i=1}^{n}\sigma_i(t,t,\cdot)dW_i(t) \\ dP(t,T) = & \left[ r(t) - \sum_{i=1}^{n}\phi_i(t)a_i(t,T,\cdot) \right] P(t,T)dt + \sum_{i=1}^{n} a_i(t,T,\cdot)P(t,T)dW_i(t), \\ dZ(t,T) = & -\sum_{i=1}^{n}\phi_i(t)a_i(t,T,\cdot)Z(t,T)dt + \sum_{i=1}^{n} a_i(t,T,\cdot)Z(t,T)dW_i(t). \end{aligned}$$

We then form the new set of processes

$$\tilde{W}_i(t) = W_i(t) - \int_0^t \phi_i(s)ds, \quad (i = 1, 2, \cdots, n).$$

By use of Girsanov's theorem these become Wiener processes under the equivalent measure $\tilde{\mathbb{P}}$. Thus the forgoing set of equations become

$$dP(t,T) = r(t)P(t,T)dt + \sum_{i=1}^{n} a_i(t,T,\cdot)P(t,T)d\tilde{W}_i(t), \tag{25.82}$$

$$dZ(t,T) = \sum_{i=1}^{n} a_i(t,T,\cdot)Z(t,T)d\tilde{W}_i(t), \tag{25.83}$$

$$\begin{aligned} dr = \Bigg[ & f_2(0,t) + \frac{\partial}{\partial t}\sum_{i=1}^{n}\int_0^t \sigma_i(v,t,\cdot)\int_v^t \sigma_i(v,y,\cdot)dydv \\ & + \sum_{i=1}^{n}\int_0^t \frac{\partial \sigma_i}{\partial t}(v,t,\cdot)d\tilde{W}_i(v) \Bigg] dt + \sum_{i=1}^{n}\sigma_i(t,t,\cdot)d\tilde{W}_i(t). \end{aligned} \tag{25.84}$$

Again under $\tilde{\mathbb{P}}$ the relative bond price $Z(t,T)$ is a martingale, so that

$$Z(t,T) = \tilde{\mathbb{E}}_t\left[Z(T,T)\right], \tag{25.85}$$

which in terms of the bond price becomes

$$P(t,T) = \tilde{\mathbb{E}}_t\left[\exp\left(-\int_t^T r(y)dy\right)\right], \tag{25.86}$$

which of course is the same as (25.29). The difference here is the $\tilde{\mathbb{E}}_t$ is generated by (25.84) which in its turn is driven by the $n$ independent noise terms $\tilde{W}_i(t)$. Thus if we were to use simulation to directly evaluate (25.86) we would need to simulate $n$ independent sequences of normal random variables in order to simulate a path for $r(t)$. Furthermore if the $\sigma_i(t,T)$ depend on other factors, such as discrete tenor forward rates, then the processes for these (under $\tilde{\mathbb{P}}$) would have to be simulated as well.

## 25.9 Relating Heath–Jarrow–Morton to Hull–White Two-Factor Models

We have already seen in Sect. 25.7.1 that the Hull–White extended Vasicek model can be derived (far more simply) in the Heath–Jarrow–Morton framework once an appropriate form for the volatility function is chosen. In this section we show that the Hull–White two-factor model can be obtained as a special case of the multi-factor Heath–Jarrow–Morton model.

First we recall the Hull–White two-factor model Hull and White (1994), as summarised by Rebonato (1998). The instantaneous spot rate is assumed to follow the process

$$dr = [\theta(t) + h(t) - ar(t)]dt + \gamma_1 dz_1, \tag{25.87}$$

where the additional term $h(t)$ in the drift satisfies

$$dh = -b\,h(t)dt + \gamma_2 dz_2. \tag{25.88}$$

Here $z_1$, $z_2$ are correlated Wiener processes, i.e.

$$\mathbb{E}[dz_1\ dz_2] = \rho dt.$$

First we note that we may reexpress (25.87) and (25.88) in terms of the independent Wiener processes $w_1$, $w_2$ as

$$dr = [\theta(t) + h(t) - ar(t)]dt + \gamma_1\sqrt{1-\rho^2}\,dw_1 + \gamma_1\rho\,dw_2, \tag{25.89}$$

$$dh = -b\,h(t)\,dt + \gamma_2 dw_2. \tag{25.90}$$

We consider a two-factor Heath–Jarrow–Morton model with volatility specifications

$$\sigma_i(t,T) = \bar{\sigma}_i e^{-\lambda_i(T-t)}, \tag{25.91}$$

where $\bar{\sigma}_i$ and $\lambda_i$ are constants, for $i = 1, 2$. Accordingly, the forward rate dynamics are expressed as

$$f(t,T) = f(0,T) + \int_0^t \alpha(v,T)dv + \int_0^t \sigma_1(v,T)dW_1(v) + \int_0^t \sigma_2(v,T)dW_2(v),$$

which under the risk-neutral measure (see Sect. 25.8) becomes

$$f(t,T) = f(0,T) + \sum_{i=1}^{2} \int_0^t \sigma_i(v,T) \int_v^T \sigma_i(v,y)dy\,dv + \sum_{i=1}^{2} \int_0^t \sigma_i(v,T)d\tilde{W}_i(v).$$

The spot rate process under the risk-neutral measure satisfies

$$r(t) = f(0,t) + \sum_{i=1}^{2} \int_0^t \sigma_i(v,t) \int_v^t \sigma_i(v,y)dy\,dv + \sum_{i=1}^{2} \int_0^t \sigma_i(v,t)d\tilde{W}_i(v), \tag{25.92}$$

or

$$\begin{aligned} dr = \Bigg[ & f_2(0,t) + \frac{\partial}{\partial t} \sum_{i=1}^{2} \int_0^t \sigma_i(v,t) \int_v^t \sigma_i(v,y)dy\,dv \\ & + \sum_{i=1}^{2} \int_0^t \frac{\partial \sigma_i}{\partial t}(v,t)d\widetilde{W_i}(v) \Bigg] dt + \sum_{i=1}^{2} \sigma_i(t,t)d\tilde{W}_i(t). \end{aligned} \tag{25.93}$$

The volatility functions (25.91) have the property

$$\frac{\partial \sigma_i(t,T)}{\partial T} = -\lambda_i \sigma_i(t,T).$$

Thus the stochastic differential equation (25.93) for the spot rate becomes

$$dr = \left[ f_2(0,t) + \frac{\partial}{\partial t} \sum_{i=1}^{2} S_i(t) - \sum_{i=1}^{2} \lambda_i x_i(t) \right] dt + \sum_{i=1}^{2} \bar{\sigma}_i d\tilde{W}_i(t), \tag{25.94}$$

where we set

$$S_i(t) = \bar{\sigma}_i^2 \int_0^t e^{-\lambda_i(t-v)} \int_v^t e^{-\lambda_i(y-v)}dy\,dv,$$

and

$$x_i(t) = \int_0^t \bar{\sigma}_i e^{-\lambda_i(t-v)} d\tilde{W}_i(v),$$

where $i = 1, 2$. The variables $x_i(t)$ satisfy the stochastic differential equations

$$\begin{aligned} dx_i &= \bar{\sigma}_i e^{-\lambda_i(t-t)} d\tilde{W}_i(t) - \int_0^t \lambda_i \bar{\sigma}_i e^{-\lambda_i(t-v)} d\tilde{W}_i(v) dt \\ &= -\lambda_i x_i(t) dt + \bar{\sigma}_i d\tilde{W}_i(t). \end{aligned} \tag{25.95}$$

The system (25.94), (25.95) for $r(t)$, $x_1(t)$, and $x_2(t)$ is the Markovian system that generates the probability distribution for $\tilde{\mathbb{E}}_t$ (i.e. this is the system we would simulate if we use Monte-Carlo simulation to evaluate $\tilde{\mathbb{E}}_t$). To obtain the link with the Hull–White two-factor model note that with the volatility functions (25.91) the stochastic integral equation for $r(t)$, see Eq. (25.92), becomes

$$r(t) = f(0,t) + \sum_{i=1}^{2} S_i(t) + x_1(t) + x_2(t). \tag{25.96}$$

This last equation may be used to eliminate $x_1(t)$ in (25.94), thus

$$x_1(t) = r(t) - f(0,t) - \sum_{i=1}^{2} S_i(t) - x_2(t),$$

which upon substitution into (25.94) yields

$$dr = [f_2(0,t) + \lambda_1 f(0,t) + \mathbf{S}(t) + (\lambda_1 - \lambda_2) x_2(t) - \lambda_1 r(t)]dt + \sum_{i=1}^{2} \bar{\sigma}_i d\tilde{W}_i(t), \tag{25.97}$$

where we have defined

$$\mathbf{S}(t) = \sum_{i=1}^{2} \left[ \frac{\partial}{\partial t} S_i(t) + \lambda_1 S_i(t) \right], \tag{25.98}$$

and $x_2(t)$ is driven by Eq. (25.95) with $i = 2$, viz

$$dx_2 = -\lambda_2 x_2(t) dt + \bar{\sigma}_2 d\tilde{W}_2(t). \tag{25.99}$$

To fully obtain the correspondence with the Hull–White two-factor model in Eqs. (25.89), (25.90) set

$$\begin{aligned} \theta(t) &= f_2(0,t) + \lambda_1 f(0,t) + \mathbf{S}(t), \\ a &= \lambda_1, \\ b &= \lambda_2, \end{aligned} \tag{25.100}$$

so that we are now dealing with the system

$$dr = [\theta(t) + (a-b)x_2(t) - ar(t)]dt + \bar{\sigma}_1 d\tilde{W}_1(t) + \bar{\sigma}_2 d\tilde{W}_2(t), \quad (25.101)$$

$$dx_2 = -b\,x_2(t)dt + \bar{\sigma}_2 d\tilde{W}_2(t). \quad (25.102)$$

If we set

$$h(t) = (a-b)x_2(t)$$

then Eq. (25.102) becomes

$$dh = -bh(t)dt + (a-b)\bar{\sigma}_2 d\tilde{W}_2(t). \quad (25.103)$$

The system (25.101) and (25.103) for $r(t)$ and $h(t)$ is equivalent to the Hull–White two-factor system (25.89), (25.90) if we set

$$\bar{\sigma}_1 = \gamma_1\sqrt{1-\rho^2}, \quad \bar{\sigma}_2 = \gamma_1\rho \quad \text{and} \quad (a-b)\bar{\sigma}_2 = \gamma_2, \quad (25.104)$$

from which the parameters of the Hull–White two-factor model may be related to the parameters of the two-factor Heath–Jarrow–Morton model via

$$\rho = \frac{\bar{\sigma}_2}{\sqrt{\bar{\sigma}_1^2 + \bar{\sigma}_2^2}}, \quad \gamma_1 = \sqrt{\bar{\sigma}_1^2 + \bar{\sigma}_2^2}, \quad \gamma_2 = (\lambda_1 - \lambda_2)\bar{\sigma}_2, \quad b = \lambda_2. \quad (25.105)$$

It is certainly instructive to understand how the Hull–White class of models can be derived within the Heath–Jarrow–Morton framework. However, the biggest advantage is that the $\theta(t)$ function in the stochastic differential equation (25.101) for $r(t)$ is automatically calibrated to the initially observed forward curve $f(0,t)$.

## 25.10 The Covariance Structure Implied by the Heath–Jarrow–Morton Model

Another important issue to be considered is the analysis of the statistical properties of the evolution of the forward rates and yields under the jump-diffusion framework. As we have mentioned before one factor models allow for only parallel shifts of the yield curve, so bond prices and forward rates of all maturities are perfectly correlated. Multi dimensional models, on the other hand, impose a correlation structure between forward rates of different maturities which based on empirical studies shows an exponentially decaying behavior. Rebonato (1998) provides an interesting discussion on forward rate correlations and examines the patterns observed in financial markets. Here we seek to understand the effect on the forward rate correlation structure implied by the assumptions concerning in the forward rate dynamics.

### 25.10.1 The Covariance Structure of the Forward Rate Changes

Under the risk neutral measure, the changes of the forward rate follow the dynamics

$$df(t,T) = \sum_{i=1}^{n} \sigma_i(t,T,\omega(t))\zeta_i(t,T,\omega(t))dt + \sum_{i=1}^{n} \sigma_i(t,T,\omega(t))d\tilde{W}_i(t). \tag{25.106}$$

Thus

$$\tilde{\mathbb{E}}_0[df(t,T)] = \sum_{i=1}^{n} \sigma_i(t,T,\omega(t))\zeta_i(t,T,\omega(t))dt. \tag{25.107}$$

Denote $T_1$ and $T_2$ two different maturities, then the covariance of the changes on the forward rate is calculated as

$$\begin{aligned}
&\mathrm{cov}_0[df(t,T_1), df(t,T_2)] \\
&= \tilde{\mathbb{E}}_0[(df(t,T_1) - \tilde{\mathbb{E}}_0[df(t,T_1)])(df(t,T_2) - \tilde{\mathbb{E}}_0[df(t,T_2)])] \\
&= \tilde{\mathbb{E}}_0\left[\sum_{i=1}^{n} \sigma_i(t,T_1,\omega(t))d\tilde{W}_i(t) \cdot \sum_{i=1}^{n} \sigma_i(t,T_2,\omega(t))d\tilde{W}_i(t)\right].
\end{aligned}$$

From the independence of the Wiener increments it readily follows that

$$\mathrm{cov}_0[df(t,T_1), df(t,T_2)] = \sum_{i=1}^{n} \sigma_i(t,T_1,\omega(t))\sigma_i(t,T_2,\omega(t))dt, \tag{25.108}$$

and the variance of the forward rate changes $df(t,T_h)$ $(h = 1,2)$ as

$$\mathrm{var}_0[df(t,T_h)] = \sum_{i=1}^{n} \sigma_i^2(t,T_h,\omega(t))dt. \tag{25.109}$$

The correlation coefficient between the forward rates changes $df(t,T_1)$ and $df(t,T_2)$ is then evaluated as

$$\rho(t,T_1,T_2) = \frac{\mathrm{cov}_0[df(t,T_1), df(t,T_2)]}{\sqrt{\mathrm{var}_0[df(t,T_1)]}\sqrt{\mathrm{var}_0[df(t,T_2)]}}, \tag{25.110}$$

where $\mathrm{cov}_0[df(t,T_1), df(t,T_2)]$ and $\mathrm{var}_0[df(t,T_h)]$, $(h = 1,2)$ are defined above. To demonstrate these results, we assume the volatility functions are of the form

$$\sigma_i(s,t) = \sigma_{0i} e^{-\kappa_{\sigma i}(t-s)}, \qquad i = 1,\ldots,n, \tag{25.111}$$

where the $\sigma_{0i}$, $\kappa_{\sigma i}$ are constant. Then the covariance (25.108) between $df(t,T_1)$ and $df(t,T_2)$ is calculated as

$$\mathrm{cov}_0[df(t,T_1), df(t,T_2)] = \sum_{i=1}^{n} \sigma_{0i}^2 e^{-\kappa_{\sigma i}(T_1+T_2-2t)}, \tag{25.112}$$

and the correlation coefficient between the forward rates changes $df(t,T_1)$ and $df(t,T_2)$ is evaluated as

$$\rho(t,T_1,T_2) = \frac{\sum_{i=1}^{n} \sigma_{0i}^2 e^{-\kappa_{\sigma i}(T_1+T_2-2t)}}{\sqrt{\mathrm{var}_0[df(t,T_1)]}\sqrt{\mathrm{var}_0[df(t,T_2)]}}, \tag{25.113}$$

where the variance of the forward rate changes $df(t,T_h)$ $(h = 1,2)$ is

$$\mathrm{var}_0[df(t,T_h)] = \sum_{i=1}^{n} \sigma_{0i}^2 e^{-2\kappa_{\sigma i}(T_h-t)}. \tag{25.114}$$

### *25.10.2 The Covariance Structure of the Forward Rate*

The forward rate under the risk neutral measure is given by

$$\begin{aligned} f(t,T) = f(0,T) &+ \sum_{i=1}^{n} \int_0^t \sigma_i(s,T,\omega(s))\zeta_i(s,T,\omega(s))ds \\ &+ \sum_{i=1}^{n} \int_0^t \sigma_i(s,T,\omega(s))d\tilde{W}_i(s). \end{aligned} \tag{25.115}$$

Thus

$$\tilde{\mathbb{E}}_0[f(t,T)] = f(0,T) + \sum_{i=1}^{n} \int_0^t \sigma_i(s,T,\omega(s))\zeta_i(s,T,\omega(s))ds. \tag{25.116}$$

Denote $T_1$ and $T_2$ two maturities then the covariance of the forward rates $f(t, T_1)$ and $f(t, T_2)$ is calculated as

$$\begin{aligned}
&\mathrm{cov}_0[f(t, T_1), f(t, T_2)] \\
&= \tilde{\mathbb{E}}_0[(f(t, T_1) - \tilde{\mathbb{E}}_0[f(t, T_1)])(f(t, T_2) - \tilde{\mathbb{E}}_0[f(t, T_2)])] \\
&= \tilde{\mathbb{E}}_0\left[\sum_{i=1}^{n}\int_0^t \sigma_i(s, T_1, \omega(s))d\tilde{W}_i(s) \cdot \sum_{i=1}^{n}\int_0^t \sigma_i(s, T_2, \omega(s))d\tilde{W}_i(s)\right].
\end{aligned} \tag{25.117}$$

Using the result

$$\mathbb{E}_0[\int_0^t \sigma_i(s, T_1)d\tilde{W}_i(s)\int_0^t \sigma_i(s, T_2)d\tilde{W}_i(s)] = \int_0^t \sigma_i(s, T_1)\sigma_i(s, T_2)ds,$$

and the covariance is given by

$$\mathrm{cov}_0[f(t, T_1), f(t, T_2)] = \sum_{i=1}^{n}\int_0^t \sigma_i(s, T_1)\sigma_i(s, T_2)ds. \tag{25.118}$$

Considering again the volatility functions of the form

$$\sigma_i(s, t) = \sigma_{0i}e^{-\kappa_{\sigma i}(t-s)}, \qquad i = 1, \ldots, n, \tag{25.119}$$

then

$$\int_0^t \sigma_i(s, T_1)\sigma_i(s, T_2)ds = \frac{\sigma_{0i}^2 e^{-\kappa_{\sigma i}(T_1+T_2)}}{2\kappa_{\sigma i}}(e^{2\kappa_{\sigma i}t} - 1). \tag{25.120}$$

The covariance between the forward rates $f(t, T_1)$ and $f(t, T_2)$ becomes

$$\mathrm{cov}_0[f(t, T_1), f(t, T_2)] = \sum_{i=1}^{n}\frac{\sigma_{0i}^2 e^{-\kappa_{\sigma i}(T_1+T_2)}}{2\kappa_{\sigma i}}(e^{2\kappa_{\sigma i}t} - 1), \tag{25.121}$$

and the correlation coefficient $\rho(t, T_1, T_2)$ between the forward rates $f(t, T_1)$ and $f(t, T_2)$ is evaluated as

$$\frac{\sum_{i=1}^{n}\frac{\sigma_{0i}^2 e^{-\kappa_{\sigma i}(T_1+T_2)}}{2\kappa_{\sigma i}}(e^{2\kappa_{\sigma i}t} - 1)}{\sqrt{\mathrm{var}_0[f(t, T_1)]}\sqrt{\mathrm{var}_0[f(t, T_2)]}}, \tag{25.122}$$

where the variance of the forward rate $f(t, T_h)$ $(h = 1, 2)$ is

$$\mathrm{var}_0[f(t, T_h)] = \sum_{i=1}^{n}\frac{\sigma_{0i}^2 e^{-2\kappa_{\sigma i}T_h}}{2\kappa_{\sigma i}}(e^{2\kappa_{\sigma i}t} - 1). \tag{25.123}$$

## 25.11 Appendix

### Appendix 25.1 Proof of Proposition 25.1

Recall that $r(t)$ satisfies the stochastic integral equation (25.23) and $f(t,\tau)$ satisfies the stochastic integral equation (25.26) with $T$ set equal to $\tau$. We assume the forward rate volatility specifications

$$\sigma\left(v,t,\omega(v)\right)=\bar{\sigma}e^{-\lambda(t-v)}g\left(r(v),f(v,\tau)\right)$$

and set

$$\begin{aligned}\sigma^*\left(v,t,\omega(v)\right)&=\sigma\left(v,t,\omega(v)\right)\int_v^t\sigma\left(v,s,\omega(v)\right)ds\\&=\bar{\sigma}^2e^{-\lambda(t-v)}g\left(r(v),f(v,\tau)\right)\int_v^t e^{-\lambda(s-v)}g\left(r(v),f(v,\tau)\right)ds\\&=\bar{\sigma}^2g^2\left(r(v),f(v,\tau)\right)e^{-\lambda(t-v)}\left(\frac{1-e^{-\lambda(t-v)}}{\lambda}\right).\end{aligned}$$

Note that the first integral term in Eq. (25.23) can be written

$$\begin{aligned}\int_0^t\sigma^*\left(v,t,\omega(v)\right)dv&=\bar{\sigma}^2\int_0^t g^2\left(r(v),f(v,\tau)\right)e^{-\lambda(t-v)}\frac{\left(1-e^{-\lambda(t-v)}\right)}{\lambda}dv\\&=\frac{e^{-\lambda t}\bar{\sigma}^2}{\lambda}\int_0^t g^2(r(v),f(v,\tau))e^{\lambda v}dv\\&\quad-\frac{e^{-2\lambda t}\bar{\sigma}^2}{\lambda}\int_0^t g^2(r(v),f(v,\tau))e^{2\lambda v}dv\\&\equiv\frac{e^{-\lambda t}}{\lambda}I(t;\lambda)-\frac{e^{-2\lambda t}}{\lambda}I(t;2\lambda).\end{aligned}$$

Next note that the second integral in Eq. (25.23) may be written as

$$\begin{aligned}\int_0^t\sigma\left(v,t,\omega(v)\right)d\tilde{W}(v)&=\bar{\sigma}\int_0^t e^{-\lambda(t-v)}g\left(r(v),f(v,\tau)\right)d\tilde{W}(v)\\&=\bar{\sigma}e^{-\lambda t}\int_0^t g\left(r(v),f(v,\tau)\right)e^{\lambda v}d\tilde{W}(v)\\&\equiv e^{-\lambda t}J(t;\lambda).\end{aligned}$$

Similarly the first integral term in Eq. (25.26) can be written

$$\int_0^t \sigma^*(v,\tau,\omega(v))\,dv = \bar{\sigma}^2 \int_0^t g^2(r(v), f(v,\tau))\, e^{-\lambda(\tau-v)} \frac{\left(1-e^{-\lambda(\tau-v)}\right)}{\lambda} dv$$
$$= \frac{e^{-\lambda\tau}}{\lambda} I(t;\lambda) - \frac{e^{-2\lambda\tau}}{\lambda} I(t;2\lambda).$$

The second integral term in Eq. (25.26) may be similarly treated, so that

$$\int_0^t \sigma(v,\tau,\omega(v))d\tilde{W}(v) = \bar{\sigma}\int_0^t e^{-\lambda(\tau-v)} g(r(v), f(v,\tau))d\tilde{W}(v)$$
$$= \bar{\sigma}e^{-\lambda\tau}\int_0^t e^{\lambda v} g(r(v), f(v,\tau))d\tilde{W}(v)$$
$$\equiv e^{-\lambda\tau} J(t;\lambda).$$

We may thus write the stochastic integral equations for $r(t)$ and $f(t,\tau)$ in terms of the integrals $I(t;\lambda)$, $I(t;2\lambda)$ and $J(t;\lambda)$ as

$$r(t) = f(0,t) + \frac{e^{-\lambda t}}{\lambda} I(t;\lambda) - \frac{e^{-2\lambda t}}{\lambda} I(t;2\lambda) + e^{-\lambda t} J(t;\lambda), \tag{25.124}$$

$$f(t,\tau) = f(0,\tau) + \frac{e^{-\lambda\tau}}{\lambda} I(t;\lambda) - \frac{e^{-2\lambda\tau}}{\lambda} I(t;2\lambda) + e^{-\lambda\tau} J(t;\lambda). \tag{25.125}$$

We note that Eqs. (25.124) and (25.125) can be re-expressed as

$$r(t) - f(0,t) + \frac{e^{-2\lambda t}}{\lambda} I(t;2\lambda) = e^{-\lambda t}\left[\frac{I(t;\lambda)}{\lambda} + J(t;\lambda)\right],$$
$$f(t,\tau) - f(0,\tau) + \frac{e^{-2\lambda\tau}}{\lambda} I(t;2\lambda) = e^{-\lambda\tau}\left[\frac{I(t;\lambda)}{\lambda} + J(t;\lambda)\right].$$

We may combine the above equations to express $I(t;2\lambda)$ as a function of $r(t)$ and $f(t,\tau)$, i.e.,

$$I(t;2\lambda) = \frac{\lambda e^{\lambda\tau}}{e^{-\lambda t} - e^{-\lambda\tau}}\left[f(t,\tau) - f(0,\tau)\right] - \frac{\lambda e^{\lambda t}}{e^{-\lambda t} - e^{-\lambda\tau}}\left[r(t) - f(0,t)\right] \tag{25.126}$$

Finally we note that

$$\psi(t) = \int_0^t \sigma^2(v,t,\omega(v))\,dv = \bar{\sigma}^2\int_0^t e^{-2\lambda(t-v)} g^2(r(v), f(v,\tau))\,dv$$
$$= \bar{\sigma}^2 e^{-2\lambda t}\int_0^t e^{2\lambda v} g^2(r(v), f(v,\tau))\,dv = e^{-2\lambda t} I(t;2\lambda).$$

Thus we finally have

$$\psi(t) = \lambda\alpha(t,\tau)\,[r(t) - f(0,t)] - \lambda e^{-\lambda(t-\tau)}\alpha(t,\tau)\,[f(t,\tau) - f(0,\tau)]\,, \tag{25.127}$$

where we set

$$\alpha(t,\tau) \equiv \frac{e^{-\lambda t}}{e^{-\lambda\tau} - e^{-\lambda t}}.$$

## Appendix 25.2 Proof of Proposition 25.2

It is readily verified that the manipulations that led to Eq. (25.125) of Appendix 25.1 are equally valid for $t$ set to a general maturity $T$. Thus (25.126) holds for $t$ set to $T$, i.e.,

$$\begin{aligned} I(t;2\lambda) &= \frac{\lambda e^{\lambda T}}{e^{-\lambda t}-e^{-\lambda T}}\,[f(t,T)-f(0,T)] - \frac{\lambda e^{\lambda t}}{e^{-\lambda t}-e^{-\lambda T}}\,[r(t)-f(0,t)] \\ &= e^{2\lambda t}\psi(t). \end{aligned}$$

Substituting the expression for $\psi(t)$ we find that

$$\begin{aligned} I(t;2\lambda) &= \lambda e^{2\lambda t}\left(\alpha(t,\tau)[r(t) - f(0,t)] - e^{-\lambda(t-\tau)}\alpha(t,\tau)[f(t,\tau) - f(0,\tau)]\right) \\ &= \lambda\left(\frac{e^{\lambda T}}{e^{-\lambda t} - e^{-\lambda T}}[f(t,T) - f(0,T)] - \frac{e^{\lambda t}}{e^{-\lambda t} - e^{-\lambda T}}[r(t) - f(0,t)]\right). \end{aligned}$$

On rearranging

$$\begin{aligned} \frac{e^{\lambda T}}{e^{-\lambda t} - e^{-\lambda T}}[f(t,T) - f(0,T)] = &\frac{e^{\lambda t}}{e^{-\lambda t} - e^{-\lambda T}}[r(t) - f(0,t)] \\ &+e^{2\lambda t}\alpha(t,\tau)[r(t) - f(0,t)] \\ &-e^{2\lambda t}e^{-\lambda(t-\tau)}\alpha(t,\tau)[f(t,\tau) - f(0,\tau)], \end{aligned}$$

from which

$$\begin{aligned} f(t,T) - f(0,T) = [r(t) - f(0,t)]&\left(\frac{e^{\lambda t}}{e^{\lambda T}} + \frac{e^{2\lambda t}(e^{-\lambda t} - e^{-\lambda T})}{e^{\lambda T}}\alpha(t,\tau)\right) \\ &- \frac{e^{2\lambda t}e^{-\lambda(t-T)}}{e^{\lambda T}}\alpha(t,\tau)(e^{-\lambda t} - e^{-\lambda T})[f(t,\tau) - f(0,\tau)]. \end{aligned} \tag{25.128}$$

Consider the following:

(i)

$$
\begin{aligned}
&\frac{e^{\lambda t}+e^{2\lambda t}(e^{-\lambda t}-e^{-\lambda T})}{e^{\lambda T}}\alpha(t,\tau)=\frac{e^{\lambda t}}{e^{\lambda T}}+\frac{e^{2\lambda t}(e^{-\lambda t}-e^{-\lambda T})e^{-\lambda t}}{e^{\lambda T}(e^{-\lambda\tau}-e^{-\lambda t})}\\
&=\frac{e^{\lambda t}(e^{-\lambda\tau}-e^{-\lambda t})+e^{\lambda t}(e^{-\lambda t}-e^{-\lambda T})}{e^{\lambda T}(e^{-\lambda\tau}-e^{-\lambda t})}=\frac{e^{\lambda t}(e^{-\lambda\tau}-e^{-\lambda T})}{e^{\lambda T}(e^{-\lambda\tau}-e^{-\lambda t})}\\
&=\frac{e^{2\lambda t}}{e^{2\lambda T}}\frac{e^{-\lambda t}}{e^{-\lambda\tau}-e^{-\lambda t}}\frac{e^{-\lambda\tau}-e^{-\lambda T}}{e^{-\lambda T}}=e^{-2\lambda(T-t)}\frac{\alpha(t,\tau)}{\alpha(T,\tau)}
\end{aligned}
$$

(ii)

$$
\begin{aligned}
&e^{2\lambda t}e^{-\lambda(t-\tau)}\alpha(t,\tau)\frac{(e^{-\lambda t}-e^{-\lambda T})}{e^{\lambda T}}=e^{2\lambda t-\lambda t+\lambda\tau}\alpha(t,\tau)\frac{(e^{-\lambda t}-e^{-\lambda T})}{e^{2\lambda T}e^{-\lambda T}}\\
&=\frac{e^{\lambda t}e^{\lambda\tau}e^{-\lambda t}(e^{-\lambda t}-e^{-\lambda T})}{e^{2\lambda T}e^{-\lambda T}(e^{-\lambda\tau}-e^{-\lambda t})}=\frac{e^{2\lambda\tau}}{e^{2\lambda T}}\frac{e^{-\lambda\tau}}{-(e^{-\lambda t}-e^{-\lambda\tau})}\frac{(e^{-\lambda t}-e^{-\lambda T})}{e^{-\lambda T}}\\
&=-e^{-2\lambda(T-\tau)}\frac{\alpha(\tau,t)}{\alpha(T,t)}.
\end{aligned}
$$

Hence Eq. (25.128) can be rewritten

$$
\begin{aligned}
f(t,T)-f(0,T)&=e^{-2\lambda(T-t)}\frac{\alpha(t,\tau)}{\alpha(T,\tau)}[r(t)-f(0,t)]\\
&\quad-e^{-2\lambda(T-\tau)}\frac{\alpha(\tau,t)}{\alpha(T,t)}[f(t,\tau)-f(0,\tau)]
\end{aligned}
$$

where

$$
\alpha(\theta_1,\theta_2)\equiv\frac{e^{-\lambda\theta_1}}{e^{-\lambda\theta_2}-e^{-\lambda\theta_1}}.
$$

We have thus proved Proposition 25.2.

## Appendix 25.3 Details of the Infinitesimal Generator $\mathscr{K}$

We recall the following result from Sect. 5.4 concerning the infinitesimal generator of an n dimensional Ito process. In our application we set

$$
\begin{aligned}
X_1&\equiv f(t,\tau),\\
a_1&\equiv\sigma^2(t,\tau,\omega(t))\frac{(e^{\lambda(\tau-t)}-1)}{\lambda},
\end{aligned}
$$

$$\begin{aligned}
\sigma_{11} &\equiv \sigma_1 \equiv \sigma(t, \tau, \omega(t)),\\
X_2 &\equiv r(t),\\
a_2 &\equiv f_2(0,t) + \lambda f(0,t) + \psi(t) - \lambda r(t),\\
\sigma_{21} &\equiv \sigma_r \equiv \sigma(t, t, \omega(t)).
\end{aligned}$$

Thus the matrix $S$ assumes the form

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_r \\ \sigma_1\sigma_r & \sigma_r^2 \end{bmatrix}.$$

Using the foregoing expression for $S$ the expression for the operator $\mathscr{K}$ in Eq. (25.67) is readily derived.

## Appendix 25.4 Proof of Proposition 25.3

Using the relationship

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right)$$

and Eq. (25.26) for the forward rate $f(t,s)$ we obtain for the bond price the expression

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left[-\left(\int_t^T \int_0^t \sigma^*(v,s,\cdot)dvds + \int_t^T \int_0^t \sigma(v,s,\cdot)d\tilde{W}(v)ds\right)\right],$$

where

$$\begin{aligned}
\sigma(v,T,\cdot) &= \bar{\sigma}e^{-\lambda(T-v)}g(r(v), f(v,\tau))\\
\sigma^*(v,T,\cdot) &= \sigma(v,T,\cdot)\int_v^T \sigma(v,s,\cdot)ds\\
&= \bar{\sigma}^2 g^2(r(v), f(v,\tau))e^{-\lambda(T-v)}\int_v^T e^{-\lambda(s-v)}ds.
\end{aligned}$$

Set

$$\begin{aligned}
I &= \int_t^T \int_0^t \sigma^*(v,s,\cdot)dvds + \int_t^T \int_0^t \sigma(v,s,\cdot)d\tilde{W}(v)ds\\
&\equiv I_1 + I_2\\
&= \int_0^t \int_t^T \sigma^*(v,s,\cdot)dsdv + \int_0^t \int_t^T \sigma(v,s,\cdot)dsd\tilde{W}(v),
\end{aligned}$$

where we have interchanged the order of integration to obtain the last equality. Next note that

$$\begin{aligned}
\int_t^T \sigma^*(v,s,\cdot)ds &= \sigma\left(r(v), f(v,\tau)\right) \int_t^T e^{-\lambda(s-v)} \int_v^s e^{-\lambda(y-v)} \sigma\left(r(v), f(v,\tau)\right) dyds \\
&= \sigma\left(r(v), f(v,\tau)\right) \int_t^T e^{-\lambda(s-v)} \left\{ \int_v^t e^{-\lambda(y-v)} \sigma\left(r(v), f(v,\tau)\right) dy \right. \\
&\quad \left. + \int_t^s e^{-\lambda(y-v)} \sigma\left(r(v), f(v,\tau)\right) dy \right\} ds \\
&= \bar{\sigma}^2 g^2(r(v), f(v,\tau)) \int_t^T e^{-\lambda(s-v)} ds \int_v^t e^{-\lambda(y-v)} dy \\
&\quad + \bar{\sigma}^2 g^2(r(v), f(v,\tau)) \int_t^T e^{-\lambda(s-v)} \int_t^s e^{-\lambda(y-v)} dyds \\
&= \bar{\sigma}^2 g^2(r(v), f(v,\tau)) e^{-\lambda(t-v)} \left( \int_t^T e^{-\lambda(s-t)} ds \right) \int_v^t e^{-\lambda(y-v)} dy \\
&\quad + \bar{\sigma}^2 g^2(r(v), f(v,\tau)) e^{-2\lambda(t-v)} \int_t^T e^{-\lambda(s-t)} \int_t^s e^{-\lambda(y-t)} dyds \\
&= \sigma^*(v,t,\cdot)\beta(t,T) + \sigma^2(v,t,\cdot)\alpha(t,T),
\end{aligned}$$

where

$$\begin{aligned}
\beta(t,T) &= \int_t^T e^{-\lambda(s-t)} ds = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right), \\
\alpha(t,T) &= \int_t^T e^{-\lambda(s-t)} \int_t^s e^{-\lambda(y-t)} dyds = \frac{1}{2}\beta^2(t,T),
\end{aligned}$$

i.e. we have shown that

$$\int_t^T \sigma^*(v,s,\cdot)ds = \beta(t,T)\sigma^*(v,t,\cdot) + \frac{1}{2}\beta^2(t,T)\sigma^2(v,t,\cdot).$$

Next consider

$$\begin{aligned}
\int_t^T \sigma(v,s,\cdot)ds &= \int_t^T e^{-\lambda(s-v)} \sigma\left(r(v), f(v,\tau)\right) ds \\
&= \sigma\left(r(v), f(v,\tau)\right) e^{-\lambda(t-v)} \left( \int_t^T e^{-\lambda(s-t)} ds \right),
\end{aligned}$$

i.e. we have shown that

$$\int_t^T \sigma(v,s,\cdot)ds = \sigma(v,t,\cdot)\beta(t,T).$$

Returning to the expressions for $I_1$, $I_2$ we can now write

$$I_1 = \int_0^t \left[\beta(t,T)\sigma^*(v,t,\cdot) + \frac{1}{2}\beta^2(t,T)\sigma^2(v,t,\cdot)\right]dv,$$

and

$$I_2 = \int_0^t \beta(t,T)\sigma(v,t,\cdot)d\tilde{W}(v),$$

so that

$$\begin{aligned} I &= \frac{1}{2}\beta^2(t,T)\int_0^t \sigma^2(v,t,\cdot)dv \\ &+\beta(t,T)\left[\int_0^t \sigma^*(v,t,\cdot)dv + \int_0^t \sigma(v,t,\cdot)d\tilde{W}(v)\right]. \end{aligned}$$

However we note from Eq. (25.23), for the instantaneous spot rate $r(t)$, that

$$\int_0^t \sigma^*(v,t,\cdot)dv + \int_0^t \sigma(v,t,\cdot)d\tilde{W}(v) = r(t) - f(0,t).$$

Hence

$$I = \frac{1}{2}\beta^2(t,T)\int_0^t \sigma^2(v,t,\cdot)dv + \beta(t,T)\left[r(t) - f(0,t)\right].$$

Recalling the definition of the subsidiary stochastic variable $\psi(t)$ we can finally write

$$I = \frac{1}{2}\beta^2(t,T)\psi(t) + \beta(t,T)\left[r(t) - f(0,t)\right].$$

Hence the expression for the bond price may be written as in Proposition 25.3.

## 25.12 Problems

**Problem 25.1** Show that the Hull–White model can be obtained within the Heath–Jarrow–Morton framework by setting

$$\sigma(t,T) = \bar{\sigma}e^{-k(T-t)},$$

where $\bar{\sigma}, k$ are constants.

**Problem 25.2** The Heath–Jarrow–Morton model takes as its starting point a stochastic differential equation for the instantaneous forward rate of the form

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t),$$

and from this determines the stochastic dynamics of the instantaneous spot rate spot rate $r(t)$ and pure discount bond price $P(t,T)$.

Suppose instead we take as the starting point a stochastic differential equation for $P(t,T)$ of the form

$$\frac{dP(t,T)}{P(t,T)} = \beta(t,T)dt + \delta(t,T)dW(t).$$

Determine the corresponding stochastic dynamics for $r(t)$ and $f(t,T)$.

Express in terms of $\beta(t,T)$ and $\delta(t,T)$ the Heath–Jarrow–Morton drift restriction that guarantees no riskless arbitrage opportunities between bonds of different maturities.

**Problem 25.3** In Sect. 23.6 we considered the volatility function

$$\sigma(t,T) = \bar{\sigma}e^{-\lambda(T-t)}$$

and showed how this allowed the system dynamics to be Markovianised.

Now consider the volatility function

$$\sigma(t,T) = [\sigma_0 + \sigma_1(T-t)]e^{-\lambda(T-t)}.$$

Show the system dynamics can be Markovianised in this case. In particular obtain the stochastic differential equations for the bond price and the instantaneous spot interest rate.

*Hint*: You will need to obtain a linked stochastic differential equation system for

$$Z_1(t) = \int_0^t (t-v)e^{-\lambda(t-v)}dW(v),$$

and

$$Z_0(t) = \int_0^t e^{-\lambda(t-v)}dW(v).$$

**Problem 25.4** The Ho–Lee model is obtained within the Heath–Jarrow–Morton framework by setting

$$\sigma(t,T) = \bar{\sigma},$$

where $\bar{\sigma}$ is a constant. Show that

$$f(t,T) = r(t) + f(0,T) - f(0,t) + \bar{\sigma}^2 t(T-t).$$

By obtaining the dynamics for $r(t)$ under the risk neutral measure, show also that

$$r(t) = f(0,t) + \frac{1}{2}\sigma^2 t^2 + \bar{\sigma}\tilde{W}(t).$$

Hence, show that for this model the bond price is given by

$$P(t,T) = \exp[-a(t,T) - (T-t)r(t)],$$

where

$$a(t,T) = \ln\frac{P(0,t)}{P(0,T)} - (T-t)f(0,t) + \frac{1}{2}\bar{\sigma}^2 t(T-t)^2.$$

**Problem 25.5 Computational Problem**—Consider the Heath–Jarrow–Morton model with the volatility function

$$\sigma(t,T) = \sigma_0 e^{-\lambda(T-t)}.$$

We know in this case that the dynamics for the instantaneous spot rate are given by [Eq. (25.54)].

Take $\sigma_0 = 0.02$ and $\lambda = 0.6$. Assume also that the initial forward curve is given by

$$f(0,T) = 0.08 - 0.03e^{-1.5T}.$$

Consider the bond pricing formula [Eq. (25.29)]. Write a program to calculate the bond price by simulating the stochastic differential equation for $r(t)$ from 0 to $t$ and performing the $\tilde{\mathbb{E}}_t$ operation by simulating a large number of paths from $t$ to $T$. This will give the bond price conditional on the value of $r(t)$ that has been obtained.

You can check the accuracy of your algorithm (and hence choose appropriate $\Delta t$ and number of paths) by using the fact that when $t = 0$ we have the exact solution

$$P(0,T) = \exp\left(-\int_0^T f(0,s)ds\right).$$

Use this to check the accuracy for $T = 0.5, 1.0, 1.5$ and 2.0.

Then use the simulation procedure to calculate $P(0.5, 1.0)$, $P(0.5, 1.5)$ and $P(0.5, 2.0)$.

Note that the evaluation of $\rho(t, T)$ will be conditional on the interest rate $r(t)$. Obtain $r(t)$ by simulating from 0 to $t$ and be sure to specify the value of $r(t)$ that you are using.

Check the accuracy of these approximations by using the exact bond-pricing formula (here you need to refer to Sect. 23.4.2, but use the $\theta(t)$ that arises in the Heath–Jarrow–Morton model).

# Chapter 26
# The LIBOR Market Model

**Abstract** The modifications to the Heath-Jarrow-Morton framework to cater for market quoted rates such as LIBOR rates were carried out by Brace and Musiela (Math Finance 4(3):259–283, 1994) (henceforth BM). In this chapter, we first describe the BM parameterisation of the Heath–Jarrow–Morton model, and then we outline the choice of volatility functions that produces lognormal dynamics for LIBOR rates. We also discuss the pricing of interest rate caps and swaptions in this framework. In the final section, we summarise the earlier effort to price an interest rate caplet when the forward rate dynamics are Gaussian (i.e. the volatility function is only time dependent).

## 26.1 Introduction

The Heath–Jarrow–Morton model provides the most general framework for the analysis of interest rate derivatives—it calibrates automatically to the currently observed forward curve and by appropriate choice of the forward rate volatilities it can generate a wide range of specific models. However from the point of view of practical implementations, the Heath–Jarrow–Morton model still requires some further development since the underlying forward rate is an instantaneous rate, whereas market quoted rates are for some discrete time period. A good example would be LIBOR rates that are quoted for periods such as 3-months or 6-months. Such rates are typically the reference rate for the interest rate caps and floors described in Chap. 21. The developments that led to what is now known as the LIBOR market model were reported in Brace and Musiela (1994) and Brace et al. (1997) (henceforth BGM), where it was in particular shown how to choose the volatility functions (and a change of measure) so that LIBOR rates follow log-normal processes. It was this latter assumption that allowed BGM to obtain Black–Scholes type formulae for the value of interest caps and floors—these formulae are of the same form as those obtained for the Merton model of stochastic interest rates in Chap. 20 and the Hull–White model in Sect. 23.7.1. These results brought together a soundly based theory of interest rate dynamics and the market practice of using Black's model to price interest rate caps, that prior to BGM had been frowned upon by finance academics. The theoretical developments advanced



C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5_26

the market practice by demonstrating how to properly specify the volatility function and so to better calibrate the resulting models to market data. It should be pointed out that similar results were also derived by Miltersen et al. (1997). Expressing the forward rate dynamics in terms of discrete rates also overcame the bothersome feature of the Heath–Jarrow–Morton model that log-normal instantaneous forward rates are badly behaved (they can go to infinity in finite time). This feature is not shared by discrete forward rates, which are well behaved as shown by Sandmann and Sondermann (1997) and by BGM.

## 26.2 The Brace–Musiela Parameterisation of the Heath–Jarrow–Morton Model

We recall the Heath–Jarrow–Morton stochastic integral equation for the instantaneous forward rates, under the risk-neutral measure, namely[1]

$$f(t,T) = f(0,T) + \int_0^t \alpha(v,T)dv + \int_0^t \sigma(v,T)d\tilde{W}(v), \tag{26.1}$$

where $t$ the time at which the rate is quoted is regarded as variable and the maturity time $T$ at which the rate applies is fixed. Note that (26.1) is arbitrage free when (see (25.26))

$$\alpha(v,T) = \sigma(v,T)\int_v^T \sigma(v,s)ds. \tag{26.2}$$

BM consider the constant period ahead forward rate

$$r(t,x) \equiv f(t,t+x), \qquad x > 0. \tag{26.3}$$

Thus $r(t,x)$ is the rate an investor can contract at time $t$ for instantaneous borrowing/lending at time $t+x$, where $x$ is fixed e.g. $x$ equals 3 months. The time line for $r(t,x)$ is displayed in Fig. 26.1 and should be contrasted with Fig. 22.13 in Chap. 22. Setting $T = t + x$ in (26.1) we have

$$f(t,t+x) = f(0,t+x) + \int_0^t \alpha(v,t+x)dv + \int_0^t \sigma(v,t+x)d\tilde{W}(v). \tag{26.4}$$

BM further define

$$\tau(t,x) \equiv \sigma(t,t+x), \tag{26.5}$$

[1]For notational convenience, we omit the dependence of the drift and volatility function to the path dependent quantities $\omega$.

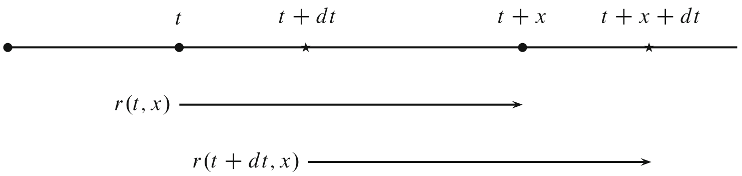


**Fig. 26.1** The time line for the instantaneous constant period ahead forward rate $r(t, x)$

and

$$a(t, x) \equiv \alpha(t, t + x),$$

though the latter quantity drops out of the preference free representation. With this notation we can write

$$\sigma(v, t + x) = \sigma(v, v + (t + x - v)) = \tau(v, t + x - v),$$

and

$$\alpha(v, t + x) = \alpha(v, v + (t + x - v)) = a(v, t + x - v).$$

The reason for this apparently roundabout notation will become apparent when the Heath–Jarrow–Morton drift restriction is adapted to the BM notation in (26.7) below. With the above identifications and noting that

$$f(0, t + x) = r(0, t + x),$$

then (26.4) can be written in terms of BM notation as

$$r(t, x) = r(0, t+x) + \int_0^t a(v, t+x-v)dv + \int_0^t \tau(v, t+x-v)d\tilde{W}(v). \quad (26.6)$$

In order to obtain the arbitrage free dynamics for $r(t, x)$ we use the drift restriction (26.2), hence

$$\begin{aligned} a(t, x) \equiv \alpha(t, t + x) &= \sigma(t, t + x) \int_t^{t+x} \sigma(t, s)ds \\ &= \sigma(t, t + x) \int_0^x \sigma(t, t + y)dy \\ &= \tau(t, x) \int_0^x \tau(t, y)dy \quad (26.7) \\ &= \tau(t, x)\psi(t, x), \quad (26.8) \end{aligned}$$

where we define the integrated volatility function[2]

$$\psi(t,x) = \int_0^x \tau(t,y)dy. \tag{26.9}$$

With this notation we obtain

$$a(t,x) = \frac{\partial}{\partial x}\left(\frac{1}{2}\psi^2(t,x)\right), \tag{26.10}$$

and

$$\tau(t,x) = \frac{\partial}{\partial x}\psi(t,x). \tag{26.11}$$

These results will be useful in later calculations. We have already mentioned the notation to define the BM volatility function at (26.5). This results in the function $\psi(t,x)$ not having time $t$ appearing in the second argument. This turns out to be crucial in order to carry out the volatility transformations in Sect. 26.4. The arbitrage free dynamics of $r(t,x)$ are given by (recall Eq. (26.6))

$$r(t,x) = r(0,t+x)+\int_0^t a(v,t+x-v)dv+\int_0^t \tau(v,t+x-v)d\tilde{W}(v), \tag{26.12}$$

with $a(t,x)$ given by (26.8). Equation (26.12) is the transformation to BM notation of (25.26) for the instantaneous Heath–Jarrow–Morton forward rate under $\tilde{\mathbb{P}}$.

However this equation is in the form of a stochastic integral equation, we also need to express it in the form of a stochastic differential equation. First of all by differentiating (26.12) with respect to $x$, we note that

$$\begin{aligned}\frac{\partial}{\partial x}r(t,x) = r_2(t,x) = r_2(0,t+x) + \int_0^t a_2(v,t+x-v)dv \\ + \int_0^t \tau_2(v,t+x-v)d\tilde{W}(v).\end{aligned} \tag{26.13}$$

Now we form the differential with respect to time in Eq. (26.12), so that

$$\begin{aligned}dr(t,x) = \Big[r_2(0,t+x) + a(t,x) + \int_0^t a_2(v,t+x-v)dv \\ + \int_0^t \tau_2(v,t+x-v)d\tilde{W}(v)\Big]dt + \tau(t,x)d\tilde{W}(t).\end{aligned} \tag{26.14}$$

[2] Actually BGM use $\sigma$, but here we prefer to use $\psi$, given that $\sigma$ is used for the Heath–Jarrow–Morton volatility function.

By making use of (26.13) the drift term above can be written more compactly, thus (26.14) may be expressed as

$$dr(t,x) = \left[\frac{\partial}{\partial x} r(t,x) + a(t,x)\right] dt + \tau(t,x) d\tilde{W}(t). \tag{26.15}$$

By use of (26.10) and (26.11), Eq. (26.15) may be written in the alternative form

$$dr(t,x) = \frac{\partial}{\partial x}\left[\left(r(t,x) + \frac{1}{2}\psi^2(t,x)\right) dt + \psi(t,x) d\tilde{W}(t)\right], \tag{26.16}$$

which is essentially BGM's equation (1.1).[3]

The price at time $t$ of a pure discount bond maturing at time $T$ and the Heath–Jarrow–Morton forward rate are related by

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right),$$

which in terms of the BM forward rate notation becomes

$$P(t,T) = \exp\left(-\int_t^T r(t,s-t)ds\right). \tag{26.17}$$

We shall refer to the quantity $P(t,T)$ as the Heath–Jarrow–Morton bond price and emphasise that it matures at a *fixed date* ahead. This will be in contrast to the BM bond price, which matures at a *fixed period* ahead. By changing the variable $u$ to $u = s - t$ in Eq. (26.17), we have that

$$P(t,T) = \exp\left(-\int_t^T r(t,s-t)ds\right) = \exp\left(-\int_0^{T-t} r(t,u)du\right). \tag{26.18}$$

## 26.3 The LIBOR Process

We introduce the $\delta$-period forward LIBOR rate $L(t,x)$ related to the BM instantaneous forward rate $r(t,x)$ according to

$$1 + \delta L(t,x) = \exp\left(\int_x^{x+\delta} r(t,u)du\right). \tag{26.19}$$

---

[3]Note, BGM allow for a vector of noise processes.

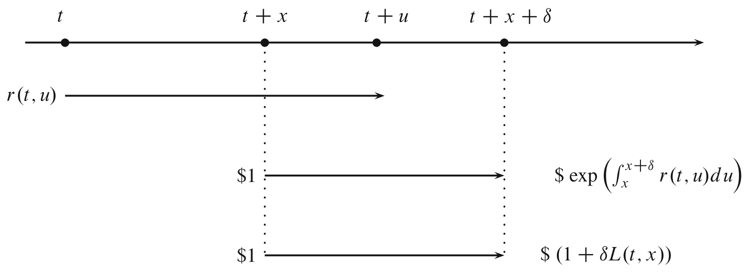


**Fig. 26.2** The time line for the definition of the $\delta$-period LIBOR rate

We refer the reader to the time line in Fig. 26.2 for the economic intuition behind the definition of $L(t,x)$. This is the simply compounded rate that an investor can contract at time $t$ for borrowing/lending over the period $(t+x, t+x+\delta)$. On the RHS of Eq. (26.19) is the value to which \$1 accumulates if it is invested at the continuously compounded BM instantaneous forward rate over the period $(t+x, t+x+\delta)$. Equating (26.19) merely states that the simply compounded LIBOR rate must be set so as to be equivalent to investing at the continuously compounded rate. Strictly speaking we should use a notation such as $L(t,x,\delta)$ or $L_\delta(t,x)$ to denote this rate in order to emphasise the dependence on the compounding period $\delta$, but we will use the simpler notation $L(t,x)$.

Note that according to (26.18), the definition (26.19) relates LIBOR rates and bond prices as

$$\begin{aligned}1+\delta L(t,x) &= \exp\left(\int_x^{x+\delta} r(t,u)du\right) = \exp\left(\int_0^{x+\delta} r(t,u)du - \int_0^x r(t,u)du\right)\\ &= \frac{P(t,t+x)}{P(t,t+x+\delta)}. \end{aligned} \tag{26.20}$$

We determine next the stochastic differential equation for $L(t,x)$. First we use the dynamics (26.6) to evaluate the quantity

$$\begin{aligned}V(t,x) &= \int_x^{x+\delta} r(t,u)du\\ &= \int_x^{x+\delta} r(0,t+u)du + \int_x^{x+\delta}\int_0^t a(v,t+u-v)dvdu\\ &\quad + \int_x^{x+\delta}\int_0^t \tau(v,t+u-v)d\tilde{W}(v)du\end{aligned}$$

$$= \int_x^{x+\delta} r(0, t+u)du + \int_0^t \int_x^{x+\delta} a(v, t+u-v)dudv$$
$$+ \int_0^t \int_x^{x+\delta} \tau(v, t+u-v)dud\tilde{W}(v).$$

Taking the differential with respect to $t$ we obtain

$$dV = \left[ \int_x^{x+\delta} r_2(0, t+u)du + \int_0^t \int_x^{x+\delta} a_2(v, t+u-v)dudv \right.$$
$$\left. + \int_0^t \int_x^{x+\delta} \tau_2(v, t+u-v)dud\tilde{W}(v) + \int_x^{x+\delta} a(t, u)du \right] dt \tag{26.21}$$
$$+ \left( \int_x^{x+\delta} \tau(t, u)du \right) d\tilde{W}(t).$$

Next we replace $x$ by $u$ in (26.13) and integrate to obtain

$$\int_x^{x+\delta} r_2(t, u)du = \int_x^{x+\delta} r_2(0, t+u)du + \int_x^{x+\delta} \int_0^t a_2(v, t+u-v)dvdu$$
$$+ \int_x^{x+\delta} \int_0^t \tau_2(v, t+u-v)d\tilde{W}(v)du,$$

which after performing the integration on the LHS becomes

$$r(t, x+\delta) - r(t, x) = \int_x^{x+\delta} r_2(0, t+u)du + \int_0^t \int_x^{x+\delta} a_2(v, t+u-v)dudv$$
$$+ \int_0^t \int_x^{x+\delta} \tau_2(v, t+u-v)dud\tilde{W}(v).$$

This last result may be used to rewrite (26.21) as

$$dV = \left[ r(t, x+\delta) - r(t, x) + \int_x^{x+\delta} a(t, u)du \right] dt \tag{26.22}$$
$$+ \left( \int_x^{x+\delta} \tau(t, u)du \right) d\tilde{W}(t).$$

We note from (26.10) and (26.11) that

$$\int_x^{x+\delta} a(t,u)du = \frac{1}{2}\left(\psi^2(t,x+\delta) - \psi^2(t,x)\right),$$

and

$$\int_x^{x+\delta} \tau(t,u)du = \psi(t,x+\delta) - \psi(t,x).$$

Thus (26.22) may finally be written as

$$dV = \mu_V dt + \sigma_V d\tilde{W}(t), \tag{26.23}$$

where

$$\begin{aligned} \mu_V &= r(t,x+\delta) - r(t,x) + \frac{1}{2}\left(\psi^2(t,x+\delta) - \psi^2(t,x)\right), \\ \sigma_V &= \psi(t,x+\delta) - \psi(t,x). \end{aligned} \tag{26.24}$$

Since $L(t,x) = \delta^{-1}(e^{V(t,x)} - 1)$ we may apply Ito's lemma to determine the stochastic differential equation satisfied by $L(t,x)$, thus

$$dL = \delta^{-1}e^V(\mu_V + \frac{1}{2}\sigma_V^2)dt + \delta^{-1}e^V\sigma_V d\tilde{W}(t).$$

From (26.24) we have that

$$\mu_V + \frac{1}{2}\sigma_V^2 = r(t,x+\delta) - r(t,x) + \psi(t,x+\delta)(\psi(t,x+\delta) - \psi(t,x)),$$

and as $1 + \delta L(t,x) = e^{V(t,x)}$, we find that

$$\begin{aligned} dL = \delta^{-1}(1+\delta L(t,x))\,[r(t,x+\delta) - r(t,x) + \psi(t,x+\delta)\,(\psi(t,x+\delta) \\ -\psi(t,x))]\,dt + \delta^{-1}(1+\delta L(t,x))(\psi(t,x+\delta) - \psi(t,x))d\tilde{W}(t). \end{aligned} \tag{26.25}$$

Equation (26.25) together with the stochastic differential equation (26.14) (or (26.15)) for $r(t,x)$ form a two-dimensional stochastic dynamic system for $r(t,x)$ and $L(t,x)$. In general we would expect this system to be non-Markovian unless we make special assumptions about the volatility function $\tau(t,x)$ of the kind we investigated in Chap. 25.

BGM prefer to express (26.25) in a slightly different form by observing from (26.19) that

$$\begin{aligned}\frac{\partial}{\partial x}L(t,x) &= \delta^{-1}\exp\left(\int_x^{x+\delta} r(t,u)du\right)(r(t,x+\delta)-r(t,x))\\ &= \delta^{-1}e^{V(t,x)}(r(t,x+\delta)-r(t,x))\\ &= \delta^{-1}(1+\delta L(t,x))(r(t,x+\delta)-r(t,x)),\end{aligned}$$

use of which yields the equation for the LIBOR rate $L(t,x)$ as

$$\begin{aligned}dL = &\left[\frac{\partial}{\partial x}L(t,x)\right.\\ &\left.+\delta^{-1}(1+\delta L(t,x))\psi(t,x+\delta)(\psi(t,x+\delta)-\psi(t,x))\right]dt\\ &+\delta^{-1}(1+\delta L(t,x))(\psi(t,x+\delta)-\psi(t,x))d\tilde{W}(t).\end{aligned} \tag{26.26}$$

Because of the appearance of the $\frac{\partial}{\partial x}L(t,x)$ in the drift, Eq. (26.26) is in fact a stochastic partial differential equation and is quite a complicated mathematical object. A great deal of the BGM article is devoted to demonstrating that Eq. (26.26) yields a well defined stochastic process.

## 26.4 Lognormal LIBOR Rates

BGM then ask the question: "What volatility function would result in the LIBOR process equation (26.26) having a lognormal volatility structure?" The answer is that $\psi(t,x)$ must satisfy a relationship

$$\delta^{-1}(1+\delta L(t,x))(\psi(t,x+\delta)-\psi(t,x)) = \gamma(t,x)L(t,x), \tag{26.27}$$

where $\gamma(t,x)$ is some function of time and maturity. Equation (26.27) may be re-expressed as

$$\psi(t,x+\delta)-\psi(t,x) = \frac{\delta L(t,x)}{1+\delta L(t,x)}\gamma(t,x) \tag{26.28}$$

which defines $\psi(t,x)$ for all $x \geq \delta$, provided $\psi(t,x)$ is defined on the initial interval $0 \leq x < \delta$. To see how $\psi(t,x)$ is built up, consider the maturity time line in Fig. 26.3.

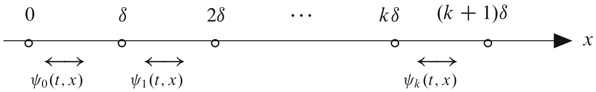


**Fig. 26.3** The sequential determination of the $\psi_k(t,x)$

We use $\psi_k(t,x)$ to denote $\psi(t,x)$ on the interval $k\delta \le x < (k+1)\delta$ and assume $\psi_0(t,x)$ is given. Then on $\delta \le x < 2\delta$ we have

$$\psi_1(t,x) = \psi_0(t,x-\delta) + \frac{\delta L(t,x-\delta)}{1+\delta L(t,x-\delta)}\gamma(t,x-\delta),$$

whilst on $2\delta \le x < 3\delta$ we have

$$\begin{aligned}\psi_2(t,x) &= \psi_1(t,x-\delta) + \frac{\delta L(t,x-\delta)}{1+\delta L(t,x-\delta)}\gamma(t,x-\delta)\\ &= \psi_0(t,x-2\delta) + \frac{\delta L(t,x-2\delta)}{1+\delta L(t,x-2\delta)}\gamma(t,x-2\delta)\\ &\quad + \frac{\delta L(t,x-\delta)}{1+\delta L(t,x-\delta)}\gamma(t,x-\delta),\end{aligned}$$

and we see that in general on $k\delta \le x < (k+1)\delta$ holds

$$\psi_k(t,x) = \psi_0(t,x-k\delta) + \sum_{j=1}^{k} \frac{\delta L(t,x-j\delta)}{1+\delta L(t,x-j\delta)}\gamma(t,x-j\delta). \tag{26.29}$$

To finally recover the Heath–Jarrow–Morton volatility function we recall from Eq. (26.11) that

$$\tau(t,x) = \frac{\partial}{\partial x}\psi(t,x).$$

With the above choice of volatility function the stochastic differential equation (26.26) for $L(t,x)$ may be written[4]

[4]Note that by use of Eq. (26.28) the second term in the drift of Eq. (26.26) becomes $\psi(t,x+\delta)L(t,x)\gamma(t,x)$. Then use again of Eq. (26.28) written as

$$\psi(t,x+\delta) = \psi(t,x) + \frac{\delta L(t,x)}{1+\delta L(t,x)}\gamma(t,x)$$

yields the result.

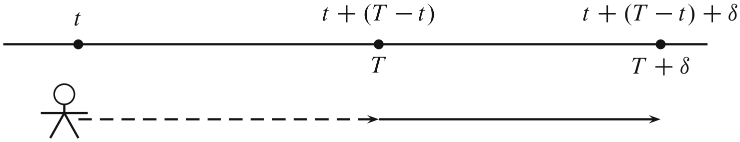


**Fig. 26.4** The time line for the rate $K(t, T)$

$$dL = \left[\frac{\partial}{\partial x}L(t,x) + L(t,x)\gamma(t,x)\psi(t,x) + \frac{\delta L^2(t,x)\gamma^2(t,x)}{1+\delta L(t,x)}\right]dt + \gamma(t,x)L(t,x)d\tilde{W}(t), \tag{26.30}$$

which has a lognormal volatility structure. Whilst we have succeeded in finding a volatility structure that yields log-normal dynamics for the LIBOR rate, we need to keep in mind that the "cost" of this choice is a very complicated drift term. However as we shall see the other technique upon which BGM rely, namely the change of measure, makes irrelevant how complicated the drift term may be. The dynamics (26.30) can further be simplified by considering the process

$$K(t,T) = L(t, T-t), \quad (0 \le t \le T). \tag{26.31}$$

We can obtain the appropriate time line for $K(t,T)$ by setting $x = T - t$ in Fig. 26.2, as shown in Fig. 26.4.

Hence we see that $K(t,T)$ is the rate an investor can contract at time $t$ for borrowing/lending between $T$ and $T + \delta$. Thus $K(T,T) = L(T,0)$ is the rate an investor can contract at time $T$ for borrowing/lending between $T$ and $T + \delta$. By taking differentials we see that $K(t,T)$ satisfies

$$dK(t,T) = dL(t, T-t) - L_2(t, T-t)dt, \tag{26.32}$$

where $L_2$ denotes partial derivative with respect to the second argument. In deriving this last result we make use of the fact that $L(t,\cdot)$ is a smooth function of its second argument, a result which is proven by BGM. Using (26.32) in conjunction with (26.30) (with $x$ set to $T - t$) we find that the dynamics of $K(t,T)$ are given by

$$dK = K(t,T)\gamma(t,T-t)\left[\psi(t,T-t) + \frac{\delta K(t,T)}{1+\delta K(t,T)}\gamma(t,T-t)\right]dt + K(t,T)\gamma(t,T-t)d\tilde{W}(t). \tag{26.33}$$

By use of (26.28) with $x = T - t$, this last equation reduces to

$$dK = K(t,T)\gamma(t,T-t)\left[\psi(t,T+\delta-t)dt + d\tilde{W}(t)\right]. \tag{26.34}$$

We define the new Wiener process $W^T(t)$ by

$$dW^T(t) = \psi(t, T - t)dt + d\tilde{W}(t),$$

i.e.

$$W^T(t) = W(t) + \int_0^t \psi(s, T - s)ds. \tag{26.35}$$

Let us consider the Radon–Nikodym derivative of the $T$-forward measure. By direct application of (25.41), the associated Radon–Nikodym derivative is given by

$$\xi(0, T) = \frac{1}{A(T)P(0, T)}.$$

This quantity can be calculated using the expression for the bond price in the Appendix 26.1 (see Eq. (26.84)),

$$\frac{P(t, T)}{A(t)} = P(0, T)\exp\left[-\frac{1}{2}\int_0^t \psi^2(s, T - s)ds - \int_0^t \psi(s, T - s)d\tilde{W}(s)\right]$$

so that, by setting $t = T$ in the last equation, we obtain

$$\xi(0, T) = \exp\left[-\frac{1}{2}\int_0^T \psi^2(s, T - s)ds - \int_0^T \psi(s, T - s)d\tilde{W}(s)\right]. \tag{26.36}$$

Given the expression for the Radon–Nikodym derivative between $\tilde{\mathbb{P}}$ and $\mathbb{P}^T$ in Eq. (26.36), we see that $W^T(t)$ will be a standard Wiener process under $\mathbb{P}^T$. Thus we can re-express (26.34) as

$$dK = K(t, T)\gamma(t, T - t)dW^{T+\delta}(t), \tag{26.37}$$

which implies that $K(t, T)$ is lognormally distributed with zero drift under $\mathbb{P}^{T+\delta}$.

The pricing of interest rate derivatives such as caps, floors and swaptions requires the modelling of a series of forward LIBOR rates. Typically, forward rates reset at certain reset dates during the life of the interest rate derivatives. If $(T_0, T_n)$ is the tenor and $T_j = T_0 + j\delta$, for $j = 0, 1, \ldots, n - 1$, are the reset dates then we denote as $L_j(t) := L(t, T_j - t) = K(t, T_j)$ the $\delta$-period forward LIBOR rate that resets at time $T_j$. By introducing the notation $\gamma_j(t) = \gamma(t, T_j - t)$ and

$\psi_j(t) = \psi(t, T_j - t)$ and by using (26.34), the risk-neutral dynamics of the rate $L_j(t)$ would be expressed as[5]

$$dL_j = L_j(t)\gamma_j(t)\left[\psi_{j+1}(t)dt + d\tilde{W}(t)\right]. \tag{26.38}$$

Similarly from (26.37), the dynamics of the rate $L_j(t)$ under the $T_{j+1}$-forward measure would be[6]

$$dL_j = L_j(t)\gamma_j(t)dW^{T_{j+1}}(t). \tag{26.39}$$

Consequently we have shown that the rate $L_j(t)$ is lognormally distributed with zero drift under the $T_{j+1}$-forward measure, i.e.,

$$\mathbb{E}_t^{T_{j+1}}[L_j(T_j)] = L_j(t). \tag{26.40}$$

Equation (26.40) also implies that the forward LIBOR rate $L_j(t)$ is a martingale under the $T_{j+1}$-forward measure.

## 26.5 Pricing Caps

Consider a forward cap on a unit principal amount settled in arrears at times $T_{j+1}$, for $j = 0, 1, \cdots, n-1$. For the caplet over $(T_j, T_{j+1})$, the associated forward LIBOR rate $L_j(T_j)$ resets at time $T_j$ and the cash flow at time $T_{j+1}$ is $\delta(L_j(T_j) - E)^+$, where $E$ is the exercise rate of the caplet (see Fig. 26.5). Thus the value $C_{pl}(t, T_j)$ of this caplet under the risk-neutral measure can be expressed as

$$C_{pl}(t, T_j) = \delta\tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T_{j+1})}\left(L_j(T_j) - E\right)^+\right], \tag{26.41}$$

[5] Note that

$$\psi(t, T_j + \delta - t) = \psi(t, T_{j+1} - t) = \psi_{j+1}(t).$$

[6] Note that

$$dW^{T_j+\delta}(t) = dW^{T_{j+1}}(t).$$

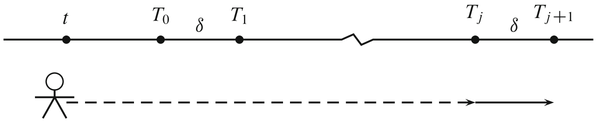


**Fig. 26.5** The time line for the caplet that resets at time $T_j$

where $A(t)$ is the money market account (25.10). We switch to the measure with numeraire the price of the bond maturing at time $T_{j+1}$ and use $\mathbb{E}_t^{T_{j+1}}$ to denote expectation with respect to this measure (see Eq. (20.16) in Sect. 20.2)

$$C_{pl}(t, T_j) = \delta \mathbb{E}_t^{T_{j+1}} \left[ \frac{P(t, T_{j+1})}{P(T_{j+1}, T_{j+1})} \left( L_j(T_j) - E \right)^+ \right],$$

so that

$$C_{pl}(t, T_j) = \delta P(t, T_{j+1}) \mathbb{E}_t^{T_{j+1}} \left[ \left( L_j(T_j) - E \right)^+ \right]. \tag{26.42}$$

Since the drift in Eq. (26.39) is zero, then (26.42) implies that we have a standard Black–Scholes model with interest rate being zero. Thus the expectation in (26.42) can be expressed as

$$L_j(t) \mathcal{N}(h(t, T_j)) - E \mathcal{N}(h(t, T_j) - \zeta(t, T_j)),$$

where

$$h(t, T_j) = \frac{\ln(\frac{L_j(t)}{E}) + \frac{1}{2}\zeta^2(t, T_j)}{\zeta(t, T_j)}, \tag{26.43}$$

and

$$\zeta^2(t, T_j) = \int_t^{T_j} \gamma_j^2(s) ds. \tag{26.44}$$

We have also used the result that when the diffusion term is time varying the volatility of the Black–Scholes formula is replaced by the integrated volatility (26.44), as we have seen in several previous chapters. Note that another common representation of this result involves the average volatility

$$\xi^2(t, T_j) = \frac{1}{T_j - t} \zeta^2(t, T_j), \tag{26.45}$$

that re-expresses (26.43) as

$$h(t, T_j) = \frac{\ln(\frac{L_j(t)}{E}) + \frac{1}{2}\xi^2(t, T_j)(T_j - t)}{\xi(t, T_j)\sqrt{T_j - t}}. \quad (26.46)$$

In order to implement the model we would need to specify some functional form for $\psi_0(t, x)$ and $\gamma(t, x)$. For implementation details we refer the reader to Brigo and Mercurio (2006) and Brace (2007).

## 26.6 Pricing Swaptions

### 26.6.1 Swaps

An interest rate swap (IRS) is an agreement between two parties to exchange floating-rate payments for fixed-rate payments on prespecified future dates. The party who agrees to pay the fixed rate and to receive the floating rate holds a "payer IRS" while the party who agrees to receive the fixed rate and to pay the floating rate holds a "receiver IRS". Note that in an IRS net payments are made. Consider a payer IRS on a unit principal amount settled in arrears at times $T_{j+1}$, $j = 0, 1, \cdots, n-1$. Thus the tenor of the swap is $(T_0, T_n)$. At every reset date $T_{j+1}$, the fixed cash flow paid is $\delta R$, where $R$ is the fixed rate and the floating cash-flow received is $\delta L_j(T_j)$, where the floating rate $L_j(T_j)$ resets at $T_j$ and pays at $T_{j+1}$. Thus the floating rates reset at dates $T_0, T_1, \ldots T_{n-1}$ and pay at dates $T_1, T_2, \ldots T_n$. Figure 26.6 illustrates the cash flow of a payer IRS. The solid-line arrows represents the fixed cash flows paid and the dashed-line arrows represents the floating cash flow received. Figure 26.7 depicts the net cash flow of a payer IRS with the initial net cash flow at $T_0$ being zero by definition. Thus for a payer IRS, the net cash flow paid at $T_{j+1}$ is $\delta(L_j(T_j) - R)$.

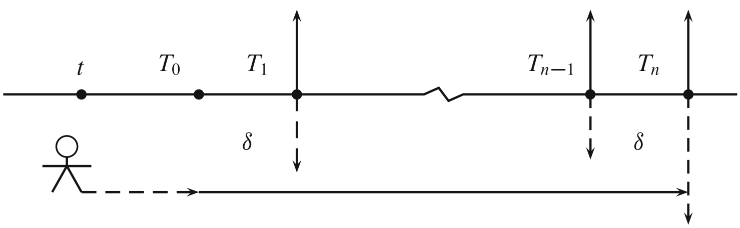


**Fig. 26.6** The cash flow for a payer IRS

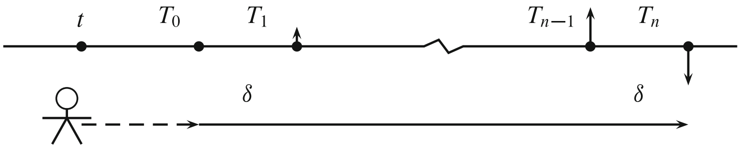


**Fig. 26.7** Net cash flow for a payer IRS

The value at time $t \leq T_0$ of the payer IRS is denoted as $IRS_p(t, T_0, T_n, R)$ and it can be evaluated under the $T_{j+1}$-forward measure by using the price of the bond maturing at time $T_{j+1}$ as the numeraire, i.e.

$$\begin{aligned} IRS_p(t, T_0, T_n, R) &= \delta \sum_{j=0}^{n-1} \mathbb{E}_t^{T_{j+1}} \left[ \frac{P(t, T_{j+1})}{P(T_{j+1}, T_{j+1})} \left( L_j(T_j) - R \right) \right] \\ &= \delta \sum_{j=0}^{n-1} P(t, T_{j+1}) \mathbb{E}_t^{T_{j+1}} [L_j(T_j)] - \delta R \sum_{j=0}^{n-1} P(t, T_{j+1}). \end{aligned} \tag{26.47}$$

By using the fact that $L_j(t)$ is a martingale under the $T_{j+1}$-forward measure (see Eq. (26.40)), we express the value of the payer IRS (26.47) as

$$IRS_p(t, T_0, T_n, R) = \delta \sum_{j=0}^{n-1} P(t, T_{j+1}) L_j(t) - \delta R \sum_{j=0}^{n-1} P(t, T_{j+1}). \tag{26.48}$$

The fixed rate of the contract that makes the IRS a fair contract at time $t$ is called the forward swap rate for the tenor $(T_0, T_n)$ and it is denoted as $R_{o,n}(t)$. This implies that the condition $IRS_p(t, T_0, T_n, R_{o,n}(t)) = 0$ should be satisfied. By setting the expression (26.48) of the payer IRS equal to zero, the forward swap rate $R_{o,n}(t)$ may be expressed as

$$R_{o,n}(t) = \sum_{j=0}^{n-1} w_j(t) L_j(t), \tag{26.49}$$

with

$$w_j(t) = \frac{P(t, T_{j+1})}{\sum_{k=0}^{n-1} P(t, T_{k+1})}. \tag{26.50}$$

According to Eq. (26.49), the forward swap rate is a weighted average of the forward rates over the tenor of the swap, as $0 < w_j(t) \leq 1$ and $\sum_{j=0}^{n-1} w_j(t) = 1$.[7]

Another expression of the payer IRS value can be obtained by using the relationship (26.20) between forward LIBOR rates and bond prices. Accordingly, we have that

$$L_j(t) = \frac{1}{\delta}\Big(\frac{P(t,T_j)}{P(t,T_{j+1})} - 1\Big). \tag{26.51}$$

Thus the value of the payer IRS (26.48) is simplified to

$$\begin{aligned} IRS_p(t,T_0,T_n,R) &= \sum_{j=0}^{n-1}(P(t,T_j) - P(t,T_{j+1})) + P(t,T_n) \\ &\quad - \delta R \sum_{j=0}^{n-1} P(t,T_{j+1}) - P(t,T_n) \\ &= P(t,T_0) - \delta R \sum_{j=0}^{n-1} P(t,T_{j+1}) - P(t,T_n). \end{aligned} \tag{26.52}$$

From (26.52), it is clear that the value of an IRS does not depend on the volatility or the correlations of the underlying forward rates, a well known property of IRS pricing. By setting the IRS price (26.52) equal to zero, the forward swap rate is expressed as

$$R_{o,n}(t) = \frac{P(t,T_0) - P(t,T_n)}{\delta \sum_{j=0}^{n-1} P(t,T_{j+1})}. \tag{26.53}$$

By re-arranging (26.53) we have that

$$P(t,T_0) - P(t,T_n) = \delta R_{o,n}(t) \sum_{j=0}^{n-1} P(t,T_{j+1}). \tag{26.54}$$

[7] This result can lead to a useful approximation. Empirically the variability of the $w_j$'s is relatively small compared to the variability of the $L$'s. By approximating the $w_j$'s by $w_j(0)$ then the volatility of the swap rate can be estimated by the volatility of forward rates by using the approximation

$$R_{o,n}(t) \approx \sum_{j=0}^{n-1} w_j(0) L_j(t).$$

Substitution of (26.54) into the value of the payer IRS (26.52) leads to

$$IRS_p(t, T_0, T_n, R) = \delta \sum_{j=0}^{n-1} P(t, T_{j+1})(R_{o,n}(t) - R). \tag{26.55}$$

This representation will be used in the next section to price swaptions.

Note that by using (26.53), the forward swap rate can be expressed solely in terms of the underlying forward LIBOR rates. By dividing the numerator and the denominator of Eq. (26.53) by $P(t, T_0)$ we obtain

$$R_{o,n}(t) = \frac{1 - \frac{P(t,T_n)}{P(t,T_0)}}{\delta \sum_{j=0}^{n-1} \frac{P(t,T_{j+1})}{P(t,T_0)}}. \tag{26.56}$$

From (26.51) we have that

$$\frac{1}{1 + \delta L_j(t)} = \frac{P(t, T_{j+1})}{P(t, T_j)}, \tag{26.57}$$

therefore for $j = 0, 1, \ldots, n-1$ the ratio $P(t, T_{j+1})/P(t, T_0)$ can be expressed as

$$\frac{P(t, T_{j+1})}{P(t, T_0)} = \prod_{m=0}^{j} \frac{P(t, T_{m+1})}{P(t, T_m)} = \prod_{m=0}^{j} \frac{1}{1 + \delta L_m(t)} = \frac{1}{\prod_{m=0}^{j}(1 + \delta L_m(t))}. \tag{26.58}$$

Then the forward swap rate is expressed as a function of the underlying LIBOR rates $L_m(t)$ as

$$R_{o,n}(t) = \frac{1 - \prod_{m=0}^{n-1} \frac{1}{1+\delta L_m(t)}}{\delta \sum_{j=0}^{n-1} \prod_{m=0}^{j} \frac{1}{1+\delta L_m(t)}}. \tag{26.59}$$

## 26.6.2 Swaptions

A European payer swaption is an option giving the right to enter a payer IRS at a given future time, which is the swaption's maturity. The swaption maturity is usually the first reset date of the underlying IRS, which is $T_0$ in our setting. Thus the payoff of the payer swaption at its maturity $T_0$ is

$$\Omega_n(T_0) = \max\left[IRS_p(T_0, T_0, T_n, R), 0\right].$$

In order to price the swaption we express the price of an option on a swap as a price of an option on the corresponding forward swap rate. Indeed, the payoff of the swaption can be expressed as the payoff of a (call) option on the forward swap rate, by using the expression Eq. (26.55) for the value of the payer IRS. Thus the payoff of the swaption at $T_0$ is given by

$$\Omega_n(T_0) = \delta \sum_{j=0}^{n-1} P(T_0, T_{j+1})(R_{o,n}(T_0) - R)^+. \tag{26.60}$$

The fixed rate $R$ of the underlying IRS becomes the swaption strike. When the forward swap rate at maturity $T_0$ of the swaption is greater than the IRS fixed rate $R$ then the payer swaption will be exercised.

One may believe that by assuming lognormal dynamics for the forward swap rate under the appropriate martingale measure will allow us to price the swaptions. Note that the forward swap rate and its associated forward rates cannot simultaneously be lognormal. To proceed, an approximation will be employed. Equation (26.60) implies the payoff of the swaption can be expressed as the product of an option on the forward swap rate and the annuity

$$C_n(t) = \delta \sum_{j=0}^{n-1} P(t, T_{j+1}), \tag{26.61}$$

for $t \leq T_0$. Using this annuity as a new numeraire we switch to a new equivalent martingale measure namely the forward swap measure $\mathbb{P}^S$, where the value of the swaption $\Omega_n(t)$ at time $t$ satisfies[8]

$$\frac{\Omega_n(t)}{C_n(t)} = \mathbb{E}^{\mathbb{S}}\Big[\frac{\Omega_n(T_0)}{C_n(T_0)}\Big] = \mathbb{E}^{\mathbb{S}}\big[(R_{o,n}(T_0) - R)^+\big]. \tag{26.62}$$

The Radon–Nikodym derivative of the new measure $\mathbb{P}^S$ with respect to the risk-neutral measure is $\tilde{\mathbb{P}}$, where

$$\frac{d\mathbb{P}^S}{d\tilde{\mathbb{P}}} = \xi^Q(0,t) = \frac{C_n(t)/C_n(0)}{A(t)/A(0)} = \frac{\delta}{C_n(0)} \sum_{j=0}^{n-1} \frac{P(t, T_{j+1})}{A(t)}. \tag{26.63}$$

By using Eq. (26.49) that expresses the forward swap rate as a function of the underlying forward LIBOR rates and the risk-neutral dynamics (26.38) of the

[8]From Eq. (26.53), the forward swap rate times the annuity $C_n(t)$ is equivalent to a bond portfolio consisting of a long bond with maturity $T_0$ and a short bond with maturity $T_n$. Thus the forward swap rate evolves as the value of this bond portfolio denominated by the annuity $C_n(t)$ and becomes a martingale under the forward swap measure $\mathbb{P}^S$.

forward LIBOR rates and by applying the Ito's lemma we derive the dynamics of the forward swap rate $R_{o,n}(t)$ under the risk-neutral measure as, see Appendix 26.2 for details.

$$dR_{o,n} = \sum_{j=0}^{n-1} \frac{\partial R_{o,n}(t)}{\partial L_j(t)} \gamma_j(t) L_j(t) \Big( d\tilde{W}(t) - \sum_{m=0}^{n-1} w_m(t) \psi_{m+1}(t) dt \Big). \tag{26.64}$$

We define the new Wiener process $W^S(t)$ so that

$$dW^S(t) = d\tilde{W}(t) - \sum_{m=0}^{n-1} w_m(t) \psi_{m+1}(t) dt,$$

where given the expression for the Radon–Nikodym derivative between $\tilde{\mathbb{P}}$ and $\mathbb{P}^S$ in Eq. (26.63) we see that $W^S(t)$ is a standard Wiener process under $\mathbb{P}^S$. Thus we can re-express Eq. (26.64) as

$$dR_{o,n} = \sum_{j=0}^{n-1} \frac{\partial R_{o,n}(t)}{\partial L_j(t)} \gamma_j(t) L_j(t) dW^S(t). \tag{26.65}$$

These dynamics are not lognormal and we will use an approximation to obtain lognormal dynamics. From Eq. (26.65) we have that for all $\nu \geq t$

$$\begin{aligned} dR_{o,n} &= R_{o,n}(\nu) \sum_{j=0}^{n-1} \frac{\partial R_{o,n}(\nu)}{\partial L_j} \frac{L_j(\nu)}{R_{o,n}(\nu)} \gamma_j(\nu) dW^S(\nu) \\ &\approx R_{o,n}(\nu) \sum_{j=0}^{n-1} \frac{\partial R_{o,n}(t)}{\partial L_j} \frac{L_j(t)}{R_{o,n}(t)} \gamma_j(\nu) dW^S(\nu) \\ &= R_{o,n}(\nu) \sum_{j=0}^{n-1} \varpi_j \gamma_j(\nu) dW^S(\nu), \end{aligned} \tag{26.66}$$

with

$$\varpi_j = \frac{\partial R_{o,n}(t)}{\partial L_j} \frac{L_j(t)}{R_{o,n}(t)}.$$

The $R_{o,n}(t)$ is now lognormally distributed with zero drift under $\mathbb{P}^S$. This approximate process was obtained by using the frozen coefficient technique that allows to relax the dependence of state variables to the model coefficients. This is the well known lognormal forward swap model. Thus Black's formula for swaptions can be employed, by using the expectation at time $t$ under the measure $\mathbb{P}^S$

$$\begin{aligned}\Omega_n(t) &= C_n(t)\mathbb{E}_t^S[\frac{\Omega_n(T_0)}{C_n(T_0)}] = C_n(t)\mathbb{E}_t^S[(R_{o,n}(T_0) - R)^+] \\ &= C_n(t)(R_{o,n}(t)\mathcal{N}(d_1) - R\mathcal{N}(d_2)),\end{aligned} \tag{26.67}$$

where

$$d_1(t,T) = \frac{\ln(\frac{R_{o,n}(t)}{R}) + \frac{1}{2}\sigma_n^2(T_0 - t)}{\sigma_n\sqrt{T_0 - t}}, \tag{26.68}$$

and

$$d_2(t,T) = d_1(t,T) - \sigma_n\sqrt{T_0 - t}. \tag{26.69}$$

with $\sigma_n$ the variances of $\ln R_{o,n}(t)$ computed as

$$\sigma_n^2 = \frac{1}{T_0 - t}\int_t^{T_0} \sum_{j,k=0}^{n-1} \varpi_j \varpi_k \gamma_j(s)\gamma_k(s)ds. \tag{26.70}$$

Note that for $n = 1$, $R_{o,1}(t) = L_1(t)$ and the swaption pricing formula equation (26.67) is reduced to the caplet pricing formula equation (26.78).

## 26.7 Pricing a Caplet Under Gaussian Forward Rate Dynamics

We present next the original BG derivation for pricing caplets. There is now probably less interest in this model by practitioners due to the dominance of the LIBOR market model, but it nevertheless constitutes an important model.

We recall the relation between the yield-to-maturity, $y(t,T_1)$, and the bond price $P(t,T_1)$ over $(t,T_1)$, namely

$$P(t,T_1) = e^{-y(t,T_1)(T_1-t)},$$

which we rewrite as

$$\ln P(t,T_1)^{-1} = y(t,T_1)(T_1 - t). \tag{26.71}$$

Using the approximation

$$\ln x \simeq -1 + x,$$

we obtain

$$-1 + P(t, T_1)^{-1} = y(t, T_1)(T_1 - t).$$

Setting

$$f(t) \equiv y(t, T_1),$$

and rearranging we obtain

$$P^{-1}(t, T_1) = 1 + f(t)(T_1 - t),$$

which becomes the relationship used by Brace–Musiela at $t = T$, namely

$$P^{-1}(T, T_1) = 1 + f(T)\delta.$$

From Eq. (26.71) we see that the stochastic differential equation followed by $y(t, T_1)$ (i.e. $f(t)$) is essentially the one followed by $\ln P(t, T_1)$ which is given by Eq. (25.25) of our discussion on the Heath–Jarrow–Morton framework. In fact

$$df = \left[ \frac{f(t) - r + \frac{1}{2}\sigma_B^2(t, T_1)}{T_1 - t} \right] dt - \frac{\sigma_B(t, T_1)}{T_1 - t} d\tilde{W}(t),$$

where

$$\sigma_B(t, T_1) = -\int_t^{T_1} \sigma(t, u) du.$$

In this version of their model Brace–Musiela assume that the volatility of the instantaneous forward rate is proportional to the yield-to-maturity, that is

$$\sigma(t, u) = \tilde{\sigma} f(t),$$

from which

$$\sigma_B(t, T_1) = -\tilde{\sigma}(T_1 - t) f(t).$$

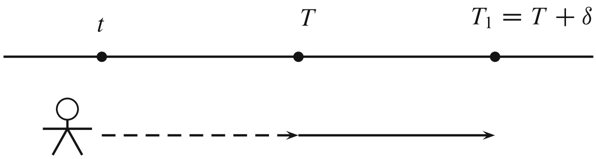


**Fig. 26.8** The time line for a caplet on a BM forward rate

Hence

$$df = \left[\frac{f(t) - r + \frac{1}{2}\tilde{\sigma}^2(T_1 - t)^2 f(t)^2}{T_1 - t}\right] dt + \tilde{\sigma} f(t) d\tilde{W}(t).$$

An application of Girsanov's theorem will yield

$$df = \tilde{\sigma} f(t) dW^*,$$

where $W^*$ is a new Brownian motion under a new measure. In fact,

$$dW^*(t) = \left[\frac{1}{\tilde{\sigma}}\left(\frac{f(t) - r + \frac{1}{2}\tilde{\sigma}^2(T_1 - t)^2 f(t)^2}{T_1 - t}\right) dt + d\tilde{W}(t)\right].$$

We denote with $C_{pl}(t, T)$ the value at time $t$ of a caplet maturing at $T$ on a BM "forward rate" maturing at $T_1 = T + \delta$.

See Fig. 26.8 for the appropriate time-line. Recalling the notation

$$A(t) = \exp\left(\int_0^t r(s)ds\right)$$

where $r(s)$ is the instantaneous spot interest rate, (i.e. using Heath–Jarrow–Morton notation) and noting that settlement on the caplet is in arrears, so that

$$C_{pl}(t, T) = \delta R \tilde{\mathbb{E}}_t \left[\frac{A(t)}{A(T_1)}(f(T) - k)^+\right]$$

where $R$ is the principal amount and $k$ is cap rate. Since $(f(T) - k)^+$ is a payoff dependent on $P(t, T_1)$ we are able to price according to the Heath–Jarrow–Morton framework using $\tilde{\mathbb{E}}$. We can re-express the value of the caplet in terms of $P(T, T_1)$ since

$$f(T) = \frac{1}{\delta}\left(\frac{1}{P(T, T_1)} - 1\right).$$

Thus

$$C_{pl}(t) = R\tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T_1)}\left(\frac{1}{P(T,T_1)} - (1+\delta k)\right)^+\right]$$
$$= R(1+\delta k)\tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T_1)P(T,T_1)}\left(\frac{1}{1+\delta k} - P(T,T_1)\right)^+\right]. \quad (26.72)$$

This looks very close to a put option (maturity $T$) written on a zero coupon bond maturing at time $T_1$, we just need to adjust the discounting with respect to $T$ rather than $T_1$. Thus

$$C_{pl}(t) = R(1+\delta k)\tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T)P(T,T_1)}\left(\frac{1}{1+\delta k} - P(T,T_1)\right)^+ \tilde{\mathbb{E}}_T\left(\frac{A(T)}{A(T_1)}\right)\right], \quad (26.73)$$

where we have used the (telescope) law of iterated expectations. However, we have that

$$P(T,T_1) = \tilde{\mathbb{E}}_T\left[\exp\left(-\int_T^{T_1} r(s)ds\right)\right] = \tilde{\mathbb{E}}_T\left[\frac{A(T)}{A(T_1)}\right].$$

Hence

$$C_{pl}(t) = R(1+\delta k)\tilde{\mathbb{E}}_t\left[\frac{A(t)}{A(T)}\left(\frac{1}{1+\delta k} - P(T,T_1)\right)^+\right]. \quad (26.74)$$

This is now the payoff on a put option (maturity $T$) written on a pure discount bond (maturity $T_1$). Using the forward measure we can re-express the caplet value as

$$C_{pl}(t) = R(1+\delta k)P(t,T)\mathbb{E}_t^T\left[\left(\frac{1}{1+\delta k} - P(T,T_1)\right)^+\right], \quad (26.75)$$

where $\mathbb{E}_t^T$ is the expectation calculated under the forward measure. In order to calculate expectations under the forward measure we need to determine the dynamics of

$$F_T(t,T_1) = \frac{P(t,T_1)}{P(t,T)},$$

since in terms of this quantity Eq. (26.75) becomes

$$C_{pl}(t) = R(1+\delta k)P(t,T)\mathbb{E}_t^T\left[\left(\frac{1}{1+\delta k} - F_T(T,T_1)\right)^+\right].$$

We already know the dynamics of $P(t,T)$ and $P(t,T_1)$ under $\tilde{W}$ hence by Ito's lemma we find that (again we recall the result in Sect. 6.6)

$$\begin{aligned} dF_T(t,T_1) = {} & F_T(t,T_1)\int_T^{T_1}\sigma(t,v)dv\int_t^T\sigma(t,v)dvdt \\ & - F_T(t,T_1)\left(\int_T^{T_1}\sigma(t,v)dv\right)d\tilde{W}(t), \end{aligned} \tag{26.76}$$

so that

$$dF_T(t,T_1) = -F_T(t,T_1)\int_T^{T_1}\sigma(t,v)dv\left[\left(-\int_t^T\sigma(t,v)dv\right)dt + d\tilde{W}(t)\right]. \tag{26.77}$$

If we define a new Wiener process

$$W^T(t,T) = -\int_0^t\int_s^T\sigma(s,v)dvds + \tilde{W}(t),$$

so that

$$dW^T(t,T) = -\int_t^T\sigma(t,v)dvdt + d\tilde{W}(t),$$

then by Girsanov's theorem $W^T(t,T)$ is a Wiener process under the forward measure $\mathbb{P}_T$. Hence we reexpress Eq. (26.76) as

$$dF_T(t,T_1) = -\zeta(t)F_T(t,T_1)dW^T(t,T),$$

where

$$\zeta(t) = \int_T^{T_1}\sigma(t,v)dv.$$

The distribution for $\mathbb{P}_T$ will be the Fokker–Plank equation (or Kolmogorov equation) associated with this stochastic differential equation. The result will be equivalent to using Black's formula for futures options but using $\zeta(t)$ as the volatility function. Thus

$$C_{pl}(t) = R(1+\delta k)P(t,T)\left[\frac{1}{1+\delta k}\mathcal{N}(-h+\bar{\zeta}\sqrt{T-t}) - P(T,T_1)\mathcal{N}(-h)\right], \tag{26.78}$$

where

$$h = \frac{\ln((1+\delta k)P(t,T_1)) + \frac{1}{2}\bar{\zeta}^2(T-t)}{\bar{\zeta}\sqrt{T-t}}, \qquad \bar{\zeta}^2(T-t) = \int_t^T \zeta^2(s)ds.$$

## 26.8 Appendix

### Appendix 26.1 Bond Price Dynamics in BM Notation

In this section we derive the Heath–Jarrow–Morton bond price dynamics in terms of the BM notation. To this end we consider the expression involving the integral of the BM forward rate in (26.17). Thus in Eq. (26.6) set $x = u - t$ and integrate with respect to $u$ to obtain

$$\begin{aligned}\int_t^T r(t,u-t)du &= \int_t^T \left[r(0,u) + \int_0^t a(s,u-s)ds + \int_0^t \tau(s,u-s)dW(s)\right]du \\ &= \int_t^T r(0,u)du + \int_0^t \int_t^T a(s,u-s)duds \\ &\quad + \int_0^t \int_t^T \tau(s,u-s)dudW(s) \quad \text{(using Fubini's theorem version III)}\end{aligned}$$

Making the change of variable $v = u - t$ in the integral on the LHS and the change of variable $k = u - s$ in the last two integrals on the RHS we obtain

$$\begin{aligned}\int_0^{T-t} r(t,v)dv &= \int_t^T r(0,u)du + \int_0^t \left(\int_{t-s}^{T-s} a(s,k)dk\right)ds \\ &\quad + \int_0^t \int_{t-s}^{T-s} \tau(s,k)dkdW(s).\end{aligned} \tag{26.79}$$

Consider the second term on the right-hand side of Eq. (26.79), we know that in the arbitrage free economy and under the equivalent martingale measure

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds.$$

Hence from Eq. (26.7) we recall that in the BGM notation this condition can be expressed as

$$a(t,x) = \tau(t,x)\psi(t,x) = \frac{\partial}{\partial x}\left(\frac{1}{2}\psi^2(t,x)\right). \tag{26.80}$$

Therefore the inner integral in the second term of (26.79) becomes

$$\begin{aligned}\int_{t-s}^{T-s} a(s,k)dk &= \frac{1}{2}\int_{t-s}^{T-s}\frac{\partial}{\partial k}\left(\psi(s,k)^2\right)dk\\ &= \frac{1}{2}\left[\psi^2(s,T-s)-\psi^2(s,t-s)\right]\\ &= \frac{1}{2}\left[\left(\int_0^{T-s}\tau(s,v)dv\right)^2-\left(\int_0^{t-s}\tau(s,v)dv\right)^2\right].\end{aligned}$$

Equation (26.79) then can be expressed as

$$\begin{aligned}\int_0^{T-t} r(t,v)dv = \int_t^T r(0,u)du + \frac{1}{2}\int_0^t &\left[\left(\int_0^{T-s}\tau(s,v)dv\right)^2\right.\\ &\left.-\left(\int_0^{t-s}\tau(s,v)dv\right)^2\right]ds\\ &+\int_0^t\left(\int_0^{T-s}\tau(s,u)du-\int_0^{t-s}\tau(s,u)du\right)d\tilde{W}(s),\end{aligned} \tag{26.81}$$

where we now use $\tilde{W}(t)$, the Wiener process under the equivalent martingale measure.

Next consider calculation of the money market account. First note that in BM notation, the instantaneous spot interest rate is given by

$$r(t) = r(t,0).$$

Thus setting $x = 0$, $t = s$ and replacing the running integration variables by $u$ in Eq. (26.6) (but with $a(u, s-u)$ chosen according to the arbitrage free relation (26.80)), the stochastic integral equation for the instantaneous spot interest rate is given by

$$r(s) = r(s,0) = r(0,s) + \int_0^s a(u,s-u)du + \int_0^s \tau(u,s-u)d\tilde{W}(u).$$

Hence

$$\int_0^t r(s)ds = \int_0^t r(0,s)ds + \int_0^t \int_0^s a(u,s-u)duds + \int_0^t \int_0^s \tau(u,s-u)d\tilde{W}(u)ds$$
$$= \int_0^t r(0,s)ds + \frac{1}{2}\int_0^t \left(\int_0^{t-s} \tau(s,u)du\right)^2 ds + \int_0^t \int_0^{t-s} \tau(s,u)dud\tilde{W}(s), \tag{26.82}$$

where the last equality follows from manipulations[9] similar to those leading up to Eq. (26.81). Next we determine the relative bond price

$$\frac{P(t,T)}{A(t)} = \frac{\exp\left(-\int_0^{T-t} r(t,u)du\right)}{\exp\left(\int_0^t r(s,0)ds\right)}$$
$$= \exp\left(-\int_0^{T-t} r(t,u)du - \int_0^t r(s,0)ds\right)$$
$$= \exp\left(-\int_0^{T-t} r(t,u)du - \int_0^t r(s)ds\right).$$

Taking logs and making use of (26.81) and (26.82) the last equation becomes

$$\ln\left[\frac{P(t,T)}{A(t)}\right] = -\int_t^T r(0,u)du - \int_0^t r(0,s)ds - \frac{1}{2}\int_0^t \left(\int_0^{T-s} \tau(s,v)dv\right)^2 ds$$
$$-\int_0^t \left(\int_0^{T-s} \tau(s,u)du\right) d\tilde{W}(s),$$

[9]Note that

$$\int_0^t \int_0^s a(u,s-u)duds = \int_0^t \int_u^t a(u,s-u)dsdu = \int_0^t \left(\int_0^{t-u} a(u,y)dy\right) du,$$

and

$$\int_0^t \int_0^s \tau(u,s-u)d\tilde{W}(u)ds = \int_0^t \int_u^t \tau(u,s-u)dsd\tilde{W}(u) = \int_0^t \int_0^{t-u} \tau(u,y)dyd\tilde{W}(u).$$

thus

$$\frac{P(t,T)}{A(t)} = \exp\left[-\int_0^T r(0,s)ds - \frac{1}{2}\int_0^t \left(\int_0^{T-s} \tau(s,v)dv\right)^2 ds - \int_0^t \left(\int_0^{T-s} \tau(s,u)du\right) d\tilde{W}(s)\right] \quad (26.83)$$

$$= P(0,T)\exp\left[-\frac{1}{2}\int_0^t \psi^2(s,T-s)ds - \int_0^t \psi(s,T-s)d\tilde{W}(s)\right]. \quad (26.84)$$

Note that Eq. (26.83) is of the form[10]

$$\frac{P(t,T)}{A(t)} = \exp\left(-\alpha_0 - \int_0^t \alpha_1(s)ds + \int_0^t \alpha_2(s)d\tilde{W}(s)\right).$$

The application of Ito's lemma yields

$$d\left[\frac{P(t,T)}{A(t)}\right] = \frac{P(t,T)}{A(t)}\left(-\alpha_1(t) + \frac{1}{2}\alpha_2^2(t)\right)dt + \frac{P(t,T)}{A(t)}\alpha_2(t)d\tilde{W}(t).$$

Since (see the definitions in footnote 7)

$$-\alpha_1(t) + \frac{1}{2}\alpha_2^2(t) = 0,$$

the last equation shows that $\frac{P(t,T)}{A(t)}$ is a martingale under the equivalent martingale measure. Next consider Eq. (26.82) with $t$ set to $T$, that is

$$\int_0^T r(s)ds = \int_0^T r(0,s)ds + \frac{1}{2}\int_0^T \left(\int_0^{T-s} \tau(s,u)du\right)^2 ds + \int_0^t \int_0^{T-s} \tau(s,u)dud\tilde{W}(s). \quad (26.85)$$

Consider the calculation of the expectation at time $t$ of this quantity. If we assume that the volatility function $\sigma(t,T)$ is dependent only on $t$ and $T$ (i.e. is independent of any path dependent quantity such as $r(t)$ or $f(t,T)$) then the right hand side

[10] Here $\alpha_0 = \int_o^T r(0,s)ds$, $\alpha_1(s) = \frac{1}{2}\psi^2(s,T-s)$ and $\alpha_2(s) = \psi(s,T-s)$.

of Eq. (26.85) is normally distributed. Furthermore we can derive that the mean is given by

$$\begin{aligned} m &\equiv \tilde{\mathbb{E}}_t \left( \int_t^T r(s)ds \right) \\ &= \int_0^T r(0,s)ds + \frac{1}{2} \int_0^T \left( \int_0^{T-s} \tau(s,u)du \right)^2 ds + \int_0^t \int_0^{T-s} \tau(s,u)dud\tilde{W}(s). \end{aligned}$$

The appearance of the last term integrated with respect to the Wiener may appear unusual, but we note that it is only integrated from 0 to $t$, the time at which the expectation is being calculated. At that point in time, this term is in the past and so represents a known or realised quantity. We then calculate the variance as

$$v^2 = \tilde{\mathbb{E}}_t \left[ \left( \int_t^T \int_0^{T-s} \tau(s,u)dud\tilde{W}(s) \right)^2 \right] = \int_t^T \left( \int_0^{T-s} \tau(s,u)du \right)^2 ds.$$

Using the result (6.47) we have

$$\tilde{\mathbb{E}}_t \left[ \exp \left( - \int_t^T r(s)ds \right) \right] = \exp \left( -m + \frac{1}{2} v^2 \right).$$

## Appendix 26.2 Forward Swap Rate Dynamics

Recall from (26.38) that the risk-neutral dynamics of $L_j(t)$ are

$$dL_j = L_j(t)\gamma_j(t)\psi_{j+1}(t)dt + L_j(t)\gamma_j(t)d\tilde{W}(t). \tag{26.86}$$

for $j = 0, 1, \ldots, n-1$. From (26.49) and (26.50) we have that

$$R_{o,n}(t) = \sum_{j=0}^{n-1} w_j(t)L_j(t), \tag{26.87}$$

with

$$w_j(t) = \frac{P(t, T_{j+1})}{\sum_{k=0}^{n-1} P(t, T_{k+1})} = \frac{\delta P(t, T_{j+1})}{C_n(t)}. \tag{26.88}$$

An application of the Ito's lemma derives the dynamics of the forward swap rate $R_{o,n}(t)$ under the risk-neutral measure as

$$dR_{o,n} = \Big[\sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\psi_{j+1}(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)} + \frac{1}{2}\sum_{j=0}^{n-1}\sum_{h=0}^{n-1} L_j(t)\gamma_j(t)L_h(t)\gamma_h(t)\frac{\partial^2 R_{o,n}(t)}{\partial L_j(t)L_h(t)}\Big]dt + \sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}d\tilde{W}(t). \tag{26.89}$$

Note that from (26.88) we obtain that for $m > j$

$$\frac{\partial w_m(t)}{\partial L_j(t)} = \delta\frac{\frac{\partial P(t,T_{m+1})}{\partial L_j(t)}C_n(t) - P(t,T_{m+1})\frac{\partial C_n(t)}{\partial L_j(t)}}{C_n^2(t)}. \tag{26.90}$$

By using (26.58) we know that

$$P(t,T_{m+1}) = \frac{P(t,T_0)}{\prod_{\kappa=0}^{m}(1+\delta L_\kappa(t))}. \tag{26.91}$$

thus for $j \le m$

$$\frac{\partial P(t,T_{m+1})}{\partial L_j(t)} = -\frac{\delta}{1+\delta L_j(t)}\frac{P(t,T_0)}{\prod_{\kappa=0}^{m}(1+\delta L_\kappa(t))} = -\frac{\delta}{1+\delta L_j(t)}P(t,T_{m+1}). \tag{26.92}$$

In addition, from (26.61) we have that (for $m \ge j$)

$$\frac{\partial C_n(t)}{\partial L_j(t)} = \delta\sum_{k=0}^{n-1}\frac{\partial P(t,T_{k+1})}{\partial L_j(t)} = \delta\sum_{m=j}^{n-1}\frac{-\delta}{1+\delta L_j(t)}P(t,T_{m+1}) = \delta C_n(t)\sum_{m=j}^{n-1}\frac{-\delta}{1+\delta L_j(t)}\frac{P(t,T_{m+1})}{C_n(t)} = -\frac{\delta C_n(t)}{1+\delta L_j(t)}\sum_{m=j}^{n-1}w_m(t). \tag{26.93}$$

By substituting (26.92) and (26.93) into (26.90) we have that for $m \geq j$

$$\frac{\partial w_m(t)}{\partial L_j(t)} = -\frac{\delta}{1+\delta L_j(t)}\Big[\frac{\delta P(t,T_{m+1})}{C_n(t)} - \frac{\delta P(t,T_{m+1})}{C_n(t)}\sum_{k=j}^{n-1} w_k(t)\Big]$$
$$= -\frac{\delta w_m(t)}{1+\delta L_j(t)}\Big[1 - \sum_{k=j}^{n-1} w_k(t)\Big] = -\frac{\delta w_m(t)}{1+\delta L_j(t)}\sum_{k=0}^{j-1} w_k(t). \tag{26.94}$$

From (26.87) the partial derivative of $R_{o,n}(t)$ with respect to $L_j(t)$ can be expressed as

$$\frac{\partial R_{o,n}(t)}{\partial L_j(t)} = w_j(t) + \sum_{m=0}^{n-1}\frac{\partial w_m(t)}{\partial L_j(t)}L_m(t),$$

where by using (26.94)[11]

$$\frac{\partial R_{o,n}(t)}{\partial L_j(t)} = w_j(t) - \frac{\delta}{1+\delta L_j(t)}\sum_{m=0}^{n-1} w_m(t)\Big[\mathbf{1}_{m\geq j} - \sum_{k=j}^{n-1} w_k(t)\Big]L_m(t)$$
$$= w_j(t) - \frac{\delta}{1+\delta L_j(t)}\Big[\sum_{m=j}^{n-1} w_m(t)L_m(t) - (\sum_{k=j}^{n-1} w_k(t))(\sum_{m=0}^{n-1} w_m(t)L_m(t))\Big]$$
$$= w_j(t) - \frac{\delta}{1+\delta L_j(t)}\sum_{m=j}^{n-1} w_m(t)(L_m(t) - R_{o,n}(t)) \tag{26.95}$$
$$= w_j(t) + \frac{\delta}{1+\delta L_j(t)}\sum_{m=0}^{j-1} w_m(t)(L_m(t) - R_{o,n}(t)). \tag{26.96}$$

[11]Note that

$$\sum_{m=0}^{j-1} w_m(t)(L_m(t) - R_{o,n}(t)) + \sum_{m=j}^{n-1} w_m(t)(L_m(t) - R_{o,n}(t))$$
$$= \sum_{m=0}^{n-1} w_m(t)(L_m(t) - R_{o,n}(t) = \sum_{m=0}^{n-1} w_m(t)L_m(t) - \sum_{m=0}^{n-1} w_m(t)R_{o,n}(t)$$
$$= R_{o,n}(t) - 1 \times R_{o,n}(t) = 0.$$

For $h \geqslant j$, the second order derivative of $R_{o,n}(t)$ may be expressed as

$$\begin{aligned}\frac{\partial^2 R_{o,n}(t)}{\partial L_j(t)\partial L_h(t)} &= \frac{\partial}{\partial L_j(t)}\Big(\frac{\partial R_{o,n}(t)}{\partial L_h(t)}\Big)\\ &= \frac{\partial w_h(t)}{\partial L_j(t)} - \frac{\delta}{1+\delta L_h(t)}\sum_{m=j}^{n-1}\Big[\frac{\partial w_m(t)}{\partial L_j(t)}(L_m(t) - R_{o,n}(t))\\ &\quad - w_m(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\Big]. \end{aligned} \tag{26.97}$$

Using (26.94)–(26.96), we have that for $h \geqslant j$

$$\begin{aligned}\frac{\partial^2 R_{o,n}(t)}{\partial L_j(t)\partial L_h(t)} &= -\frac{\delta w_h(t)}{1+\delta L_j(t)}\sum_{k=0}^{j-1} w_k(t)\\ &\quad - \frac{\delta}{1+\delta L_h(t)}\Big[-\sum_{m=j}^{n-1}\frac{\delta w_m(t)}{1+\delta L_j(t)}\sum_{k=0}^{j-1} w_k(t)(L_m(t) - R_{o,n}(t))\\ &\quad - \frac{\partial R_{o,n}(t)}{\partial L_j(t)}\sum_{m=0}^{h-1} w_m(t)\Big]\\ &= -\frac{\delta w_h(t)}{1+\delta L_j(t)}\sum_{k=0}^{j-1} w_k(t)\\ &\quad - \frac{\delta}{1+\delta L_h(t)}\Big[\sum_{k=0}^{j-1} w_k(t)\Big(\frac{\partial R_{o,n}(t)}{\partial L_j(t)} - w_j(t)\Big)\\ &\quad - \frac{\partial R_{o,n}(t)}{\partial L_j(t)}\sum_{m=0}^{h-1} w_m(t)\Big]\\ &= \sum_{k=0}^{j-1} w_k(t)\Big[-\frac{\delta w_h(t)}{1+\delta L_j(t)} + \frac{\delta w_j(t)}{1+\delta L_h(t)}\Big] - \frac{\partial R_{o,n}(t)}{\partial L_j(t)}\frac{\delta}{1+\delta L_h(t)}\\ &\quad \times\Big[\sum_{k=0}^{j-1} w_k(t) - \sum_{m=0}^{h-1} w_m(t)\Big]. \end{aligned} \tag{26.98}$$

For $j = h$, expression (26.98) reduces to zero. Finally by using the expression (26.98) for the second order derivative of $R_{o,n}(t)$[12]

$$\sum_{j=0}^{n-1}\sum_{h=0}^{n-1} L_j(t)\gamma_j(t)L_h(t)\gamma_h(t)\frac{\partial^2 R_{o,n}(t)}{\partial L_j(t)L_h(t)}$$
$$= \sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\sum_{h=j+1}^{n-1} L_h(t)\gamma_h(t)\frac{\partial^2 R_{o,n}(t)}{\partial L_j(t)L_h(t)}$$
$$= -\sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\sum_{h=j+1}^{n-1}\frac{\delta L_h(t)\gamma_h(t)}{1+\delta L_h(t)}\sum_{m=j}^{h-1} w_m(t), \tag{26.99}$$

thus the drift of (26.89) becomes

$$\sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\Big(\psi_{j+1}(t) - \sum_{h=j+1}^{n-1}\frac{\delta L_h(t)\gamma_h(t)}{1+\delta L_h(t)}\sum_{m=j}^{h-1} w_m(t)\Big)$$
$$= \sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\Big(\sum_{h=0}^{j}\frac{\delta L_h(t)\gamma_h(t)}{1+\delta L_h(t)} - \sum_{m=j}^{n-1} w_m(t)\sum_{h=j+1}^{m}\frac{\delta L_h(t)\gamma_h(t)}{1+\delta L_h(t)}\Big)$$
$$= \sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\Big(-\sum_{m=j}^{n-1} w_m(t)\sum_{h=0}^{m}\frac{\delta L_h(t)\gamma_h(t)}{1+\delta L_h(t)}\Big)$$
$$= -\sum_{j=0}^{n-1} L_j(t)\gamma_j(t)\frac{\partial R_{o,n}(t)}{\partial L_j(t)}\sum_{m=0}^{n-1} w_m(t)\psi_{m+1}(t). \tag{26.100}$$

This result derives the dynamics (26.64).

---

[12]Note that

$$\sum_{j=0}^{n-1}\sum_{h=0}^{n-1}\Big[-\frac{\delta w_h(t)}{1+\delta L_j(t)} + \frac{\delta w_j(t)}{1+\delta L_h(t)}\Big] = -\sum_{j=0}^{n-1}\frac{\delta\sum_{h=0}^{n-1} w_h(t)}{1+\delta L_j(t)} + \sum_{h=0}^{n-1}\frac{\delta\sum_{j=0}^{n-1} w_j(t)}{1+\delta L_h(t)}$$
$$= -\sum_{j=0}^{n-1}\frac{\delta}{1+\delta L_j(t)} + \sum_{h=0}^{n-1}\frac{\delta}{1+\delta L_h(t)} = 0.$$

## 26.9 Problems

**Problem 26.1** Explain why a payer swaption has a value that is always smaller than the value of the corresponding cap contract.

**Problem 26.2** Consider a forward contract on a zero-coupon bond $P(t, T_1)$ with maturity $T_1$ that at time $T < T_1$ exchanges the bond price $P(T, T_1)$ for $F_T(t, T_1)$. Show that the forward price $F_T(t, T_1)$ satisfies the relation

$$F_T(t, T_1) = \frac{P(t, T_1)}{P(t, T)}.$$

Given that the bond price follows the risk-neutral dynamics (recall Eq. (25.21))

$$\frac{dP(t, T)}{P(t, T)} = r(t)dt + \sigma_B(t, T)d\tilde{W}(t), \tag{26.101}$$

show that $F_T(t, T_1)$ is a $P^T$-martingale and its reciprocal $1/F_T(t, T_1)$ is a $P^{T_1}$-martingale.

**Problem 26.3** Using the relationship equation (26.49) between the forward swap rate and the forward rates

$$R_{o,n}(t) = \sum_{i=0}^{n-1} w_i(t)L_i(t),$$

where

$$w_i(t) = \frac{P(t, T_{i+1})}{\sum_{k=0}^{n-1} P(t, T_{k+1})},$$

show that

$$\frac{\partial R_{o,n}(t)}{\partial L_i(t)} = w_i(t) + \frac{\delta}{1 + \delta L_i(t)} \sum_{k=0}^{i-1} w_k(t)(L_k(t) - R_{o,n}(t)).$$

**Problem 26.4** **Computational Problem**—Assume that the initial forward curve is given by

$$L_i(0) = 0.08 - 0.04e^{-0.2\times T_i}.$$

Under the market model, forward swap rates are assumed to follow lognormal dynamics and assume that the volatility of the forward swap rate is 15 %. Swap payments are made semi-annually.

**Table 26.1** Pricing swaptions

| Maturity | Tenor | Strike | Annuity | MC price | Black's price |
|---|---|---|---|---|---|
| 1 | 0.5 | | | | |
| 5 | 0.5 | | | | |
| 10 | 0.5 | | | | |
| 1 | 5 | | | | |
| 5 | 5 | | | | |
| 10 | 5 | | | | |

Use the information in Table 26.1. The maturity of the swaption is given in the first column and the tenor of the underlying swap is given in the second column.

(a) Use Eq. (26.59) to calculate the prevailing forward swap rates. For ATM swaption, the swap rates become the strike rates. Record your answers in the third column of Table 26.1.
(b) Using Eq. (26.58), compute the annuity term Eq. (26.61). Record your answers in the fourth column of Table 26.1.
(c) Price ATM swaptions with maturities and tenor of the underlying swap as provided in Table 26.1.

(i) Consider the swaption pricing formula (26.71). Write a program to calculate the swaption price by simulating a stochastic differential equation of the type (26.77) for $R_{o,n}(t)$ from 0 to $T_0$ to obtain $R_{o,n}(T_0)$ and by simulating a large number of paths (use 100,000 paths) to compute the $\mathbb{E}_t^S[(R_{o,n}(T_0) - R)^+]$. Record your answers in the fifth column of the table.
(ii) You can check the accuracy of your algorithm by comparing the simulated swaption prices to the swaption prices obtained by Black's formula equation (26.67). Record your answers in the sixth column of the table.

**Problem 26.5 Computational Problem**—Consider the parameter specifications of Problem equation (26.4), where instead of a 15 % constant forward swap rate volatility, the volatility of the forward rates has the following functional form

$$\gamma_i(t) := \gamma(t, T_i - t) = 0.15 - 0.02e^{-0.15(T_i - t)}.$$

Answer questions (a)–(c).

# References


Abramowitz, M., & Stegun, I. A. (1970). The impact of jump risks on nomminal interest rates and foreign exchange rates. *Review of Quantitative Finance and Accounting, 2*, 17–31.

Ahn, C. M., & Thompson, H. E. (1988). Jump-diffusion processes and the term structure of interest rates. *Journal of Finance, 43*, 155–174.

Akgiray, V., & Booth, G. G. (1988). Mixed diffusion-jump processes modelling of exchange rate movements. *The Review of Economics and Statistics, 70*, 631–637.

Amin, K. I., & Morton, A. (1994). Implied volatility functions in arbitrage-free term structure models. *Journal of Financial Economics, 35*, 141–180.

Amin, K. I., & Ng, V. K. (1993). Option valuation with systematic stochastic volatility. *Journal of Finance, 48*, 881–910.

Astrom, K. J. (1970). *Introduction to stochastic control theory*. New York: Academic Press.

Barone-Adesi, G., & Whaley, R. (1987) Efficient analytic approximation of American option values. *Journal of Finance, 42*, 301–320.

Bates, D. (1996). Jumps and stochastic volatility: The exchange rate processes implicit in Deutschemark options. *Review of Financial Studies, 9*(1), 69–107.

Baxter, M., & Rennie, A. (1996). *Financial calculus-an introduction to derivative pricing*. Cambridge: Cambridge University Press.

Bhar, R., & Chiarella, C. (1997a). Transformation of Heath-Jarrow-Morton models to Markovian systems. *European Journal of Finance, 3*, 1–26.

Bhar, R., & Chiarella, C. (1997b). Interest rate futures: Estimation of volatility parameters in an arbitrage-free framework. *Applied Mathematical Finance, 4*(4), 181–199.

Bhar, R., Chiarella, C., El-Hassan, N., & Zheng, X. (2000). Reduction of forward rate dependent HJM models to markovian form: Pricing European bond options. *The Journal of Computational Finance, 3*(3), 47–72.

Bhar, R., Chiarella, C., & Tô, T. D. (2004). *Estimating the volatility structure of an arbitrage-free interest rate model via the futures markets*. Working Paper.

Bhar, R., & Hunt, B. (1993). Predicting the short-term forward interest rate structure using a parsimonious model. *Review of Futures Markets, 12*, 577–590.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*, 167–179.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–659.

Blattberg, R., & Gonedes, N. (1974). A comparison of the stable and Student distributions as statistical models for stock prices. *Journal of Business, 47*, 244–280.




C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Brace, A. (2007). *Engineering BGM*. Boca Raton: Chapman & Hall/CRC.

Brace, A., Gatarek, D., & Musiela, M. (1997). The market model of interest rate dynanics. *Mathematical Finance, 7*(2), 127–155.

Brace, A., & Musiela, M. (1994). A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton. *Mathematical Finance, 4*(3), 259–283.

Bremaud, P. (1981). *Point process and queues: Martingale dynamics.* New York: Springer.

Brennan, M. J., & Schwartz, E. S. (1979). A continuous-time approach to the pricing of bonds. *Journal of Banking Finance, 3*, 135–155.

Brenner, R. J., Harjes, P. H., & Kroner, K. F. (1996). Another look of models of the short-term interest rate. *Journal of Financial and Quantitative Analysis, 31*, 85–107.

Brigo, D., & Mercurio, F. (2006). *Interest rate models theory and practice* (2nd ed.). Berlin: Springer.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance, 47*, 1209–1227.

Cheyette, O. (1992). Term structure dynamics and mortgage valuation. *Journal of Fixed Income, 1*, 28–41.

Chiarella, C., & El-Hassan, N. (1996). A preference free partial differential equation for the term structure of interest rates. *Financial Engineering and the Japanese Markets, 3*, 217–238.

Chiarella, C., & El-Hassan, N. (1998). *Pricing American interest rate options in a Heath-Jarrow-Morton framework using method of lines*. Quantitative Finance Research Group, School of Finance and Economics, UTS.

Chiarella, C., & Kwon, O. K. (1999). *Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model*. Quantitative Finance Research Group, School of Finance and Economics, UTS (Research Paper 5).

Chung, K. L., & Williams, R. J. (1990). *Introduction to stochastic integration* (2nd ed.). Boston: Birkhäusen.

Cont, R., & Tankov, P. (2004). *Financial modelling with jump processes*. *CRC financial mathematics series*. Boca Raton: Chapman & Hall.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985a). A theory of the term structure of interest rates. *Econometrica, 53*(2), 385–406.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985b). An intertemporal general equilibrium model of asset prices. *Econometrica, 53*(2), 363–384.

Cox, J. C., & Ross, S. A. (1976a). A survey of some new results in financial options pricing theory. *Journal of Finance, 31*, 382–402.

Cox, J. C., & Ross, S. A. (1976b). The valuation of options for alternative stochastic processes. *Journal of Finance, 3*, 145–166.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*, 229–263.

de Jager, G. (1995). *Option Pricing Properties and Techniques Using Binomial and Multinomial Lattices*. PhD thesis, School of Finance and Economics, UTS.

Derman, E., & Kani, I. (1994). Riding on a smile. *Risk, 7*(2), 32–39.

Duffie, D., & Kan, R. (1996). A yield-factor model of interrest rates. *Mathematical Finance, 6*(4), 379–406.

Dupire, B. (1994). Pricing with a smile. *Risk, 7*(1), 18–20.

Eberlein, E. (2001). Application of generalized hyperbolic Lévy motions to finance. In O. E. Barndorff-Nielsen, T. Mikosch, & S. Resnick (Eds.), *Lévy processes: Theory and applications*. Boston: Birkhäusser.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*, 383–417.

Fama, E., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*, 246–273.

Feller, W. (1951). Two singular diffusion processes. *Annals of Mathematics, 54*, 173–182.

Feller, W. (1966). Infinitely divisible distributions and bessel functions associated with random walks. *SIAM Journal of Applied Mathematics, 14*, 864–875.
Flesaker, B. (1993). Testing the Heath-Jarrow-Morton/Ho-Lee model of interest rate contingent claims pricing. *Journal of Financial Quantitative Analysis, 28*, 483–495.
French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*, 3–30.
Gardiner, C. W. (1985). *Handbook of stochastic methods for physics, chemistry and the natural sciences* (2nd ed.). Berlin: Springer.
Garman, M. B., & Kohlhagen, S. W. (1983). Foreign currency option values. *Journal of International Money and Finance, 2*, 231–237.
Gatheral, J. (2008). Consistent modeling of spx and vix options. In *The Fifth World Congreaa of the Bachelier Finance Society*, London.
Geske, R. (1979). The valuation of compound options. *Journal of Financial Econometrics, 7*, 63–82.
Geske, R., & Johnson, H. (1984). The American put valued analytically. *Journal of Financial, 39*, 1511–24.
Gihman, I. I., & Skorohod, A. V. (1972). *Stochastic differential equations*. Berlin: Springer.
Gihman, I. I., & Skorohod, A. V. (1979). *The theory of stochastic processes*. New York: Springer.
Greenberg, M. D. (1971). *Application of greens function in science and engineering*. Englewood Cliffs: Prentice-Hall.
Greenberg, M. D. (1978). *Foundations of applied mathematics*. Englewood Cliffs: Prentice-Hall.
Hagan, P. S., Kumar, D., Lesniewski, A. S., & Woodward, D. E. (2002). Managing smile risk. *WILMOTT Magazine*, 84–108.
Harrison, M. J. (1990). *Brownian motion and stochastic flow systems*. Malabar, FL: Robert E. Krieger Publishing Co.
Harrison, M. J., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory, 20*, 381–408.
Harrison, M. J., & Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications, 11*, 215–260.
Heath, D., Jarrow, R., & Morton, A. (1990). Bond pricing and the term structure of interest rates: A discrete time approximation. *Journal of Financial and Quantitative Analysis, 25*(4), 419–440.
Heath, D., Jarrow, R., & Morton, A. (1992a). Bond pricing and the term structure of interest rates: A new methodology for continent claim valuations. *Econometrica, 60*(1), 77–105.
Heath, D., Jarrow, R., Morton, A., & Spindel, M. (1992b). Easier done than said. *Risk, 5*(9), 77–80.
Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.
Ho, L. C., Cadle, J., & Theobald, M. (2001). Estimation and hedging with a one-factor Heath-Jarrow-Morton model. *Journal of Derivatives, 8*(4), 49–61.
Ho, T. S. Y., & Lee, S. B. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance, XLI*, 1011–1029.
Hobson, G., & Rogers, L. (1998). Complete models with stochastic volatiltity. *Mathematical Finance, 8*(1), 27–48.
Horsthemke, W., & Lefever, R. (1984). *Noise-induced transitions - Theory and applications in physics, chemistry and biology*. *Series in synergetics* (Vol. 15). Berlin: Springer.
Hull, J. (2000). *Options, futures and other derivatives* (4th ed.). Boston: Prentice-Hall.
Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance, 42*, 281–299.
Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies, 3*(4), 573–592.
Hull, J., & White, A. (1994). Numerical procedures for implementing term structure models II: Two-factor models. *Journal of Derivatives, 2*, 37–48.
Inui, K., & Kijima, M. (1998). A Markovian framework in multi-factor Heath-Jarrow-Morrton models. *Journal of Financial and Quantitative Analysis, 33*, 423–440.

Jarrow, R., & Madan, D. (1995). Option pricing using the term structure of interest rates to hedge systematic discontinuities in asset returns. *Mathematical Finance, 5*(4), 311–336.

Johnson, H., & Shanno, D. (1987). Option pricing when the variance is changing. *Journal of Financial and Quantitative Analysis, 22*, 143–151.

Jones, E. P. (1984). Option arbitrage and strategy with large price changes. *Journal of Financial Economics, 13*, 91–113.

Keim, D. B. (1986). The CAPM and equity return regularities. *Financial Analysts Journal, 42*, 19–34.

Kijima, M. (2002). *Stochastic processes with applications to finance*. Boca Raton: CRC Press.

Kolodner, I. I. (1956). Free boundary problem for the heat equation with applications to problems of change of phase. *Communications in Pure and Applied Mathematics, 9*, 1–31.

Krylov, N. V. (1995). *Introduction to the theory of diffusion processes*. *Translations of mathematical monographs* (Vol. 142). Providence: American Mathematical Society.

Kucera, A., & Ziogas, A. (2001). *American options: An extension and elaboration of mckean's solution*. Working Paper, QFRG, School of Finance and Economics, UTS.

Kushner, H. J. (1967). *Stochastic stability and control*. New York: Academic Press.

Lamperti, J. W. (1996). *Probability: A survey of the mathematical theory* (2nd ed.). New York: Wiley

Leroy, S. F. (1989). Efficient capital markets and martingales. *Journal of Economic Literature, XXVII*, 1583–1621.

Lighthill, M. J. (1980). *An introduction to Fourier analysis and generalised functions*. Cambridge: Cambridge University Press.

Lo, A. W., & Mackinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies, 1*, 41–66.

Macmillan, L. W. (1986). Analytic approximation for the American put option. *Advances in Futures and Options Research, 1*(A), 119–139.

Malliaris, A. G., & Brock, W. A. (1982). *Stochastic methods in economics and finance*. Amsterdam: North-Holland.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance, 33*, 177–186.

Mckean, H. P. (1965). A free boundary problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review, 6*(2), 32–39.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*, 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*, 125–144.

Merton, R. C. (1982). On the mathematics and economics assumptions of continuous-time models. In W. F. Sharpe & C. M. Cootner (Eds.), *Financial economics; essays in honor of Paul Cootner*. Englewood Cliffs: Prentice-Hall.

Merton, R. C., & Samuelson, P. A. (1974). Fallacy of the log-normal approximation of optimal portfolio decision making over many periods. *Journal of Financial Economics, 1*, 67–94.

Miltersen, K., Sandmann, K., & Sondermann, D. (1997). Closed form solutions for term structure derivatives with log-normal interest rates. *Journal of Finance, 52*, 409–430.

Musiela, M., & Rutkowski, M. (1997). *Martingale methods in financial modelling*. New York: Springer.

Neftci, S. N. (2000). *An introduction to the mathematics of financial derivatives* (2nd ed.). New York: Academic Press.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.

Oksendal, B. (2003). *Stochastic differential equations* (6th ed.). New York: Springer.

Rebonato, R. (1998). *Interest-rate option models* (2nd ed.). Chichester: Wiley.

Ritchken, P., & Chuang, Y. (1999). Interest rate option pricing with volatility humps. *Review of Derivatives Research, 3*, 237–262.

Ritchken, P., & Sankarasubramanian, L. (1995). Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance, 5*, 55–72.
Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 343–362.
Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.
Samuelson, P. A. (1965). Proof that property anticipated prices fluctuate randomly. *Industrial Management Review, 6*, 41–49.
Samuelson, P. A. (1973). Mathematics of speculative price. *SIAM Review, 15*(1), 1–42.
Sandmann, K., & Sondermann, D. (1997). A note on the stability of lognormal interest rate models and the pricing of eurodollar futures. *Mathematical Finance, 7*, 119–128.
Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation and an application. *Journal of Financial and Quantitative Analysis, 22*, 419–438.
Stein, E. M., & Stein, J. C. (1991). Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies, 4*, 727–752.
Stratonovich, R. L. (1963). *Topics in the theory of random noise* (Vols. 1 and 2). New York: Gordon and Brach.
Sundaran, R. K. (1997). Equivalent martingale measures and risk-neutral pricing: An expository note. *Journal of Derivatives,* Fall, 85–98.
Thorp, E. O., & Kassouf, S. T. (1967). *Beat the market*. New York: Random House.
Vasicek, O. (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics, 5*, 177–188.
West, G. (2005). Calibration of the sabr model in illiquid markets. *Applied Mathematical Finance, 12*(4), 371–385.
West, K. D. (1988). Bubbles, fads and stock price volatility: A partial evaluation. *Journal of Finance, 43*, 639–656.
Wiggins, J. (1987). Option values under stochastic volatility. *Journal of Financial Economics, 19*, 351–372.
Wilmott, P., Dewynne, J., & Howison, S. (1993). *Option pricing: Mathematical models and computation*. Oxford: Oxford Financial Press.
Xu, G., & Taylor, S. (1994). The magnitude of implied volatility smiles: Theory and empirical evidence for exchange rates. *Review of Futures Markets, 13*, 355–380.
Zhang, P. G. (1997). *Exotic options*. Singapore: World Scientific.

# Index




C. Chiarella et al., *Derivative Security Pricing*, Dynamic Modeling and Econometrics in Economics and Finance 21,
DOI 10.1007/978-3-662-45906-5