The HJM Model for Interest Rate Evolution

by

Saurav Sen

Submitted to the Faculty of Mathematical Sciences
on 21 May 1998, in partial fulølment of the
requirements for the degree of
M.Sc. in Mathematical Modelling and Numerical Analysis

Abstract

The Heath-Jarrow-Morton (HJM) model provides a technically rigorous framework for the evolution of the entire term structure of interest rates. By utilising information about the entire yield curve, it is able to posit a spot rate process that is independent of the market price of risk in the risk-adjusted measure. Thus it enables bonds to be priced independently of the market's perception of risk. This paper develops the model from basics, providing a thorough survey of the necessary mathematical and economic ideas that HJM utilises. It explores the HJM framework in detail, and compares it with existing term structure models. Finally, it outlines recent theoretical developments and points towards areas of possible future research.

Contents

Acknowledgments

I would like to thank

- Dr. Sam Howison for his motivational help, for introducing me to financial modelling, and for asking me all the awkward questions on HJM that I hoped nobody would;
- Nandita Narain, who persuaded me that Mathematics was a Good Thing;
- Freyan Panthaki and Jamie Wilson, for tennis and food, and for enduring their forced tutorials in HJM with brave smiles and polite interest;
- Douglas Adams, for the Hitch-Hiker's Guide to the Galaxy
- Calvin and Hobbes; and
- my family, for their constant support.

Chapter 1

An Overview

The Heath-Jarrow-Morton (HJM) framework provides a robust, elegant and technically rigorous model for the term structure of interest rates. Since its introduction in 1992, it has quickly established itself as a benchmark.

1.1 How HJM Works

HJM takes as given an initial yield curve, reÆecting an initial term structure of interest rates, and models the evolution of the entire yield curve over time. By allowing several independent Brownian motions to inÆuence the randomness of this evolution, HJM allows for imperfect correlation among the movements of diœerent parts of the yield curve. In addition, it provides a relationship that must exist between the drift and volatility of forward rates to ensure the absence of arbitrage opportunities. However, the feature most frequently cited as the principal contribution of the HJM framework is that it allows bonds and other interest rate derivatives to be priced independently of the market's perception if risk. In this sense, HJM does for bond pricing what the Black-Scholes equation does for option pricing, i.e. provide a valuation formula that is independent of market risk.

1.2 Other Approaches

1.2.1 Spot Rate Models

Earlier models, which study the evolution of the spot rate of interest, allow bond prices to be expressed as solutions of partial diœerential equations (PDEs). They can be made sophisticated enough to capture several real features of spot rates, such as mean-reversion and positivity, while still giving relatively simple closed-form bond price formulae. Multifactor models allow the spot rate to be dependent on several underlying random processes, which could be economic variables such as inÆation, employment etc.

1.2.2 Drawbacks of Spot Rate Models

One major drawback of spot rate models is the ıinversionȷ of the yield curve that is needed to infer the parameters of the spot rate process. It is for this reason that researchers often seek yield curves that are aŒne in the spot rate - a complicated function may not be invertible. This is a severe restriction of spot rate models. Another problem is that one-factor spot rate models try to capture the dynamics of the entire yield curve from the behaviour of just one point on it. From an empirical perspective too, it has been shown that none of the established spot rate models øts historical data well.

1.3 HJM and Spot Rate Models

1.3.1 Problems Overcome by HJM

By incorporating an initial term structure of interest rates by construction, HJM bypasses the problem of inverting the yield curve completely. Since it models the behaviour of the whole yield curve and uses multiple stochastic factors, it captures a richer variety of the dynamics of interest rate movements. Spot rates can be inferred from forward rates, and so HJM models are in a sense more general than spot rate models. HJM models can be made strongly path-dependent, while most spot rate models are Markov diœusions, to preserve tractability. All existing spot rate models are in fact special cases of appropriately chosen HJM models.

1.3.2 Problems With HJM

However, there are a number of issues that HJM leaves unresolved. For one, HJM allows spot rates to become negative with positive probability. No ınaturalȷ restrictions of the model that ensure positivity of interest rates are known. Since the yield curve sits in an inønite-dimensional state space and HJM models its evolution as a whole, there are in general no PDE formulations for bond prices under the HJM framework. Formulation in terms of inønite-dimensional stochastic PDEs is an active area of research. There are also proposed extensions of the model that replace the ønitely many independent Brownian motions with a Brownian sheet. From an implementation perspective, HJM has to be simulated using computationally expensive Monte-Carlo methods because of the strong path-dependence of HJM models. On the empirical side, some work has been published on Markovian HJM models, but not much on general non-Markovian models.

1.4 Goals

This paper seeks to understand how HJM works. It is self-contained, and develops all the necessary mathematical and economic theory before analysing the model itself. Chapter 2 provides an overview of the interest rate market, and outlines the methodology of pricing contingent claims, by arbitrage and by expectation. Chapter 3 is the mathematical ımeatȷ of the paper, and introduces all the mathematical tools needed to implement the pricing strategy given in chapter 2. Chapter 4 provides an explicit recipe for pricing contingent claims, based on the insights of chapter 2 and the mathematical tools of chapter 3. It introduces the ønancial concepts necessary to complete the picture. Chapters 5 and 6 deal exclusively with the HJM framework. Since the single-factor version of the model is easy to understand and the extension from one to many factors is both technically and conceptually uncomplicated, chapter 5 is devoted to exploring 1-factor HJM models. Chapter 6 studies the general, multifactor HJM framework. Chapter 7 is a survey of existing term structure models. It outlines the strengths and weaknesses of the various models currently used to explain the dynamics of interest rates. Chapter 8 assesses HJM, providing a critique on the model and surveying the work that has been done since the model was formulated.

Thus the aim of this paper is to provide an exhaustive summary of term structure modelling in the context of the HJM framework, and point towards areas of future research.

Chapter 2

Bonds and Interest Rates

Don't Panic.

2.1 Objectives

- Deøne bonds and introduce the interest rate market.
- Compare methodologies for pricing contingent claims, i.e. risk-neutral pricing vs. arbitrage pricing.
- Describe the idea and signiøcance of a risk-neutral measure and a market measure.
- Present a general recipe for valuing contingent claims.
- Deøne yields and the term structure of interest rates.
- Deøne spot rates and forward rates.
- Present an overview of modelling considerations.

2.2 Bonds

2.2.1 The Interest Rate Market

The most basic interest rate contract is an agreement to pay some money now in exchange for the promise of a usually larger amount later. The value of such a contract depends both on its length or maturity, and on random Æuctuations in the interest rate market, where the price of money is set. The dependence of interest rates on maturity is called the term structure of interest rates. The randomness of the interest rate market opens up the possibility of ønancial instruments that derive their value from the future value of money. A good interest rate model should therefore approximate both the initial term structure as well as its evolution over time, and hence price interest rate derivatives.

2.2.2 Zero Coupon Bonds

A zero-coupon bond (ZCB) is an instrument that pays no dividends between the time it is bought and the time it matures, when it pays out a known amount, which we conventionally take to be a dollar. A ZCB requires two numbers to describe it. One is its length T and the other is the ratio of the initial investment to the ønal payoœ. A dollar at time T is worth $P(0,T)$ at time zero. The promise of a dollar at time T can also be regarded as an asset, which will have some value $P(t,T)$ at each time $0 < t < T$. Thus $P(t,T)$ is the price of the ZCB at time t.

What does the function $P(t,T)$ look like? First, we expect bonds with maturities that are close together to move roughly in step, i.e. to be well-correlated. For example, we expect the short-run behaviour of 25 and 30-year maturity bonds to be similar, but they might behave very diœerently from that of a bond that expires in 2 years. Second, since a bond is an interest rate based derivative, we expect random Æuctuations in interest rates over time to be manifested in the evolution of bond prices as well. In essence, this means that we expect to be a smooth function of T, but a stochastic function of t.

2.3 Pricing a Bond

2.3.1 No-Arbitrage

What is the current worth of a contract that guarantees a dollar at maturity? Suppose interest rates were øxed and constant r between now and maturity. Then

$$P(t,T) = e^{-r(T-t)}$$

If the contract was selling higher, we could sell it, invest the above amount into a ZCB to be assured of a dollar at T, and pocket the diœerence. Similarly, if it was traded lower, we could buy and realise a riskless proøt. One of our basic assumptions is that it ought to be impossible to make this sort of riskless proøt. This is called the principle of no-arbitrage.

2.3.2 Pricing by Expectation

Now suppose interest rates were known, but not constant. If $t = t_0 < t_1 < .. < t_n = T$ and the rate was r_1 in $[t, t_1]$, r_2 in $(t_1, t_2]$ and so on, then we would have

$$P(t, T) = e^{-\sum_{i=1}^{n} r_i (t_i - t_{i-1})}$$

In general, if we knew in advance what the rate of interest was going to be, then the price of a ZCB at time t would be given by

$$P(t, T) = e^{-\int_t^T r(s) ds}$$

In reality of course we do not know what these rates are going to be. Thus $r(t)$ is a random quantity and our formula should be modiøed to

$$P(t, T) = \mathbb{E}_t \left[e^{-\int_t^T r(s) ds} \right]$$

This means that the ZCB price is the expected value of the quantity in square brackets. The subscript means that the expectation is computed using all information till time t.

2.3.3 Comments

A couple of comments are in order. First, the interest rate $r(t)$ will, in later sections, be shown to be the spot rate of interest. This is the rate of borrowing money at the present time and returning it instantaneously.

The second point is about the distribution of $r(t)$. We have not speciøed anywhere the underlying probability density of the random quantity $r(t)$. In general, diœerent investors will have diœerent opinions about this. Thus the expectation is computed given not only all the information till time t, but also the investor's perception of how interest rates are likely to move. Hence the price may not be unique.

This is the feature that allows a market to exist: diœerent investors have diœerent perceptions about how an asset is going to move, and so will prefer to either buy or sell depending on whether they think the asset will go up or down.

2.3.4 The Rational Expectations Hypothesis

The rational expectations hypothesis allows us to get around the problem of non-uniqueness. According to it, the present value of an uncertain quantity is its expected future value, where expectations are computed in an unbiased manner, using all information available till the current time. The formula above incorporates the latter but not the former.

In the light of rational expectations, bond prices are given by

$$P(t,T) = \mathbb{E}_t^* \left[e^{-\int_t^T r(s)ds} \right]$$

where $\mathbb{E}^*$ represents expectations with respect to some appropriately deønes risk-adjusted probability distribution for future values of $r(t)$. This leads onto an important insight about risk and probability distributions, discussed next.

2.4 Risk and Probability

2.4.1 The Market Measure

If we have money to invest, we could put it in a bank and let it grow there. Money in a bank grows at the riskless rate $r(t)$. This is the highest rate of return that can be risklessly guaranteed. Another possibility is to invest the money in, say, a stock or a bond. Since holding one of these involves a certain amount of risk, we expect the average return on a stock or bond to be higher than $r(t)$.

In principle at least, we can observe the average rate of growth of a bank deposit and the average rate of growth of the risky asset (e.g. a stock or bond). We assume that the random movements of both these processes (money in a bank and risky asset) are governed by the same underlying set of probabilities. This underlying measure, of which the observed movements of assets are a particular set of outcomes, is called the market measure and will be denoted through out this paper by $\mathbb{Q}$. In a sense, the market measure characterises our actual observations about the randomness of, say, stock prices.

Under the market measure, the average rate of growth of money in a bank is $r(t)$ and that of a stock is, say, $\alpha(t)$ which (we expect) is higher than $r(t)$.

2.4.2 The Risk-Adjusted Measure

The market measure is simply a frame of reference. Is there another set of probabilities under which the stock grows at $r(t)$? That is to say, can we shift our frame of reference (or change the relative weights of various possible outcomes) so that in the new frame, the stock (a risky asset) grows at the riskless rate? We would like this new frame to be in some sense equivalent to the market measure. This means that they should both agree on what events are possible, but the relative likelihoods of the same event might be diœerent under the two measures. If such a measure exists, it is called the risk-adjusted or risk-neutral or risk-free measure, and will be denoted by $\mathbb{Q}^*$. We shall dwell on the existence and uniqueness of $\mathbb{Q}^*$ in greater detail in later chapters when the necessary mathematical tools have been developed, but for now all we seek is a heuristic idea of risk-adjusted probability.

Thus the risk-adjusted measure $\mathbb{Q}^*$ is that measure, if it exists, equivalent to the market measure $\mathbb{Q}$, under which a risky asset appears to grow at a riskless rate.

2.4.3 Market Price of Risk

In the market measure, the diœerence between the return on a stock and the return on a bank deposit reÆects the amount of risk inherent in holding the stock. Market price of risk is the sum of money that quantiøes this risk. A person investing in stocks is willing to take on some risk in order to achieve high rates of growth. A person investing in a bank is not. The return on a stock, stripped of the market price of risk, is equal to the return on a bank deposit. Thus market price of risk is an asset-speciøc quantity, reÆecting the fact that diœerent ønancial instruments are associated with diœerent amounts of risk.

In a sense that will be made rigorous in later chapters, market price of risk provides the link between $\mathbb{Q}$ and $\mathbb{Q}^*$. Observed stock prices are governed by $\mathbb{Q}$. By factoring out the market price of risk, we can observe a process governed by $\mathbb{Q}^*$. This observation allows us to develop a strategy for pricing claims.

2.4.4 A Recipe for Pricing Contingent Claims

We want to be able to price bonds. All our observations are made in the market measure. We ørst factor out the market price of risk in order to place ourselves in a risk-neutral world. Once we are in this setting, we use rational expectations to ønd the present value of a future claim.

2.5 Interest Rate Terminology

2.5.1 Yields

Consider a continuous-trading economy. If interest rate is constant, say r, then by the principle of no-arbitrage, a bond $P(t,T)$ that guarantees a dollar at maturity is worth $e^{-r(T-t)}$ at time t. If it was trading higher, we could sell the contract, invest of the proceeds to guarantee a dollar later, and lock in the diœerence as a riskless proøt. Similarly, if it was trading lower we would buy and make a proøt. Note that in this case, the expectation-based and arbitrage-based prices coincide, as they should, because after all a bond is a tradable instrument. The interest rate can be explicitly recovered as

$$r = -\frac{\log P(t,T)}{T-t}$$

In practice, interest rates are not constant, but this quantity is useful, so we deøne

$$Y(t,T) = -\frac{\log P(t,T)}{T-t} \tag{2.1}$$

to be the yield on the T-maturity bond. The yield curve is a plot of yield as a function of T. The mapping of bond price to yield is 1-1 and so no information is lost.

The shape of the yield curve reÆects the market's opinion on future interest rates. An increasing yield curve implies a possible rise in interest rates, and yield curves are frequently increasing functions of maturity. However, there is no canonical form for them and they may take a variety of shapes. If market opinion is that interest rates are likely to fall, then the yield curve may become “inverted”, with the yield on long-dated bonds lower than that on short-dated ones. A good interest rate model should be able to cope with all these possibilities.

2.5.2 Spot Rates

The yield curve gives an idea of the rate of borrowing for each term length, but what is the cost of borrowing now for an amount to be returned instantaneously? The yield for a period $[t, t+dt]$ is

$$Y(t, t+dt) = -\frac{\log P(t, t+dt) - \log P(t, t)}{dt}$$

This follows since $\log P(t, t) = 0$. If the left-hand side converges as $dt \to 0$, it defines a quantity called the spot rate, $r(t)$. Thus

$$r(t) = -\frac{\partial}{\partial T} \log P(t, t) \tag{2.2}$$

The spot rate is an important process in the interest rate market and many models are based exclusively on its behaviour, with bond prices extrapolated from it. A spot rate model allows us to express bond prices as solutions of partial differential equations and can be used to derive yield curves that are affine in the spot rate. However, there is some loss of information involved in the mapping from the discount bond price $P(t, T)$ to the spot rate $r(t)$. Ideally, we would like a natural extension of the concept of a spot rate that entails no loss of information but still preserves the idea of instantaneity. Since the spot rate is the present cost of borrowing for a contract which matures immediately, a possible candidate for the mapping we seek is the present rate for a contract which starts at a later time T and matures immediately. This leads on to the concept of forward rates.

2.5.3 Forward Rates

Consider a forward contract, i.e. a contract struck at time t, to make a payment at a later date T_1 and receive a payment in return at an even later date T_2. If at time t we buy one T_1-bond and sell k units of the T_2-bond, then the combined value of these assets at t is

$$V(t) = P(t, T_1) - kP(t, T_2)$$

To give the contract zero initial value, we must choose

$$k = \frac{P(t,T_1)}{P(t,T_2)}$$

This choice ensures that the contract must have value zero throughout, and in particular at time T_1. This means that

$$V(T_1) = P(T_1,T_2) - k = 0$$

which in turn means that k must be the time T_1-price of a T_2-maturity bond. The corresponding forward yield is

$$Y(T_1,T_2;t) = -\frac{\log k}{T_2 - T_1} = -\frac{\log P(t,T_2) - \log P(t,T_1)}{T_2 - T_1}$$

The third argument in Y signiøes the fact that the contract is struck at time t, for the trading interval $[T_1, T_2]$. If we now choose T_1 and T_2 close together, say $T_1 = T$ and $T_2 = T + \Delta T$ then the forward yield converges to a quantity that represents the rate of borrowing instantaneously at time T, when the contract is struck at time t. This is called the forward rate at time t for instantaneous borrowing at time T. Thus we have the deønition

$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T) \tag{2.3}$$

As expected, the instantaneous forward rate for borrowing at the current time coincides with the previous deønition of spot rates.

$$r(t) = f(t,t)$$

2.5.4 Equivalence of Bond Prices, Yields and Forward Rates

Unlike spot rates, the transformations between bond price, yield and forward rate are all 1-1 and hence knowing any of them is equivalent to knowing all of them. If we know $Y(t,T)$ then $P(t,T)$ is given by

$$P(t,T) = e^{-Y(t,T)(T-t)}$$

Thus knowing Y implies knowing P. If we know $P(t,T)$ then

$$f(t,T) = -\frac{\partial}{\partial T} \log P(t,T)$$

Thus knowing P implies knowing f. Finally, if we know f then

$$Y(t,T) = \frac{1}{T-t} \int_t^T f(t,s)\, ds$$

Thus knowing f implies knowing Y. Hence knowing any one of the three is equivalent to knowing them all.

2.5.5 The Term Structure of Interest Rates

The yield curve at any date tells us what the market is willing to pay for loans of various lengths of time made on that day. An upward sloping yield curve reÆects the fact that the market is willing to pay more interest on a loan of longer duration. This dependence of interest rate on the length of a loan is called the term structure of interest rates. The term structure of interest rates is described by the yield curve

$$\{Y(t,T) : T \in [t,\tau]\}$$

or equivalently by ZCB prices (which are also called discount factors).

2.6 Endnotes

For modelling purposes, specifying the behaviour of bond prices, forward rates and yields are equivalent and so we need to model only one of them. Spot rate models such as the Vasicek and Cox-Ingersoll-Ross (CIR) models specify the behaviour of spot rates and then deduce the evolution of forward rates from them. They assume that the term structure is a function of the spot rate (usually an aŒne function) and then use the observed yield curve to ønd the parameters of the spot rate process. This inversion of the term structure is diŒcult if we assume more complicated functional relationships between spot rates and yields, and is often not possible. In fact, there is no reason why the yield curve should have a given functional relationship with the spot rate.

Another more serious problem is that all spot rate models assume an endogenously speciøed functional form for the market price of risk, even under the risk-adjusted measure. The HJM framework models forward rates and by doing so derives a set of conditions under which the spot rate process is independent of the market price of risk in the risk-adjusted measure.

Given that spot rate modelling is fraught with assumptions about risk preferences and that there is loss of information involved in the mapping from bond prices to spot rates, modelling the forward rate might be a better idea. Ho & Lee (HL) were the ørst to model forward rates directly, and did so in a discrete-time framework. The HJM class of models extends the HL framework to a continuous-time setting.

So now we turn to the question of how best to model forward rates. As discussed earlier, we know that term structure is smooth but its evolution is rough. The HJM approach takes as given an entire initial forward rate curve (this is usually smooth), and then allows it to evolve stochastically over time. This preserves the smoothness of $f(t, T)$ as a function of T and its random nature as a function of t. In deterministic functions, smoothness is intimately connected with diœerentiability, so the ørst step is to set up a calculus-like set of rules for stochastic functions.

Once this is done, we are in a position to investigate the behaviour of forward rates in detail. HJM postulate that the forward rate evolves with two components - one deterministic and the other random. More on the precise nature of the random component will be said later. The next step is to ønd restrictions that must be applied to these components so that interest rates behave in a sensible manner. The last task is to price interest rate derivatives from the model.

Chapter 3

Stochastic Calculus

...something almost but not quite entirely unlike tea.

3.1 Objectives

- Deøne stochastic processes in one and many dimensions, and the various terms associated with them.
- Introduce and study Brownian motion.
- Develop the relevant fundamentals of Ito calculus. These include

  Representation of stochastic processes in integral and diœerential forms.

  Ito integration, with emphasis on the following integrals:

$$\int_a^b X(t)dt \quad \text{and} \quad \int_a^b h(t)dW(t)$$

  Ito's formula for stochastic diœerentiation.

  The product rule.

- Deøne the Radon-Nikodym derivative and state its basic properties.
- Investigate how change of measure aœects a stochastic process.

- BrieÆy discuss the Cameron-Martin-Girsanov theorem and the Martingale Representation Theorem and their relevance to bond pricing.

3.2 Deønitions

3.2.1 Stochastic Processes

A stochastic process $X(t)$ is a sequence of random variables indexed by the parameter t, which is usually interpreted as time and may be measured discretely or continuously. Thus at each time t, the value of $X(t)$ is distributed according to some probability density function. In this way, the notion of an underlying probability distribution is implicitly tied in with that of a stochastic process. For each t, we can assign a probability to the likelihood of the process taking values in a given set, i.e. we can say for example

$$\Pr[a < X(t) < b] = p(a, b; t)$$

What is the measure being used to calculate this probability? If the process is normally distributed at each instant in time, then we will get a diœerent value for than if the process is, say, uniformly distributed. Later in the chapter, we shall systematically analyse the eœects of change of measure on a stochastic process.

The set of values $\{X(t) \mid t \in I\}$, where I is an index set, is called a realisation of the process or random walk. For example, in ønancial markets, the price of a bond $P(t, T)$ as a function of t is a stochastic process, and actual prices are realisations of the process. The set of all possible values of a process is called the state space.

Example of a stochastic process

3.2.2 Filtrations

A øltration $\mathcal{F}(t)$ is the history of the process till time t. An n-dimensional stochastic process has n components $X(t) = [X_1(t), ..., X_n(t)]^{\mathrm{T}}$ and in this case the øltration $\mathcal{F}(t)$ represents the combined histories of all these processes till the current time.

3.2.3 Claims

In a ønancial context, a claim C is a function of the state space of a stochastic process at some time horizon T. Thus a process is deøned for all times $0 < t < T$ but a claim is a random variable deøned only at the terminal time.

3.2.4 An Example

For example, consider a European call option, which is a contract that gives its holder the right (without obligation) to buy a speciøed number of shares at a speciøed price (called the strike price) at a speciøed date in the future (called the expiry date). If the stock is trading higher than the strike price at expiry, the holder can exercise the option and buy at a lower price. If the stock price is higher, the holder can choose not to exercise the option and may buy from the market directly.

Thus if S_T denotes stock price at time T, $V(t)$ the price of the option at time t and E the strike price, then the ønal payoœ is given by the function

$$C_T = \max\{S_T - E, 0\}$$

The payoœ diagram, as it is called, is shown below.

Payoœ diagram for a European call.

In the language of processes and claims, $V(t)$ is a stochastic process for $0 < t < T$, and C_T is a claim.

3.2.5 Conditional Expectation

The conditional expectation operator $\mathbb{E}_{\mathbb{Q}}[\cdot \mid \mathcal{F}(t)]$ extends the concept of expectation to two parameters - a probability measure $\mathbb{Q}$ and a øltration $\mathcal{F}(t)$. The conditional expectation of a claim C given a øltration $\mathcal{F}(t)$ is written $\mathbb{E}_{\mathbb{Q}}[C \mid \mathcal{F}(t)]$and is the expectation measured across the latter half of all paths whose starting point is determined by $\mathcal{F}(t)$, as shown below.

Conditional expectation.

3.2.6 Processes and Claims

For each time t, we can deøne a process $C(t) := \mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(t)\right]$ and in this way form a process from a claim. Note that $\mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(0)\right] = \mathbb{E}_{\mathbb{Q}}[C]$, the usual expectation, and $\mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(T)\right] = C$. Why are we interested in forming processes from claims? In ønancial markets, the ønal payoœ is a claim C, and we want to ønd the present value of this claim. In other words, the claim has some value at each time till it matures. We want to construct a portfolio and follow a strategy that tracks this value. In other words, we want to ønd a process whose value at any time is the value of the claim at this time. What is the value it should track? The answer is $\mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(t)\right]$. Thus if we know a claim, we care about forming a process from it so that we know what value our synthetic approximation (i.e. our portfolio) should track over time.

3.2.7 Previsible Processes

A previsible process is a one with the same state space and probability measure as $X(t)$, but whose current values can be predicted from the history of the process till just before the present time, i.e. from the øltration $\mathcal{F}(t_-)$. In discrete time, a process φ on the same tree is previsible if $\varphi(i)$ can be determined from the øltration $\mathcal{F}(i-1)$. In the context of markets, suitably chosen previsible process will play the role of trading strategies, where we have to base the next step in our strategy only on information available to us till the present time.

3.2.8 Martingales

A special class of previsible processes, called martingales, are of particular interest in the modelling of ønancial markets. A process $X(t)$ is called a martingale under a probability measure $\mathbb{Q}$ if

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left[|X(t)|\right] &< \infty \text{ for all } t \geq 0, \text{ and} \\ \mathbb{E}_{\mathbb{Q}}\left[X(t) \mid \mathcal{F}(s)\right] &= X(s) \text{ for } s < t \end{aligned}$$

This means that if we know the value of the process at a given time, then the expected future value of the process is just equal to the value we know. Rewritten, it also means that the process has no bias or drift up or down with respect to the probability measure $\mathbb{Q}$ under the expectation operator. Such a measure is called a martingale measure for the process.

For any claim C, the process $C(t) = \mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(t)\right]$ is always a $\mathbb{Q}$-martingale. This is an immediate consequence of the ıtower lawȷ, which is a statement of the fact that

$$\mathbb{E}_{\mathbb{Q}}\ \left[\left[C \mid \mathcal{F}(t)\right] \mid \mathcal{F}(s)\right] = \mathbb{E}_{\mathbb{Q}}\left[C \mid \mathcal{F}(s)\right]$$

whenever $s < t$. This is true, because the conditioning ørstly upon history till time t and then on an earlier time s is the same as conditioning directly from the time s. The information set from s to the present is larger than that from t to the present.

3.3 Brownian Motion

3.3.1 Discrete Approximation

A stochastic process widely used in ønancial modelling is Brownian motion. Consider a sequence of random variables $\{W_n(t)\}_{n=1}^{\infty}$ such that

- $W_n(0) = 0$ for all n;
- time steps are of size $1/n$;
- up and down jumps are equal and of size $1/\sqrt{n}$; and

- the probability measure $\mathbb{Q}$ is uniformly equal to 1/2 everywhere on the tree.

Discrete Brownian motion.

One way of constructing this sequence iteratively is to deøne, for an IID sequence $\{X_i\}$ of binomial random variables taking values +1 or -1 with equal probability 1/2,

$$W_n\left(\frac{i}{n}\right) = W_n\left(\frac{i-1}{n}\right) + \frac{X_i}{\sqrt{n}} = \frac{X_1 + ... + X_n}{\sqrt{n}}$$

Then it is easy to see that

$$W_n(t) = \sqrt{t}\left(\frac{\sum_{i=1}^{nt} X_i}{\sqrt{nt}}\right)$$

Using the central limit theorem (CLT), the term in brackets converges to the standard normal distribution $N(0,1)$, and so $W_n(t)$ converges in distribution to $N(0,t)$. In addition, since each step is independent of the previous ones, we can use the CLT again and claim that $W_n(t+s) - W_n(t) \sim N(0,s)$ independent of the history of the process till time t.

It can be formally shown that the marginals, conditional marginals and the distributions of these processes converge as $n \to \infty$. The limiting process is called a Brownian motion.

3.3.2 Deønition

Given a probability measure $\mathbb{Q}$, a process $W = \{W(t) : t \geq 0\}$ is called a $\mathbb{Q}$-Brownian motion if

- $W(0) = 0$;
- $W(t)$ is a continuous function of the time-like parameter t;
- at time t, $W(t)$ is normally distributed under the measure $\mathbb{Q}$, with mean zero and variance t, i.e. $W(t) \sim N(0,t)$ under $\mathbb{Q}$. This means that at any time t,

$$\Pr_{\mathbb{Q}}[a < W(t) < b] = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}} dx$$

  and

- under $\mathbb{Q}$, $W(t+s) - W(t) \sim N(0,s)$ independent of the øltration $\mathcal{F}(t)$, the history of the process till time t.

3.3.3 Comments

It is important to emphasise that we always refer to an underlying probability distribution when referring to a Brownian motion. Thus a $\mathbb{Q}$-Brownian motion is a Brownian motion under the measure $\mathbb{Q}$ and a $\mathbb{Q}^*$-Brownian motion is one under $\mathbb{Q}^*$.

Brownian motion is also called a driftless Wiener process or a 1-dimensional Gaussian process. It is everywhere continuous but nowhere diœerentiable. It looks uniformly rough at any magniøcation, and thus is a fractal. Using the intermediate value theorem and continuity of Brownian motion, it can be shown that once a Brownian motion hits a particular value, it immediately hits it again inønitely often and then again from time to time. Also, it hits every real number eventually, with probability 1.

3.3.4 Covariance Function for a Brownian Motion

Since at each time t, a Brownian motion is a random variable, we can deøne its covariance function, which in this case is

$$\Gamma(s,t) = \mathbb{E}_{\mathbb{Q}}[W(s)W(t)]$$

In general, the covariance function is deøned as

$$\Gamma_X(s,t) = \mathbb{E}[(X(s) - \mathbb{E}X(s))(X(t) - \mathbb{E}X(t))]$$

and measures the linear relationship of the two random variables $X(t)$ and $X(s)$. It will be positive when $X(t)$ and $X(s)$ tend to have the same sign with high probability, and negative if they have opposite signs with high probability. Since a Brownian motion has mean zero, the covariance function has a simpler form.

If $s < t$ then

$$\begin{aligned} \Gamma(s,t) &= \mathbb{E}_{\mathbb{Q}}\left[W(s)W(t)\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[(W(s)(W(t)-W(s))+W(s)^2\right] \\ &= \mathbb{E}_{\mathbb{Q}}W(s)\mathbb{E}_{\mathbb{Q}}\left[W(t)-W(s)\right]+\mathbb{E}_{\mathbb{Q}}\left[W(s)^2\right] \\ &= 0+s \\ &= s \end{aligned}$$

where the third line follows from the fact that the increment in a Brownian motion is independent of the past history. Since the roles of s and t can be switched, we have the result

$$\Gamma(s,t) = \min\{s,t\}$$

The covariance function plays an important role in stochastic integration, discussed later.

3.3.5 Brownian Motion in Several Dimensions

The concept of a Brownian motion is easily extended to several dimensions by considering the vector $W(t) = [W_1(t), ..., W_n(t)]^{\mathrm{T}}$ where each component behaves like a 1-dimensional Brownian motion and all components are independent or correlated with a known correlation matrix speciøed.

3.3.6 Brownian Motion With Drift

A Brownian motion with drift is one whose mean changes with time. For example,

$$X(t) = \sigma W(t) + \mu t$$

where σ and μ are constants, is a Brownian motion with constant drift coeŒcient μ.

3.4 Ito Calculus

The idea behind Newtonian diœerentiation is that zooming in on a diœerentiable curve eventually makes it look like a straight line at suŒcient magniøcation. The slope of this line is then deøned as the derivative at that point. For a stochastic process, this technique does not work because Brownian motion is a fractal and so zooming in on it will not make it look any smoother. Motivated by this, we try to use the self-similarity of a Brownian motion as a building block for deøning a calculus-like set of rules for stochastic processes.

3.4.1 Representation of Stochastic Processes

As a starting point, we set out the a convenient form with which to represent stochastic processes. If $X(t)$ is a stochastic process, then its increment has a Newtonian term dt and an increment in the Brownian term, $dW(t)$. The stochastic processes we shall study are those which can be represented as

$$X(t) = X(0) + \int_0^t \sigma_s dW(s) + \int_0^t \alpha_s ds \tag{3.1}$$

where the volatility σ_t and the drift α_t can depend on the whole history $\mathcal{F}(t)$ of the process till time t, as well as on $X(t)$ itself. Such processes, which depend on the øltration $\mathcal{F}(t)$ are called $\mathcal{F}(t)$-adapted processes. The representation (3.1) is essentially unique, which means that two processes represented by the same equation diœer only on a set of measure zero. This is a consequence of the Doob-Meyer decomposition of semimartingales. For the deønition to make sense, we insist on the technical condition

$$\int_0^t \left(\sigma_s^2 + |\alpha_s|\right) ds < \infty$$

which ensures that the integrals remain bounded. As we shall show later,

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\int_0^t \sigma_s dW(s)\right)^2\right] = \int_0^t \mathbb{E}_{\mathbb{Q}}\left[\sigma_s^2\right] ds$$

so that the technical condition is essentially a boundedness condition.

In diœerential form, (3.1)can be written as

$$dX(t) = \sigma_t dW(t) + \alpha_t dt \tag{3.2}$$

In the special case when σ_t and α_t are deterministic and depend on the Brownian motion only through $X(t)$, (3.2) is called the stochastic diœerential equation (SDE) for $X(t)$. If a solution to (3.2) exists, it is called a diœusion.

In several dimensions, the integrand in (3.1) is replaced by the standard inner product of n-vectors, so that a stochastic process in n dimensions is written as

$$X(t) = X(0) + \sum_{j=1}^{n} \int_0^t \sigma_j(s) dW_j(s) + \int_0^t \alpha_s ds \tag{3.3}$$

or

$$dX(t) = \sum_{j=1}^{n} \sigma_j(t) dW_j(t) + \alpha_t dt \tag{3.4}$$

where

$$\int_0^t \left(\sum_{j=1}^{n} \sigma_j(s)^2 + |\alpha_s|\right) ds < \infty$$

The total volatility of the process is given by

$$\sigma_{\text{TOT}} = \left(\sigma_1(t)^2 + ... + \sigma_n(t)^2\right)^{1/2}$$

In other words, the variance is equal to the sum of the variance of the components, which follows from the independence of the underlying Brownian motions.

3.4.2 Stochastic Integration

To recover stochastic processes from their SDE representations, we need to assign meaning to the following symbols:

$$\int_a^b X(t)dt \quad \text{and} \quad \int_a^b h(t)dW(t)$$

where $X(t)$ is a regular (suŒciently well-behaved) stochastic process. These integrals can be deøned in a manner analogous to Riemann integrals in Newtonian calculus, i.e. by approximating the integrals with discrete sums and then deøning the limits of these sums (if they exist) to be the values of the respective integrals. This section presents two general results, and then examines the particular case of a Brownian motion more closely.

First, we note that the integrals themselves are random variables. Unlike Newtonian calculus, the integral of a stochastic process has no deønite value, but rather a probability distribution function.

Integrating a Browonian Motion over Time

Consider $\int_a^b X(t)dt$. Form a partition $a = t_0 < t_1 < ... < t_n = b$ and let $t_{i-1} < \xi_i < t_i$. Let I_n denote the ønite sum

$$I_n = \sum_{i=1}^{n} (t_i - t_{i-1}) X(\xi_i) \tag{3.5}$$

If, as $n \to \infty$ in such a way that $\max |t_i - t_{i-1}| \to 0$, the sequence I_n converges in mean square to I (say), then we deøne I to be the integral over $[a, b]$ of the stochastic process $X(t)$, that is to say

$$I := \int_a^b X(t)dt$$

Since the integral is a random variable, we need to specify the sense in which the sequence converges. Convergence in mean square means that

$$\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}}\left[|I_n - I|^2\right] = 0.$$

Mean square convergence implies convergence in distribution.

There is a general result which says that a regular stochastic process $X(t)$ with covariance function $\Gamma(s,t)$ is integrable over $[a,b]$ if and only if the double integral

$$\int_a^b \int_a^b \Gamma(s,t)dsdt \tag{3.6}$$

exists, and in this case

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\int_a^b X(t)dt\right)\left(\int_a^b X(s)ds\right)\right] = \int_a^b \int_a^b \Gamma(s,t)dsdt \tag{3.7}$$

Since we shall primarily be concerned with Brownian motion, we look at this result in some detail for the special case when $X(t)$ is a Brownian motion.

Let I_n be deøned as in (3.5), but with $X(t)$ replaced by $W(t)$. From the deønition of mean square convergence, $I_n \longrightarrow_{m.s.} I$ if and only if $\mathbb{E}_{\mathbb{Q}}[I_n I_p] = \mathbb{E}_{\mathbb{Q}}[I^2]$, where I_p is the sum corresponding to the partition $a = s_0 < s_1 < ... < s_n = b$, and $s_{i-1} < \eta_i < s_i$. We have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[I_n I_p] &= \mathbb{E}_{\mathbb{Q}}\left[\left(\sum_{i=1}^{n}(t_i - t_{i-1})\,W(\xi_i)\right)\left(\sum_{j=1}^{p}(s_j - s_{j-1})\,W(\eta_j)\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\sum_{i=1}^{n}\sum_{j=1}^{p}(t_i - t_{i-1})(s_j - s_{j-1})\,W(\xi_i)\,W(\eta_j)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{p}(t_i - t_{i-1})(s_j - s_{j-1})\,\mathbb{E}_{\mathbb{Q}}\left[W(\xi_i)\,W(\eta_j)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{p}(t_i - t_{i-1})(s_j - s_{j-1})\,\Gamma(\xi_i, \eta_j)
\end{aligned}$$

and if the limit of this quantity exists as $n \to \infty$, $p \to \infty$, $\max|t_i - t_{i-1}|$ and $\max|s_i - s_{i-1}|$, then it is precisely the double integral in (3.6). This proves both our claims.

Assuming convergence is uniform, and since the expectation operator is linear, we have

$$\begin{aligned} \lim \mathbb{E}_{\mathbb{Q}}[I_n] &= \mathbb{E}_{\mathbb{Q}}[\lim I_n] \quad \text{or} \\ \mathbb{E}_{\mathbb{Q}}\left[\int_a^b X(t)dt\right] &= \int_a^b \mathbb{E}_{\mathbb{Q}}[X(t)]dt \end{aligned} \tag{3.8}$$

As an example, consider the integral of a Brownian motion, $I(0,T) = \int_0^T W(t)dt$. Using (3.8) it is easy to see that $\mathbb{E}_{\mathbb{Q}}[I(0,T)] = 0$. From (3.7) we have

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[I(0,T)^2] &= \int_0^T \int_0^T \min\{s,t\} ds dt \\ &= \int_0^T \left\{ \int_0^s t dt + \int_s^T s dt \right\} ds \\ &= \int_0^T \left\{ Ts - \frac{1}{2}s^2 \right\} ds \\ &= \frac{1}{3}T^3 \end{aligned}$$

Since the sum of normal distributions is normal, we conclude that

$$I(0,T) = \int_0^T W(t)dt = W\left(\frac{1}{3}T^3\right) \sim N\left(0, \frac{1}{3}T^3\right) \tag{3.9}$$

Integrating over Brownian Increments

The other important stochastic integral that we shall encounter is the one we deøne next:

$$J := \int_a^b h(t)dW(t)$$

As before, we deøne the integral to be the limit, if it exists, of the sequence

$$J_n = \sum_{j=0}^{n-1} h\left(\xi_j\right)\left(t_{j+1} - t_j\right) \tag{3.10}$$

Since each term of the summation is normally distributed with mean zero, the integral itself must be normally distributed with mean zero. There is a theorem which states that

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\int_a^b h(t)dW(t)\right)\left(\int_a^b g(s)dW(s)\right)\right] = \int_a^b \mathbb{E}_{\mathbb{Q}}\left[h(t)g(t)\right]dt \tag{3.11}$$

Though the proof of this result in general is beyond the scope of this paper, the basic outline is as follows: ørst, we prove the result for a class of functions called step functions, which essentially cover all functions whose exact integral is given by (3.10). These functions are dense in the class of integrable functions, so that the result (3.11) for step functions can be continuously extended to the whole class. The proof of (3.11) for step functions is sketched below.

A function $h(.)$ is called a step function if there exists a partition $a = t_0 < t_1 < ... < t_n = b$ with associated random variables $h_0, ..., h_{n-1}$ such that

$$h(t) = \sum_{i=1}^{n} h_{i-1} \mathcal{X}_{[t_{i-1}, t_i)}(t)$$

where $\mathcal{X}$ is the indicator function. For these functions, the integral is deøned as

$$J(h) = \sum_{i=0}^{n} h_{i-1}(W(t_i) - W(t_{i-1}))$$

For two diœerent step functions, we can represent the integral using the same partition but diœerently chosen random variables. If h and g are step functions then

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}} \left[\left(\int_a^b h(t) dW(t) \right) \left(\int_a^b g(s) dW(s) \right) \right] \\ = \; & \mathbb{E}_{\mathbb{Q}} \left[\sum_{i,j=1}^{n} h_{i-1} g_{j-1} (W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1})) \right] \\ = \; & \mathbb{E}_{\mathbb{Q}} \left[\sum_{i,j=1}^{n} h_{i-1} g_{j-1} (W(t_i) - W(t_{i-1}))(W(t_j) - W(t_{j-1})) \middle| \mathcal{F}(\max\{t_i, t_j\}) \right] \end{aligned}$$

If $i \neq j$ then we can use independence of Brownian increments and say that the expression must equal zero, since the increments themselves have mean zero. If $i = j$ then the inner expectation becomes

$$\begin{aligned} & h_{i-1} g_{i-1} \mathbb{E}_{\mathbb{Q}} \left[|W(t_i) - W(t_{i-1})|^2 \middle| \mathcal{F}(t_i) \right] \\ = \; & h_{i-1} g_{i-1} \mathbb{E}_{\mathbb{Q}} \left[|W(t_i) - W(t_{i-1})|^2 \right] \\ = \; & h_{i-1} g_{i-1} (t_i - t_{i-1}) \end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[\left(\int_a^b h(t)dW(t)\right)\left(\int_a^b g(s)dW(s)\right)\right] &= \sum_{i=1}^{n} \mathbb{E}_{\mathbb{Q}}\left[h_{i-1}g_{i-1}\right](t_i - t_{i-1}) \\
&= \int_a^b \mathbb{E}_{\mathbb{Q}}\left[h(t)g(t)\right]dt
\end{aligned}$$

This result gives us a neat alternative to deriving (3.9). We can write

$$\begin{aligned}
I(0,T) &= \int_0^T W(t)dt = \int_0^T \left(\int_0^t dW(s)\right)dt \\
&= \int_0^T \left(\int_s^T dt\right)dW(s) = \int_0^T (T-s)dW(s)
\end{aligned}$$

and now, using (3.11),

$$\mathbb{E}_{\mathbb{Q}}\left[I(0,T)^2\right] = \int_0^T (T-s)^2 ds = \frac{1}{3}T^3$$

The two results on stochastic integration derived in this section are the ones we shall need to use in this paper.

3.4.3 Stochastic Dioeerentiation: Ito's Formula

Suppose $h(x) = x^2$. Then under Newtonian rules, we expect $dh(x) = 2xdx$, but does this hold even if the argument is stochastic, for example a Brownian motion? If it did, then we should have

$$d(W(t)^2) = 2W(t)dW(t)$$

which means that

$$\int_0^t dW(t)^2 = W(t)^2 = 2\int_0^t W(t)dW(t)$$

We approximate the last integral by discretising it, so that

$$2\int_0^t W(t)dW(t) = 2\sum_{i=0}^{n-1} W\left(\frac{it}{n}\right)\left\{W\left(\frac{(i+1)t}{n}\right) - W\left(\frac{it}{n}\right)\right\}$$

Now each increment is normally distributed with mean zero, so the sum itself is normally distributed with mean zero. However, the integral, which if we are right should be $W(t)^2$, has mean t. Thus the assumption that cannot be right. Clearly, we need to look at stochastic diœerentials and integrals in a slightly diœerent way from Newtonian ones.

What went wrong? Consider a Taylor expansion

$$dh(W(t)) = h'(W(t))dW(t) + \frac{1}{2}h''(W(t))dW(t)^2 + ...$$

Newtonian diœerentiation approximates a derivative with one term of a Taylor series, because higher order terms decay much faster than the leading order term. Is this necessarily true for stochastic diœerentials as well? To answer this, consider the second term in the series above. We can write

$$\int_0^t dW(t)^2 = \sum_{i=0}^{n-1} \left\{ W\left(\frac{(i+1)t}{n}\right) - W\left(\frac{it}{n}\right) \right\}^2$$

Deøne

$$Z(n,i) = \frac{W\left(\frac{(i+1)t}{n}\right) - W\left(\frac{it}{n}\right)}{\sqrt{\frac{t}{n}}}$$

so that

$$\int_0^t dW(t)^2 = t\sum_{i=1}^{n} \frac{Z(n,i)^2}{n}$$

Since for each n, the sequence $Z(n,1)$, $Z(n,2)$, ... is a sequence of IID normals $N(0,1)$, we can use the weak law of large numbers and say that the distribution of each term in the sum on the right converges to the mean of each $Z(n,i)^2$, namely 1. Thus

$$\int_0^t dW(t)^2 = t$$

This means that the second term cannot be ignored! Higher order terms can, though, and so we have

$$dh(W(t)) = h'(W(t))dW(t) + \frac{1}{2}h''(W(t))dt + \text{ h.o.t.}$$

The generalisation of this result is called Ito's lemma or Ito's formula. If $X(t)$ is a stochastic process such that

$$dX(t) = \sigma_t dW(t) + \alpha_t dt$$

and $h \in \mathcal{C}^2$, i.e. h is a deterministic twice continuously diœerentiable function, then $Y(t) :=$

$h(X(t))$ is also a stochastic process and

$$dY(t) = \sigma_t h'(X(t))dW(t) + \left\{ \alpha_t h'(X(t)) + \frac{1}{2}\sigma_t^2 h''(X(t)) \right\} dt \tag{3.12}$$

In n dimensions, the formulation is as follows: if $X(t)$ is given by the SDE

$$dX(t) = \sum_{i=1}^{n} \sigma_i(t) dW(t) + \alpha_t dt$$

and $Y(t) := h(X(t))$ then

$$dY(t) = \sum_{i=1}^{n} \sigma_i(t) h'(X(t)) dW_i(t) + \left\{ \alpha_t h'(X(t)) + \frac{1}{2} \sum_{i=1}^{n} \sigma_i(t)^2 h''(X(t)) \right\} dt \tag{3.13}$$

3.4.4 The Product Rule

Consider two stochastic processes adapted to the same Brownian motion, which means they can be written as

$$\begin{aligned} dX &= \sigma dW + \alpha dt \\ dY &= \rho dW + \mu dt \end{aligned}$$

where all the coeŒcients are $\mathcal{F}(t)$-adapted. We are interested in an expression for $d(XY)$. Note that XY can be written as

$$XY = \frac{1}{2} \left\{ (X+Y)^2 - X^2 - Y^2 \right\}$$

We can apply Ito's lemma directly to X^2 and Y^2 (take $h(x) = x^2$). As for $X + Y$, the SDE that it satisøes is

$$d(X+Y) = (\sigma + \rho)\, dW + (\alpha + \mu)\, dt$$

and we can apply Ito's lemma as before to this as well. Since $d(XY)$ is $O(dt)$, comparing coeŒcients of dt in the resulting equation gives the expression

$$d(XY) = X dY + Y dX + \sigma\rho dt \tag{3.14}$$

In the multidimensional case, if $W_1, ..., W_n$ are independent Brownian motions and

$$\begin{aligned} dX &= \sum \sigma_i dW_i + \alpha dt \\ dY &= \sum \rho_i dW_i + \mu dt \end{aligned}$$

then a similar exercise gives the product rule in n dimensions as follows:

$$d(XY) = X dY + Y dX + \left(\sum \rho_i \sigma_i\right) dt \tag{3.15}$$

This has an important consequence, while considering stochastic processes adapted to two independent Brownian motions. We can view the two Brownian motions as components of a 2-dimensional Brownian motion, and then recover the two processes by choosing coeŒcients appropriately. Explicitly, in a two-Brownian-motion case, we have $n = 2$, $\sigma_1 = 1$, $\sigma_2 = 0$, $\rho_1 = 0$, $\rho_2 = 1$ and so

$$d(XY) = X dY + Y dX \tag{3.16a}$$

3.5 The Radon-Nikodym Derivative

We are now in a position to address a question that had persistently cropped up earlier. Implicit in the deønition of a stochastic process was the concept of an underlying probability measure. How does the process change if the underlying measure is changed?

3.5.1 The Discrete Radon-Nikodym Derivative

The ørst step is to ønd a way to represent change of measure. Consider a discrete binomial process shown below.

Radon-Nikodym on a discrete tree.

Under this measure $\mathbb{Q}$ described by the p_i, the probability π_i of reaching each node i can be calculated. For example, $\pi_6 = p_1(1 - p_2)$. Suppose now we have a diœerent probability measure $\mathbb{Q}^*$, described by the numbers p_i^*, then we shall have a corresponding set of probabilities associated with each node, say π_i^*. Comparing these two measures, the relative likelihood of reaching a particular node is given by π_i^*/π_i. At each node, this ratio deønes a quantity called the Radon-Nikodym derivative of $\mathbb{Q}^*$ with respect to $\mathbb{Q}$ at that node. Thus, we deøne the discrete Radon-Nikodym derivative or likelihood ratio to be

$$\frac{d\mathbb{Q}}{d\mathbb{Q}}(i) = \frac{\pi_i^*}{\pi_i}$$

This is well deøned except if some of the p_i are zero. Also, if some of the p_i^* are zero, we lose information about subsequent nodes in that measure. So we need to restrict our deønition to measures that exactly agree on what events are possible, which is a crude way of deøning equivalent probability measures. More formally, two probability measures and are equivalent if and only if

$$\mathbb{Q}(\omega) = 0 \Longleftrightarrow \mathbb{Q}^*(\omega) = 0$$

for all suitably chosen measurable sets ω in the sample space.

3.5.2 Radon-Nikodym and Expectations

One of the beneøts of deøning Radon-Nikodym derivatives is that if we are given expectations with respect to one measure, we can deduce expectations under the other.

$$\mathbb{E}_{\mathbb{Q}^*}[X] = \sum_j \pi_j^* x_j = \sum_j \pi_j \frac{\pi_j^*}{\pi_j} x_j = \mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{Q}^*}{d\mathbb{Q}} X\right] \tag{3.17}$$

3.5.3 Processes from Radon-Nikodym Derivatives

Given a discrete stochastic process, the Radon-Nikodym derivative at each time step is a random variable. For example in the situation described above, the Radon-Nikodym derivative at the end of the ørst time step is given by

$$\left.\frac{d\mathbb{Q}^*}{d\mathbb{Q}}\right|_{\text{time=1}} = \begin{cases} \frac{p_1^*}{p_1} & \text{at node 3} \\ \frac{1-p_1^*}{1-p_1} & \text{at node 2} \end{cases}$$

As the process evolves, so does the Radon-Nikodym derivative. Following our general motivation to derive processes from claims, can we associate a process when we know the Radon-Nikodym derivative at some time horizon T? We can after all think of this Radon-Nikodym derivative as a claim. If we are on the tree where the measure is given by $\mathbb{Q}$, we can construct a process by taking the derivative at each step, and moving forward in time, i.e. we let $\zeta(t)$ be the derivative evaluated till the time t. Another way of looking at this is the conditional expectation of the T-horizon Radon-Nikodym derivative, just as we formed a process from a claim in the opening section of this chapter.

$$\zeta(t) = \mathbb{E}_{\mathbb{Q}}\left[\left.\frac{d\mathbb{Q}^*}{d\mathbb{Q}}\right| \mathcal{F}(t)\right]$$

The process $\zeta(t)$ represents the amount of change of measure so far till time t along the current path. If we want to know $\mathbb{E}_{\mathbb{Q}^*}[X(t)]$, it is just $\mathbb{E}_{\mathbb{Q}}[\zeta(t)X(t)]$.

What if we want $\mathbb{E}_{\mathbb{Q}^*}[X(t) \mid \mathcal{F}(s)]$ for $0 < s < t < T$? This is precisely the situation we will encounter later, when we have a prediction about future values of a ønancial asset in a measure $\mathbb{Q}^*$, given information at some time prior to maturity, and we want to translate this prediction into one in terms of the actual market measure $\mathbb{Q}$. We need the amount of change in measure between the two times, which is just $\zeta(t)/\zeta(s)$, the change to time t less the change to time s.

In other words,

$$\mathbb{E}_{\mathbb{Q}^*}[X(t) \mid \mathcal{F}(s)] = \zeta(s)^{-1}\mathbb{E}_{\mathbb{Q}}[\zeta(t)X(t) \mid \mathcal{F}(s)] \tag{3.18}$$

3.5.4 The Continuous Radon-Nikodym Derivative

Now we turn to the continuous case. The Radon-Nikodym derivative is the relative likelihood of a particular path under two diœerent equivalent measures. One way of calculating this relative likelihood is to discretise the path and calculate a discrete analogue, and then let the mesh become dense in the time interval under consideration.

Consider a time interval $[0, T]$. Form a partition $0 = t_0 < t_1 < ... < t_n = T$. Let $\omega = \omega(t)$ be a path. The idea is to sample points from the path and ønd the joint likelihood of these points under each measure. If the points are $\{x_i = \omega(t_i)\}_{i=0}^{n}$, then we ask: Given a measure $\mathbb{Q} = p(\cdot, t)$, what is the likelihood of all these states being attained? If we have just one point, the answer is $p(x, t)$. If we have two points, we want, crudely, the probability that x_1 is attained at t_1 and x_2 at t_2. Call this $p(x_1, x_2; t_1, t_2)$. Similarly, if we sample many points on the given path, we get a joint likelihood function

$$\mathbb{Q}_n(\omega) = p(x_1, ..., x_n; t_1, ..., t_n) \tag{3.19}$$

which represents the probability, under the measure $\mathbb{Q}$ of attaining the states $\{x_i\}$on the path $\omega(t)$ at times $\{t_i\}$. If we allow the mesh $\{t_i\}$ to become dense in $[0, T]$, then (3.19) provides us a good approximation to $\mathbb{Q}(\omega)$, the likelihood ω of under $\mathbb{Q}$.

Continuous Radon-Nikodym derivative.

Now we do the same thing with a diœerent probability measure $\mathbb{Q}^*$, to get $\mathbb{Q}^*(\omega)$, the likelihood of ω under $\mathbb{Q}^*$. The ratio of these two likelihoods is called the continuous Radon-Nikodym derivative of the measure $\mathbb{Q}^*$ with respect to the measure $\mathbb{Q}$ at ω, which in this case is a point in the sample space (consisting of all possible paths).

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}}(\omega) = \lim_{m,n\to\infty} \frac{p^*(y_1, ..., y_m; s_1, ..., s_m)}{p(x_1, ..., x_n; t_1, ..., t_n)} = \lim_{m,n\to\infty} \frac{\mathbb{Q}^*_n(\omega)}{\mathbb{Q}_n(\omega)}$$

An elegant way of looking at the Radon-Nikodym derivative is as follows. Two measures are equivalent if there exists a function φ such that for any measurable set Ω, we can write

$$\mathbb{Q}^*(\not\leq) = \int_{\omega\in\Omega} \varphi(\omega) d\mathbb{Q}(\omega)$$

The function φ is precisely the continuous Radon-Nikodym derivative. This is a reøned statement of the following intuitive concept:

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}}(\omega) = \lim_{\Omega\to\{\omega\}} \frac{\mathbb{Q}^*(\omega)}{\mathbb{Q}(\omega)}$$

It can be shown that the continuous Radon-Nikodym derivative also satisøes (3.17) and (3.18).

3.6 Change of Measure and Brownian Motion

3.6.1 An Example

This section illustrates the ideas of the previous sections with a concrete example. Let $W(t)$ be a $\mathbb{Q}$-Brownian motion. What will it look like if we change the underlying probability measure? Suppose $\mathbb{Q}^*$ is a measure equivalent to $\mathbb{Q}$. Equivalent measures can be characterised using the Radon-Nikodym derivative of one measure with respect to the other. Thus, suppose $\mathbb{Q}^*$ is described by the following Radon-Nikodym derivative:

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = e^{-\gamma W(t) - \frac{1}{2}\gamma^2 t} \tag{3.20}$$

What does $W(t)$ look like under $\mathbb{Q}^*$? To answer this question, we look at the distribution of $W(t)$ under the new measure.

3.6.2 Characterising Normal Random Variables

A characterisation of normal random variables using moment generating functions that is particularly useful in situations like this is as follows:

$$X \sim N_{\mathbb{Q}}(\mu, \sigma^2) \Longleftrightarrow \mathbb{E}_{\mathbb{Q}}\left[e^{\theta X}\right] = e^{\theta\mu + \frac{1}{2}\sigma^2\theta^2} \tag{3.21}$$

3.6.3 Example–Continued

Since $W(t) \sim N(0, t)$ under $\mathbb{Q}$, we have, using (3.17), (3.20) and (3.21)

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}^*}\left[e^{\theta W(t)}\right] &= \mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{Q}^*}{d\mathbb{Q}} e^{\theta W(t)}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[e^{-\gamma W(t) - \frac{1}{2}\gamma^2 t + \theta W(t)}\right] \\
&= e^{-\frac{1}{2}\gamma^2 t}\mathbb{E}_{\mathbb{Q}}\left[e^{(\theta-\gamma)W(t)}\right] \\
&= e^{-\frac{1}{2}\gamma^2 t + \frac{1}{2}(\theta-\gamma)^2 t} \\
&= e^{-\theta\gamma t + \frac{1}{2}\theta^2 t}
\end{aligned}$$

so that $W(t) \sim N(-\gamma t, t)$ under $\mathbb{Q}^*$, which looks like a Brownian motion with a drift term added. In fact, if we define $W^*(t) = W(t) + \gamma t$ then it is easy to see that $W^*(t)$ is a $\mathbb{Q}^*$-Brownian motion. This change of measure has thus had the effect of converting a Brownian motion into a Brownian motion with drift. Changing measure amounts to changing the relative likelihood of a path being chosen. In the original measure, we expect the Brownian motion to have mean zero. This does not mean that it cannot have mean zero in the second measure, only that the driftless path is less likely.

3.7 The Cameron-Martin-Girsanov Theorem

3.7.1 Motivation

From the previous section, it is clear that a change in measure can change the drift of a Brownian motion. The Cameron-Martin-Girsanov theorem (CMG) asserts that this is all a change of measure can achieve. Changing measure amounts to changing the drift of a Brownian motion, and conversely a Brownian motion with drift can be regarded as a driftless Brownian motion in some suitably chosen measure.

3.7.2 The Theorem

If $W(t)$ is a $\mathbb{Q}$-Brownian motion and $\gamma(t)$ is an $\mathcal{F}$-previsible process that satisfies the boundedness condition

$$\mathbb{E}_{\mathbb{Q}}\left[e^{\frac{1}{2}\int_0^T \gamma(t)^2 dt}\right] < \infty$$

then there exists a measure $\mathbb{Q}^*$ such that

- $\mathbb{Q}^*$ is equivalent to $\mathbb{Q}$;
- $\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \exp\left\{-\int_0^T \gamma(t) dW(t) - \frac{1}{2}\int_0^T \gamma(t)^2 dt\right\}$; and
- $W^*(t) = W(t) + \int_0^s \gamma(s) ds$ is a $\mathbb{Q}^*$-Brownian motion.

This means that within constraints, if we want to change a Brownian motion into a Brownian motion with drift, then we can find a measure $\mathbb{Q}^*$ that allows us to do so.

3.7.3 CMG Converse

The CMG has a converse as well. If $W(t)$ is a $\mathbb{Q}$-Brownian motion and $\mathbb{Q}^*$ is a measure equivalent to $\mathbb{Q}$ then there exists an $\mathcal{F}$-previsible process $\gamma(t)$ such that

$$W^*(t) = W(t) + \int_0^t \gamma(s)ds$$

is a $\mathbb{Q}^*$-Brownian motion. Additionally, the Radon-Nikodym derivative of $\mathbb{Q}^*$ with respect to $\mathbb{Q}$ is given by

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \exp\left\{-\int_0^T \gamma(t)dW(t) - \frac{1}{2}\int_0^T \gamma(t)^2 dt\right\}$$

3.7.4 Illustrating CMG

In the context of the stochastic processes we have encountered so far, CMG gives us an eŒcient tool for controlling the drift of a process. Suppose we are given a stochastic process

$$dX = \sigma dW + \mu dt$$

in the measure $\mathbb{Q}$, and we want to ønd a measure $\mathbb{Q}^*$ such that the drift coeŒcient in the new measure is ν. We write

$$dX = \sigma\left(dW + \frac{\mu - \nu}{\sigma}dt\right) + \nu dt$$

and if we now set $\gamma(t) = (\mu(t) - \nu(t))/\sigma(t)$ and the boundedness condition of the CMG theorem is satisøed, then we can quote CMG and say that the required measure does indeed exist. Under this measure,

$$W^*(t) = W(t) + \int_0^t \frac{\mu(s) - \nu(s)}{\sigma(s)}ds$$

is a Brownian motion. Thus in the measure $\mathbb{Q}^*$ the process is

$$dX = \sigma dW^* + \nu dt$$

In particular, if we set $\nu(t)$ to be zero, then under minor constraints we can change any stochastic process to a driftless one.

3.7.5 CMG in Several Dimensions

The CMG in several dimensions is not hard to write down. Let $W(t) = [W_1(t), ..., W_n(t)]^{\mathrm{T}}$ be an n-dimensional $\mathbb{Q}$-Brownian motion. Suppose $\gamma(t) = [\gamma_1(t), ..., \gamma_n(t)]^{\mathrm{T}}$ is an n-dimensional $\mathcal{F}$-previsible process such that the boundedness condition

$$\mathbb{E}_{\mathbb{Q}} \exp\left[\left(\frac{1}{2}\|\gamma(s)\|^2_{\mathcal{L}^2(0,t)}\right)\right] < \infty$$

is satisøed, and we set

$$W_i^*(t) = W_i(t) + \int_0^t \gamma_i(s)ds$$

for $i = 1, ..., n$. Then there exists a measure $\mathbb{Q}^*$ equivalent to $\mathbb{Q}$, determined by

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \exp\left\{-\sum_{i=1}^n \int_0^t \gamma_i(s)dW_i(s) - \frac{1}{2}\sum_{i=1}^n \|\gamma_i(s)\|^2_{\mathcal{L}^2(0,t)}\right\}$$

for $0 < t < T$, such that $W^*(t) = [W_1^*(t), ..., W_n^*(t)]^{\mathrm{T}}$ is an n-dimensional $\mathbb{Q}^*$-Brownian motion.

The converse also holds, and is exactly analogous to the one-dimensional case.

3.8 The Martingale Representation Theorem

3.8.1 Motivation

Two results that are central to the modelling of ønancial derivatives in general and to interest rates in particular are the CMG, discussed in the previous section, and the martingale representation theorem (MRT), discussed here. The essence of the MRT is that any martingale can be expressed in terms of another with a suitable change of scale.

3.8.2 The Theorem

If is $M(t)$ a $\mathbb{Q}$-martingale process whose volatility $\sigma(t)$ satisøes the additional condition that it is (with probability 1) always non-zero, and $N(t)$ is any other $\mathbb{Q}$-martingale, then there exists an essentially unique $\mathcal{F}$-previsible process $\varphi(t)$ such that

$$\int_0^T \varphi(t)^2\sigma(t)^2 dt < \infty$$

with probability 1, and $N(t)$ can be represented in terms of $M(t)$ as

$$N(t) = N(0) + \int_0^t \varphi(s) dM(s)$$

3.8.3 Characterising Martingales

Two characterisations of martingale processes are particularly useful for our purposes. The first asserts that under suitable technical constraints, any driftless stochastic process is a martingale. If

$$dX = \sigma dW + \mu dt$$

and the volatility coefficient satisfies the boundedness condition

$$\mathbb{E}_{\mathbb{Q}} \left[||\sigma||_{\mathcal{L}^2(0,T)} \right] < \infty$$

then we have the result:

$$X(t) \text{ is a martingale} \iff X(t) \text{ is driftless, i.e. } \mu = 0$$

If the boundedness condition fails, a driftless process may not be a martingale. Such processes are called local martingales.

3.8.4 Exponential Martingales

The second result characterises exponential martingales. If $dX = \sigma X dW$ for some $\mathcal{F}$-previsible process σ then

$$\mathbb{E}_{\mathbb{Q}} \left[\exp \left(\frac{1}{2} \int_0^T \sigma(s)^2 ds \right) \right] < \infty \Longrightarrow X(t) \text{ is a } \mathbb{Q}\text{-martingale.} \tag{3.22}$$

3.8.5 MRT in Several Dimensions

The MRT can be extended to several dimensions as well. Here, instead of the volatility coefficient, we have a volatility matrix $\sigma = [\sigma_{ij}(t)]_{i,j=1}^n$. The j-th component of an n-dimensional

$\mathbb{Q}$-martingale $M(t)$ can be written in terms of an n-dimensional $\mathbb{Q}$-Brownian motion as

$$dM_j(t) = \sum_{i=1}^{n} \sigma_{ij}(t) dW_i(t)$$

and the process $M(t)$ itself can be represented as $dM(t) = \sigma dW$. If the volatility matrix is non-singular with probability 1 and $N(t)$ is any 1-dimensional $\mathbb{Q}$-martingale, then there exists an essentially unique n-dimensional $\mathcal{F}$-previsible process $\varphi(t) = [\varphi_1(t), ..., \varphi_n(t)]^{\mathrm{T}}$ satisfying

$$\int_0^T \left(\sum_{j=1}^{n} \sigma_{ij}(t)\varphi_j(t) \right)^2 dt < \infty$$

and the process $N(t)$ can be written as

$$N(t) = N(0) + \sum_{j=1}^{n} \int_0^t \varphi_j(s) dM_j(s)$$

3.9 Endnotes

What has the purpose of this tedious exercise been? The reason for going through stochastic calculus in some detail was to be able to understand CMG and the MRT. Using CMG we change measure in the ønancial world to factor the risk preferences of an economic agent out, thus placing the agent in a risk-neutral economy. It will turn out that in such a world, quantities that can be traded in line with the principle of no-arbitrage are precisely those which can be represented as martingales in the risk-neutral world. We then use the MRT. Once we have valued our ønancial instruments, we once again can use CMG to factor in an investors risk preferences and quote a value appropriate to those preferences. This process will be more fully outlined in subsequent chapters. For instruments that are traded, we need to ønd synthetic constructions that track the value of these traded instruments over their lifetimes. For those that are not, we observe the prices of tradable instruments whose value derives from the underlying non-tradables, and thus track the non-tradable. These are the last pieces that we need to complete the picture.

Chapter 4

Prelude to HJM

There is a theory which states that if ever anyone discovers exactly what the universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable.

There is another theory which states that this has already happened.

4.1 Objectives

- Prepare the ground for a general approach to risk-neutral contingent claim pricing.
- Introduce the concept of tradable quantities, numeraires, self-ønancing portfolios and replicating portfolios.
- Demonstrate contingent claim valuation using replicating portfolios, CMG and MRT.
- Illustrate the procedure by using it to price stock options.
- Show the equivalence of this formulation and the Black-Scholes framework.
- Review market completeness and its consequences.
- Investigate change of numeraires.

4.2 The Market

Consider a continuous-trading economy and a trading interval $[0, \tau]$. Assume that a continuum of default-free bonds trades in this economy, with expiries on every trading date in the time interval. This means a set of bonds

$$\{P(t, T) : 0 < t < T < \tau\}$$

Uncertainty in the economy is characterised by a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\mathcal{F}$ is the øltration generated by n independent Brownian motions

$$\{W_1(t), ..., W_n(t)\}$$

initialised at zero. In the 1-factor world, a single Brownian motion generates all the uncertainty.

A note about the probability space. $\mathbb{Q}$ is called the market measure. It reÆects the market's prediction of the future distribution of asset prices and reÆects the market's perception to risk as a whole. In this paper, we will primarily be concerned with the market's perception of how bond prices will move, since these are the tradable instruments that we shall consider.

4.3 Trading Strategies

4.3.1 Portfolios

Intuitively, a portfolio is a collection of ønancial assets held by an investor at a given time. Portfolios are dynamic since the investor can buy or sell any amount of these at will. In a simple 1-factor case, a portfolio is a vector process (ϕ, ψ) that is bounded, $\mathcal{F}$-previsible and non-negative. It consists of ϕ units of the T-expiry bond $P(t, T)$ and ψ units of the numeraire. A numeraire is a basic security relative to which the value of others can be judged. Usually, we take it to be a cash or money market account $B(t)$, started with unit amount and rolling over at the riskless spot rate of interest $r(t)$. Thus our numeraire is

$$B(t) = e^{\int_0^t r(s)ds}$$

4.3.2 Self-Financing Portfolios: Discrete Time

Consider a discrete-time framework and a trading strategy that involves buying a portfolio Π_i at time i, which consists of ϕ_{i+1} units of the T-bond $P_i = P(t_i, T)$ and ψ_{i+1} units of the cash bond $B_i = B(t_i)$. Note that since the processes are previsible, we can get their values at $i+1$ from their values at i. The value of this portfolio at time $i-1$ is

$$V_{i-1} = \phi_i P_{i-1} + \psi_i B_{i-1}$$

If this strategy is held over the next time tick, then its value changes to

$$\phi_i P_i + \psi_i B_i \tag{4.1}$$

since the securities will have changed values but the amounts held remain unchanged. The ønancing gap D_i is the diœerence between (4.1) and V_i, and is the extra cash that needs to be injected in order to buy the portfolio Π_i at time i and so keep the trading strategy going. Thus we have

$$\begin{aligned} D_i &= V_i - \phi_i P_i - \psi_i B_i \\ &= V_i - V_{i-1} + V_{i-1} - \phi_i P_i - \psi_i B_i \\ &= \Delta V_i + \phi_i P_{i-1} + \psi_i B_{i-1} - \phi_i P_i - \psi_i B_i \\ &= \Delta V_i - \phi_i \left(P_i - P_{i-1}\right) - \psi_i \left(B_i - B_{i-1}\right) \\ &= \Delta V_i - \phi_i \Delta P_i - \psi_i \Delta B_i \end{aligned}$$

If the ønancing gap were zero (i.e. $D_i = 0$), then no external inÆows would be needed to sustain the portfolio. In other words, the portfolio would be self-ønancing. Thus the condition for a portfolio to be self-ønancing is

$$\Delta V_i = \phi_i \Delta P_i + \psi_i \Delta B_i \tag{4.2}$$

This means that all changes in the value of the portfolio are entirely due to changes in the value of the securities held.

4.3.3 Self-Financing Portfolios: Continuous Time

In the continuous case, the strategy would be to hold a portfolio $\Pi(t) = (\phi(t), \psi(t))$ whose value at time t is

$$V(t) = \phi(t)P(t,T) + \psi(t)B(t)$$

If we hold this portfolio for the interval $[t, t+dt]$ then at $t+dt$ its value is

$$\phi(t)P(t+dt,T) + \psi(t)B(t+dt)$$

The ønancing gap is therefore given by

$$\begin{aligned} D(t) &= V(t+dt) - \phi(t)P(t+dt,T) - \psi(t)B(t+dt) \\ &= dV(t) - \phi(t)\left(P(t+dt,T) - P(t,T)\right) - \psi(t)\left(B(t+dt) - B(t)\right) \\ &= dV(t) - \phi(t)dP(t,T) - \psi(t)dB(t) \end{aligned}$$

The portfolio is self-ønancing if and only if the ønancing gap is zero, and so in continuous time the condition for a portfolio to be self-ønancing is, analogous to (4.2) in discrete time, given by

$$dV(t) = \phi(t)dP(t,T) + \psi(t)dB(t) \tag{4.3}$$

i.e. the change in the portfolios value comes only from changes in values of the securities held in it.

4.3.4 The Discounted Bond Price Process

Given a bond price process $P(t,T)$ the discounted bond price process is deøned to be the bond price normalised by the numeraire. If the numeraire is a money market account, then the discounted bond price process is $Z(t,T) = B(t)^{-1}P(t,T)$. Consider a portfolio $\Pi(t)$ as above. In terms of the discounted bond price process, we can write

$$H(t) = B(t)^{-1}V(t) = \phi(t)Z(t,T) + \psi(t)$$

By considering $Z(t,T)$ as a security in itself and following exactly the same procedure as above, the self-ønancing condition can also be written as

$$dH(t) = \phi(t) dZ(t,T) \quad (4.4)$$

As before, changes in the value of the portfolio are derived entirely from changes in the prices of the securities in it.

4.3.5 Replicating Portfolios

Given a claim C, a replicating portfolio for the claim is one which is self-ønancing and whose terminal value is C. Thus $\Pi(t) = (\phi(t), \psi(t))$ is a replicating portfolio for C if

- $dV(t) = \phi(t)dP(t,T) + \psi(t)dB(t)$; and
- $V(T) = C$.

Why do we care about replicating portfolios? If one exists, we can set it up at some initial time and be sure that it will pay oœ C at expiry without any external inÆows of cash. The value of the replicating portfolio at any time prior to expiry is exactly the arbitrage-free price at that time. If the traded price was lower, we could buy the (lower-priced) contract that guaranteed C at expiry and short the replicating portfolio, thus locking in a riskless proøt. Similarly, an arbitrage opportunity would exist if the traded price was higher than the value of the replicating portfolio. Thus replicating portfolios allow us to synthesize a contingent claim.

4.4 Claim Valuation

We now have the tools ready for valuing a contingent claim. From the paragraph above, the value of a contingent claim at any time is equal to the value of the replicating portfolio at that time. This section presents a recipe for valuation of contingent claims, using CMG, MRT and the idea of a replicating portfolio. To illustrate the basic ideas, we shall use a 1-factor model. The extension to several factors is conceptually straightforward.

4.4.1 Change of Measure

Suppose we are given a market consisting of a bond $P(t,T)$, a numeraire $B(t)$ and some contingent claim C. This is a 1-factor version of the framework outlined at the beginning of this chapter. The ørst step to valuation is to form the discounted bond price process $Z(t,T) = B(t)^{-1}P(t,T)$. We then use CMG to ønd a probability measure $\mathbb{Q}^*$ equivalent to the market measure $\mathbb{Q}$ such that $Z(t,T)$ is a $\mathbb{Q}^*$-martingale. For reasons that will become clear in subsequent chapters, $\mathbb{Q}^*$ is the risk-neutral measure. Form the process

$$H(t) = \mathbb{E}_{\mathbb{Q}^*}\left[B(T)^{-1}C \middle| \mathcal{F}(t)\right]$$

which, by construction, is a $\mathbb{Q}^*$-martingale. Using MRT, there exists an $\mathcal{F}$-previsible process $\phi(t)$ such that

$$dH(t) = \phi(t)dZ(t,T)$$

4.4.2 Constructing a Replicating Portfolio

Our trading strategy now is to hold a portfolio $\Pi(t) = (\phi(t), \psi(t))$ for $0 < t < T < \tau$ consisting of

- $\phi(t)$ units of the T-bond $P(t,T)$; and
- $\psi(t) = H(t) - \phi(t)Z(t,T)$ units of the cash bond (numeraire) $B(t)$.

We need to check that this recipe is correct, i.e. that we have indeed created a replicating portfolio for the claim. The value of the portfolio at time t is

$$\begin{aligned} V(t) &= \phi(t)P(t,T) + \psi(t)B(t) \\ &= \phi(t)P(t,T) + (H(t) - \phi(t)Z(t,T))\,B(t) \\ &= H(t)B(t) \end{aligned}$$

Now since $B(t)$ is a zero-volatility process and $H(t)$ is stochastic, we can use the chain rule as in Newtonian calculus and write

$$\begin{aligned} dV &= HdB + BdH \\ &= (\psi + \phi Z)\, dB + \phi B dZ \\ &= \phi\,(ZdB + BdZ) + \psi dB \\ &= \phi d(BZ) + \psi dB \\ &= \phi dP + \psi dB \end{aligned}$$

Thus the portfolio is self-financing. To check that it is replicating, we need to check that its terminal value is C. Since

$$H(T) = \mathbb{E}_{\mathbb{Q}^*}\left[B(T)^{-1}C\middle|\, \mathcal{F}(T)\right] = B(T)^{-1}C$$

we have $V(T) = H(T)B(T) = C$. Thus the portfolio is replicating.

4.4.3 A Claim Valuation Formula

From what we have outlined above, the current value of the contingent claim must be given by

$$\begin{aligned} V(t) &= B(t)\mathbb{E}_{\mathbb{Q}^*}\left[B(T)^{-1}C\middle|\, \mathcal{F}(t)\right] \\ &= \mathbb{E}_{\mathbb{Q}^*}\left[Ce^{-\int_t^T r(s)ds}\middle|\, \mathcal{F}(t)\right] \qquad (4.5) \end{aligned}$$

This is just what we would expect from the rational expectations hypothesis.

4.5 Pricing Stock Options

To get an idea of how the ideas described so far work in practice, consider the pricing of stock options. In a simplistic sense, an option may be described as a claim defined at some future date, whose value is a function of the stock price at that time. One such instrument, the European call option, has already been introduced in chapter 3. This section presents the Black-Scholes model, which is the benchmark for option pricing.

4.5.1 Black-Scholes

The Black-Scholes framework posits the existence of a deterministic riskless interest rate r and speciøes a model for the behaviour of stock prices $S(t)$ under the market measure $\mathbb{Q}$ as follows:

$$S(t) = S_0 e^{\sigma W(t) + \left(\mu - \frac{1}{2}\sigma^2\right)t} \tag{4.6}$$

where σ is the constant stock volatility and μ is a constant. The quantity $\left(\mu - \frac{1}{2}\sigma^2\right)$ is the (constant) stock drift parameter. The SDE for $S(t)$ is

$$dS(t) = \sigma S(t) dW(t) + \mu S(t) dt \tag{4.7}$$

This is why we take an apparently contrived drift term in (4.6)¯conventionally, μ represents the drift of dS/S. Black and Scholes considered a portfolio consisting of one option $V = V(S,t)$ and short a number Δ of the stock S. Using Ito's lemma, we have

$$dV = \sigma S V_S dW + \left\{ V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \mu S V_S \right\} dt \tag{4.8}$$

Thus for the portfolio $\Pi = V - \Delta S$, we have

$$d\Pi = \sigma S \left(V_S - \Delta\right) dW + \left\{ V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \mu S \left(V_S - \Delta\right) \right\} dt \tag{4.9}$$

If we choose $\Delta = V_S$, we can knock out the random term above and so make the portfolio instantaneously riskless. Thus the return on the portfolio must be $r\Pi dt$ and substituting in various quantities yields the celebrated Black-Scholes equation

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + r S V_S - r V = 0 \tag{4.10}$$

4.5.2 Option Pricing by Expectation

Can we follow the technique developed in this chapter to ønd an option pricing formula consistent with the Black-Scholes price?

The ingredients we need are a stock price process, a numeraire and a claim to replicate.

Our strategy will be to use the stock and the numeraire and construct a portfolio that replicates the option. As we shall see later, the original Black-Scholes approach uses the option and the asset (stock) to replicate a cash bond numeraire.

The stock price is given in the market measure $\mathbb{Q}$ by the process

$$S(t) = S_0 e^{\sigma W(t) + \left(\mu - \frac{1}{2}\sigma^2\right)dt}$$

which corresponds to the SDE

$$dS = \sigma S dW + \mu S dt$$

The numeraire is chosen as a cash bond starting with dollar initial investment and rolling over at the riskless rate of interest, i.e.

$$B(t) = e^{rt}$$

Note that interest rates are assumed constant. This numeraire satisøes the ODE (a special case of an SDE!)

$$dB = rBdt$$

The ørst step is to construct the discount stock price process

$$Z(t) = B(t)^{-1} S(t)$$

The SDE that Z obeys is

$$dZ = \{\sigma dW + (\mu - r)\, dt\}\, Z \tag{4.11}$$

Next, we use CMG to ønd a measure $\mathbb{Q}^*$ equivalent to the market measure $\mathbb{Q}$ such that Z is a martingale under the new measure $\mathbb{Q}^*$. We can rewrite (4.11) as

$$dZ = \sigma Z \left(dW + \frac{\mu - r}{\sigma} dt \right)$$

and then set

$$\gamma(t) = \frac{\mu - r}{\sigma}$$

which clearly satisøes the boundedness conditions for CMG, since it is deøned using constants only. Now we can use CMG to show the existence of a probability measure $\mathbb{Q}^*$ equivalent to the market measure $\mathbb{Q}$ such that

$$\begin{aligned} W^*(t) &= W(t) + \int_0^t \gamma(s)ds \\ & \quad W(t) + \frac{\mu - r}{\sigma}t \end{aligned}$$

is a Brownian motion under $\mathbb{Q}^*$. In this measure, we can write the SDE for Z as

$$dZ = \sigma Z dW^*$$

which implies, from the characterisation of exponential martingales (3.22), that Z is a martingale under the measure $\mathbb{Q}^*$.

If C is the option payoœ at expiry, then the discounted claim is deøned as $B(T)^{-1}C$. Form the process

$$H(t) = \mathbb{E}_{\mathbb{Q}^*}\left[B(T)^{-1}C \middle| \mathcal{F}(t) \right]$$

which, by construction, is a $\mathbb{Q}^*$-martingale. Using the MRT, there exists an $\mathcal{F}$-previsible process $\phi(t)$ such that

$$dH = \phi dZ$$

We are now ready to state our trading strategy. Construct a portfolio (ϕ, ψ) consisting of

- $\phi(t)$ units of the stock $S(t)$; and
- $\psi(t) = H(t) - \phi(t)Z(t)$ units of the cash numeraire $B(t)$

at time t.

It is a routine veriøcation to check that we do indeed have a self-ønancing replicating portfolio for the claim C. The price at time t of the option that pays C at expiry must therefore be

$$\begin{aligned} V(t) &= B(t)H(t) \\ &= e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}^*}[C \mid \mathcal{F}(t)] \end{aligned} \tag{4.12}$$

4.5.3 Equivalence of Formulations

For a European call option, it can be shown that solving the PDE (4.10) under appropriate boundary conditions gives the same answer as solving (4.12). In fact, both the formulations are equivalent.

In the original formulation, Black & Scholes considered the hedged portfolio $V - \Delta S$ consisting of one option and short a number Δ of the stock. They then chose Δ such that the resulting portfolio grew like a riskless cash bond. Their portfolio thus replicated a cash bond. The choice of Δ that knocked out risk from the portfolio was found to be $\Delta = V_S(S,t)$.

In our formulation, we are long the cash bond and the stock, and we try to replicate the option. The asset price random walk is given by

$$dS = \sigma S dW^* + rSdt$$

Let $V(S,t)$ be the value of the replicating portfolio we constructed above. Using Ito's lemma,

$$dV = \sigma S V_S dW^* + \left\{ V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S \right\} dt \tag{4.13}$$

But from the self-ønancing condition, we have

$$\begin{aligned} dV &= \phi dS + \psi dB \\ &= \phi(\sigma S dW^* + rSdt) + r\psi B dt \\ &= \phi\sigma S dW^* + r(\phi S + \psi B)\,dt \\ &= \phi\sigma S dW^* + rVdt \end{aligned} \tag{4.14}$$

Since these represent the same quantity and SDE representations are unique, comparing the stochastic terms in (4.13) and (4.14) must match, giving $\phi = V_S$. Thus the number of stocks that we need to hold is the option delta. In the original Black-Scholes model, the strategy was to be long an option and short a number delta of the stock. In this formulation, it appears that we need to be long delta of the stock. The apparent discrepancy can be reconciled by noting that the Black-Scholes hedged portfolio tries to replicate a cash bond using an option and a stock, whereas here we use the cash bond and the stock to try and replicate the option.

Matching drift terms of the two SDE representations gives

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

which is precisely the Black-Scholes PDE. This shows the equivalence of the two formulations discussed in this chapter.

4.6 Arbitrage-Free Complete Markets

4.6.1 Complete Markets

A market is arbitrage˘free if there is no way of making a riskless proøt. An arbitrage opportunity would be a trading strategy that starts at zero initial value and terminates at some positive value. No such opportunities exist in an arbitrage- free market. A market is called complete if any possible contingent claim can be hedged by trading with a self-ønancing portfolio of securities.

4.6.2 Uniqueness of the Risk-Adjusted Measure

Suppose we have a market of securities (bonds in our case, but the analysis holds in general as well) and a numeraire. Harrison and Pliska showed that this market is arbitrage-free if and only if there exists a probability measure equivalent to the market measure under which all the discounted bond prices (i.e. the bond prices normalised by the numeraire) are martingales. In this case, they show that the market is complete if and only if this equivalent martingale measure is unique.

Suppose there exists a potential arbitrage opportunity in the self-ønancing strategy outlined above. From (4.4), the self-ønancing condition on the portfolio is

$$dH(t) = \phi(t)dZ(t,T)$$

Suppose the strategy starts from zero value but has a non-negative terminal payoœ $V(T) \geq 0$.

Then we have

$$\mathbb{E}_{\mathbb{Q}^*}[H(T)] = \mathbb{E}_{\mathbb{Q}^*}[H(T) \mid \mathcal{F}(0)] = H(0) = V(0) = 0$$

since H is a martingale under the risk-adjusted measure. The fact that the $\mathbb{Q}^*$-expectation of $H(T)$ is zero means that $H(T)$ can only be zero, so that $V(T)$ too can only be zero. Thus if an equivalent martingale measure $\mathbb{Q}^*$ exists, then arbitrage opportunities cannot.

Market completeness means we can hedge using tradables. If we can hedge, there can be at most one equivalent martingale measure $\mathbb{Q}^*$. To see this, suppose two such measures $\mathbb{Q}^*$ and $\mathbb{Q}^0$ exist. If A is an event in the øltration, consider the contingent claim given by the Arrow-Debreu price

$$C = B(T)\mathcal{X}_A$$

where $\mathcal{X}$ is the indicator function. This means the payoœ is $B(T)$ if A occurs and zero otherwise. The complete markets assumption implies that this claim can be replicated, which means, as before, that there exists a discounted strategy such that

$$dH(t) = \phi(t)dZ(t,T)$$

Now $Z(t,T)$ is a martingale under both $\mathbb{Q}^*$ and $\mathbb{Q}^0$ and so $H(t)$ is too. Thus we must have

$$H(0) = \mathbb{E}_{\mathbb{Q}^*}[H(T)] = \mathbb{E}_{\mathbb{Q}^0}[H(T)]$$

Since $H(T)$ is simply the indicator function of the event A, we have $\mathbb{E}_{\mathbb{Q}^*}[A] = \mathbb{E}_{\mathbb{Q}^0}[A]$ for all events A and so the measures $\mathbb{Q}^*$ and $\mathbb{Q}^0$ are identical. Thus market completeness implies that the equivalent martingale measure (EMM) is unique.

4.7 Change of Numeraire

Suppose we have a market of bonds $P(t,T_1), ..., P(t,T_n)$ where $0 < T_i < \tau$ and let their common domain be $[0,T]$ where $T = \min T_i$. We consider these bonds on the common domain and two other instruments $B(t)$ and $G(t)$, one of which is chosen as a numeraire. Let $P_i(t)$ denote the bond $P(t, T_i)$ on the common domain. Thus we have the set of securities

$$\{B(t), G(t), P_1(t), ..., P_n(t)\}$$

If we choose $B(t)$ to be the numeraire, then we seek an EMM $\mathbb{Q}^B$ such that $B(t)^{-1}P_i(t)$ and $B(t)^{-1}G(t)$ are $\mathbb{Q}^B$-martingales. If on the other hand we choose $G(t)$ to be the numeraire, then we seek an EMM $\mathbb{Q}^G$ such that $G(t)^{-1}P_i(t)$ and $G(t)^{-1}B(t)$ are $\mathbb{Q}^G$-martingales. We can actually ønd the Radon-Nikodym derivative of $\mathbb{Q}^G$ with respect to $\mathbb{Q}^B$.

Recall from (3.18) that for any process $X(t)$,

$$\zeta(s)\mathbb{E}_{\mathbb{Q}^G}[X(t) \mid \mathcal{F}(s)] = \mathbb{E}_{\mathbb{Q}^B}[\zeta(t)X(t) \mid \mathcal{F}(s)]$$

It follows that if $X(t)$ is a $\mathbb{Q}^G$-martingale then

$$\zeta(s)X(s) = \mathbb{E}_{\mathbb{Q}^B}[\zeta(t)X(t) \mid \mathcal{F}(s)]$$

and so $X(t)$ is a $\mathbb{Q}^B$-martingale as well. The canonical $\mathbb{Q}^G$-martingales are

$$1, G^{-1}B, G^{-1}P_1, ..., G^{-1}P_n$$

and similarly the canonical $\mathbb{Q}^G$-martingales are

$$1, B^{-1}G, B^{-1}P_1, ..., B^{-1}P_n$$

Each corresponding pair has common ratio $\zeta(t) = B(t)^{-1}G(t)$. Thus the Radon-Nikodym derivative of $\mathbb{Q}^G$ with respect to $\mathbb{Q}^B$ is simply the ratio of the numeraire $G(t)$ to the numeraire $B(t)$:

$$\frac{d\mathbb{Q}^G}{d\mathbb{Q}^B} = \frac{G(t)}{B(t)} \ , \ 0 < t < T$$

We should check that contingent claim valuation remains unchanged under a change in numeraire. The price of a claim C under the measure $\mathbb{Q}^G$ is

$$V^G(t) = G(t)\mathbb{E}_{\mathbb{Q}^G}\left[G(T)^{-1}C \middle| \mathcal{F}(t)\right]$$

Using the Radon-Nikodym result that

$$\mathbb{E}_{\mathbb{Q}^G}[C \mid \mathcal{F}(T)] = \zeta(t)^{-1}\mathbb{E}_{\mathbb{Q}^B}[\zeta(T)C \mid \mathcal{F}(t)]$$

we have

$$\begin{aligned} V^G(t) &= \zeta(t)^{-1}G(t)\mathbb{E}_{\mathbb{Q}^G}\left[G(T)^{-1}C \middle| \mathcal{F}(t)\right] \\ &= B(t)^{-1}\mathbb{E}_{\mathbb{Q}^B}\left[B(T)^{-1}C \middle| \mathcal{F}(t)\right] \end{aligned}$$

which is the value of the claim under the measure $\mathbb{Q}^B$. Thus the two prices agree, as we hoped they would.

4.8 Endnotes

We now have the tools with which to study HJM models. Chapter 3 provided the mathematical grounding in terms of CMG, Radon-Nikodym derivatives and the MRT. This chapter has shown how we can apply these to price contingent claims. The next chapter deals with 1-factor HJM models, and uses the results of this chapter and the last one to price contingent claims under the HJM framework.

Chapter 5

Single-Factor HJM

...like having your brains smashed out by a slice of lemon wrapped round a large gold brick.

5.1 Objectives

- Introduce HJM through its single-factor version.
- Stipulate a forward rate process.
- Follow the analysis of chapter 3 and set up the model under a martingale measure for the discounted bond price.
- Deduce the drift restriction on forward rates that ensures the absence of arbitrage opportunities.
- Link the EMM to market price of risk.
- Use HJM to price bonds.

5.2 Introduction

HJM models study the evolution of an initial yield curve. The general model allows several independent stochastic factors to inÆuence random movements of the yield curve. The single-factor version allows only one such component, but since the transition from one to many factors is conceptually and technically easy, it is useful to study the simpler 1-factor model in detail before moving on to the general model.

We start by considering, as in the previous chapter, a continuous-trading economy, a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ characterising uncertainty in the economy, and a T-expiry bond $P(t,T)$. HJM models the behaviour of the forward rate $f(t,T)$ for $0 < t < T$. It starts with a given initial forward rate curve $f(0,T)$ and volatility structure and then imposes successively restrictive conditions on the volatility and drift of the forward rate SDE so that CMG and MRT can be used.

5.3 1-Factor HJM: The Model

5.3.1 The Forward Rate Process

Given an initial forward rate curve (also called an initial yield curve or an initial term structure) $f(0,T)$, the forward rate $f(t,T)$ for each maturity $0 < T < \tau$ is assumed to evolve under the market measure $\mathbb{Q}$ as

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T,\omega)\, dW(s) + \int_0^t \alpha(s,T,\omega)\, ds \tag{5.1}$$

or in diœerential form

$$d_t f(t,T) = \sigma(t,T,\omega)\, dW(t) + \alpha(t,T,\omega)\, dt \tag{5.2}$$

Thus the forward rate evolution is governed by a deterministic term and a single stochastic term. In the next chapter, we shall allow the stochastic term to be multidimensional. The volatilities and drifts are $\mathcal{F}(t)$-adapted processes in general. The argument ω is a state variable. A state variable is a quantity that gives us some information about a portion of the economy at that time. For example, we could take the state variable to be the forward rate itself.

In this 1-factor model, a single Brownian motion generates all the randomness of forward rate movements, and so the incremental changes of all forward rates and thus all yields and all bond prices are perfectly correlated. This obvious drawback is overcome by multifactor models, which will be discussed in later chapters.

From (5.1), the spot rate process is given by

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma(s,t,\omega)\,dW(s) + \int_0^t \alpha(s,t,\omega)\,ds \tag{5.3}$$

5.3.2 Technical Conditions - 1

The HJM framework imposes several technical constraints on the drift and volatility functions, to ensure that we can indeed perform the manipulations that we need to. These will be presented in two sets, to make for easy reading.

- σ and α are $\mathcal{F}$-previsible processes.
- $\int_0^T \sigma(t,T,\omega)^2\,dt$ and $\int_0^T |\alpha(t,T,\omega)|\,dt$ are ønite.
- The initial forward rate curve is deterministic and $\int_0^T |f(0,u)|\,du$ is ønite.
- $\int_0^T \int_0^u |\alpha(s,u,\omega)|\,ds du < \infty$.
- $\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left|\int_0^u \sigma(s,u,\omega)\,dW(s)\right| du\right] < \infty$.

The ørst three of these ensure that the forward rate given by (5.1) is well-deøned. The next two allow us to use a stochastic analogue of Fubini's theorem.

5.3.3 Bond Price SDE

We have a family of processes for the forward and spot rates of interest. Now we need a tradable quantity and a numeraire. As shown in the last chapter, choice of numeraire is arbitrary, but algebraic convenience points to a canonical numeraire, the money market account or cash bond starting with dollar investment and rolling over at the spot rate of interest. Thus our numeraire of choice is

$$B(t) = \exp\left\{\int_0^t r(s)ds\right\}$$

The technical conditions above ensure that this is a well-behaved process. For the tradable instrument, we choose the T-expiry bond $P(t,T)$.

Next, we are interested in the dynamics of the bond price process. If we deøne

$$a(t,T) = -\int_t^T \sigma(t,s,\omega)\,ds \text{ ; and} \tag{5.4}$$

$$b(t,T) = -\int_t^T \alpha(t,s,\omega)\,ds + \frac{1}{2}a(t,T)^2 \tag{5.5}$$

then the bond price SDE is

$$d_tP(t,T) = a(t,T)\,P(t,T)\,dW(t) + \{r(t) + b(t,T)\}\,P(t,T)\,dt \tag{5.6}$$

To see this, observe that

$$\begin{aligned}\log P(t,T) &= -\int_t^T f(t,s)\,ds = -\int_t^T f(0,s)\,ds - \int_t^T \int_0^t \sigma(u,s,\omega)\,dW(u)\,ds \\ &\quad - \int_t^T \int_0^t \alpha(u,s,\omega)\,duds\end{aligned}$$

Now use the stochastic analogue of Fubini's theorem to interchange the order of integration

$$\begin{aligned}\log P(t,T) &= -\int_0^T f(0,s)\,ds + \int_0^t f(0,s)\,ds \\ &\quad - \int_0^t \int_t^T \sigma(u,s,\omega)\,dsdW(u) - \int_0^t \int_t^T \alpha(u,s,\omega)\,dsdu\end{aligned}$$

Adding and subtracting the same terms from inside the double-integrals gives

$$\begin{aligned}\log P(t,T) &= -\int_0^T f(0,s)\,ds + \int_0^t f(0,s)\,ds \\ &\quad + \int_0^t \int_u^t \sigma(u,s,\omega)\,dsdW(u) + \int_0^t \int_u^t \alpha(u,s,\omega)\,dsdu \\ &\quad - \int_0^t \int_u^T \sigma(u,s,\omega)\,dsdW(u) - \int_0^t \int_u^T \alpha(u,s,\omega)\,dsdu\end{aligned}$$

In the third and fourth terms, we can integrate over the same triangle in two ways, and write (informally)

$$\int_0^t \int_u^t ...dsdu = \int_0^t \int_0^s ...duds$$

Thus we can write

$$\begin{aligned}
\log P(t,T) &= \log P(0,T) \\
&+ \int_0^t \left\{ f(0,s) + \int_0^s \sigma(u,s,\omega)\, dW(u) + \int_0^s \alpha(u,s,\omega)\, du \right\} ds \\
&- \int_0^t \int_u^T \sigma(u,s,\omega)\, dsdW(u) - \int_0^t \int_u^T \alpha(u,s,\omega)\, dsdu
\end{aligned}$$

Using the deønitions of $r(t)$, $a(t,T)$ and $b(t,T)$ from (5.3), (5.4) and (5.5) respectively gives

$$\begin{aligned}
\log P(t,T) &= \log P(0,T) + \int_0^t a(u,T)\, dW(u) \\
&- \frac{1}{2}\int_0^t a(u,T)^2\, du + \int_0^t \{r(u) + b(u,T)\}\, du
\end{aligned}$$

A simple application of Ito's formula now gives the bond price SDE as (5.6).

The discounted bond price is $Z(t,T) = B(t)^{-1} P(t,T)$. From the above discussion, it is easy to see that the SDE for this process is given by

$$d_t Z(t,T) = Z(t,T)\{a(t,T)\, dW(t) + b(t,T)\, dt\} \tag{5.7}$$

5.3.4 Technical Conditions - 2

Now we look for a martingale measure. In order to ensure that things go smoothly, we need to lay down a few more technical constraints.

- There exists an $\mathcal{F}$-previsible process $\gamma(t)$ such that

$$b(t,T) + a(t,T)\gamma(t) = 0$$

  for all $0 < t < T$.

- The process $a(t,T)$ is nonzero for almost all t and for every maturity T.
- $\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\exp\left(\int_0^T \gamma(t)^2\,dt\right)\right] < \infty$.
- $\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{2}\exp\left(\int_0^T \{\gamma(t) - a(t,T)\}^2\,dt\right)\right] < \infty$.

5.3.5 Change of Measure

We have an SDE for the discounted process Z in the market measure $\mathbb{Q}$ and in keeping with the strategy outlined in chapter 3, we now seek an equivalent measure $\mathbb{Q}^*$ such that Z is a $\mathbb{Q}^*$-martingale. Recall that the SDE for Z is

$$d_t Z(t,T) = Z(t,T)\{a(t,T)\,dW(t) + b(t,T)\,dt\}$$

For a given maturity date T, suppose there exists a function $\gamma_T(t)$ such that

$$b(t,T) + a(t,T)\gamma_T(t) = 0$$

If $a(t,T) \neq 0$ then we can write

$$-\gamma_T(t) = \frac{b(t,T)}{a(t,T)} \tag{5.8}$$

Now we can rewrite (5.7) as

$$\begin{aligned} d_t Z(t,T) &= a(t,T)\,Z(t,T)\left\{dW + \frac{b(t,T)}{a(t,T)}dt\right\} \\ &= a(t,T)\,Z(t,T)\{dW - \gamma_T(t)\,dt\} \end{aligned}$$

Applying Girsanov's theorem, we see the existence of a unique equivalent probability measure $\mathbb{Q}^*$ such that

$$W^*(t) = W(t) - \int_0^t \gamma_T(s)\, ds$$

is a $\mathbb{Q}^*$-Brownian motion. Under this new measure, the SDE for Z is

$$d_t Z(t,T) = a(t,T)\, Z(t,T)\, dW^*(t) \tag{5.9}$$

The conditions on $a(t,T)$ and the characterisation of exponential martingales (3.22) ensure that Z is a $\mathbb{Q}^*$-martingale.

Before moving on to the pricing of contingent claims, two points are worthy of note. The ørst is the restriction on the drift parameter of the forward rate process that we have implicitly imposed in order to ensure the absence of arbitrage, and the second is the introduction of a market price of risk through the change of measure transformation. The next two sections brieÆy discuss each of these in turn.

5.3.6 Drift Restriction

If the measure is unique, then since it is deøned through the ratio b/a, we must have, for all expiries S, T in the trading interval,

$$-\gamma_T(t) = \frac{b(t,T)}{a(t,T)} = \frac{b(t,S)}{a(t,S)} = -\gamma_S(t)$$

which means that $\gamma_T(t)$ is independent of T. Thus we may write it as simply $\gamma(t)$, satisfying the technical conditions above. From the ørst of these, we have

$$b(t,T) + a(t,T)\,\gamma(t) = 0$$

Substituting for each term using the deønitions gives

$$-\int_t^T \alpha(t,s,\omega)\, ds + \frac{1}{2} a(t,T)^2 - \gamma(t) \int_t^T \sigma(t,s,\omega)\, ds = 0$$

Partial dicerentiation with respect to T gives

$$-\alpha(t,T,\omega)+a(t,T)\{-\sigma(t,T,\omega)\}-\gamma(t)\sigma(t,T,\omega)=0$$

or

$$\alpha(t,T,\omega)=-\sigma(t,T,\omega)\left\{\gamma(t)-\int_t^T\sigma(t,s,\omega)\,ds\right\} \tag{5.10}$$

Equation (5.10) is the restriction on the drift parameter of the HJM evolution that is necessary for the absence of arbitrage. This is so because it is the condition needed to ensure uniqueness of the martingale measure $\mathbb{Q}^*$, and a unique EMM implies an arbitrage-free complete market, as discussed in the previous chapter.

5.3.7 Market Price of Risk

The transformation from market to risk-neutral probability measure was characterised by the function $\gamma(t)$. This section shows how $-\gamma(t)$ can be interpreted as the market price of risk. Consider the bond price SDE under the market measure

$$d_tP(t,T)=a(t,T)P(t,T)\,dW(t)+\{r(t)+b(t,T)\}P(t,T)\,dt$$

If we write $\gamma(t)=-b(t,T)/a(t,T)$ then we can write this as

$$\begin{aligned}d_tP(t,T)&=a(t,T)P(t,T)\{dW(t)-\gamma(t)\,dt\}+r(t)P(t,T)\,dt\\&=a(t,T)P(t,T)\,dW^*(t)+r(t)P(t,T)\,dt\end{aligned}$$

Thus the bond price evolves like a riskless bank account with random perturbations. The point to note is that the deterministic term contains only the spot rate of interest, and not r plus some ıriskyȷ term. Transforming the equation using γ thus makes the asset grow on average like a riskless bank deposit, or like the cash numeraire.

If we write $\mu(t,T)=r(t)+b(t,T)$ then the bond price SDE under the market measure can be written as

$$d_tP(t,T)=a(t,T)P(t,T)\,dW(t)+\mu(t,T)P(t,T)\,dt$$

The quantity $\{\mu(t,T) - r(t)\}/a(t,T)$ measures the market return over and above the risk-free rate, normalized by volatility. This is the market price of risk. But $\{\mu(t,T) - r(t)\}/a(t,T)$ is simply $\gamma(t)$ and so we label it the market price of risk.

This analysis holds for instruments other than bonds as well, and provides a neat alternate derivation of the Black-Scholes equation. Consider the pricing of stock options. Prior to 1973, the approach to option pricing was as follows. Consider the option $V = V(S,t)$ as above, and then set $\mathbb{E}_{\mathbb{Q}}[dV] = rVdt$. This gives us an equation similar to the Black-Scholes equation (4.10), but subtly dicerent. Explicitly, we get

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \mu S V_S - rV = 0$$

The immediate problem with this is that estimating the drift μ is diŒcult. In order to get rid of the drift term, we need to change measure, so that we can use rational expectations and take expectations with respect to the risk-neutral measure.

How do we reconcile the two? Suppose we consider the riskless measure $\mathbb{Q}^*$. Under this measure, the asset must grow on average at the riskless rate. This means that the drift of $S(t)$ must be r. Now we use techniques from Chapter 3 to write the stock price SDE (4.7) as follows

$$\begin{aligned} dS &= \sigma S\left(dW + \frac{\mu - r}{\sigma}dt\right) + rSdt \\ &= \sigma S dW^* + rSdt \end{aligned}$$

If we now use the equation $\mathbb{E}_{\mathbb{Q}^*}[dV] = rVdt$, then it is easy to see that we shall end up with the Black-Scholes equation.

In the market measure, the asset price SDE drifts at a rate μ, whereas in a riskless world, it drifts at the riskless rate r. The dicerence between the two drifts normalised by the volatility of the asset price gives us a measure of the discrepancy between the market measure and the risk-neutral measure. This quantity, $(\mu - r)/\sigma$ measures the excess return over the riskless rate that accrues when we are in the market measure. For this reason, it is called the market price of risk.

Black and Scholes' contribution was to equate the return on a hedged portfolio rather than just the option to the riskless return. By choosing Δ appropriately, they ensure that the portfolio instantaneously remains free of risk. They utilised the extra degree of freedom acœorded by the option delta to equate not just the expected return, but the actual return on their hedged portfolio to the risk-free return.

5.3.8 Pricing Contingent Claims Using HJM

This section will follow discussions in earlier chapters about contingent claim pricing under a martingale measure. Most of the theory has been covered in chapter 4. Consider a claim C at some time horizon T. Form the discounted claim $B(T)^{-1} C$ and work in the risk-neutral measure $\mathbb{Q}^*$. Form the process

$$H(t) = \mathbb{E}_{\mathbb{Q}} \left[B(T)^{-1} C \middle| \mathcal{F}(t) \right]$$

By construction, this is a $\mathbb{Q}^*$-martingale. Using the MRT, there exists a previsible process ϕ such that

$$dH(t) = \phi(t)\, d_t Z(t, T)$$

The strategy is as before, to hold a portfolio which at time t consists of

- $\phi(t)$ units of the T-bond $P(t, T)$; and
- $\psi(t) = H(t) - \phi(t) Z(t, T)$ units of the cash bond $B(t)$.

As before, we can show that this is a self-ønancing replicating portfolio, and so its value at any time $t < T$ is given by

$$V(t) = \mathbb{E}_{\mathbb{Q}^*} \left[C e^{-\int_t^T r(s) ds} \middle| \mathcal{F}(t) \right]$$

In particular if we have a bond that pays \$1 with certainty at expiry, then using this formula we can write the time-t price of this bond as

$$P(t, T) = \mathbb{E}_{\mathbb{Q}^*} \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}(t) \right]$$

under the risk-adjusted measure.

The spot rate process in $\mathbb{Q}^*$ is given by

$$r(t) = f(0,t) + \int_0^t \sigma(s,t,\omega)\, dW^*(s) + \int_0^t \sigma(s,t,\omega) \left\{ \int_s^t \sigma(s,u,\omega)\, du \right\} ds$$

This is where HJM contributes, since the expression for the spot rate of interest is independent of the market price of risk under the risk-neutral measure. In most of the established spot rate models, a speciøed market price of risk prevails even under the risk-neutral measure. This is a subtle point: we do not get rid of market price of risk by mere virtue of the fact that we are in a risk-neutral measure. HJM uses information about the market to actually factor out the market price of risk from the spot rate process and hence from the formula for bond price.

5.4 Endnotes

To review, this chapter has developed a 1-factor HJM model. Though it is a simpliøed version of the general framework, it does capture some characteristics of the n-factor version. Most notably, it shows how the spot rate process is independent of the market price of risk in an HJM setting under the risk-neutral measure.. This allows us to price bonds using no information other than what the market provides. The extension to several factors is conceptually straightforward, and is discussed in the next chapter.

Chapter 6

The General HJM Framework

Of course in these enlightened days no one believes a word of it.

6.1 Objectives

- Motivate the need to study a multifactor HJM model.
- Develop the general HJM model analogous to the 1-factor case.
- Price contingent claims using this general model.
- Comment on the strengths and weaknesses of the model.

6.2 Introduction

The aim of the HJM class of models is to provide a framework for pricing interest rate sensitive contingent claims, taking as given the price of zero-coupon bonds (ZCBs). This chapter studies the full-blown multifactor version of the HJM framework. Keeping in mind that the price of interest rate based derivatives is characterised by the term structure of interest rates, HJM takes as given an initial yield curve. This is also called an initial forward rate curve and is in 1-1 correspondence with a given set of ZCB prices. The model studies how this curve evolves over time. Thus HJM models the evolution of forward rates. As we shall see later, the only other input needed is a set of volatility functions, or a term structure of volatilities. Given these two parameters, HJM can be used to price all interest rate sensitive contingent claims.

The multifactor version of HJM has three powerful advantages. First, it imposes a stochastic structure directly on the evolution of forward rates. In this way it avoids modelling spot rates altogether. As we saw in chapter 2, we can deduce spot rates from forward rates but not vice versa, since there is loss of information in the transformation $f(t,T) \to r(t) = f(t,t)$.

Secondly, since it does not model spot rates, it does not require an ıinversion of term structureȷ. If we are given a spot rate model, we can express bond price as the solution to a PDE. We then infer the parameters of the spot rate process by comparing solutions of this equation with observed ZCB prices. This will be more fully discussed in the next chapter. By modelling forward rates directly, the HJM framework can incorporate any initial yield curve by construction and so does not require an inversion of term structure to øt the parameters.

Thirdly, it allows multiple independent stochastic factors to inÆuence the evolution of the yield curve. This feature has enabled HJM models to be used to price instruments that derive their value from imperfect correlation among the movements of diœerent parts of the yield curve. The presence of several stochastic factors implies that HJM models can be calibrated so that bond prices for a particular maturity may be strongly correlated with some factors and weakly with others. Single-factor models cannot deal with this level of complexity.

HJM is an example of a state-space model. The state variable in HJM is the forward rate. Spot rate models take the spot rate as this state variable. The process driving a state variable might be one or many dimensional. This is called the dimension of the state space. Frequently, the dimension of the state space is identiøed with the number of Brownian factors driving the process, though the former is generally larger than the latter. Single-factor models attempt to explain the dynamics of interest rates using just one explanatory variable. Using this approach, the greatest sophistication that can be achieved while still retaining computational tractability is to have time-dependent parameters. An alternative to increasing the explanatory power of a model is to increase the number of underlying state variables. This is what multifactor models seek to do.

One-factor models imply a perfect correlation between movements of diœerent parts of the yield curve. Multifactor models allow the yield curve to move in a space of greater complexity (i.e. in a space of higher dimension) and so can be used to price instruments that derive their value from imperfect correlations among diœerent parts of the yield curve.

Oœsetting these advantages are two crucial diŒculties. The ørst is that it is often diŒcult to disentangle the individual eœects of several stochastic factors. Some underlying state variables such as the rate of inÆation are not easily observed. The other drawback is computational. Increasing the number of factors necessitates the use of cumbersome numerical recipes. With increases in computing power, this second problem is now diminishing in magnitude for models that consider two or three factors. Larger models however are still stuck with this intractability.

The idea behind pricing remains the same. We observe the market probabilities, and then change probability measure to ønd an EMM for the discounted bond price. Thus we factor out the risk. Now that we are in a risk-neutral world, there exists only one price for the contingent claim. We now use rational expectations to price bonds and other interest rate derivatives.

The framework remains the same as before. To recap, we have a continuous trading economy with trading interval $[0, \tau]$ and a continuum of default-free bonds $P(t, T)$ trading in the economy, with expiries for each trading date $T \in [0, \tau]$ prior to expiry. Uncertainty in the economy is characterised by a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\mathcal{F}$ is the øltration generated by n independent Brownian motions

$$\{W_1(t), ..., W_n(t)\}$$

initialised at zero. The bonds $P(t, T)$ will play the role of tradables and the money market account $B(t)$ will be the numeraire.

6.3 The HJM Model

6.3.1 The Forward Rate Process

For a øxed but arbitrary $T \in [0, \tau]$, the forward rate $f(t, T)$ for a T-expiry contingent claim evolves according to the stochastic process

$$f(t, T) = f(0, T) + \sum_{i=1}^{n} \int_0^t \sigma_i(s, T, \omega)\, dW_i(s) + \int_0^t \alpha(s, T, \omega)\, ds \tag{6.1}$$

The total instantaneous square volatility of $f(t,T)$ is $\sum \sigma_i^2$ and the covariance of the increments of two forward rates $f(t,T)$ and $f(t,S)$ is

$$\Gamma(S,T) = \sum_{i=1}^{n} \sigma_i(t,T,\omega)\,\sigma_i(t,S,\omega)$$

In 1-factor models, and so increments of forward rates of all maturities are perfectly correlated. From (6.1), the forward rate SDE is

$$d_t f(t,T) = \sum_{i=1}^{n} \sigma_i(t,T,\omega)\,dW_i(t) + \alpha(t,T,\omega)\,dt \tag{6.2}$$

and the spot rate process is

$$r(t) = f(0,t) + \sum_{i=1}^{n} \int_0^t \sigma_i(s,t,\omega)\,dW_i(s) + \int_0^t \alpha(s,t,\omega)\,ds \tag{6.3}$$

6.3.2 Technical Conditions - 1

As in chapter 5, we need technical conditions to ensure that these are well-deøned. Also, we would like conditions that allow us to use a stochastic analogue of Fubini's theorem, and ensure that the numeraire is well-behaved. These conditions are stated below.

- For each i, σ_i and α are $\mathcal{F}$-previsible processes.
- $\int_0^T \sigma_i(t,T,\omega)^2\,dt$ is ønite for all i and so is $\int_0^T |\alpha(t,T,\omega)|\,dt$.
- The initial forward rate curve is deterministic and $\int_0^T |f(0,u)|\,du$ is ønite.
- $\int_0^T \int_0^u |\alpha(s,u,\omega)|\,dsdu < \infty$.
- $\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left|\int_0^u \sigma_i(s,u,\omega)\,dW(s)\right| du\right] < \infty$ for all i.

6.3.3 The Bond Price SDE

In a manner exactly following the analysis of chapter 5, it can be shown that the process for the bond price satisøes

$$\begin{aligned}\log P(t,T) &= \log P(0,T) + \sum_{i=1}^{n} \int_0^t a_i(s,T)\, dW_i(s) \\ &\quad - \frac{1}{2} \sum_{i=1}^{n} \int_0^t a_i(s,T)^2\, ds + \int_0^t \{r(s) + b(s,T)\}\, ds\end{aligned}$$

where

$$\begin{aligned}a_i(t,T) &= -\int_t^T \sigma_i(t,s,\omega)\, ds \text{ ; and} \\ b(t,T) &= -\int_t^T \alpha(t,s,\omega)\, ds + \frac{1}{2} \sum_{i=1}^{n} a_i(t,T)^2\end{aligned}$$

A straightforward application of Ito's lemma now yields the bond price SDE under the market measure as

$$d_t P(t,T) = \sum_{i=1}^{n} a_i(t,T)\, P(t,T)\, dW_i(t) + \{r(t) + b(t,T)\}\, P(t,T)\, dt \tag{6.4}$$

Thus the SDE for the discounted bond price process Z is

$$d_t Z(t,T) = Z(t,T) \left[\sum_{i=1}^{n} a_i(t,T)\, dW_i(t) + r(t)dt \right] \tag{6.5}$$

6.3.4 Market Price of Risk

We now seek an EMM for the discounted bond price process. Since we are working in a multifactor world, we need a multifactor version of the CMG. The ørst step is to ønd a transformation to the risk-neutral world. We then try and ensure that such a measure is unique, and this will be our EMM.

For a given set of expiry dates $0 < T_1 < ... < T_n < \tau$, we assume that there exist well-behaved solutions $\gamma_i(\cdot;T_1,...,T_n)$ to the system of simultaneous equations

$$\begin{bmatrix} b(t,T_1) \\ \vdots \\ b(t,T_n) \end{bmatrix} + \begin{bmatrix} a_1(t,T_1) & \cdots & a_n(t,T_1) \\ \vdots & & \vdots \\ a_1(t,T_n) & \cdots & a_n(t,T_n) \end{bmatrix} \begin{bmatrix} \gamma_1(t;T_1,..,T_n) \\ \vdots \\ \gamma_n(t;T_1,..,T_n) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6.6}$$

As before $-\gamma_i(t;T_1,..,T_n)$ can be interpreted as the market price of risk associated with the i-th random factor. We assume that the matrix $A(t) = [a_i(t,T_j)]_{i,j}$ is non-singular. This ensures that the solutions to the above system are unique.

6.3.5 Technical Conditions - 2

In this section, we list the technical conditions that ensure smooth changes of measure and applications of CMG and MRT.

First, we insist that $\int_0^{T_1} \gamma_i(s;T_1,..,T_n)^2\, ds$ is ønite for all i. Next, we assume

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\sum_{i=1}^{n}\int_0^{T_1}\gamma_i(s;T_1,..,T_n)\,dW_i(s) -\frac{1}{2}\sum_{i=1}^{n}\int_0^{T_1}\gamma_i(s;T_1,..,T_n)^2\,ds\right\}\right] = 1$$

and ønally

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{\sum_{i=1}^{n}\int_0^{T_1}(a_i(s,T_j)+\gamma_i(s;T_1,..,T_n))\,dW_i(s) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_1}(a_i(s,T_j)+\gamma_i(s;T_1,..,T_n))^2\,ds\right\}\right] = 1$$

6.3.6 Existence and Uniqueness of an EMM

We can now use the multivariate version of CMG. Given a set of expiries as before, there exists a probability measure $\mathbb{Q}^*_{T_1,..,T_n}$ equivalent to the market measure $\mathbb{Q}$ such that the discount bond prices

$$\{Z(t,T_1), ..., Z(t,T_n)\}$$

are martingales with respect to the measure $\mathbb{Q}^*_{T_1,..,T_n}$. CMG identiøes this measure through the Radon-Nikodym derivative

$$\begin{aligned} \frac{d\mathbb{Q}^*_{T_1,..,T_n}}{d\mathbb{Q}} &= \exp\left\{\sum_{i=1}^{n}\int_0^{T_1}\gamma_i(s;T_1,..,T_n)\,dW_i(s)\right. \\ &\left.-\frac{1}{2}\sum_{i=1}^{n}\int_0^{T_1}\gamma_i(s;T_1,..,T_n)^2\,ds\right\} \end{aligned} \tag{6.7}$$

Under this measure

$$W_i^{*T_1,...,T_n} = W_i(t) - \int_0^t \gamma_i(s;T_1,..,T_n)\,ds \tag{6.8}$$

are independent Brownian motions. Nonsingularity of the matrix $A(t)$ ensures that the measure is unique. We need to show now that this measure is independent of the choice of discounted bonds.

Suppose $\mathbb{Q}^*$, deøned though $\mathbb{Q}^* = \mathbb{Q}^*_{T_1,..,T_n}$ for any increasing set of maturities T_i in the trading interval, is the unique EMM for the discount bond prices $Z(t,T)$ where $T \in [0,\tau]$ and $t \in [0,T_1]$. Since the measure is deøned through $\gamma_i(t;T_1,..,T_n)$, it (the measure $\mathbb{Q}^*$) is independent of the T_i if and only if the market price of risk $-\gamma_i(t;T_1,..,T_n)$ is too. This means that for any other increasing set of n maturities S_i, we must have

$$\gamma_i(t;T_1,..,T_n) = \gamma_i(t;S_1,..,S_n) = \gamma_i(t) \tag{6.9}$$

Thus for each volatility factor, we have an associated market price of risk $\gamma_i(t)$ which is independent of the vector of bonds chosen to deøne it. From (6.6), we have

$$b(t,T) + \sum_{i=1}^{n} a_i(t,T)\gamma_i(t) = 0$$

Substituting for a_i and b and diœerentiating partially with respect to T gives the drift restriction

$$\alpha(t,T,\omega) = -\sum_{i=1}^{n} \sigma_i(t,T,\omega)\left\{\gamma_i(t) - \int_t^T \sigma_i(t,s,\omega)\,ds\right\} \tag{6.10}$$

Equation (??) is called the standard ønance condition. It ensures the absence of arbitrage, by ensuring that we get the same market price of risk no matter what vector of bonds we use to derive it. Thus (6.10) can be interpreted as the restriction on the drift analogous to (5.10) of an HJM evolution that is necessary to prevent arbitrage opportunities. It is also the restriction necessary to ensure a unique EMM.

6.3.7 Pricing Contingent Claims

We would now like to use a suitable version of the MRT to ønd a replicating strategy for a given claim and thus price contingent claims using familiar results. Since we have an n-factor model, we need n bonds and the numeraire in order to hedge. An advantage of HJM, as shown by our preceding discussion, is that these bonds may be chosen pretty much as we like.

Suppose we have a claim C expiring at some time T. We choose bonds with expiries after this date. Thus we have

$$0 < T < T_1 < ... < T_n < \tau$$

We already know that each discounted bond price process $Z(t,T_i)$ is a martingale under the common EMM $\mathbb{Q}^*$. Now we form the process

$$H(t) = \mathbb{E}_{\mathbb{Q}^*}\left[B(T)^{-1}C\middle|\mathcal{F}(t)\right]$$

which, by construction is a martingale under $\mathbb{Q}^*$. We use the multivariate version of MRT, discussed in chapter 3, to claim the existence of unique processes $\phi_i(t)$ such that

$$dH(t) = \sum_{i=1}^{n} \phi_i(t)\,dZ(t,T_i)$$

Consider a strategy $(\phi_1(t), ..., \phi_n(t), \psi(t))$, which involves holding, at time t,

- $\phi_i(t)$ units of the bond $P(t,T_i)$; and

- $\psi(t) = H(t) - \sum_{i=1}^{n} \phi_i(t) Z(t, T_i)$ units of the cash numeraire $B(t)$.

It is a routine exercise to check that this is a self-ønancing replicating portfolio for the claim, and thus the present value of the claim is

$$V(t) = \mathbb{E}_{\mathbb{Q}^*}\left[C e^{-\int_t^T r(s)ds} \middle| \mathcal{F}(t) \right]$$

In a risk-neutral setting, the spot rate process r is given by

$$\begin{aligned} r(t) &= f(0,t) + + \sum_{i=1}^{n} \int_0^t \sigma_i(s,t,\omega)\, dW_i^*(s) \\ &\quad + \int_0^t \sigma_i(s,u,\omega) \left\{ \int_s^t \sigma_i(s,u,\omega)\, du \right\} ds \end{aligned}$$

6.4 Endnotes

HJM has gained enormous popularity because it does for øxed-income derivatives what Black-Scholes does for option pricing, i.e. provide a term structure that is independent of the risk preferences of an economic agent. The market prices of risk drop out of the expression for the spot rate, and are replaced by an expression involving the volatilities across diœerent maturities of the forward rates, or a ıterm structure of volatilitiesȷ. Thus contingent claims can be priced independent of the market price of risk under the risk-adjusted measure.

The next chapter discusses the implications of this and provides a critique of the HJM framework, especially as regards where it stands among other models of the term structure of interest rates.

Chapter 7

Other Term Structure Models

The courageous Spaceman Spiœ, interplanetary explorer extraordinaire, lands on yet another bizarre planet.

7.1 Objectives

- Describe 1-factor spot rate models.
- Show how PDE formulations can be derived.
- Outline the main established spot rate models.
- Describe the multifactor spot rate framework.
- State the Feynman-Kac formulas for PDE formulation of multifactor models.
- Outline two examples of multifactor spot rate models.

7.2 Introduction

The term structure of interest rates has been modelled for over 20 years now. How does HJM compare with the models that have been in use for all this time? The crucial diœerence between the older models and HJM is that the older ones are all spot rate models. The one notable exception is the model of Ho & Lee (HL), which is formulated in discrete time and seeks to study the dynamics of the whole yield curve (i.e. is a forward rate model). HJM is a continuous-time, multivariate generalisation of HL. Thus, in order to see how HJM compares with older models, this chapter presents a brief review of spot rate modelling.

7.3 Spot and Forward Rates

How are spot and forward rates related? An interesting alternative viewpoint to that outlined in chapter 2 is as follows. Consider a bond $P(t,T)$. In terms of forward rates, we have

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \tag{7.1}$$

In terms of the spot rate, we have

$$P(t,T) = \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}\right] \tag{7.2}$$

where the subscript t is a shorthand for the expectation conditional on $\mathcal{F}(t)$. Expanding both these expressions in Taylor series gives

$$\begin{aligned} P(t,T) &= 1 - \int_t^T f(t,s)\,ds + \frac{1}{2}\left(\int_t^T f(t,s)\,ds\right)^2 - \frac{1}{6}\left(-\int_t^T f(t,s)\,ds\right)^3 + ... \\ &= 1 - \mathbb{E}_t\left[\int_t^T r(s)\,ds\right] + \frac{1}{2}\mathbb{E}_t\left[\left(\int_t^T r(s)\,ds\right)^2\right] - ... \end{aligned}$$

Thus, to first order at least, we have

$$\int_t^T f(t,s)\,ds = \mathbb{E}_t\left[\int_t^T r(s)\,ds\right] = \int_t^T \mathbb{E}_t[r(s)]\,ds$$

which gives the intuitively appealing result

$$f(t,T) = \mathbb{E}_t[r(T)]$$

This amounts to saying that the forward rate at time t for any future time T is just the expected value of the spot rate at T, conditional on information available at time t. This fits in nicely into a no-arbitrage argument. If the market expected future spot rates to be any different, then why would anyone enter into a forward contract at the rate $f(t,T)$? It is important to note here that the expectation is computed under the market measure. Note also that $\mathbb{E}_t[r(T)]$ gives us a recipe for getting a ıprocess from a claimȷ so that as t evolves, we get a forward rate evolution.

Since we want to see how HJM (a forward rate model) compares with spot rate models, the next few sections are devoted to a brief survey of spot rate modelling of the term structure.

7.4 One-Factor Spot Rate Models

7.4.1 Generalities

A one-factor spot rate model takes the spot rate itself as a 1-dimensional state variable, and speciøes an SDE for its evolution. In the market measure $\mathbb{Q}$ this takes the form

$$dr(t) = \rho(r(t),t)dW(t) + \mu(r(t),t)dt \tag{7.3}$$

Note that this covers all Markovian (or Gaussian) spot rate models. If $-\gamma(t)$ is the market price of risk, then it is easy to see that the corresponding SDE in the risk-adjusted measure $\mathbb{Q}^*$ is of the form

$$dr(t) = \rho(r(t),t)\,dW^*(t) + \nu(r(t),t)\,dt \tag{7.4}$$

where $\nu = \mu + \gamma\sigma$.

Considerable eœort has been expended by researchers in trying to make spot rate models capture the characteristics of observed spot rates. For example, the Vasicek model outlined later incorporates mean-reversion, while the Cox-Ingersoll-Ross (CIR) model additionally ensures that interest rates stay non-negative. However, none of these models has any economic basis and it has been shown that are not empirically sound. Moreover, by being Markovian, they have low predictive power. This is a serious setback, because even if they can be adjusted to match historical observations, there is no guarantee that they will predict well. Since bond prices depend on future distributions of spot rates, this is a problem. There is evidence to show that AR(1) processes, which are essentially of the form

$$y_t = \beta y_{t-1} + \varepsilon_t$$

where $\beta < 1$ and ε_t is some error term, are much better approximators to interest rates than Markov processes. Thus interest rates have path-dependence which Gaussian spot rate models of the type above fail to capture.

So why do we bother with these models? For all their drawbacks, these models allow bond prices to be expressed as the solutions of PDEs. Thus they can be made fairly complicated and will still give closed form bond price formulas. They are analytically tractable, giving yield curves that are aŒne in the spot rate (an aŒne function is a constant plus a linear function). In addition, they capture observed characteristics of interest rates such as mean reversion and positivity.

One of the main drawbacks of forward rate models is that they allow interest rates to go negative with positive probability. Also, it is diŒcult to get closed-form solutions for bond prices for realistic models. There seems to be a trade-oœ between how realistic the model is and how simple the bond price formulas derived from it are.

7.4.2 PDE Formulation

Suppose we have an instrument whose value depends on the spot rate and calendar time. As an example, consider a bond $P(r, t, T)$. Suppressing the dependence on T, we can use Ito's lemma and (7.3) to write

$$\begin{aligned} dP &= P_r dr + P_t dt + \frac{1}{2} P_{rr} dr^2 + ... \\ &= \rho P_r dW + \left(P_t + \frac{1}{2}\rho^2 P_{rr} + \mu P_r \right) dt \end{aligned}$$

so that

$$\mathbb{E}_{\mathbb{Q}}[dP] = \left(P_t + \frac{1}{2}\rho^2 P_{rr} + \mu P_r \right) dt$$

We might be tempted to equate this to $rPdt$ but we must be careful. r is the riskless rate of return. In the riskless measure, the drift is actually $\mu - \gamma$. So if we want to equate the return to the riskless return, we must use the risk-adjusted drift. In other words, we write

You are using the ”gather” environment in a style in which it is not defined.
“Bbb–E″˙–“Bbb–Q″ˆ–*″″“left[dP“right] =“left(P˙–t″+“frac–1″–2″“rho
ˆ–2″P˙–rr″+“left(“mu +“gamma “sigma “right) P˙–r″“right) dt=rPdt“text– ; or″
“nonumber ““
P˙–t″+“frac–1″–2″“rho ˆ–2″P˙–rr″+“left(“mu +“gamma “sigma “right) P˙–r″-rP=0
“label–1-spot BPE″

We can solve this under appropriate boundary conditions. By calibrating the calculated bond prices with the observed ZCB prices, it is possible to infer the parameters of the spot rate process. This is what inversion of term structure means.

The next few sections very brieÆy outline some of the established 1-factor spot rate models.

7.4.3 The Vasicek Model

In the risk-neutral measure, the speciøcation of the spot rate SDE is given as

$$dr = \sigma dW^* + (\theta - \alpha r)\, dt \tag{7.5}$$

This is historically one of the earliest models for the term structure of interest rates. It assumes a constant market price of risk. The diœusion process proposed by Vasicek is a mean-reverting Ornstein-Uhlenbeck process. The spot rate is deøned as the strong solution to the SDE (7.5), where σ, θ and α are strictly positive constants. It is well-known that the solution to (7.5) is a Markov process with continuous sample paths and Gaussian increments. It allows for negative interest rates, which is an undesirable feature. The counter-argument often oœered is that for appropriate values of the constants, the probability of negative spot rates is very small. Vasicek does not claim that this is the best model for actual spot rate behaviour.

Without going into details, the Vasicek SDE permits representations of bond prices in the form $e^{-A(t,T)-rB(t,T)}$, thus giving an aŒne yield curve. It can also be shown that if we set up a Vasicek model for expiries far into the future, the mean-reversion becomes severely diminished, so that the model is essentially an unrestricted diœusion.

7.4.4 The Cox-Ingersoll-Ross Model

The CIR model studies the evolution of a single state variable, the spot rate, in a general equilibrium framework. Heath, Jarrow and Morton show how this can be equivalently formulated in a no-arbitrage setting as well. The spot rate process proposed in the CIR model is

$$dr = \sigma\sqrt{r}dW^* + (\theta - \alpha r)\,dt$$

where α, θ and σ are strictly positive constants. The market price of risk is proportional to $\sqrt{r/\sigma}$. Due to the presence of the square root term, the CIR spot rate takes only non-negative values. It can go to zero, but it never dips below zero. In fact it can be shown that for $2\theta \geq \sigma^2$, spot rates stay strictly positive. In this way, it rules out negative interest rates. Like the Vasicek model, it also has a mean-reverting drift term. There appears to be modest empirical support for the CIR model.

This too, like the Vasicek model, belongs to the class of aŒne yield curves, but closed form solutions are not necessarily ıprettyȷ.

7.4.5 Longstaœ's Model

Longstaœ modiøed the CIR model by postulating the following dynamics for the spot rate process

$$dr = \sigma\sqrt{r}dW^* + \lambda\left(\theta - \alpha\sqrt{r}\right)dt$$

which is also referred to as a double square-root (DSR) process. Closed-form solutions to the bond price equation exist, but in this case they are of the form

$$P(r,t,T) = e^{-A(t,T)-rB(t,T)-\sqrt{r}C(t,T)}$$

for explicitly known functions A, B and C.

The yield curve is thus a non-linear function of the spot rate. Also, the bond price is not a monotonically decreasing function of the spot rate. This makes valuation of bond options more complicated. An empirical comparison of the Longstaœ model, done by Longstaœ in 1989, suggests that it outperforms CIR in most circumstances.

7.4.6 The Hull-White Model

Both the Vasicek and the CIR models are special cases of the following mean-reverting diœusion process

$$dr = \sigma r^{\beta} dW^{*} + (\theta - \alpha r)\, dt$$

where $0 < \beta < 1$ is a constant. The Hull-White model is a generalisation of this process, where $\beta \geq 0$ is a constant and all other parameters are time-dependent.

$$dr = \sigma(t) r^{\beta} dW^{*} + (\theta(t) - \alpha(t) r)\, dt$$

If $\beta = 0$, we get the generalised Vasicek model and if $\beta = 0.5$ we get the generalised CIR model. The most important feature of the general Hull-White approach is the possibility of an exact øt to a given initial term structure and in some cases to a term structure of volatilities.

7.4.7 The Black-Derman-Toy Model

The BDT model derives from a discrete-time model of the term structure. At each node, it involves choosing spot rates and the corresponding transition probabilities so that the model matches not only the initial term structure, but also the initial volatility structure of forward rates. In continuous time, BDT is a special case of the more general lognormal model

$$d \log r = \sigma(t) dW^{*} + (\theta(t) - \alpha(t) \log r)\, dt \tag{7.6}$$

for deterministic functions $\sigma(t)$, $\theta(t)$ and $\alpha(t)$. The BDT model corresponds to the speciøcation $\alpha(t) = -\sigma'(t)/\sigma(t)$. The lognormal model speciøed in (7.6) has also been studied by Black & Karasinski, who postulate that it øts the yield curve, the volatility curve and the cap curve. Hogan & Weintraub showed that the dynamics of (7.6) lead to inønite prices for Eurodollar futures. To overcome this, Sandmann & Sondermann proposed a focus on eœective annual rates rather than simple rates over shorter periods, but this will not be covered here.

7.5 Multifactor Spot Rate Models

7.5.1 Generalities

All the one-factor models surveyed in the previous section took a state variable, the spot rate, and modelled its evolution over time. They attempted to explain the dynamics of interest rates using just one explanatory variable. Using this approach, the greatest sophistication that can be achieved while still retaining computational tractability is to have time-dependent parameters. An alternative to increasing the explanatory power of a model is to increase the number of underlying state variables. This is what multifactor models seek to do. They assume that the spot rate itself is a function of several underlying state variables.

From the theoretical point of view, a multifactor process is based on the speciøcation of a multidimensional Markov process $X = (X_1, ..., X_n)^{\mathrm{T}}$ which is deøned as the strong solution to the vector SDE

$$dX(t) = \sigma(X(t)) \cdot dW^*(t) + \mu(X(t))dt \tag{7.7}$$

where $X(t)$, $W(t)$ and $\mu(.)$ take values in $\mathbb{R}^n$, and $\sigma(.)$ takes values in $\mathbb{R}^{n\times n}$, the set of all $n \times n$ matrices with real entries. As is common in existing literature, state variables here are identiøed with stochastic factors, so that the dimensionality of the Brownian motion is equal to the number of underlying state variables. Generally, the latter are larger than the former. Given the SDE (7.7) the spot rate is deøned by a function

$$r(t) = g(X(t))$$

for some deterministic function $g : \mathbb{R}^n \to \mathbb{R}$.

As a special case, we could postulate that $X(t)$ satisøed the following linear SDE with bounded, time-dependent coeŒcients

$$dX(t) = \sigma(t) \cdot dW^*(t) + (a(t) + b(t)X(t))dt$$

It is common to then set either

$$r(t) = \frac{1}{2}|X(t)|^2 : \text{a squared Gauss-Markov (SGM) process, or}$$

$$r(t) \quad = \quad \lambda \cdot X(t) \text{ for some } \lambda \in \mathbb{R}^n\text{: a Gaussian process.}$$

From a practical viewpoint, it is important to identify multifactor models which lead to yields that are aŒne in the spot rate. DuŒe and Kan show that a multifactor model has this property if and only if the coeŒcients of the process for $X(t)$ and the function g are aŒne in some suitable sense.

7.5.2 The Feynman-Kac Formula

Given a setting described in the previous section, it is frequently convenient to exploit the Feynman-Kac formula to ønd a PDE representation for bond prices. In a risk-neutral world, the ZCB price is determined by the formula

$$P(X,t,T) = \mathbb{E}_{\mathbb{Q}^*}\left[\exp\left(-\int_t^T g(X(s))ds\right)\middle| X(t)\right]$$

The function P deøned above satisøes the PDE

$$P_t + \frac{1}{2}\text{Trace}\left[P_{xx}\sigma\sigma^{\mathrm{T}}\right] + P_x\mu + gP = 0 \tag{7.8}$$

with boundary data $P(x,T,T) = 1$. The 1-factor version of this result is precisely the bond pricing equation (??). If we seek a term structure that is aŒne in the spot rate, we look for solutions of the form

$$P = e^{-A(t,T)-B(t,T)\cdot x}$$

to (7.8). DuŒe and Kan consider a special case of this solution, i.e. one of the form

$$P(x,t,T) = \exp\left(u(T-t) + v(T-t)\cdot x\right)$$

where $u(.) = (u_1(.),...,u_n(.))^{\mathrm{T}}$ and $v(.)$ are suitably well-behaved and satisfy the boundary conditions $v(0) = 0$ and $u_i(0) = 0$ for $i = 1,...,n$. They show that the functions u and v must satisfy the Riccati-type diœerential equations

$$u_i''(s) \quad = \quad c_i + k_i\cdot u(s) + u(s)^{\mathrm{T}} q_i u(s), \quad i = 1,...,n$$

$$v'(s) \quad = \quad c_0 + k_0 \cdot u(s) + u(s)^{\mathrm{T}} q_0 u(s)$$

where $c_i \in \mathbb{R}$, $k_i \in \mathbb{R}$ and $q_i \in \mathbb{R}^n$ are constants given in terms of the coeŒcients deøning the aŒne functions μ, $\sigma\sigma^{\mathrm{T}}$ and g. Indeed, DuŒe and Kan show that the term structure is aŒne in the spot rate if and only if these functions are.

7.5.3 The Generalised Multifactor CIR Model

A simple example of an aŒne model that is not Gaussian is the multifactor CIR model, where the Markov process corresponding to (7.7) is

$$dX_i(t) = \sigma_i \sqrt{X_i(t)} dW^* + (\theta_i - \alpha_i X_i(t))\, dt \tag{7.9}$$

for positive constants σ_i, θ_i and α_i, with the spot rate given by $r(t) = X_1(t) + ... + X_n(t)$. In terms of the function g discussed above, we have $g(x) = \sum x_i$. This model has been extended by Heston (1991), Longstaœ & Schwartz (1992), Chen & Scott (1992), Pearson & Sun (1994) and Jamshidian (1995). Restrictions apply: for interest rates to remain positive, we must have $\theta_i > \alpha_i^2/2$ as shown by Ikeda and Wantanabe (1981). Chen (1994) studies another aŒne model, this time with a three-dimensional state space.

7.5.4 DuŒe & Kan's General AŒne Model

DuŒe & Kan (1996) study the general aŒne case, where after a linear change of variable the SDE for the underlying Markov process is

$$dX_i(t) = \sigma(t) \cdot dW^* + (\alpha X(t) + \theta)\, dt$$

with

$$\sigma_{ij}(t) = \lambda_{ij} \sqrt{m_{ij} + p_{ij} \cdot X(t)}$$

where λ_{ij}, m_{ij}, p_{ij}, $\theta \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}^{n \times n}$. In this case, the state space is

$$\{x \in \mathbb{R}^n : m_{ij} + p_{ij} \cdot X(t) > 0,\ i, j = 1, ..., n\}$$

Restrictions analogous to those for the CIR model apply, but are more complicated.

7.6 Endnotes

This chapter has outlined most of the established spot rate models. By assuming that the spot rate process lives in a finite-dimensional state space, these models lead to bond prices that are solutions of PDEs. Spot rate models are versatile enough to allow for closed form bond price formulae even while capturing fairly sophisticated and realistic qualitative features of the dynamics of interest rates. The main drawback is that they all assume a specific functional form for the market price of risk, even under the risk-neutral measure. In addition, they generally match empirical observations poorly. Since they are Gaussian, they have low predictive power too. Apart from this, they have to be calibrated to fit the yield curve, and this inversion of term structure severely restricts the generality of spot rate models.

Chapter 8

Assessing the HJM Framework

Why does Man create? Is it Man's purpose on earth to express himself, to bring form to thought, and to discover meaning in experience?

Or is it just something to do when he's bored?

8.1 Objectives

- Review the ideas of HJM.
- Show how all 1-factor spot rate models are special cases of 1-factor HJM models.
- Illustrate this with an example.
- Show how spot rates are path-dependent under HJM.
- Study a case when they are not path-dependent.
- Discuss PDE formulations under HJM.
- Survey the state of aœairs in HJM-type modelling.

8.2 Introduction

Chapters 5 and 6 treated the HJM model in some detail. This chapter aims to provide a critique of the HJM framework. First, let us step back and review the broad ideas in the model.

HJM is a state space model. It takes the forward rate as a state variable and speciøes a process for its evolution over time. The inputs to the model are an initial term structure of interest rates $\{f(0,T)\}_T$ and a term structure of volatilities. Using these quantities, HJM arrives at a spot rate process that is independent of the market price of risk. Once this process is known, it uses rational expectation to ønd bond prices. While deriving this formulation, the framework also imposes restrictions on the drift of the forward rate process that ensure the absence of arbitrage opportunities.

8.3 HJM and Spot Rate Models

8.3.1 General Considerations

All the 1-factor models discussed in the previous chapter are special cases of appropriately chosen HJM models. To see this, consider a spot rate model in a risk-adjusted measure,

$$dr(t) = \rho(r(t), t)\, dW^*(t) + \nu(r(t), t)\, dt$$

Bond prices are given by

$$P(t,T) = e^{-\int_t^T f(t,s)ds} = \mathbb{E}_{\mathbb{Q}^*}\left[\left. e^{-\int_t^T r(s)ds} \right| \mathcal{F}(t) \right]$$

Thus

$$\begin{aligned} \int_t^T f(t,s)\, ds &= -\log \mathbb{E}_{\mathbb{Q}^*}\left[\left. e^{-\int_t^T r(s)ds} \right| \mathcal{F}(t) \right] \\ &= g(r(t), t, T) \end{aligned}$$

where

$$g(x,t,T) = -\log \mathbb{E}_{\mathbb{Q}^*}\left[\left. e^{-\int_t^T r(s)ds} \right| r(t) = x \right] \tag{8.1}$$

Thus, given a spot rate model, the implied forward rate is given by

$$f_1(t,T) = \frac{\partial}{\partial T} g(r(t), t, T)$$

From this, we can use Ito's lemma to deduce df_1. We get

$$df_1(t,T) = \rho \frac{\partial^2 g}{\partial x \partial T} dW^* + \left\{ \nu \frac{\partial^2 g}{\partial x \partial T} + \frac{\partial^2 g}{\partial t \partial T} + \frac{1}{2}\rho^2 \frac{\partial^3 g}{\partial x^2 \partial T} \right\} dt$$

Comparing this with a risk-adjusted HJM evolution

$$df(t,T) = \sigma(t,T,\omega)\, dW^*(t) + \sigma(t,T,\omega) \left(\int_t^T \sigma(t,s,\omega)\, ds \right) dt$$

we can equate the volatility coefficients to get

$$\sigma(t,T,\omega) = \rho(r(t),t) \frac{\partial^2 g}{\partial x \partial T}(r(t),t,T);\ \text{and} \tag{8.2}$$

$$\int_t^T \sigma(t,s,\omega)\, ds = \rho(r(t),t) \frac{\partial g}{\partial x}(r(t),t,T) \tag{8.3}$$

Thus we can find an HJM model corresponding to all Markovian spot rate models.

Suppose we don't know for sure if a model is Markovian. It can be shown that such a model is a special case of HJM as well. Given an HJM evolution, the discounted bond price SDE is

$$d_t Z(t,T) = a(t,T)\, Z(t,T)\, dW^*(t)$$

Using Ito's lemma, it is easy to check that the solution of this SDE is

$$Z(t,T) = Z(0,T) \exp\left\{ \int_0^t a(s,T)\, dW^*(s) - \frac{1}{2}\int_0^t a(s,T)^2\, ds \right\}$$

By definition, we also have

$$\begin{aligned} Z(t,T) &= B(t)^{-1} P(t,T) = P(t,T)\, e^{-\int_0^t r(s)ds} \\ &= e^{-\int_0^t r(s)ds - g(r(t),t,T)} \end{aligned}$$

Comparing these two expressions for Z gives

$$\begin{aligned} g(r(t),t,T) &= -\int_0^t a(s,T)\, dW^*(s) - \int_0^t \left\{ r(s) - \frac{1}{2} a(s,T)^2 \right\} ds \\ &\quad - \log P(0,T) \end{aligned}$$

Using the deønition of $a(t,T)$, this can be rewritten as

$$g(r(t),t,T) = \int_0^t \int_s^T \sigma(s,u,\omega)\,du\,dW^*(s) - \int_0^t \left\{ r(s) - \frac{1}{2}\left(\int_s^T \sigma(s,u,\omega)\,du\right)^2 \right\} ds$$
$$- \log P(0,T)$$

Since the implied forward rate f_1 is the partial derivative of g with respect to T, we get

$$f_1(t,T) = f(0,T) + \int_0^t \sigma(s,T,\omega)\,dW^*(s) + \int_0^t \sigma(s,T,\omega)\left\{\int_t^T \sigma(s,u,\omega)\,du\right\} ds$$

which is precisely the risk-neutral HJM evolution. Thus all spot rate models are special cases of HJM. The next section illustrates this with an example.

8.3.2 Vasicek in Terms of HJM

Under the risk-neutral measure, the Vasicek SDE is

$$dr(s) = \sigma dW^*(s) + (\theta - \alpha r(s))\,ds$$

for $s \in [t,T]$ and σ, θ and α constants. If we are given $r(t) = x$ then

$$dr + \alpha r ds = \sigma dW^* + \theta ds$$

If we multiply throughout by the integrating factor $e^{\alpha s}$ then we can write this as

$$d(e^{\alpha s} r) = \sigma e^{\alpha s} dW^* + \theta e^{\alpha s} ds$$

Integrating both sides from t to s and using $r(t) = x$ gives

$$re^{\alpha s} - xe^{\alpha t} = \sigma \int_t^s e^{\alpha u} dW^*(u) + \frac{\theta}{\alpha}\left(e^{\alpha s} - e^{\alpha t}\right)$$

This can be rearranged to write the solution as

$$r(s) = \frac{\theta}{\alpha} + e^{-\alpha(s-t)}\left(x - \frac{\theta}{\alpha}\right) + \sigma e^{-\alpha s} \int_t^s e^{\alpha u} dW^*(u) \qquad (8.4)$$

To get the HJM model corresponding to this spot rate model, we need to ønd the correct function g as in (8.1). First, we simplify (8.4). Consider the third term on the right hand side. Using (3.11), and writing $J = \int_t^s e^{\alpha u} dW^*(u)$, we have

$$\mathbb{E}_{\mathbb{Q}^*}\left[J^2\right] = \int_s^t e^{2\alpha u} du = \frac{1}{2\alpha}\left(e^{2\alpha s} - e^{2\alpha t}\right)$$

Thus

$$J = W^*\left(\frac{e^{2\alpha s} - e^{2\alpha t}}{2\alpha}\right)$$

so that the third term in (8.4) can be rewritten as

$$\sigma W^*\left(\frac{1 - e^{-2\alpha(s-t)}}{2\alpha}\right)$$

To ønd g, we need to ønd $\int_t^T r(s)\,ds$.

$$\begin{aligned}\int_t^T r(s)\,ds &= \frac{\theta}{\alpha}(T-t) - \left(x - \frac{\theta}{\alpha}\right)\frac{1}{\alpha}\left(e^{-\alpha(T-t)} - 1\right) \\ &\quad + \sigma\int_t^T W^*\left(\frac{1 - e^{-2\alpha(s-t)}}{2\alpha}\right) ds\end{aligned}$$

Substitution gives

$$\begin{aligned}g(x,t,T) &= \frac{\theta}{\alpha}(T-t) - \left(x - \frac{\theta}{\alpha}\right)\frac{1}{\alpha}\left(e^{-\alpha(T-t)} - 1\right) \\ &\quad - \log \mathbb{E}_{\mathbb{Q}^*}\left[\exp\left\{-\sigma\int_t^T W^*\left(\frac{1 - e^{-2\alpha(s-t)}}{2\alpha}\right) ds\right\}\right]\end{aligned}$$

The ørst and third terms are independent of x and so

$$\begin{aligned}\frac{\partial g}{\partial x} &= -\frac{1}{\alpha}\left(e^{-\alpha(T-t)} - 1\right) \\ \frac{\partial^2 g}{\partial x^2} &= e^{-\alpha(T-t)}\end{aligned}$$

Now, using (8.2) and (8.3),we can get the corresponding HJM parameters

$$\begin{aligned}\sigma_{\text{HJM}}(t,T,\omega) &= \sigma e^{-\alpha(T-t)} \\ \alpha_{\text{HJM}}(t,T,\omega) &= \frac{\sigma^2}{\alpha}e^{-\alpha(T-t)}\left(e^{-\alpha(T-t)}-1\right)\end{aligned}$$

8.4 Path-Dependence of HJM Spot Rates

The general HJM framework gives forward and spot rates that are strongly path dependent. This section illustrates path-dependence of spot rates under a general 1-factor HJM model. Consider the evolution of the spot rate under HJM, given by the equation

$$r(t) = f(0,t) + \int_0^t \sigma(s,t,\omega)\, dW^*(s) + \int_0^t \alpha(s,t,\omega)\, ds$$

Let us temporarily suppress the dependence on ω. If we work in a discrete-time framework starting at time 0, then

$$\begin{aligned} r(1) &= f(0,1) + \sigma(0,1)\,\delta W^*(0) + \alpha(0,1)\,\delta t \\ r(2) &= f(0,2) + \sigma(0,2)\,\delta W^*(0) + \sigma(1,2)\,\delta W^*(1) \\ &\quad + \{\alpha(0,2) + \alpha(1,2)\}\,\delta t \end{aligned}$$

Thus the increment in r is given by $\delta r(1) = r(2) - r(1)$, which is

$$\begin{aligned} \delta r(1) &= f(0,2) - f(0,1) + \{\sigma(0,2) - \sigma(0,1)\}\,\delta W^*(0) + \sigma(1,2)\,\delta W^*(1) \\ &\quad + \{\alpha(0,2) + \alpha(1,2) - \alpha(0,1)\}\,\delta t \end{aligned}$$

If we assume that the drift is time-homogeneous (i.e. the volatility is time-homogeneous too) then $\alpha(1,2) = \alpha(0,1)$ and so

$$\delta r(1) = f(0,2) - f(0,1) + \{\sigma(0,2) - \sigma(0,1)\}\,\delta W^*(0) + \sigma(1,2)\,\delta W^*(1) \tag{8.5}$$

$$+\alpha(0,2)\,\delta t \tag{8.6}$$

In general, this is not the case and we will need the entire evolution of the parameters, i.e. we will not be able to discard $\alpha(1,2)$ and $\alpha(0,1)$ above. However, let us stay with the simpler time-homogeneous case for the present.

The forward rate increment over the first time period is

$$\delta f(1,2) = \sigma(1,2)\,\delta W^*(1) + \alpha(1,2)\,\delta t$$

and therefore

$$\begin{aligned} \delta r(1) - \delta f(1,2) &= f(0,2) - f(0,1) + \{\sigma(0,2) - \sigma(0,1)\}\,\delta W^*(0) \\ &\quad + \{\alpha(0,2) - \alpha(0,1)\}\,\delta t \end{aligned}$$

Thus the difference in drifts between the forward rate and the spot rate at time 1 stem from three factors

- A deterministic term linked to the shape of the yield curve at time zero, i.e. the term $f(0,2) - f(0,1)$;
- The difference in the drifts, at time 0, of the forward rates maturing at times 1 and 2 respectively, i.e. the coefficient of δtin the equation above; and
- A term depending on past Brownian increments, $\{\sigma(0,2) - \sigma(0,1)\}\,\delta W^*(0)$.

Let us carry forward the analysis leading to $\delta r(1)$. We can write

$$\begin{aligned} r(3) &= f(0,3) + \sigma(0,3)\,\delta W^*(0) + \sigma(1,3)\,\delta W^*(1) + \sigma(2,3)\,\delta W^*(2) \\ &\quad + \{\alpha(0,3) + \alpha(1,3) + \alpha(2,3)\}\,\delta t \end{aligned}$$

From this we get, on multiplying and dividing by δt,

$$\begin{aligned} \delta r(3) &= r(3) - r(2) \\ &= \left\{ \frac{f(0,3) - f(0,2)}{\delta t} \right. \\ &\quad + \frac{\sigma(0,3) - \sigma(0,2)}{\delta t}\delta W^*(0) + \frac{\sigma(1,3) - \sigma(1,2)}{\delta t}\delta W^*(1) \end{aligned}$$

$$+\left(\frac{\alpha(0,3)-\alpha(0,2)}{\delta t}+\frac{\alpha(1,3)-\alpha(1,2)}{\delta t}\right)\delta t+\alpha(2,3)\Bigg\}\delta t$$
$$+\sigma(2,3)\,\delta W^{*}(2)$$

Now we are in a position to generalise this to an arbitrary time step as follows:

$$\begin{aligned}
\delta r(n) &= r(n+1)-r(n)\\
&= \left\{\frac{f(0,n+1)-f(0,n)}{\delta t}+\sum_{i=0}^{n-1}\frac{\sigma(i,n+1)-\sigma(i,n)}{\delta t}\delta W^{*}(i)\right.\\
&\quad\left.+\sum_{i=0}^{n-1}\frac{\alpha(i,n+1)-\alpha(i,n)}{\delta t}\delta t+\alpha(n,n+1)\right\}\delta t\\
&\quad+\sigma(n,n+1)\,\delta W^{*}(n)
\end{aligned}\tag{8.7}$$

The drift term itself has a term which depends on realisations of the Brownian motion, and so (8.7) clearly shows that the process for the spot rate is in general path-dependent. Notice that path-dependence is exhibited even in the special time-homogeneous case. Thus the general framework is very strongly path-dependent.

It is now an easy matter to move to the continuous case, by allowing $\delta t \to 0$. This gives

$$dr(t) = \left\{\frac{\partial f(0,t)}{\partial t}+\int_0^t\frac{\partial\sigma(s,t)}{\partial t}dW^{*}(s)+\int_0^t\frac{\partial\alpha(s,t)}{\partial t}ds\right\}dt \tag{8.8}$$
$$+\sigma(t,t)\,dW^{*}(t) \tag{8.9}$$

Notice that for the HJM approach, we have because of the relation between drift and volatility imposed by the model. The above equation clearly shows that the spot rate process is non-Markovian in general.

8.5 Markovian Spot Rates Under HJM

As the previous section demonstrated, the HJM approach generally produces models in which the spot rate is path dependent. Therefore, a discrete-time approximation within the HJM framework is usually less eŒcient than the path-independent case, since the number of operations rises exponentially with the number of steps. The aim of this section is to ønd conditions under which spot rates are Markovian in an HJM setting.

For simplicity, consider a 1-factor HJM setting. The extension to several factors is straightforward. It will be shown now that spot rates given by an HJM evolution are Markovian if and only if there exist functions p and q such that the HJM volatility $\sigma(t,T,\omega)$ can be represented (suppressing dependence on ω) as

$$\sigma(t,T) = p(t)\,q(T) \tag{8.10}$$

Writing the spot rate process under HJM,

$$r(t) = f(0,t) + \int_0^t \sigma(s,t)\,dW^*(s) + \int_0^t \alpha(s,t)\,ds$$

it is obvious that spot rates are Markovian if and only if

$$D_t = \int_0^t \sigma(s,t)\,dW^*(s)$$

is Markovian, i.e.

$$\mathbb{E}_{\mathbb{Q}^*}\left[h(D_T)\middle|\mathcal{F}(D_t)\right] = \mathbb{E}_{\mathbb{Q}^*}\left[h(D_T)\middle|D_t\right] \tag{8.11}$$

for any bounded measurable function $h:\mathbb{R}\to\mathbb{R}$ and any $T\in[t,\tau]$. Note that

$$\begin{aligned} D_T &= \int_0^T \sigma(s,T)\,dW^*(s) \\ &= D_t + \int_0^T \sigma(s,T)\,dW^*(s) - \int_0^t \sigma(s,t)\,dW^*(s) \\ &= D_t + \int_t^T \sigma(s,T)\,dW^*(s) + \left(\int_0^t \sigma(s,T)\,dW^*(s) - \int_0^t \sigma(s,t)\,dW^*(s)\right) \end{aligned}$$

Now observe that (8.11) holds if and only if the above equation only depends on Brownian increments between t and T. Thus the term in parentheses must be zero. This amounts to saying that $J(t,T) = \int_0^t \sigma(s,T)\,dW^*(s)$ is completely determined by D_t. Treating both these quantities as random variables, this in turn is equivalent to having them perfectly correlated. Both have mean zero, so from the deønition of correlation, perfect correlation holds if and only if

$$\mathbb{E}_{\mathbb{Q}^*}\left[J\left(t,T\right)D_t\right]^2 = \mathbb{E}_{\mathbb{Q}^*}\left[J\left(t,T\right)^2\right]\mathbb{E}_{\mathbb{Q}^*}\left[D_t^2\right]$$

Using the properties of Ito integrals derived in chapter 3, this can be written as

$$\left(\int_0^t \sigma\left(s,t\right)\sigma\left(s,T\right)ds\right)^2 = \left(\int_0^t \left|\sigma\left(s,t\right)\right|^2 ds\right)\left(\int_0^t \left|\sigma\left(s,T\right)\right|^2 ds\right)$$

Written in terms of the standard inner product on $\mathcal{L}^2\left(0,t\right)$, this is just saying

$$\left\langle \sigma\left(s,t\right), \sigma\left(s,T\right)\right\rangle = \left\|\sigma\left(s,t\right)\right\| \left\|\sigma\left(s,T\right)\right\|$$

or that the two functions are collinear (i.e. each lies in the linear span of the other). Thus, the Markov property for spot rates (8.11) holds if and only if, for any $T \in [0,\tau]$ and $t \in [0,T]$ there exists a function $q\left(T\right)$ such that

$$\sigma\left(t,T\right) = q\left(T\right)\sigma\left(t,\tau\right)$$

If we now set $\sigma\left(t,\tau\right) = p\left(t\right)$, (8.10) follows immediately.

If spot rates are Markovian, then for any three time points $t < t^* < T$ we have, by deønition,

$$\begin{aligned}
a\left(t,T\right) &= -\int_t^T \sigma\left(t,s\right)ds = -p\left(t\right)\int_t^T q\left(s\right)ds \\
&= \frac{p\left(t\right)q\left(t^*\right)}{p\left(t^*\right)q\left(t^*\right)}\left[-p\left(t^*\right)\int_{t^*}^T q\left(s\right)ds - p\left(t^*\right)\int_t^{t^*} q\left(s\right)ds\right] \\
&= \frac{\sigma\left(t,t^*\right)}{\sigma\left(t^*,t^*\right)}\left[-\int_{t^*}^T \sigma\left(t^*,s\right)ds + p\left(t^*\right)\int_{t^*}^t q\left(s\right)ds\right] \\
&= \frac{\sigma\left(t,t^*\right)}{\sigma\left(t^*,t^*\right)}\left[-\int_{t^*}^T \sigma\left(t^*,s\right)ds + \int_{t^*}^t \sigma\left(t^*,s\right)ds\right] \\
&= \frac{\sigma\left(t,t^*\right)}{\sigma\left(t^*,t^*\right)}\left(a\left(t^*,T\right) - a\left(t^*,t\right)\right)
\end{aligned}$$

If in addition the volatility $\sigma\left(t,T\right)$ is time-homogeneous then

$$w\left(T-t\right) := \log\sigma\left(t,T\right) = \log p\left(t\right) + \log q\left(T\right) = p^*\left(t\right) + q^*\left(T\right)$$

Diœerentiating this ørst with respect to t and then T gives

$$w' = -p^{*\prime} = q^{*\prime}$$

Since p^* is independent of T and q^* of t, we must have $w' = K$ (constant). Solving this gives

$$w\left(T-t\right) = K\left(T-t\right) + \log L$$

so that $\sigma\left(t,T\right) = Le^{K(T-t)}$. Substituting into the deønition of $a\left(t,T\right)$ gives

$$a\left(t,T\right) = \frac{L}{K}\left(e^{K(T-t)} - 1\right)$$

8.6 Non-Existence of PDEs Under HJM

The Feynman-Kac equation outlined in the previous chapter works only when the underlying state space for the process is ønite-dimensional. Except for rather restrictive special cases (such as Sankarasubramanian, 1993), there is no ønite-dimensional state space for the HJM framework, so PDE-based methods cannot be used. Instead, the process is constructed from ørst principles in a discrete setting and Monte-Carlo simulations are used to compute the expectations necessary.

For the general state-space representation of HJM models, we take the forward rate to be the state variable. For example, Musiela (1994) takes the state space to be $\mathcal{C}^1\left[0,\infty\right)$. The current state $X\left(t\right)$ is the function mapping the maturity T to the forward rate $f\left(t,t+T\right)$. With the goal of viewing X as a Markov process satisfying an SDE, Musiela shows that the HJM framework leads to

$$dX_t\left(T\right) = \frac{\partial}{\partial T}\left\{X_t\left(T\right) + \frac{1}{2}\left\|\int_0^T \sigma\left(t,t+s\right)\right\|_{\mathcal{L}^2}^2\right\}dt + \sigma\left(t,t+T\right)dW\left(t\right) \qquad (8.12)$$

Then, taking $\sigma\left(t,t+T\right) = \varphi\left(X,T\right)$ for some function φ, we can view (8.12) as a stochastic PDE in the sense of Walsh (1994). General conditions on φ for the existence and uniqueness of Markovian solutions to (8.12) have yet to be deduced.

8.7 Further Research in HJM

Though HJM has obvious strengths, there are still questions and issues raised by it that have not been satisfactorily answered as yet. One of these is the positivity of interest rates. Under HJM, interest rates can go negative with positive probability. One explanation put forward for this is that HJM seeks to model a derived quantity (the forward rate) rather than an economic fundamental (the spot rate) and so ıloses controlȷ (Rogers). Heath, Jarrow and Morton suggest a volatility structure that ensures positive interest rates, but this is highly artiøcial and not grounded in any economic reasons. Their proposed volatility structure is essentially $\sigma(t, T, \omega) = \min\{\sigma(t, T), k\}$ for some $k > 0$. More work in this area has been done by Flesaker & Hughston (1996).

Secondly, there remains the issue of yield curve øtting and PDE formulations. It is often claimed that HJM can avoid the compromise between theory and practice that arises from yield curve calibration because it can be made to øt any initial term structure. However, the original HJM formulation allows the yield curve to move in a space generated by a ønite-dimensional Brownian motion and so restricts the movements that can be made without calibration. Recent work by Kennedy (1995) has extended the basic HJM model to allow the multidimensional Brownian motion to be replaced by a Brownian sheet for the case of deterministic volatility processes σ. It seems natural to view the HJM spot rate process as an inønite-dimensional stochastic PDE. There are however no results as yet in this vein.

The last issues are computational and empirical. In the absence of PDE formulations, computationally expensive Monte-Carlo simulations need to be used to calculate bond prices using HJM. Heath, Jarrow, Morton and Spindel (1994) claim that though the method is expensive, it yields accurate results in as little as a dozen iterations. On the empirical side, a lot of work has gone into testing parametric Markovian HJM models. However, there is good reason to believe that forward rates are non-Markovian, and very few results have been published in this area. Stanton (1997) uses a non-parametric method to extend Ait-Sahalia's (1996) estimation of spot rate models to reject virtually every established spot rate model. The method can be extended to cover the Gaussian HJM case, but this has not been done as yet.

Bibliography

[1] Ait-Sahalia, Y., 1996, Nonparametric Pricing of Interest Rate Derivative Securities, Econometrica 64, 527-560

[2] Ait-Sahalia, Y., 1996, Testing Continuous-Time Models of the Spot Interest Rate, Rev. Fin. Studies 9, 385-426.

[3] Amin, K., Jarrow, R., 1992, Pricing Options on Risky Assets in a Stochastic Interest Rate Economy.

[4] Anderson, N., Breedon, F., et al, 1996, Estimating and Interpreting the Yield Curve, Wiley, New York.

[5] Ash, R.B., Gardner, M.F., 1975, Topics in Stochastic Processes, Academic Press, New York.

[6] Ball, C.A., Torus, W.N., 1983, Bond Price Dynamics and Options, J. Fin Quant. Anal. 17, 53-66.

[7] Baxter, M.W., Rennie, A., 1996, Financial Calculus, CUP, Cambridge.

[8] Bjork, T., forthcoming, Interest Rate Theory, Springer Notes in Mathematics.

[9] Black, F., Derman, E., Toy, W., 1990, A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options, Fin. Analysts J. 46(1), 33-39.

[10] Black, F., Karasinski, P., 1991, Bond and Option Pricing when Short Rates are Lognormal, Fin. Analysts J. 47(4), 52-59.

[11] Black, F., Scholes, M., 1973, The Pricing of Options and Corporate Liabilities, J. Pol. Econ. 81, 637-654.

[12] Brace, A., Gatarek, D., Musiela, M., 1997, The Market Model of Interest Rate Dynamics, Math. Fin. 7, 127-154.

[13] Brace, A., Musiela, M., 1994, A Multufactor Gauss-Markov Implementation of Heath, Jarrow, and Morton, Math. Fin. 4, 259-283.

[14] Brennan, M.J., Schwartz, E.S., A Continuous-Time Approach to the Pricing of Bonds, J. Bank. Fin. 3, 135-155.

[15] Brown, R.H., Schaefer, S.M., 1994, The Term Structure of Real Interest Rates and the Cox, Ingersoll, and Ross Model, J. Fin. Econ. 35, 3-42.

[16] Brown, S.J., Dybvig, P.H., 1986, The Empirical Implications of the Cox, Ingersoll, and Ross Theory of the Term Structure of Interest Rates, J. Fin. 41, 617-632.

[17] Carverhill, A.P., 1994, A Simpliøed Exposition of the Heath, Jarrow, and Morton Model, Hong Kong University of Science and Technology.

[18] Chan, K.C., Karolyi, G.A., Longstaœ, F.S., Sanders, A.B., 1992, The Volatility of Short-Term Interest Rates: An Empirical Comparison to Alternative Models of the Term Structure of Interest Rates, J. Fin. 47, 1209-1227.

[19] Chen, R., 1996, Understanding and Managing Interest Rate Risks, World Scientiøc, New York.

[20] Chen, R., Scott, L., 1992, Pricing Interest Rate Options in a Two-Factor Cox- Ingersoll-Ross Model of the Term Structure, Rev. Fin. Stud. 5, 613-636.

[21] Cheyette, O., 1992, Markov Representations of the Heath-Jarrow-Morton Model, Capital Management Sciences.

[22] Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985, A Theory of the Term Structure of Interest Rates, Econometrica 53, 385-407.

[23] DaPrato, G., Zabczyk, J., 1992, Stochastic Equations in Inønite Dimensions, CUP.

[24] DuŒe, D., 1993, State Space Models of the Term Structure of Interest Rates, Phil. Trans. Roy. Soc. Series A 347, 577-586

[25] DuŒe, D., Kan, R., 1994, Multifactor Interest Rate Models, Phil. Trans. Roy. Soc. London Ser. A 347, 577-586.

[26] DuŒe, D., Kan, R., 1996, A Yield-Factor Model of Interest Rates, Math. Fin. 6, 379-406.

[27] Dumas, B., Allaz, B., 1996, Financial Securities, Chapman-Hall, London.

[28] Flesaker, B., 1993, Testing the Heath-Jarrow-Morton/Ho-Lee Model of Interest Rate Contingent Claims Pricing, J. Fin. Quant. Anal. 28, 483-495.

[29] Flesaker, B., Hughston, L., 1996, Positive Interest, Risk 9(1), 46-49.

[30] Geman, H., El Karoui, N., Rochet, J.C., 1995, Changes of Numeraire, Changes of Probability Measure and Option Pricing, App. Prob. 32.

[31] Gibbons, M.R., Ramaswamy, K., 1993, A Test of the Cox, Ingersoll, and Ross Model of the Term Structure, Rev. Fin. Stud. 6, 619-658.

[32] Goldys, B., Musiela, M., 1996, On Partial Diœerential Equations Related to Term Structure Models, Working Paper, UNSW.

[33] Harrison, D.M., Kreps, D.M., 1979, Martingales and Arbitrage in Multiperiod Securities Markets, J. Econ. Th. 20, 381-408.

[34] Harrison, D.M., Pliska, S.R., Martingales and Stochastic Integrals in the Theory of Continuous Trading, Stoch. Proc. Appl. 11, 215-260.

[35] Heath, D., Jarrow, R., Morton, A., 1992, Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation, Econometrica 60, 77-105.

[36] Heath, D., Jarrow, R., Morton, A., Spindel, R., 1992, Easier Done Than Said, Risk 5(9), 77-80.

[37] Ho, T.S.Y., Lee, S.-B., 1986, Term Structure Movements and the Pricing of Interest Rate Contingent Claims, J. Fin. 41, 1011-1029.

[38] Hogan, M., Weintraub, K., 1993, The Lognormal Interest Rate Model and Eurodollar Futures, Working Paper, Citibank, New York.

[39] Howison, S., Wilmott, P., Dewynne, J., 1995, The Mathematics of Financial Derivatives, CUP, Cambridge.

[40] Hull, J., White, A., 1990, Pricing Interest Rate Derivative Securities, Rev. Fin. Stud. 3, 573-592.

[41] Ikeda, N., Wantanabe, S., 1981, Stochastic Diœerential equations and Diœusion Processes, North-Holland, Amsterdam.

[42] Jamshidian, F., 1995, A Simple Class of Square-Root Interest Rate Models, Appl. Math. Fin. 2, 61-72.

[43] Jarrow, R., Turnbull, S., 1995, Derivative Securities, Southwestern Publishers, Cleveland.

[44] Jeœrey, A., 1995, Single-Factor Heath-Jarrow-Morton Term Structure Models Based on Markov Spot Interest Rate Dynamics, J. Fin. Quant. Anal. 30, 619-642.

[45] Karatzas, I., Shreve, S., 1988, Brownian Motion and Stochastic Calculus, Springer, New York.

[46] Karlin, S., Taylor, H.M., 1981, A First Course in Stochastic Processes, Academic Press, New York.

[47] Karlin, S., Taylor, H.M., 1981, A Second Course in Stochastic Processes, Academic Press, New York.

[48] Kennedy, D.P., 1994, The Term Structure of Interest Rates as a Gaussian Random Field, Math. Fin. 4, 247-258.

[49] Longstaœ, S.A., 1989, A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates, J. Fin. Econ. 23, 195-224.

[50] Milterstern, K., Sandmann, K., Sondermann, D., 1997, Closed-Form Solutions for Term Structure Derivatives with Lognormal Interest Rates, J. Fin. 52, 409-430.

[51] Musiela, M., Rutkowski, M., 1997, Martingale Methods in Financial Modelling, Springer, New York.

[52] Neftci, S.N., 1996, An Introduction to the Mathematics of Financial Modelling, Academic Press, New York.

[53] Nowman, K.B., 1997, Gaussian Estimation of Single-Factor Continuous-Time Models of the Term Structure of Interest Rates, J. Fin. 52(4), 1695-1706.

[54] Parzen, E., 1964, Stochastic Processes, Holden-Day, San Fransisco.

[55] Pearson, N.D., Sun, T.-S., 1994, Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model, J. Fin. 49, 1279-1304.

[56] Prabhu, N.U., 1965, Stochastic Processes, Macmillan, New York.

[57] Rebonato, R., 1997, Interest Rate Option Models, Wiley, New York.

[58] Rogers, L.C.G., 1995, Which Model for Term Structure of Interest Rate Should One Use? IMA Vol. 65, Springer, New York, pp. 93-116.

[59] Sandmann, K., Sondermann, D., 1994, On the Stability of Lognormal Interest Rate Models and the Pricing of Eurodollar Futures, Working Paper, University of Bonn.

[60] Stanton, R., 1997, A Nonparametric Model of Term Structure if Interest Rate Risk, J. Fin. 52(5), 1973-2002.

[61] Vasicek, O., 1977, An equilibrium Characterisation of the Term Structure, J. Fin. Econ. 5, 177-188.