No-arbitrage Positive Interest Rate Models

Honours Thesis

Department of Mathematics and Statistics
University of Melbourne

Jonathan Brailey

Supervisor: Prof. Kostya Borovkov

November 7, 2008
Abstract

The aim of this thesis is to present viable no-arbitrage interest rate models where interest rates remain positive using the approach developed by Flesaker and Hughston (1996), Rogers (1997) and Rutkowski (1997) and the summary in Chapter 8 of Cairns (2004). The approach of this thesis will be to first outline the mathematical structure of the model and to show that this structure exhibits no arbitrage as well as show the conditions under which interest rates remain positive. Secondly, examples of specific interest rate models will be developed using the framework. Lastly, numerical simulations of selected models will be run to demonstrate the variety of term structures that can be modelled using the framework.
Contents

1 Introduction

1.1 Interest rate modelling .............................................. 1
  1.1.1 Definitions ................................................. 2
  1.1.2 No-arbitrage and the Fundamental Theorem of Asset Pricing .. 5

1.2 No-arbitrage versus short-rate models ............................ 6
  1.2.1 Short-rate models .......................................... 6
  1.2.2 No-arbitrage models ....................................... 7

1.3 Model framework ................................................. 8
  1.3.1 No-arbitrage ............................................. 10
  1.3.2 Positive Interest Rates .................................. 13
  1.3.3 Flesaker and Hughston Approach .......................... 14

2 Examples

2.1 Model 1: Square-root process, degenerate case .................. 17
2.2 Model 2: Square-root process, quadratic case .................... 22
2.3 Model 3: Square-root process, exponential case .................. 25
2.4 Model 4: Univariate Ornstein-Uhlenbeck process .................. 29
## CONTENTS

2.5 Model 5: Squared Gaussian multivariate model ........................ 33
2.6 Discussion of examples ....................................................... 45

### 3 Simulations

3.1 Simulations of model 3 .......................................................... 49
   3.1.1 The Yield Curve ....................................................... 49
   3.1.2 The short-rate ............................................................ 51
3.2 Simulations of model 5 .......................................................... 53
   3.2.1 The Yield Curve ....................................................... 53

### 4 Concluding Remarks

4.1 Further Study and Next Steps ............................................... 56
Acknowledgements

I would like to thank

Kostya Borovkov, my supervisor for his guidance and support throughout the project;

Steve Carnie, Aihua Xia, Kaye Marion, Fima Klebaner, Catherine Forbes and Gael Martin for their challenging and interesting honours courses;

Owen Jones, John Sader, Ken Sharpe, Omar Foda, Barry Hughes, Ray Watson, Peter Taylor, Ian Gordon, Christine Mangelsdorf, Sanming Zhou and Derek Chan for their undergraduate courses which prepared me to embark on my honours year;

The Maurice Belz Honours scholarship which allowed me to focus on study instead of part-time work;

My family and friends, especially Bobbie, for their patience, understanding and support throughout the year.
Chapter 1

Introduction

In this chapter we first outlay the basic definitions required to describe the term structure of interest rates, and then characterise where the no-arbitrage positive interest approach sits in relation to other well known models. The approach itself is then developed from the main articles of Flesaker and Hughston (1996), Rogers (1997) and Rutkowski (1997). Finally the price process is shown to exhibit no-arbitrage, and the conditions under which interest rates remain positive are derived.

1.1 Interest rate modelling

In modern financial markets there are a plethora of traded instruments who’s value is contingent on some form of interest rate. Bonds, swaps, interest rate derivatives and credit instruments are but a few of the products directly related to interest rates. There are also a multitude of products which are indirectly affected by movements in
interest rates such as equities and equity derivatives, foreign exchange rates, and real estate related products. For the banks who trade these products, as well as the speculators, pension funds, hedge funds, insurance agencies and governments who buy and sell them, it is important to have a mathematical model which describes the way in which interest rates evolve over time. Such a model is necessary to value interest rate products, identify possible arbitrage opportunities in the market, and estimate the risk of a portfolio for both management and regulatory purposes.

In order to create a model for interest rates we first need a concise way of defining interest rates themselves. The first section of this chapter lays out the basic definitions necessary to construct an interest rate model.

1.1.1 Definitions

- A bond is a securitized loan, where the buyer of the bond pays an initial price $P$ in exchange for a predetermined sequence of payments. The date of the final payment made to the buyer is known as the maturity date of the bond, which we will denote as $T$. Since the bond is securitized, the bond may be traded at any time before maturity. In this thesis we will only consider bonds which have zero probability of default \(^1\).

\(^1\)Default occurs when the bond issuer is unwilling or unable to make the predetermined series of payments specified by the bond. To allow the probability of default to be greater than zero is to introduce the notion of credit risk, which is beyond the scope of this thesis.
• $P(t, T)$ is the price of a bond at time $t$ which pays 1 at time $T$. Since such a bond makes only one payment to the buyer at maturity, it is also known as a zero coupon bond (there are no intermediate payments of any kind). The zero coupon bond is an especially useful tool for describing the evolution of bond prices under a given interest rate model. The zero coupon bond which matures at time $T$ is often referred to as the $T$-bond. In what follows we will always consider $t \in [0, T]$.

• $R(t, T)$ is the yield to maturity at time $t$ of a bond which matures at time $T$. This means that if we invest 1 at time $t$ in the $T$-bond for $T - t$ years, then this will accumulate at an average rate of $R(t, T)$ over the whole period. $R(t, T)$ is related to $P(t, T)$ in the following way

$$R(t, T) = -\frac{\ln P(t, T)}{T - t}$$ (1.1)

• $r(t)$ is the instantaneous risk free rate (often referred to as the short rate), which is obtained by taking the limit

$$r(t) = \lim_{T \downarrow t} R(t, T) = R(t, t)$$ (1.2)

• $F(t, T, S)$ is the forward rate at time $t$ which applies between times $T$ and $S$, where $T < S$. This is interpreted as the yield for the period $S$ to $T$ which can be locked in at time $t$. It is defined as

$$F(t, T, S) = \frac{1}{S - T} \ln \frac{P(t, T)}{P(t, S)}$$ (1.3)
• $f(t, T)$ is the instantaneous forward rate, which is obtained by taking the limit

$$f(t, T) = \lim_{S \downarrow T} F(t, T, S) = -\frac{\partial}{\partial T} \ln P(t, T)$$ (1.4)

At any time $t \leq T$, $P(t, T)$, $f(t, T)$, and $R(t, T)$ uniquely determine each other in the following way

$$P(t, T) = \exp[-R(t, T)(T - t)] = \exp \left[ - \int_t^T f(t, s) \, ds \right]$$ (1.5)

It is common to talk about interest rates in terms of the ‘yield curve’ or equivalently the ‘term structure of interest rates’. Unfortunately there is no single agreed upon definition as to what exactly constitutes the term structure, but broadly speaking it is the relationship between the time to maturity of a bond, and the yield which will be earnt on the bond if it is held to maturity. A curve which relates any of $P(t, T)$, $f(t, T)$, $F(t, T, S)$, or $R(t, T)$ to $T$ can be referred to as representing the term structure of interest rates, since all of those quantities uniquely determine the yield on the $T$-bond.

The choices which the interest rate modeller must face do not end with the description of the term structure. As it currently stands, there is no ubiquitous approach to describing the evolution of the term structure over time. There is a wide variety of popular interest rate models which have been developed over the last few decades. Before we discuss the popular interest rate models in the literature though, we need to introduce the concept of no-arbitrage and the Fundamental Theorem of Asset Pricing.
1.1.2 No-arbitrage and the Fundamental Theorem of Asset Pricing

Arbitrage refers to the situation where a trading strategy exists that requires zero initial capital and results in a risk free profit. Opportunities to engage in arbitrage in financial markets are exceedingly rare, and as such a central building block of all interest rate models is the assumption of no-arbitrage. That is to say, the bond price process is defined in such a way that arbitrage opportunities are not present. The way no-arbitrage is incorporated into a model is via the Fundamental Theorem of Asset Pricing, which is presented below\(^2\).

We first require one more definition. Let \( B(t) \) be defined as the cash account in the following way

\[
B(t) = B(0) \exp \left[ \int_0^t r(s) \, ds \right]. \tag{1.6}
\]

This means the cash account grows according to the continuously compounded short-rate.

We can now state the Fundamental Theorem of Asset Pricing.

1. *Bond prices evolve in a way which is arbitrage free if and only if there exists a measure \( Q \), equivalent to \( P \), under which, for each \( T \), the discounted price process \( P(t, T) / B(t) \) is a martingale for all \( t : 0 < t < T \).*

\(^2\)Proof of the Fundamental Theorem of Asset Pricing is not given in this thesis, but can be found in Chapters 3 and 4 of Cairns (2004)
2. If 1 holds, then the market is complete if and only if $Q$ is the unique measure under which $P(t,T)/B(t)$ are martingales.

Part 1 in particular will be used later to show that the framework being presented in this thesis exhibits no-arbitrage.

1.2 No-arbitrage versus short-rate models

There are two main types of interest rate models, no-arbitrage models and short-rate models. It is important to note that the term no-arbitrage in this section simply refers to a general modelling approach, and not the technical condition of no-arbitrage which was defined in terms of the Fundamental Theorem of Asset Pricing in the previous section. Both modelling approaches feature bond prices which exhibit no-arbitrage, they just do so from a different starting point.

1.2.1 Short-rate models

Short-rate models generally consist of a stochastic model for the evolution of the short-rate based on assumptions about how the economy works. The Fundamental Theorem of Asset Pricing can then be used to derive a theoretical set of bond prices, which will evolve in a way which is free from arbitrage.

Successful short-rate models include the work of Vasicek (1977), Cox, Ingersoll and
Ross (1985), Black, Derman and Toy (1990), and Hull and White (1990).

In a practical context, after the short-rate model has been calibrated to historical data, the theoretical set of bond prices which are derived from the model may not match observed market prices. This may indicate the existence of an arbitrage opportunity (assuming the underlying model is correct).

In the case of pricing interest rate derivatives, the use of a short-rate model which yields bond prices that differ to those observed in the market may lead to theoretical derivative prices which are drastically different to observed derivative prices in the market. This poses a major potential problem for banks trading such securities.

1.2.2 No-arbitrage models

No-arbitrage models on the other hand use the observed term structure of interest rates as the starting point\(^3\). The evolution of future bond prices is then modelled in a way which is arbitrage free. These models have the advantage that theoretical prices will match current prices, and as such are more suitable for use in pricing derivative securities.

Successful no-arbitrage models include the work of Ho and Lee (1986), and Heath, Jarrow and Morton (1992).

\(^3\)Recall from (1.5) that the price of the \(T\)-bond for each \(T\) is one way of representing the term structure.
A major disadvantage of no-arbitrage models is that the implied short-rate from such models can have peculiar dynamics which are hard to justify in an economic context. In particular no-arbitrage models can exhibit the possibility of the short-rate becoming negative which is unrealistic and highly undesirable.

The framework presented in this thesis is of the no-arbitrage variety. The evolution of the price process \( P(t, T) \) is modelled directly, in a way which is arbitrage free. The key attribute of this framework is that unlike the no-arbitrage models mentioned above, the short-rate will remain positive almost surely as long as certain conditions regarding the stochastic processes which drive the model are met. These conditions will be detailed in the following section.

### 1.3 Model framework

The key step in developing positive interest models in the Flesaker and Hughston (1996), Rogers (1997) and Rutkowski (1997) framework is to generate a strictly positive supermartingale diffusion process which will be used to construct the model for the yield curve. Recall that a supermartingale is a process \( X(t) \) which satisfies

\[ E[X(t + s)|X(t)] \leq X(t), \quad \text{for all } s > 0. \]

In other words a supermartingale is a process which decreases on average.
Rutkowski (1997) defined the diffusion directly, in the following way. Let $(\Sigma, \mathcal{F}, \hat{P})$ be a probability triple and let $\hat{W}(t)$ be a standard Brownian motion in $\mathbb{R}^d$ under $\hat{P}$. Let $\mathcal{F}_t$ be the sigma-algebra generated by $\hat{W}(s) : 0 \leq s \leq t$.

Let $A(t)$ be a strictly positive supermartingale diffusion process in $\mathbb{R}$ adapted to $\mathcal{F}_t$ such that

$$dA(t) = A(t)[\mu_A(t)dt + \sigma_A(t)'d\hat{W}(t)]$$

(1.7)

where $\mu_A(t)$ is a scalar and $\sigma_A(t)$ is a $(d \times 1)$ vector.

Rogers (1997) went back one extra step and constructed a positive diffusion by taking functions of a general diffusion process in the following way. Let $f : \mathbb{R}^d \to \mathbb{R}^+$ be a strictly positive, twice differentiable function. Let $X(t)$ be a diffusion process in $\mathbb{R}^d$ with

$$dX(t) = \mu_X(t)dt + \sigma_X(t)d\hat{W}(t).$$

We then define

$$A(t) = \frac{e^{-\alpha t}f(X(t))}{f(X(t))}$$

(1.8)

It should be noted that Rogers’ method does not guarantee a supermartingale process. The function $f(t)$ must be chosen intelligently to ensure the resulting $A(t)$ is in-fact a supermartingale. The benefit of Rogers method is that, by choosing different forms of $f(t)$, many unique interest rate models can be generated, and as long as the resulting $A(t)$ is a supermartingale they will all feature strictly positive interest rates.

---

$^4$ $\hat{P}$ is known as the ‘pricing measure’ It is different from, but equivalent to the real world probability.
Now define

\[ P(t, T) = \frac{E_P[A(T)|\mathcal{F}_t]}{A(t)}, \quad (1.9) \]

which we use as a model for the price at time \( t \) of a zero-coupon bond which matures at time \( T \). The information contained in \( P(t, T) \) describes the entire yield curve at any given time, including the forward rate curve \( f(t, T) \) and the short-rate \( r(t) \). From our model of \( P(t, T) \) we are able to derive all related quantities of interest.

### 1.3.1 No-arbitrage

To be a plausible model for bond prices, it must be shown that (1.9) is arbitrage free.

Let

\[ D(t, T) = A(t)P(t, T) = E_P[A(T)|\mathcal{F}_t] \]

Since \( A(t) \) is a strictly positive diffusion process \( D(t, T) \) must be a strictly positive martingale. Therefore we can write

\[ dD(t, T) = D(t, T)\sigma_D(t, T) d\tilde{W}(t) \]

\(^5\text{See (1.5).}\)
and we have, by Ito’s formula,

$$dP(t, T) = d\left(\frac{D(t, T)}{A(t)}\right) = dD(t, T) \frac{1}{A(t)} + D(t, T) d\left(\frac{1}{A(t)}\right) + dD(t, T) d\left(\frac{1}{A(t)}\right)$$

$$= D(t, T) \frac{1}{A(t)} \sigma\left(\frac{1}{A(t)}\right) \tilde{W}(t) + D(t, T) d\left(\frac{1}{A(t)}\right) + dD(t, T) d\left(\frac{1}{A(t)}\right).$$

Using Ito’s formula and (1.7) we have

$$d\left(\frac{1}{A(t)}\right) = -\frac{1}{A(t)}[\mu_A(t) dt + \sigma_A(t) d\tilde{W}(t)] + \frac{1}{A(t)} \sigma_A(t)^2 dt,$$

so that

$$dP(t, T) = \frac{D(t, T)}{A(t)}[\sigma_D(t, T) d\tilde{W}(t) + \sigma_A(t)^2 dt - \mu_A(t) dt - \sigma_A(t) d\tilde{W}(t)$$

$$+ \sigma_A(t)^2 dt \sigma_D(t, T) d\tilde{W}(t) - \sigma_A(t) dt \sigma_D(t, T)]$$

$$= P(t, T)[\sigma_D(t, T) d\tilde{W}(t) + (\mu_A(t) + \sigma_A(t)^2) dt - \sigma_A(t) d\tilde{W}(t)$$

$$- \sigma_A(t) \sigma_D(t, T) dt].$$

Now use the change of measure under which the process $\tilde{W}(t)$, with

$$d\tilde{W}(t) = d\tilde{W}(t) - \sigma_A(t) dt$$

will be a Brownian motion.
We now have

\[ dP(t, T) = P(t, T)[\sigma_D(t, T)(d\tilde{W}(t) + \sigma_A(t)dt) - \mu_A(t)dt + \sigma_A(t)^2 dt \]
\[ - \sigma_A(t)(d\tilde{W}(t) + \sigma_A(t)dt) - \sigma_A(t)\sigma_D(t, T)dt] \]

\[ = P(t, T)[-\mu_A dt + (\sigma_D(t, T) - \sigma_A(t))d\tilde{W}(t)] \]

\[ = P(t, T)[r(t) dt + S(t, T)d\tilde{W}(t)], \tag{1.10} \]

where we put

\[ r(t) = -\mu_A(t), \tag{1.11} \]
\[ S(t, T) = \sigma_D(t, T) - \sigma_A(t). \tag{1.12} \]

To show that (1.10) is a martingale when discounted we first define the discounted price process

\[ Z(t, T) = \frac{P(t, T)}{B(t)}. \]

We now apply Ito’s formula which gives

\[ dZ(t) = d\left( \frac{P(t, T)}{B(t)} \right) = dP(t, T) \left( \frac{1}{B(t)} \right) + P(t, T) d\left( \frac{1}{B(t)} \right) + dP(t, T) d\left( \frac{1}{B(t)} \right). \]
In order to proceed we need to find \( \frac{1}{B(t)} \) which we find by observing that (1.6) implies that \( dB(t) = r(t)B(t) dt \), and again use Ito’s formula to find that

\[
d\left( \frac{1}{B(t)} \right) = - \left( \frac{r(t)}{B(t)} \right) dt.
\]

We can now compute

\[
dZ(t, T) = \frac{P(t, T)}{B(t)} [r(t) dt + S(t, T) d\tilde{W}] - \frac{P(t, T)}{B(t)} r(t) dt
\]

\[
= Z(t, T) S(t, T) d\tilde{W},
\]

which is a martingale. Therefore by the Fundamental Theorem of Asset Pricing (1.9) is arbitrage free.

### 1.3.2 Positive Interest Rates

It now needs to be shown that (1.9) does in fact guarantee positive interest rates.

Since we have stipulated that \( A(t) \) is a super martingale, \( E_\mathcal{F}_t[A(T)|\mathcal{F}_t] \) must be a decreasing function of \( T \).

Hence

\[
f(t, T) = - \frac{\partial P(t, T)}{\partial T} \geq 0
\]
Therefore
\[ r(t) = \lim_{T \downarrow t} f(t, T) \geq 0 \]
so that interest rates remain positive for all \( t \) and \( T \).

### 1.3.3 Flesaker and Hughston Approach

Flesaker and Hughston (1996) presented the framework in the following way. Their approach was to begin with a family of strictly positive martingales \( M(t, T) \) (considered as processes in \( t \leq T \)) and define the price process as
\[
P(t, T) = \frac{\int_{T}^{\infty} \phi(s) M(t, s) \, ds}{\int_{t}^{\infty} \phi(s) M(t, s) \, ds},
\]
where \( \phi(s) \) is a deterministic and strictly positive function.

We then have
\[
f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T)
\]
\[
= -\frac{\partial}{\partial T} \left( \ln \int_{T}^{\infty} \phi(s) M(t, s) \, ds - \ln \int_{t}^{\infty} \phi(s) M(t, s) \, ds \right)
\]
\[
= \frac{\phi(T) M(t, T)}{\int_{T}^{\infty} \phi(s) M(t, s) \, ds},
\]
which gives us

\[ r(t) = f(t, t) = \frac{\phi(t)M(t, t)}{\int_{t}^{\infty} \phi(s)M(t, s) \, ds} \]

Since both \( \phi(t) \) and \( M(t, T) \) are strictly positive, we can see that the rate process is also positive for all \( t \) and \( T \).
Chapter 2

Examples

In this chapter the no-arbitrage positive interest framework is used to develop five interest rate models. By choosing an underlying diffusion process and a suitable function $f(x)$ a broad variety of models are demonstrated. In each case the bond price process $P(t, T)$, as well as the short-rate process $r(t)$ are derived.

The first three models are all based on using the square-root process as the underlying diffusion, with varying forms for $f(x)$. The fourth model uses a univariate Ornstein-Uhlenbeck process with an exponential $f(x)$, and the final (and most ambitious) model uses a multivariate Ornstein-Uhlenbeck process with an exponential quadratic $f(x)$.

We note that the fourth model is considered in chapter 8 of Cairns(2004), and the final model is considered without derivation in both Rogers(1997) and chapter 8 of Cairns(2004).
2.1 Model 1: Square-root process, degenerate case

The square root process is a commonly used stochastic diffusion process in financial mathematics. By choosing sensible parameters we can ensure that the process possesses several features which are highly desirable to the financial modeller such as mean reversion and non-negativity. Also its distribution, moment generating function and so on are readily available in the literature (for example Dufresne (2001), Dassios and Nagaradjasarma (2006)) making its use convenient in the no-arbitrage positive interest framework being presented in this thesis.

In this section a univariate square-root process will be used as the underlying diffusion with $f(x) = x$ to generate a yield curve model under the no-arbitrage positive interest framework.

Let $X(t)$ be a univariate square root process for $t > 0$ given by

$$dX(t) = \beta X(t) \, dt + \gamma \sqrt{X(t)} \, d\hat{W}(t), \quad X(0) > 0,$$

and take

$$f(x) = x.$$

Using (1.8) we get

$$A(t) = \frac{e^{-\alpha t} f(X(t))}{f(X(0))} = \frac{e^{-\alpha t} X(t)}{X(0)},$$
and using Ito’s Lemma we find that

\[ dA(t) = A(t) \left[ (\beta - \alpha) \, dt + \frac{\gamma}{\sqrt{\exp(\alpha t) A(t) X(0)}} \, d\hat{W}(t) \right] \]

From this we can see immediately that for \( A(t) \) to be a supermartingale (and therefore to be able to guarantee a non-negative short-rate process) we require \( \alpha > \beta \). Using (1.11) we can also say that

\[ r(t) = -\mu A(t) = \alpha - \beta. \]

To find the price process we use the definition given in (1.9)

\[ P(t, T) = \frac{E^\top[A(T)|\mathcal{F}_t]}{A(t)} \]

\[ = \frac{X(0)}{e^{-\alpha t} X(t)} E^\top \left[ \frac{e^{-\alpha T} X(T)}{X(0)} \bigg| \mathcal{F}_t \right] \]

\[ = \frac{e^{-\alpha (T-t)}}{X(t)} E^\top[X(T)|X(t)]. \]

To compute the necessary conditional first moment of the square root process, as well as other similar moments throughout this chapter, Theorem 1 from Dassios and Nagaradjasarma (2006) (a similar result appears in Dufresne (2001)) is used:

**Consider processes** \( X(t), Y(t) \) **given by**

\[ dX(t) = (a + \beta X(t)) \, dt + \gamma \sqrt{X(t)} \, d\hat{W}(t) \]
and

\[ Y(t) = \int_0^t S(u) \, du, \]

where

\[ dS(t) = rS(t) \, dt + \gamma \sqrt{S(t)} \, d\hat{W}(t), \]

the joint moments of \( X(t) \) and \( Y(t) \) are given by

\[
M_{m,n}(t) = E_{\hat{P}}[Y^m(t)X^{n-m}(t)] = \sum_{j=0}^n e^{ij\beta t} \left( \sum_{i=1}^{I_{m,n}^m} \alpha_{j,i}^m \frac{t^{i-1}}{(i-1)!} \right)
\]

where

\[ I_{j}^{m,n} = \min(n + 1 - j, m + 1). \]

The coefficients \( \alpha_{j,i}^m \) are obtained by recursion through the following relations:

- For \( j \neq n - m \),

\[
\alpha_{j,i}^m = m \sum_{i' = 1}^{I_{j}^{m,n}} \frac{(-1)^{i' \cdot i - i - i' \cdot i}}{(n - m - j - \beta)^{i' \cdot i + 1}} + (n - m) \left( a + (n - m - 1) \frac{\gamma^2}{2} \right)
\]

\[
\times \sum_{i' = 1}^{I_{j}^{m,n}} \frac{(-1)^{i' \cdot i - i - i' \cdot i}}{(n - m - j - \beta)^{i' \cdot i + 1}}.
\]

- For \( j = n - m \), one has:
- For $i = 1$,

\[
\alpha_{n-m,1}^{m,n} = c_{m,n} + m \sum_{j=0}^{n} \sum_{i=1}^{I_{n-m}^{m,1,n}} \alpha_{j,i}^{m-1,n} \frac{1}{((j-n+m)-\beta)^{i-1}} + (n-m) \left( a + (n-m-1) \frac{\gamma^2}{2} \right)
\]

\[
\times \sum_{j=0}^{n-1} \sum_{i=1}^{I_{n-m}^{m,1,n}} \alpha_{j,i}^{m,n-1} \frac{1}{((j-n+m)-\beta)^{i}} ;
\]

- For $i > 1$,

\[
\alpha_{n-m,i}^{m,n} = m \alpha_{n-m,i-1}^{m-1,n} + (n-m) \left( a + (n-m-1) \frac{\gamma^2}{2} \right) \alpha_{n-m,i-1}^{m,n-1},
\]

where

\[
c_{m,n} = X(0)^n 1_{[m=0]}
\]

with initial condition

\[
\alpha_{0,1}^{0,0} = 1.
\]

For our purposes we simply set $m = 0, n = 1, a = 0$ and use (2.2) to compute the
necessary moment. We find that

\[ P(t, T) = \frac{e^{-\alpha(T-t)}}{X(t)} \mathbb{E}_P [X(T)|X(t)] \]

\[ = \frac{e^{-\alpha(T-t)}}{X(t)} X(t) e^\beta(T-t) \]

\[ = e^{(\beta-\alpha)(T-t)}. \]

We can easily find the forward rate and short-rate process from the price process:

\[ f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) = \alpha - \beta, \]

\[ r(t) = f(t, t) = \alpha - \beta, \]

which coincides with the short-rate process we obtained directly from the \( A(t) \) process.

The yield structure described by this model is bland and unrealistic. Both the price and short-rate process contain no stochastic components, leaving a purely deterministic model. Whilst this model satisfies the demands of being a no-arbitrage interest rate model with a non-negative short-rate process, there is no use for this particular model in an applied context.
2.2 Model 2: Square-root process, quadratic case

Since the univariate square-root diffusion process in conjunction with \( f(x) = x \) did not produce a stochastic result, the next two models will explore the results that are produced when the univariate square-root process is used with a slightly more elaborate \( f(x) \).

Again let \( X(t) \) be a univariate square root process given by (2.1) but this time take

\[
  f(x) = x^2.
\]

Using (1.8) we get

\[
  A(t) = \frac{e^{-\alpha t} f(X(t))}{f(X(0))} = \frac{e^{-\alpha t} X(t)^2}{X(0)^2},
\]

and using Ito’s formula we find that

\[
  dA(t) = A(t) \left[ \left( 2\beta - \alpha + \frac{\gamma^2}{X(t)} \right) dt + \frac{2\gamma}{\sqrt{X(t)}} d\hat{W}(t) \right].
\]

We can see that for \( A(t) \) to be a supermartingale (and therefore to be able to ensure a non-negative short-rate process) we require that

\[
  2\beta - \alpha + \frac{\gamma^2}{X(t)} < 0,
\]

\[
  \Rightarrow X(t) > \frac{\gamma^2}{\alpha - 2\beta}.
\] (2.3)
This represents a severe restriction on the model. As (2.3) clearly fails with a positive probability, $A(t)$ will not be a supermartingale and hence we cannot guarantee that the interest rates produced by the model will always be positive. Nevertheless, we can explore the model further to find out what dynamics for $r(t)$ it will produce.

From the $A(t)$ process we can also see that

$$r(t) = -\mu_A(t) = (\alpha - 2\beta) - \frac{\gamma^2}{X(t)}. \quad (2.4)$$

To find the bond price process we use the definition given in (1.9):

$$P(t, T) = \frac{E_{\tilde{P}}[A(T)|\mathcal{F}_t]}{A(t)}$$

$$= \frac{X^2(0)}{e^{-\alpha t}X^2(t)} E_{\tilde{P}} \left[ \frac{e^{-\alpha T}X^2(T)}{X^2(0)} | \mathcal{F}_t \right]$$

$$= \frac{e^{-\alpha(T-t)}}{X^2(t)} E_{\tilde{P}} \left[ X^2(T) | X(t) \right].$$

The second conditional moment of the square-root process is again computed using (2.2), this time setting $m = 0, n = 2, a = 0$. We find that

$$E_{\tilde{P}} \left[ X^2(T) | X(t) \right] = \frac{\gamma^2 X(t)}{\beta} \left( e^{2\beta(T-t)} - e^{\beta(T-t)} \right) + X^2(t) e^{2\beta(T-t)}.$$

So we have

$$P(t, T) = e^{-\alpha(T-t)} \left( \frac{\gamma^2}{\beta X(t)} \left( e^{2\beta(T-t)} - e^{\beta(T-t)} \right) + e^{2\beta(T-t)} \right).$$
We can now recover the short rate process directly from \( P(t, T) \) to confirm the result we have from the \( A(t) \) process:

\[
    f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T)
\]

\[
    = -\frac{\partial}{\partial T} \left[ -\alpha(T - t) + \ln \left( \frac{\gamma^2}{\beta X(t)} (e^{2\beta(T-t)} - e^{\beta(T-t)}) + e^{2\beta(T-t)} \right) \right]
\]

\[
    = \alpha - \frac{\gamma^2}{X(t)} \left( 2e^{2\beta(T-t)} - e^{\beta(T-t)} \right) + 2\beta e^{2\beta(T-t)} \cdot \frac{\gamma^2}{X(t)} (e^{2\beta(T-t)} - e^{\beta(T-t)}) + e^{2\beta(T-t)}
\]

So we have

\[
    r(t) = f(t, t) = (\alpha - 2\beta) - \frac{\gamma^2}{X(t)},
\]

which coincides with (2.4).

The model produced here by taking a quadratic function of the univariate square-root process is significantly more interesting than model considered in section 2.1. We see here that both price and short-rate process contain a stochastic component.

To achieve a non-negative short rate with a high probability we can choose \( \alpha \gg 2\beta \) so that it is unlikely \( X(t) \) will ever cross into the region where \( A(t) \) ‘ceases to be a supermartingale’. It would still not be possible to completely ensure that the short rate process remained positive.
2.3 Model 3: Square-root process, exponential case

In this section the square-root process will be used once more, but this time in conjunction with $f(x) = e^{-ax}$ that will be used to generate the $A(t)$ process.

As before, $X(t)$ is a univariate square root process given by (2.1), but now we take

$$f(x) = e^{-ax}, \quad a > 0.$$ 

Using (1.8) we get

$$A(t) = \frac{e^{-at} f(X(t))}{f(X(0))} = \frac{e^{-at} e^{-aX(t)}}{e^{-aX(0)}},$$

and using Ito’s Lemma we find that

$$dA(t) = A(t) \left[ \left( -\alpha + (\ln A(t) - aX(0) + at) \left( \beta - a \gamma^2/2 \right) \right) dt + \gamma (\ln A(t) - aX(0) + at) d\hat{W}(t) \right]$$

$$= A(t) \left[ \left( -\alpha - aX(t) \left( \beta - a \gamma^2/2 \right) \right) dt - \gamma aX(t) d\hat{W}(t) \right].$$

The condition that will make $A(t)$ a supermartingale is

$$-\alpha - aX(t) \left( \beta - a \gamma^2/2 \right) < 0,$$
which is equivalent to

\[ X(t) > -\frac{1}{a} \left[ \frac{2\alpha}{2\beta - \alpha \gamma^2} \right]. \]

Since \( X(t) > 0 \) for all \( t \) and \( \alpha, a > 0 \) this condition is clearly met when \( 2\beta - \alpha \gamma^2 > 0 \).

The \( A(t) \) process also reveals the short rate as

\[ r(t) = -\mu_A(t) = \alpha + aX(t) \left( \beta - \frac{\gamma^2 a}{2} \right). \tag{2.5} \]

We can now find the price process using the definition given in (1.9):

\[
P(t, T) = \frac{E_{\hat{P}}[A(T)|\mathcal{F}_t]}{A(t)}
= \frac{e^{-aX(0)}}{e^{-\alpha t} e^{-aX(t)}} E_{\hat{P}} \left[ \frac{e^{-aT} e^{-aX(T)}}{e^{-aX(0)}} \right] \mathcal{F}_t
= \exp \left[ -\alpha (T - t) - aX(t) \right] E_{\hat{P}}[e^{-aX(T)}|X(t)]
\]

To find the moment of the square-root process which is required to complete this calculation we refer to Dassios and Nagaradjasarma (2006) who show that the joint conditional moment generating function of \( X(t) \) and \( \int_0^T X(u) \, du \) is

\[
E \left[ e^{-\lambda X(T)} e^{-\mu \int_0^T X(u) \, du} \mid X(t) \right] = e^{-X(t)\psi}, \tag{2.6}
\]
where
\[
\psi = \frac{\lambda ((\phi + \beta) + e^{-\phi(T-t)}(\phi - \beta)) + 2\mu(1 - e^{-\phi(T-t)})}{\gamma^2 \lambda(1 - e^{-\phi(T-t)}) + (\phi - \beta) + e^{-\phi(T-t)}(\phi + \beta)}
\]
and
\[
\phi = \sqrt{\beta^2 + 2\mu \gamma^2}.
\]
By setting \(\lambda = a\) and \(\mu = 0\) we find that the moment generating function of our square root process is
\[
E_P[e^{-aX(T)|X(t)}] = \exp\left[-\frac{2X(t)\beta a}{e^{-\beta(T-t)}(2\beta - a\gamma^2) + a\gamma^2}\right],
\]
so our price process is given by
\[
P(t, T) = \exp\left[-\alpha(T - t) + aX(t) - \frac{2X(t)\beta a}{e^{-\beta(T-t)}(2\beta - a\gamma^2) + a\gamma^2}\right]
\]
\[
= \exp\left[-\alpha(T - t) + aX(t)\left(1 - \frac{2\beta}{e^{-\beta(T-t)}(2\beta - a\gamma^2) + a\gamma^2}\right)\right].
\]
From here we can verify the short rate process by computing

\[ f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) \]

\[ = \alpha - \frac{\partial}{\partial T} \left( \frac{-2a\beta X(t)}{e^{-\beta(T-t)}(2\beta - a\gamma^2) + a\gamma^2} \right) \]

\[ = \alpha + 2a\beta X(t) \left[ \frac{\beta(2\beta - \gamma^2a) e^{-\beta(T-t)}}{(e^{-\beta(T-t)}(2\beta - \gamma^2a) + \gamma^2a)^2} \right], \]

which gives us

\[ r(t) = f(t, t) = \alpha + aX(t) \left( \beta - \frac{\gamma^2a}{2} \right). \]

This coincides with the short rate process we obtained from the \( A(t) \) process directly.

The strength of this model compared with (2.4) is, of course, that the process \( A(t) \) is a supermartingale under the condition \( 2\beta - \alpha\gamma^2 > 0 \), and therefore the short rate does not become negative. The form of \( r(t) \) is also simple and it is easy to see how we could set a desired minimum value for the short-rate by choosing the magnitude of \( \alpha \).

Overall this model provides the scope for interesting dynamics in both the price process and the short rate, whilst also succeeding in ensuring a positive short rate without resorting to over-bearing restrictions on the parameters.
2.4 Model 4: Univariate Ornstein-Uhlenbeck process

In this section, a different underlying diffusion process will be considered. A univariate Ornstein-Uhlenbeck process will be used in conjunction with an exponential \( f(x) \) to produce a model for the evolution of the yield curve. Verification of the following model appears as exercise 8.1 in Cairns (2004).

Let \( X(t) \) be a univariate Ornstein-Uhlenbeck diffusion given by

\[
dX(t) = -\theta X(t) \, dt + d\hat{W}(t) \tag{2.7}
\]

and take

\[
f(x) = e^{\sigma x},
\]

where

\[
\theta > 0, \quad \sigma > 0
\]

Using (1.8) we get

\[
A(t) = e^{-\alpha t} \frac{f(X(t))}{f(X(0))} = \frac{e^{-\alpha t} e^{\sigma X(t)}}{e^{\sigma X(0)}}
\]

and using Ito’s Lemma we find that

\[
dA(t) = \left( -\alpha A(t) - \theta X(t) \sigma A(t) + \frac{1}{2} \sigma^2 A(t) \right) \, dt + \sigma A(t) \, d\hat{W}(t)
\]

\[
= A(t) \left[ \left( -\alpha + \frac{\sigma^2}{2} - \sigma \theta X(t) \right) dt + \sigma \, d\hat{W}(t) \right].
\]
We observe that for $A(t)$ to be a supermartingale one needs that

$$\mu_A(t) = -\alpha + \frac{\sigma^2}{2} - \sigma \theta X(t) > 0,$$

which means that

$$X(t) > \frac{\sigma^2 - 2\alpha}{2\sigma \theta}.$$

Since $X(t)$ evolves according to an Ornstein-Uhlenbeck process and can take any real value, we note that no choice of parameters ensure that $A(t)$ will be a supermartingale. Therefore the short rate process may indeed become negative in this model.

Using the $A(t)$ process to find the short rate process and the market price of risk we see that

$$r(t) = -\mu_A(t) = \alpha - \frac{\sigma^2}{2} + \sigma \theta X(t).$$

We can now find the price process using the definition given in (1.9):

$$P(t, T) = \frac{E_{\tilde{P}}[A(T)|\mathcal{F}_t]}{A(t)}$$

$$= \frac{e^{\sigma X(0)}}{e^{-\alpha t} e^{\sigma X(t)}} E_{\tilde{P}} \left[ \frac{e^{-\alpha T} e^{\sigma X(T)}}{e^{\sigma X(0)}} \Bigg| \mathcal{F}_t \right]$$

$$= \exp \left( -\alpha (T-t) - \sigma X(t) \right) E_{\tilde{P}} \left[ e^{\sigma X(T)} | X(t) \right].$$
Since $X(t)$ follows an Ornstein-Uhlenbeck process as described in (2.7), we can immediately say that the conditional distribution of $X(T)$ is Normal (Karatzas and Shreve (1988), pg358) with

$$E_{\hat{P}}[X(T)|X(t)] = X(t) e^{-\theta(T-t)}$$

and

$$Var_{\hat{P}}[X(T)|X(t)] = \frac{1 - e^{-2\theta(T-t)}}{2\theta}.$$ 

We can say therefore about the conditional distribution of $\sigma X(T)|X(t)$ the following:

$$\sigma X(T)|X(t) \sim N\left(\sigma X(t) e^{-\theta(T-t)}, \frac{\sigma^2[1 - e^{-2\theta(T-t)}]}{2\theta}\right).$$

Hence $e^{\sigma X(T)}|X(t)$ is distributed lognormally with

$$E_{\hat{P}}[e^{\sigma X(T)}|X(t)] = \exp\left[\sigma X(t) e^{-\theta(T-t)} + \frac{\sigma^2}{4\theta} \left(1 - e^{-2\theta(T-t)}\right)\right].$$

So we have

$$P(t, T) = \exp\left[-\alpha(T-t) - \sigma X(t) + \sigma X(t) e^{-\theta(T-t)} + \frac{\sigma^2}{4\theta} \left(1 - e^{-2\theta(T-t)}\right)\right]$$

$$= \exp\left[-\alpha(T-t) - \sigma X(t)(1 + e^{-\theta(T-t)}) + \frac{\sigma^2}{4\theta} \left(1 - e^{-2\theta(T-t)}\right)\right].$$
Solving for the short rate process, we find
\[ f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) \]
\[ = \alpha + \sigma \theta X(t) e^{-\theta(T-t)} - \frac{\sigma^2}{2} e^{-2\theta(T-t)}, \]
and so
\[ r(t) = f(t, t) = (\alpha - \frac{\sigma^2}{2}) + \sigma \theta X(t). \] (2.8)

It turns out that (2.8) is equivalent to a very famous short-rate model. If we apply Ito’s formula to (2.8) we get
\[ dr(t) = -\sigma \theta^2 X(t) \, dt + \sigma \theta \, d\hat{W}(t) \]
\[ = \theta((\alpha - \frac{1}{2}\sigma^2) - r(t)) \, dt + \sigma \theta \, d\hat{W}(t), \]
which is of the same form as the Vasicek (1977) equilibrium term structure model. The Vasicek model permits negative interest rates, and as mentioned above the inability to enforce the conditions which ensure \( A(t) \) is a supermartingale mean that we also have the possibility of negative interest rates in our formulation. It is interesting to note, however, that in this case we have been able to incorporate mean reversion in the short-rate process through the no-arbitrage positive interest framework.

This further model demonstrates the care that is needed when choosing an underlying diffusion and function \( f(x) \) to be used in the no-arbitrage positive interest framework. Whilst it may appear a simple task to enforce conditions on the parameters which will
ensure \( A(t) \) is a supermartingale, these conditions may be extremely pervasive into the structure of the model or, as in this case, simply impossible to achieve.

### 2.5 Model 5: Squared Gaussian multivariate model

In this section a multivariate model will be developed and discussed under the no-arbitrage positive interest framework. The underlying diffusion will be a multivariate Ornstein-Uhlenbeck process.

Ahn, Dittmar and Gallant (2002) present evidence that multivariate term structure models which designate the yield on a bond as a quadratic function of underlying state variables are better able to explain empirical term structure data compared with corresponding affine term structure models. The yield structure resulting from this derivation falls into the category known as Quadratic Term Structure Models (QTSMs).

Cairns (2004) summarizes in Section 8.5 the following model, but notes that much ‘tedious algebra’ is required to derive the results. The results of the model without detailed derivation are also presented in Rogers (1997). This section demonstrates the derivation in its entirety.

Let \( X(t) \) be a \( d \)-dimensional Ornstein-Uhlenbeck process \((d > 1)\) with

\[
\mathrm{d}X(t) = -BX(t) \mathrm{d}t + C \mathrm{d}W(t),
\]

(2.9)
where $B$ is a diagonal $d \times d$ matrix with $\text{diag}(\beta_1, ..., \beta_d)$, $C$ is a $d \times d$ matrix of full rank, and $\hat{W}(t)$ is a $d$-dimensional Brownian motion. The matrix $C$ is also defined such that, for $R = (\rho_{ij})_{i,j=1}^d = CC'$ (where $C'$ is the transpose of $C$), one has $\rho_{ii} = 1$ for all $i$ so that $R$ can be interpreted as a correlation matrix.

The stochastic differential equation for (2.9) can be solved using the method of ‘Variation of Constants’.

First we solve the deterministic part $dX(t) = -BX(t) dt$ of (2.9): clearly, this implies

$$\ln X(t) = -Bt + K, \text{ or}$$

$$X(t) = Ke^{-Bt}$$

Next we will assume the constant of integration $K = K(t)$ varies with time. We can use this expression to find a complete solution to (2.9) by equating coefficients and finding an expression for $K(t)$. Using Ito’s formula,
\[ dX(t) = dK(t) e^{-Bt} + K(t) d(e^{-Bt}) \]
\[ = e^{-Bt} dK(t) - B e^{-Bt} dt K(t) \]
\[ = e^{-Bt} dK(t) - B X(t) dt \]
\[ = -B X(t) dt + C d\dot{W}(t). \]

So we have \( dK(t) = C e^{Bt} d\dot{W}(t) \), and hence

\[ K(t) = K(0) C \int_0^t e^{Bs} d\dot{W}(s). \]

So our solution is

\[ X(t) = X(0) e^{-Bt} + C e^{-Bt} \int_0^t e^{Bs} d\dot{W}(s). \]

Or equivalently (due to the time homogeneity of the Ornstein-Uhlenbeck process)

\[ X(T) = X(t) e^{-B(T-t)} + C e^{-B(T-t)} \int_t^T e^{Bs} d\dot{W}(s). \quad (2.10) \]

Since \( X(t) \) follows an Ornstein-Uhlenbeck process, \( X(T) \mid X(t) \) will be distributed as a multivariate Normal random variable. The mean vector and covariance matrix of \( X(T) \mid X(t) \) will be derived here and referred to as the model is developed. First we observe that

\[ e^{-Bt} = \exp[-t \ diag(\beta_1, \ldots, \beta_d)] = \text{diag}(\ e^{-t\beta_1}, \ldots, \ e^{-t\beta_d}). \]
The conditional mean vector components are:

\[(\mu_{t,T})_i = E_{\hat{P}}[X_i(T)|F_t] \]
\[= E_{\hat{P}} \left[ X_i(t) e^{-\beta_i(T-t)} + C_i e^{-\beta_i(T-t)} \int_t^T e^{\beta_i(s)} d\hat{W}(s) | F_t \right] \]
\[= X_i(t) e^{-\beta_i(T-t)}. \]

The conditional covariances are:

\[(V_{t,T})_{ij} = Cov_{\hat{P}}[X_i(T), X_j(T) | F_t] \]
\[= E_{\hat{P}}[(X_i(T) - X_i(t) e^{-\beta_i(T-t)})(X_j(T) - X_j(t) e^{-\beta_j(T-t)}) | F_t] \]

Substituting in the expression for \(X_i(T)\) and \(X_j(T)\) from (2.10) we have

\[= E_{\hat{P}} \left[ \left( C_i e^{-\beta_i(T-t)} \int_t^T e^{\beta_i(s)} d\hat{W}(s) \right) \left( C_j e^{-\beta_j(T-t)} \int_t^T e^{\beta_j(s)} d\hat{W}(s) \right) \right]. \]

Multiplying out the vectors gives

\[= E_{\hat{P}} \left[ \left( C_{i1} e^{-\beta_i(T-t)} \int_t^T e^{\beta_i(s)} d\hat{W}_1(s) + \cdots + C_{id} e^{-\beta_i(T-t)} \int_t^T e^{\beta_d(s)} d\hat{W}_d(s) \right) \right. \]
\[\times \left. \left( C_{j1} e^{-\beta_j(T-t)} \int_t^T e^{\beta_j(s)} d\hat{W}_1(s) + \cdots + C_{jd} e^{-\beta_j(T-t)} \int_t^T e^{\beta_d(s)} d\hat{W}_d(s) \right) \right]. \]

Using the property that \(d\hat{W}_i(s) d\hat{W}_j(s) = \delta_{ij} ds\), we observe that all cross terms are equal to zero, and there no more random terms. We are left with

\[= (C_{i1} C_{j1} + C_{i2} C_{j2} + \cdots + C_{id} C_{jd}) e^{-(\beta_i + \beta_j)(T-t)} \int_t^T e^{(\beta_i + \beta_j)s} ds. \]
Recall that we defined $C$ such that $R = (\rho_{ij})_{i,j=1}^d = CC'$, giving us

$$= \rho_{ij} e^{-(\beta_i + \beta_j)(T-t)} \int_t^T e^{(\beta_i + \beta_j)s} \, ds$$

$$= \rho_{ij} \frac{1 - e^{-(\beta_i + \beta_j)(T-t)}}{\beta_i + \beta_j}.$$ 

In particular we note,

$$(V_{t,T})_{ii} = \text{Var}_P[X_i(T) \mid \mathcal{F}_t]$$

$$= 1 - e^{-2\beta_i(T-t)} \frac{1}{2\beta_i}.$$

Now let $f(x) = \exp\left(\frac{1}{2}(x - c)^t \theta (x - c)\right)$, where $c$ is a constant vector and $\theta$ is a $d \times d$ symmetric positive definite matrix.

This gives us

$$A(t) = \frac{e^{-\alpha(t)} f(X(t))}{f(X(0))}$$

$$= \exp \left[ -\alpha t + \frac{1}{2} (X(t) - c)^t \theta (X(t) - c) - \frac{1}{2} (X(0) - c)^t \theta (X(0) - c) \right].$$
The bond price process can now be found via

\[ P(t, T) = \frac{E_P[A(T) \mid \mathcal{F}_t]}{A(t)} \]

\[ = \frac{e^{\frac{1}{2}(X(0) - c)'\theta(X(0) - c)}}{e^{-\alpha t} e^{\frac{1}{2}(X(t) - c)'\theta(X(t) - c)}} E_P \left[ \frac{e^{-\alpha T} e^{\frac{1}{2}(X(T) - c)'\theta(X(T) - c)}}{e^{\frac{1}{2}(X(0) - c)'\theta(X(0) - c)}} \mid \mathcal{F}_t \right] \]

\[ = \exp \left[ -\alpha (T - t) - \frac{1}{2}(X(t) - c)'\theta(X(t) - c) \right] E_P \left[ e^{\frac{1}{2}(X(T) - c)'\theta(X(T) - c)} \mid \mathcal{F}_t \right] . \]

Recall that we have

\[ X(T) \mid X(t) \sim \text{MVN}(\mu_{t,T}, V_{t,T}), \]

where \( \text{MVN}(\mu_{t,T}, V_{t,T}) \) represents the multivariate Normal distribution with mean vector \( \mu_{t,T} \) and covariance matrix \( V_{t,T} \).

So one can represent the conditional distribution of \( X(T) \) as the law of \( MZ + \mu_{t,T} \) where \( Z \sim \text{MVN}(0, I) \) is independent of \( \mathcal{F}_t \) and \( M = V_{t,T}^{\frac{1}{2}} \).
So we have

\[
P(t, T) = \exp\left[-\alpha(T - t) - \frac{1}{2} (X(t) - c)'\theta(X(t) - c)\right] \\
\times E_P\left[\exp\left\{\frac{1}{2} (MZ - (\mu_{t,T} - c))'\theta(MZ - (\mu_{t,T} - c))\right\}\mid \mathcal{F}_t\right] \\
= \exp\left[-\alpha(T - t) - \frac{1}{2} (X(t) - c)'\theta(X(t) - c)\right] \\
\times E_P\left[\exp\left\{\frac{1}{2} (Z - M^{-1}(\mu_{t,T} - c))'M'\theta M (Z - M^{-1}(\mu_{t,T} - c))\right\}\mid \mathcal{F}_t\right] \\
= \exp\left[-\alpha(T - t) - \frac{1}{2} (X(t) - c)'\theta(X(t) - c)\right] \\
\times E_P \exp\left[\frac{1}{2} (Z - c_2)'\theta_2(Z - c_2)\right]
\]

where \( \theta_2 = M'\theta M = V_{t,T}^{1/2}\theta V_{t,T}^{1/2} = \theta V_{t,T} \), since \( \theta \) and \( V^{1/2} \) are symmetric, and \( c_2 = M^{-1}(\mu_{t,T} - c) \). Since \( Z \sim N(0, I) \) this can be expressed as

\[
\exp\left[-\alpha(T - t) - \frac{1}{2} (X(t) - c)'\theta(X(t) - c)\right] \\
\times \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \exp\left\{\frac{1}{2} (z - c_2)'\theta_2(z - c_2)\right\} \exp\left[-\frac{1}{2} z'z\right] dz \\
= \exp\left[-\alpha(T - t) - \frac{1}{2} (X(t) - c)'\theta(X(t) - c)\right] \\
\times \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \exp\left[\frac{1}{2} (z'\theta_2 z - c_2'\theta_2 z - z'\theta_2 c_2 + c_2'\theta_2 c_2 - z'z)\right] dz
\]
\[
\begin{align*}
&= \exp \left[ -\alpha(T - t) - \frac{1}{2}(X(t) - c)'\theta(X(t) - c) \right] \\
&\times \frac{e^{\frac{1}{2}c_2^'\theta_2 c_2}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \exp \left[ -\frac{1}{2}(z'(I - \theta_2)z) \right] \exp[-c_2^'\theta_2 z] \, dz \\
&= \exp \left[ -\alpha(T - t) - \frac{1}{2}(X(t) - c)'\theta(X(t) - c) \right] \\
&\times \frac{e^{\frac{1}{2}c_2^'\theta_2 c_2}}{(2\pi)^{d/2}} \sqrt{\det(I - \theta_2)^{-1}} \int_{\mathbb{R}^d} \exp \left[ -\frac{1}{2}(z'(I - \theta_2)z) \right] \exp[-c_2^'\theta_2 z] \, dz \\
&= \exp \left[ -\alpha(T - t) - \frac{1}{2}(X(t) - c)'\theta(X(t) - c) \right] \\
&\times e^{\frac{1}{2}c_2^'\theta_2 c_2} \sqrt{\det(I - \theta_2)^{-1}} E \left[ e^{-c_2^'\theta_2 Z} \right]
\end{align*}
\]

Recalling that the moment generating function for \( X \sim \text{MVN}(0, (I - \theta_2)^{-1}) \) is \( E[e^{tZ}] = \exp \left[ \frac{1}{2}t'(I - \theta_2)^{-1}t \right] \), we have

\[
\frac{e^{-\alpha(T - t)}}{\sqrt{\det(I - \theta_2)^{-1}}} \exp \left[ -\frac{1}{2}(X(t) - c)'\theta(X(t) - c) \right] \\
+ \frac{1}{2}(c_2^'\theta_2 c_2) + (c_2^'\theta_2(I - \theta_2)^{-1}\theta_2 c_2)
\]

Expanding the last two terms in the exponential and returning to our original notation gives
\[
\frac{e^{-\alpha(T-t)}}{\sqrt{\det(I - \theta^2)}} \exp \left[ -\frac{1}{2}((X(t) - c)'\theta(X(t) - c)) 
+ \frac{1}{2}(V_{t,T}^{-\frac{1}{2}}(c - \mu_{t,T}))'\theta V_{t,T}(V_{t,T}^{-\frac{1}{2}}(c - \mu_{t,T})) 
+ \frac{1}{2}(\theta V_{t,T}V_{t,T}^{-\frac{1}{2}}(c - \mu_{t,T}))'(I - \theta V_{t,T})^{-1}(\theta V_{t,T}V_{t,T}^{-\frac{1}{2}}(c - \mu_{t,T})) \right]
\]

\[
= \frac{e^{-\alpha(T-t)}}{\sqrt{\det(I - \theta^2)}} \exp \left[ -\frac{1}{2}((X(t) - c)'\theta(X(t) - c)) 
+ \frac{1}{2}(c - \mu_{t,T})'(\theta + \theta'(I - \theta V_{t,T})^{-1}\theta V_{t,T})(c - \mu_{t,T}) \right]
\]

\[
= \frac{e^{-\alpha(T-t)}}{\sqrt{\det(I - \theta^2)}} \exp \left[ -\frac{1}{2}((X(t) - c)'\theta(X(t) - c)) 
+ \frac{1}{2}(c - \mu_{t,T})'(I + \theta'(I - \theta V_{t,T})^{-1}V_{t,T})\theta(c - \mu_{t,T}) \right]
\]

\[
= \frac{e^{-\alpha(T-t)}}{\sqrt{\det(I - \theta^2)}} \exp \left[ -\frac{1}{2}((X(t) - c)'\theta(X(t) - c)) 
+ \frac{1}{2}(c - \mu_{t,T})'(I - \theta V_{t,T})^{-1}((I - \theta V_{t,T}) + \theta V_{t,T})\theta(c - \mu_{t,T}) \right]
\]

\[
= \frac{e^{-\alpha(T-t)}}{\sqrt{\det(I - \theta^2)}} \exp \left[ -\frac{1}{2}((X(t) - c)'\theta(X(t) - c)) 
+ \frac{1}{2}(c - \mu_{t,T})'(I - \theta V_{t,T})^{-1}\theta(c - \mu_{t,T}) \right]
\]

Observing that \(X(t) = \mu_{t,t}\) and trivially rearranging terms gives
\[ P(t, T) = e^{-\alpha(T-t)} \det(I - \theta V_{t,T})^{-\frac{1}{2}} \times \exp \left[ \frac{1}{2} (\mu_{t,T} - c)' (I - \theta V_{t,T})^{-1} \theta (\mu_{t,T} - c) - \frac{1}{2} (\mu_{t,t} - c)' \theta (\mu_{t,t} - c) \right], \]

which matches the result given without derivation in Rogers (1997).

We will now recover the short-rate process from \( A(t) \).

Let \( Y(t) = \ln A(t) \). Recall that we defined \( dA(t) = A(t) [\mu(t) + \sigma(t)' \hat{W}(t)] \). Using Ito’s formula gives us

\[
\begin{align*}
    dY(t) & = d \ln A(t) \\
    & = \left( \frac{1}{A(t)} \mu(t) A(t) - \frac{1}{2 A^2(t)} \sigma(t)' \sigma(t) A^2(t) \right) dt \\
    & + \frac{1}{A(t)} A(t) \sigma(t) d\hat{W}(t) \\
    & = \mu(t) dt - \frac{1}{2} \sigma(t)' \sigma(t) dt + \sigma(t)' d\hat{W}(t) \quad (2.12)
\end{align*}
\]
Recall that

\[ A(t) = \frac{e^{-\alpha(t) f(X(t))}}{f(X(0))} \]

\[ = \exp \left[ -\alpha t + \frac{1}{2}(X(t) - c)'\theta(X(t) - c) - \frac{1}{2}(X(0) - c)'\theta(X(0) - c) \right], \]

so we have

\[ \ln A(t) = -\alpha t + \frac{1}{2}(X(t) - c)'\theta(X(t) - c) - \frac{1}{2}(X(0) - c)'\theta(X(0) - c), \]

and hence

\[ d\ln A(t) = -\alpha \, dt + \frac{1}{2} d \left( (X(t) - c)'\theta(X(t) - c) \right) - 0 \]

\[ = -\alpha \, dt + \frac{1}{2} d \left( X(t)'\theta X(t) - 2X(t)'\theta c + c'\theta c \right). \]

Evaluating the second term using Ito’s formula gives

\[ d\ln A(t) = -\alpha \, dt + \frac{1}{2} \left( -2X(t)'BX(t) \, dt + tr(\theta C'C) \, dt \right. \]

\[ \left. + 2\theta X(t)C \, d\hat{W}(t) - 2(-BX(t) \, dt + C \, d\hat{W}(t))\theta c \right) \]

\[ = \left( -\alpha + \frac{1}{2} tr(\theta C'C) \right) \, dt + (X(t) - c)'\theta(-BX(t) \, dt + C \, d\hat{W}(t)). \]  

(2.13)
By equating (2.13) with (2.12) we find that

\[ \mu_A(t) = -\alpha + \frac{1}{2} tr(\theta C'C) - (X(t) - c)'\theta BX(t) \]
\[ + \frac{1}{2}(X(t) - c)'\theta CC'\theta(X(t) - c). \]

So we have

\[ r(t) = -\mu_A(t) \]
\[ = \alpha - \frac{1}{2} tr(\theta C'C) + (X(t) - c)'\theta BX(t) \]
\[ - \frac{1}{2}(X(t) - c)'\theta CC'\theta(X(t) - c) \]

\[ = \alpha - \frac{1}{2} tr(\theta C'C) + (X(t) - c)'\theta C'BX(t) - c'\theta BX(t) \]
\[ - \frac{1}{2} X(t)'\theta CC'\theta X(t) + X(t)'\theta CC'\theta c + \frac{1}{2} c'\theta CC'\theta c. \]

We can simplify the expression by completing the square in \( X(t) \) to obtain

\[ r(t) = \alpha - \frac{1}{2} tr(\theta C'C) + \frac{1}{2} c'\theta CC'\theta c \]
\[ + X(t)' \left[ \theta B - \frac{1}{2} \theta CC'\theta \right] X(t) - X(t)'(B - \theta CC')\theta c \]
Now, if we let \( S = (\theta B + B\theta - \theta CC'\theta) \) and \( v = (B - \theta CC')\theta c \) we can factorize our expression for \( r(t) \) as

\[
r(t) = \alpha - \frac{1}{2} \text{tr}(\theta C'C) + \frac{1}{2} c'\theta CC'\theta c + \frac{1}{2} (X(t) - S^{-1}v)'S(X(t) - S^{-1}v) - \frac{1}{2} v' S^{-1}v \tag{2.14}
\]

We arrive at a stochastic process for \( r(t) \) which is quadratic in the underlying state variables \( X(t) \). We are able to choose the minimum value that \( r(t) \) can attain by choosing appropriate values for the parameters in the model. For example, if we choose

\[
\alpha = \frac{1}{2} \text{tr}(\theta CC') + \frac{1}{2} c'\theta CC'\theta c + \frac{1}{2} v' S^{-1}v
\]

then the minimum possible value of \( r(t) \) will be 0.

They key feature of this model in comparison with preceding models is that it is multivariate. This allows for a far richer term structures to be described, as will be explored in the next chapter.

### 2.6 Discussion of examples

Overall the models derived in this chapter reveal several interesting features of the positive interest framework.

The most striking is the ease in which the modeller can create a situation where the conditions which are required to ensure positive interest rates are either overly invasive on the model parameters or simply impossible to satisfy. We observed that in models
two and four this was the case. In these situations nothing is gained from the positive interest approach and the modeller would be just as well off to use a more general approach such as the Heath, Jarrow, Morton (1992) framework to develop a no-arbitrage model for the term structure.

In the situations where the choice of underlying diffusion and function $f(x)$ were clever enough (or lucky enough) to produce more natural and easily satisfied restrictions on the model parameters, the positive interest approach stands out as being a useful framework for being able to build a no-arbitrage style model while still ensuring interest rates remain positive. Models three and five are examples of such situations.

It is also interesting to note the variety of features which can be built into the short-rate process via the positive interest framework. Aside from the obvious non-negativity which we are aiming for, we also saw the possibility of a mean reverting short-rate as well as a short-rate which is quadratic in the underlying state variables.

In terms of computation of bond prices, the biggest challenge (and possible restriction) is in finding the required moments of $A(t)$. To derive closed form expressions for the bond price process we need to be able to derive the moments of potentially peculiar functions of our underlying diffusion process. A combination of a complicated diffusion process with a ‘nasty’ function $f(x)$ may rule out the possibility of a closed form solution altogether.

The models which stood out as being the most successful examples of the positive
interest approach were models three and five. In the next chapter we will investigate both of these formulations numerically.
Chapter 3

Simulations

In this chapter the two successful models from Chapter 2, models three and five, will be investigated numerically via simulation to see what type of term structures they are able to describe. By varying the parameters and starting values of the underlying diffusions which drive the models we are able to generate a variety of term structure scenarios, as well as investigate their short-rate processes numerically.
3.1 Simulations of model 3

3.1.1 The Yield Curve

Figure 1

(a) $\alpha = 0.3, a = 1.3, \gamma = -0.12, \beta = 0.06, X(0) = 50$

(b) $\alpha = 6, a = 4, \gamma = -0.12, \beta = 0.079, X(0) = 0.3$

(c) $\alpha = 6, a = 3, \gamma = -0.2, \beta = -0.01, X(0) = 0.2$

(d) $\alpha = 6, a = 4, \gamma = -0.12, \beta = 0.65, X(0) = 0.5$
Figure 1 demonstrates a sample of the variety of term structures that can be generated by model 3. By varying the model parameters and the initial value of the underlying square-root process a number of realistic term structures are achieved.

Panel (a) resembles a standard upward sloping yield curve. Panel (b) demonstrates that a ‘hump’ can be created in the term structure. Panels (c) and (d) show that an inverted yield curve is also possible, the former quickly assymptoting to a limiting value and the latter receding at a much slower rate. It is also worth noting that in all cases the bulk of the ‘variation’ in the forward rate happens for small T, and the curves tend to stabilize for large T, as is generally the case with real yield curves.

What is not so obvious from figure 1, but became apparent during the creation of the plots are the limitations of the of the model with regard to creating commonly seen term structures. Due to the parameter constraints which ensure we have both a positive short rate and a positive market price of risk, along with the fact that \( X(t) \) following a square-root process can never become negative, means that it is impossible to create a ‘negative hump’ in the term structure (this would appear as the vertical reflection of panel (b)). This is not an uncommon yield curve shape seen in rate markets, where in the short term the curve is inverted but in the longer term the curve takes resumes an upward slope, and as such represents a significant shortcoming of the model.

Also due to the model being univariate, there is a limit to the number of features that can be present in the term structure. We can have at most one hump, and also the yield curve is always above the long term limiting value.
3.1.2 The short-rate

Figure 2

(a) \( \alpha = 0.3, \beta = 0.06, X(0) = 50 \)
(b) \( \alpha = 6, \beta = 0.079, X(0) = 0.3 \)
(c) \( \alpha = 6, \beta = -0.01, X(0) = 0.2 \)
(d) \( \alpha = 6, \beta = 0.65, X(0) = 0.5 \)
Figure 2 shows simulations of the short rate process of model 3, using the same sets of parameters that generated the yield curves in figure 1.

All 4 panels exhibit a strictly positive short-rate process as we expect. The initial values of each short-rate process is equal to the y-intercept from the corresponding yield curve in figure 1, and thereafter the short-rate diffuses according to the function of the square-root process described by (2.5). The amount of volatility in each process and its drift is determined by the choice of parameters in each case.
3.2 Simulations of model 5

3.2.1 The Yield Curve

![Yield Curve Graphs](image)

Figure 3
The yield curves featured in Figure 3 were generated via a 2-dimensional simulation of model 5 with parameters $c = (1, 2)$, $\theta = \begin{bmatrix} -0.8 & -0.1 \\ -0.1 & 0.2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$, $\rho_{12} = 0.4$, $\alpha = 4$, and for (a) $X(0) = (-1, 2)$, (b) $X(0) = (-0.8, -2)$, (c) $X(0) = (-3, 3)$ (d) $X(0) = (1, 0)$.

Figure 3 demonstrates a sample of the yield curves that can be generated by model 5. Similarly to the yield curves in Section 3.1, we are able to generate upward sloping, inverted and humped term structures, but in this case we observe two key improvements.

Firstly, only the $X(0)$ values had to be changed to produce dramatically different term structure models. In an applied modelling context, model parameters are often fit to long term historic data, and are not reset frequently. As opposed to the yield curves in Section 3.1 where we had to alter all the underlying parameters of the model to generate different yield curves, here we are able to retain our set of parameters, and observe markedly different term structures simply by altering the current value of the underlying diffusion process.

Secondly, due to model 5 being multivariate, there is greater richness in the yield curves that can be produced by the model. We observe in Figure 3 that the yield curves are able cross their long term limiting values for small $T$ and then slowly reapproach them for larger $T$. This is a valuable tool for the interest rate modeller, as term structures often exhibit high volatility over short term horizons whilst slowly approaching limiting values in the longer term. It is also worth noting that model 5 can produce both negative and positive humps.
Chapter 4

Concluding Remarks

The investigation of the no-arbitrage positive interest rate framework in this thesis has led to several outcomes.

By developing the framework as constructed in the main articles of Flesaker and Hugheston (1996), Rogers (1997) and Rutkowski (1997) and summarized by Chapter 8 of Cairns (2004) we have seen where the approach fits into the interest rate modelling landscape, we have observed the technical conditions under which the approach produces positive interest rate models, and we have seen how the principle of no-arbitrage pricing is applied to the framework.

By developing five models we have gained an understanding of the challenges in computing bond prices within the framework - in particular the calculation of sometimes unusual moments. We have also noted that the seemingly simple conditions under which the approach ensures positive interest rates can often be extremely hard, and in some cases impossible to meet.
By undertaking numerical analysis of the models which stood out as being the most promising of those developed we have seen the variety of term structures which can be modelled within the framework, as well as the sample dynamics of the short-rate process under those models. We have also seen the increased power and flexiblity that a multivariate interest rate possesses when compared to a univariate model.

4.1 Further Study and Next Steps

The natural next steps for the work presented in this thesis would be to calibrate models three and five to real market data and measure their ability provide explanatory and predictive power in terms of real market data in comparison to other popular interest rate models.
Bibliography


