Abstract

Affine interest rate models are becoming increasingly popular due to their analytical and computational tractability. Affine processes have an explicit closed-form log bond price formula which is a linear function of the initial value of the underlying process. Quadratic processes are, to some extent, an extension of affine models and have similar properties as affine models. This thesis compares these affine and quadratic models on a theoretical and an empirical level. For the theoretical level, this thesis explains the mathematics of affine and quadratic interest rate models. To properly compare the different classes of models, it constructs a similar framework as the well-known affine framework to describe the mathematics of quadratic models [15]. Besides the zero-coupon bond formulas, for both affine and quadratic models analytical forms for derivatives of the short rate (such as call and put options) are provided using admissible parameters and Riccati equations. Also, using the analytical bond prices, an empirical comparison is performed where some computational examples are discussed.
Preface

This thesis finishes the master ‘Stochastics and Financial Mathematics’ at the University of Amsterdam (UvA). For this thesis I wanted to experience the use of mathematics in the financial world and broaden my knowledge to the ‘Financial Mathematics’ part of my studies. RiskQuest gave me the opportunity to perform mathematics in the industry setting by combining an internship with a working student job.

First of all, I want to thank all my RiskQuest colleagues for their support: the colleagues at the office who were caring and gave me daily support, but of course also my RiskQuest supervisor Wouter van Krieken and my former colleague Frank Pardoel, with whom I discussed my thesis in more detail. I also want to thank the other colleagues Corné Ruwaard and Dick de Heus that challenged me content-wise and explained the empirical context to supplement my mostly theoretical background.

Moreover, I would particularly like to thank my supervisor Peter Spreij of the UvA for all his guidance. The mathematics was more sophisticated than I imagined, but every time I got stuck, he offered to help me and motivated me to continue. Also, he provided the opportunity to participate the winter school in Lunteren and meet Damir Filipović, who signed my book on which I had already spent many hours on.

Finally, I owe my family and friends a lot of thanks for their loving support during my whole studies and in particular during the writing of this thesis.
Contents

Introduction 6

1. Short-rate models 9
   1.1. Term-structure equation 9
   1.2. Affine term-structures 10
   1.3. Popular short-rate models 11
      1.3.1. Vasiček model 11
      1.3.2. Cox-Ingersoll-Ross (CIR) model 13
      1.3.3. Hull-White model (extended Vasiček) 14

2. Affine models 16
   2.1. Definition of an affine process 16
   2.2. Canonical state space for Affine processes 18
   2.3. Pricing in Affine models 19
      2.3.1. Fourier transform technique 22
      2.3.2. Price formula of call and put options on a bond 24
      2.3.3. Price formula of cap 25
   2.4. Examples of affine short-rate models 26
      2.4.1. Vasiček short-rate model 26
      2.4.2. CIR short-rate model 30

3. Quadratic models 33
   3.1. Definition of a quadratic process 33
   3.2. Canonical state space for quadratic models 35
   3.3. Pricing with Quadratic models 35
      3.3.1. Price formula for exponential-quadratic payoffs 43
      3.3.2. Fourier transform techniques for quadratic models 44
      3.3.3. Price formula of call and put options on a bond 46
   3.4. Example of quadratic short-rate model 46

4. Comparison of affine and quadratic models 48
   4.1. From quadratic models to affine models 49
   4.2. Riccati equations for rescaled processes 52
      4.2.1. From rescaled quadratic processes to affine processes 52

5. Financial applications of affine and quadratic models 57
   5.1. Short-rate models in practice 57
      5.1.1. Monte Carlo method on Vasiček and CIR 57
      5.1.2. Sensitivity analysis 59
   5.2. Calibration of the models 59
      5.2.1. Data set of Euribor rates 59
      5.2.2. Nonlinear least-squares method for calibration 60
      5.2.3. Calibration accuracy 63
   5.3. Pricing 66
      5.3.1. Zero-coupon bonds 66
      5.3.2. Call and Put options 67
6. Discussion

6.1. Models in practice ................................................. 69

Layman’s summary .................................................. 70

Bibliography .......................................................... 71

A. Basics ............................................................... 73

A.1. Interest rates ..................................................... 73
A.2. Itô’s lemma ....................................................... 73
A.3. Forward measures .............................................. 74
    A.3.1. Forward as numéraire ................................. 74

B. Calculations on popular short-rate models .................... 75

B.1. Vasiček short-rate model ...................................... 75
    B.1.1. Unique global solution of the Vasiček short-rate model ........... 75
    B.1.2. Explicit expression for the forward rate .................. 76

B.2. The CIR short-rate model ..................................... 79
    B.2.1. Unique global solution of the CIR short-rate model ........... 79
    B.2.2. Proof of claim ......................................... 80
    B.2.3. Explicit expression of the forward rate .................. 84

Index ..................................................................... 86
Introduction

Interest rate risk plays an important role in the financial industry. Banks and insurance companies, for example, heavily rely on interest rate risk models for managing risk [23]. There could be gained a lot in modeling and understanding the interest rates, since lately interest rates dropped below zero while this was very unlikely to happen according to most of the models. Also, the level of the interest rate is very important in the world: according to The Bank for International Settlements (BIS), the worldwide amount of debt instruments is no less than $21,288 billion [11]. Moreover, in The Netherlands alone the total amount of mortgage loans outstanding is €650 billion, which equals 95% of the gross domestic product [12]. These examples show that the interest rate level has a high impact, even on a global scale, so it is important to model them properly.

The models for the term-structure of interest rates have been studied broadly for several decades already. In 1977, Vasić [30] published one of the earliest models of the term-structure which lead to the famous Vasić model. It was the start of the introduction of many term-structure models such as the Cox-Ingersoll-Ross model [9] in 1985 and the Hull-white model [21] in 1990.

These models all have a special property that makes the models mathematically appealing: the affine property. This affine property implies an explicit closed-form log bond price formula which is a linear function of the initial value of the underlying process. Hence, for affine models there exists an analytical representation of the bond price that prevents the use of time-consuming numerical methods for derivation.

The class of affine models is a wide class of stochastic interest rate models with nice features that make them analytically tractable. In 2003, Duffie, Filipović and Schachermayer provided in their frequently cited article [15] the definition and a complete characterization of (regular) affine processes. In short, a process in the affine class is a stochastic Markov process. This process is said to be affine if the conditional characteristic function is exponential affine with respect to the initial value (see Chapter 2). Duffie et al [15] show that an affine process on the canonical state space $\mathbb{R}^m \times \mathbb{R}^n$ can be characterized by its admissible parameters and, moreover, by the so-called Riccati equations used in the representation of the conditional characteristic function. A modified form of these Riccati equations determine also the bond price.

A short-rate model is affine if it is a linear combination of an affine state space process. These affine short-rate models have promising empirical performance [1, 10] and are becoming increasingly important due to their computational tractability [15, 18]. However, the affine term structure models turn out to have shortcomings in flexibility and also, the affine term structure models do not allow nonlinearity [1, 7]. Therefore, quadratic term structure models are considered; these models might be better in modeling and maybe also in predicting interest rate risk.

Quadratic short-rate models are part of the set of polynomial models, which are an extension of the affine short-rate models. The polynomial models increasingly play an important role in finance [17]. For a polynomial model to be arbitrage-free, it can only have maximal degree of two; hence the consideration of the quadratic models [17]. Instead of only a linear combination, these models also allow an extra quadratic term of the state variable in the expression for the short-rate. For these quadratic short-rate models similar properties hold as for the affine models. For example, for quadratic models one can also characterize the process with admissible parameters and Riccati equations, although there are less restrictions
for this model. The pricing function of the zero-coupon bond changes to an expression with an extra quadratic term. Also, like the affine setting, for one-factor quadratic models the coefficients of this pricing function can be solved analytically.

Together with the fact that the additional quadratic term enables more sophisticated dynamics than the linear combination in the affine term structures [24], the similarity with the affine models on theoretical basis makes the quadratic short-rate models mathematically very interesting. Moreover, in earlier literature it has also been proven that quadratic term structure models can outperform the affine term structure models, while these quadratic models still have a good analytical tractability comparable to that of the affine term structure models [1, 7, 8, 25].

While affine models are already broadly studied, quadratic models did not have that much attention in research. In 2000, Leippold and Wu [24] mention that this lack of research could be due to the fact that for quadratic models nobody before them had identified and characterized the complete quadratic class as Duffie and Kan [16] did in 1996 for the affine class. Therefore, this thesis aims to combine the results of Leippold and Wu [24, 25] with the results of Chen et al. [7] to have a similar theoretical framework as for the affine models described in [16, 15, 18].

This thesis considers both the affine and quadratic interest rate models and explains the similarities and the differences. A theoretical and empirical study is conducted on both models, which lead to a thorough overview of the affine and quadratic models. To finish the study, also a small empirical comparison is performed in Matlab using Euribor rates from De Nederlandsche Bank (DNB) [13].

Since both models need quite sophisticated mathematical concepts, this thesis starts in Chapter 1 with a short recap of the theory of short rate models. In this chapter the affine term-structures are introduced and also three of the most famous short-rate models, the Vasiček model, the CIR model and the Hull-White model, and their characteristics are considered.

In Chapter 2, the theory of affine term-structures will be continued by studying affine models. The mathematical concepts of affine models will be explained in detail. This first includes the formal definition, theorems that state whether a model is affine in general (Section 2.1) and when they are defined on the canonical state space $\mathbb{R}_m^+ \times \mathbb{R}^n$ (Section 2.2). For the latter, the admissibility conditions of the parameters, $a, \alpha, b, \beta$, and the system of Riccati equations with solutions $\varphi$ and $\psi$, will be stated in Theorem 2.2.1; these concepts are the key elements of affine models. Secondly, the use of affine models in pricing derivatives is discussed in Section 2.3, which leads to a pricing formula for zero-coupon bonds in terms of an adjusted version of the previous mentioned system of Riccati equations with adjusted solutions $\Phi$ and $\Psi$ (see Theorem 2.3.1 and Corollary 2.3.2). Followed by that, the price formula for calls, puts, caps and floors are determined in Sections 2.3.2 and 2.3.3 by using Fourier transform techniques discussed in Section 2.3.1. The chapter concludes in Section 2.4 with pricing derivatives for two examples of affine short-rate models: the Vasiček model and the CIR model.

Chapter 3 discusses the other researched models: the quadratic models. This chapter has the same structure as the chapter about affine models and therefore starts with the formal definition of a quadratic process and a definition of an admissible parameter set, and states a theorem about the sufficient conditions for $X$ to be a quadratic process, in terms of admissibility conditions, $a, b, \beta$, and the system of Riccati equations with solutions $\varphi$, $\psi$ and $\omega$. In Section 3.3 the theory of quadratic models is continued in the pricing context. As for the affine model, here also a pricing formula is given for the zero-coupon bond in terms of adjusted versions of the system of Riccati equations with adjusted solutions $\Phi, \Psi, \Omega$. In addition, closed-form expressions are given for the Riccati equations for dimension one.

In the fourth chapter the analytical comparison is done for the models. In this chapter one can read that affine models are not simply a subset of quadratic models. Only if the
underlying process of an affine model is an Ornstein-Uhlenbeck process, then the model also satisfies the conditions of a quadratic model. This implies that the Riccati equations for that affine model according to the affine framework should be the same as the Riccati equations according to the quadratic framework. Therefore, to get more insight in the Riccati equations of the quadratic model, in this chapter the Riccati equations corresponding to the affine framework in case of an OU-process are derived from those of the quadratic model.

In practice, also the market price of risk is included in the models. Since in this thesis (almost) everything is done under the risk neutral measure, this market price of risk is assumed to be zero. However, as Leippold and Wu did in [25], one could also give the quadratic framework for rescaled processes that incorporate the market price of risk. For this process also the Riccati equations change. This is considered at the end of Chapter 4.

The last chapter shows the results of the empirical study for this thesis. It starts with an explanation of Monte Carlo methods applied for the Vasiček and CIR model. Then, after a short note on the sensitivity of such a method, it is explained how the models are calibrated to the Euribor data, supported by some graphs of the estimated parameters and the error of the calibration. By explanation, two samples are taken from the calibration on which are zoomed in by considering short-rate sample for the two sets of calibrated parameters. The chapter finalizes with a section about pricing with affine and quadratic models. In that section, zero-coupon bonds, and call and put options are priced and discussed using different models.

In conclusion, this thesis finalizes with a view on the comparison of the two different kind of models, both on theoretical and on computational level. In addition, a discussion is added about the use of these model in practice.
1. Short-rate models

In the financial industry, short-rate models are broadly used. The short rates models are used to model the behavior of the interest rates over time, and interest rates are crucial in valuating financial products [23]. A short rate, $r_t$, is also referred to as the instantaneous short rate, which cannot be directly observed but is fundamental to no-arbitrage pricing [18]. A model for a short rate consists of a drift and a diffusion term. The rate follows from the continuously compounded forward rate $R(t; T, S)$ via the continuously compounded spot rate (see Appendix A.1).

This chapter gives a short recap of the theory of short rate models needed for the rest of this thesis. The affine term-structures are introduced and also three of the most famous short-rate models, the Vasiček model, the CIR model and the Hull-White model, and their characteristics are considered.

Note that in this chapter it is assumed that the reader has some background knowledge about for example Itô processes, equivalent martingale measures, change of measures, etc. More reading material for this knowledge can for example consist of [3], [20], [22] or [27].

1.1. Term-structure equation

Consider an asset $B(t)$ which moves instanteneously with short rate $r_t$ satisfying

$$dB(t) = r_t B(t) dt, \quad B(0) = 1, \quad \text{or equivalently} \quad B(t) = e^{\int_0^t r(s) ds}.$$  

This is called a money-market account [18]. The short rate determining this money-market account is assumed to follow an Itô process

$$dr_t = b(t) dt + \sigma(t) dW_t,$$

with $W$ a $d$-dimensional Brownian motion with respect to a probability measure $P$.

In all the following, arbitrage opportunities should be avoided. In order to achieve this, assume that there exists an equivalent martingale measure $Q$. Then, according to the first fundamental theorem of asset pricing ([3, Proposition 10.8] and [18, Lemma 4.6]), the model is arbitrage-free. By the equivalent martingale measure $Q$, the bond price process $P(t, T)$ with $t \leq T$, discounted with the money-market account, $P(t, T)/B(t)$, is a $Q$-martingale, with $P(T, T) = 1$. The Girsanov’s Change of Measure Theorem (see [18, Theorem 4.6]) implies that the martingale measure $Q$ is such that the Radon-Nikodym derivative on $\mathcal{F}_t$ is of the form $dQ/dP |_{\mathcal{F}_t} = E_t(\gamma \bullet W)$, with $E(\cdot)$ the stochastic exponential [18].

Since $Q$ is an equivalent martingale measure, the bond price satisfies

$$P(t, T) = B(t)E_Q \left[ \frac{P(T, T)}{B(T)} \right] = E_Q \left[ B(t) \right] = E_Q \left[ e^{-\int_t^T r(s) ds} \right] .$$

Thus, a short-rate model $r_t$ fully describes bond prices for different maturities, referred to as the term structure. Hence, in order to specify this bond price, one needs to model the short-rate $r_t$ and find a way to derive the bond price. The following lemma, also referred to as the Feynman-Kač stochastic representation formula [18, Lemma 5.1], will be proven to be helpful in the derivation of this bond price.
Lemma 1.1.1. Let $T > 0$ and $\Phi$ be a continuous function on a closed interval with non-empty interior, $Z \subset \mathbb{R}$, and assume that $F = F(t, r) \in C^{1,2}([0, T] \times Z)$ is a solution to the boundary value problem on $[0, T] \times Z$

\[
\frac{\partial}{\partial t}F(t, r) + b(t, r)\frac{\partial}{\partial r}F(t, r) + \frac{1}{2}\sigma^2(t, r)\frac{\partial^2}{\partial r^2}F(t, r) - rF(t, r) = 0,
\]

\[F(T, r) = \Phi(r).\]  

Then $M(t) = F(t, r_t)e^{- \int_0^t r_u \, du}$, $t \leq T$, is a local martingale. If in addition either:

(a) $\mathbb{E}_Q \left[ \int_0^T |\partial_r F(t, r_t)e^{- \int_0^t r_u \, du} \sigma(t, r_t)|^2 \, dt \right] < \infty$, or

(b) $M$ is uniformly bounded,

then $M$ is a true martingale, and

\[F(t, r_t) = \mathbb{E}_Q \left[ e^{- \int_0^t r_u \, du} \Phi(r_T) \mid \mathcal{F}_t \right], \quad t \leq T.\]  

This lemma implies that, assuming all the necessary conditions of the lemma, in particular for the constant function $\Phi = 1$, $F(t, r_t) = \mathbb{E}_Q \left[ e^{- \int_0^t r_u \, du} \Phi(r_T) \mid \mathcal{F}_t \right]$ for $t \leq T$. In other words,

\[F(t, r_t) = P(t, T).\]

Therefore, if the $F(t, r_t)$ that satisfies the term-structure equation (1.1) is determined, the bond price is also determined. Also, this $F$ only depends on $t$ and $r_t$. However, finding a solution of the boundary value problem is often complicated. To make it more feasible to solve the boundary problem, it would be favorable to impose more restrictions on one of the parameters $t$ or $r_t$. One way of doing this is, is by a restriction on $r_t$ by only considering short-rate models that admit closed-form solutions; hence a closed form of the price function.

1.2. Affine term-structures

A specific class of short-rate models that admit closed-form expressions of the implied bond price, is the class that implies an affine term-structure (ATS). This is defined as follows [18].

Definition 1.2.1. Models are said to provide an affine term-structure (ATS) are models with bond prices of the following form

\[P(t, T) = e^{-A(t, T) - B(t, T)r_t},\]

where $A$ and $B$ are deterministic functions [3, Definition 24.1].

Note that it follows that also the expression of the forward rate is known: by definition of $f(t, T)$, (A.3), the following is obtained

\[f(t, T) = -\frac{\partial \log P(t, T)}{\partial T} = -\frac{\partial \log e^{-A(t, T) - B(t, T)r_t}}{\partial T} = -\frac{\partial (-A(t, T) - B(t, T)r_t)}{\partial T} = \partial_T A(t, T) + \partial_T B(t, T)r_t.\]

The following proposition of [18] proves that the affine term-structure models can be completely characterized.
Proposition 1.2.2. The short-rate model with the stochastic differential equation

\[ dr_t = (b(t_0 + t, r_t) dt + \sigma(t_0 + t, r_t) dW^*, \quad r(0) = r_{t_0}, \]  

(1.5)

provides an affine term structure if and only if the diffusion and drift terms are of the form

\[ \sigma^2(t, r) = a(t) + \alpha(t) r \quad \text{and} \quad b(t, r) = b(t) + \beta(t) r, \]  

(1.6)

for some continuous functions \( a, \alpha, b, \beta \), and the functions \( A \) and \( B \) as in Definition 1.2.1 satisfy the system of ordinary differential equations,

\[ \partial_t A(t, T) = \frac{1}{2} a(t) B^2(t, T) - b(t) B(t, T), \quad A(T, T) = 0 \]  

(1.7)

\[ \partial_t B(t, T) = \frac{1}{2} \alpha(t) B^2(t, T) - \beta(t) B(t, T) - 1, \quad B(T, T) = 0 \]  

(1.8)

for all \( t \geq T \) [18, Proposition 5.2].

Later in this chapter, it will be shown that one can state useful mathematics about affine term-structures that simplify the pricing algorithm for financial derivatives. Thus, the above proposition that can conclude whether a model has an affine term-structure, is very useful.

1.3. Popular short-rate models

In the following section three popular short-rate models are considered. Particularly these models are introduced, since they will turn out to be affine and will therefore be used later again in this thesis. The first two models, the Vasiček model (1977) and the Cox, Ingersoll and Ross model (1985), are so-called time-homogeneous models, which means that the assumed short rate dynamics only depend on constant coefficients, i.e. the drift and diffusion coefficients, [5]. The third model is an extension of the Vasiček model which is not time-homogeneous anymore. This is the extended Vasiček model: the Hull-White model (1990).

For modeling the short rate, one has to choose which short-rate model to use such that the parameters can be calibrated. There is a variety of models to choose from and they all have their own specialties. This choice can be based on for example if a model is analytically tractable, if it is mean reverting, hence whether the expected value of the short rate tends to a constant value for time \( t \) to infinity, while the variance does not explode and if it is suited for Monte Carlo simulation [5]. Below, the advantages and disadvantages per model will be discussed. For simplicity, all models are given with respect to the risk-neutral measure and therefore the Brownian motion under the risk-neutral measure, \( W \), will be used in the differential equations.

1.3.1. Vasiček model

As noted in the introduction, the Vasiček model is a time-homogeneous model. The great advantage of this model, is that the \( r_t \) is Gaussian distributed with easily obtainable mean and variance. Also, this linear model can be solved explicitly, where the stochastic differential equation is given by

\[ dr_t = (b + \beta r_t) dt + \sigma dW_t, \]

In order to determine its solution, consider the process \( Y_t = e^{-\beta t} r_t \), then applying Itô’s lemma (see A.2.1) gives

\[ dY_t = e^{-\beta t} dr_t - \beta e^{-\beta t} r_t dt \]

\[ = e^{-\beta t} ((b(t) + \beta r_t) dt + \sigma dW_t) - \beta e^{-\beta t} r_t dt \]

\[ = b(t) e^{-\beta t} dt + e^{-\beta t} \sigma dW_t, \]
thus
\[ r_t = e^{\beta t} Y_t \]
\[ = e^{\beta t} r(0) + \int_0^t b e^{\beta(t-s)} ds + \int_0^t e^{\beta(t-s)} \sigma dW_s \]
\[ = e^{\beta t} r(0) + \frac{b}{\beta} (e^{\beta t} - 1) + \int_0^t e^{\beta(t-s)} \sigma dW_s. \]

Hence \( r_t \) is a Gaussian process with
\[
\mathbb{E}_Q[r_t] = \mathbb{E}_Q \left[ e^{\beta t} r(0) + \frac{b}{\beta} (e^{\beta t} - 1) + \int_0^t e^{\beta(t-s)} \sigma dW_s \right]
\[ = e^{\beta t} r(0) + \int_0^t b e^{\beta(t-s)} ds + \sigma e^{\beta t} \mathbb{E}_Q \left[ \int_0^t e^{-\beta s} dW_s \right] \]
\[ = e^{\beta t} r(0) + \frac{-b}{\beta} e^{\beta t} \left[ \frac{1}{\beta} - 1 \right] \]
\[ = e^{\beta t} r(0) - \frac{b}{\beta} (1 - e^{\beta t}), \]

where (*) holds, since the stochastic integral \( \int_0^t e^{\beta(t-s)} \sigma dW_s \) is normally distributed with mean zero. The variance satisfies
\[
\text{Var}_Q[r_t] = \text{Var}_Q \left[ e^{\beta t} r(0) + \frac{b}{\beta} (e^{\beta t} - 1) + \int_0^t e^{\beta(t-s)} \sigma dW_s \right]
\[ = \sigma^2 e^{2\beta t} \text{Var}_Q \left[ \int_0^t e^{-\beta s} dW_s \right]
\[ = \sigma^2 e^{2\beta t} \mathbb{E}_Q \left[ \left( \int_0^t e^{-\beta s} dW_s \right)^2 \right]
\[ = \sigma^2 e^{2\beta t} \mathbb{E}_Q \left[ \int_0^t e^{-2\beta s} ds \right]
\[ = \sigma^2 e^{2\beta t} \int_0^t e^{-2\beta s} ds
\[ = \frac{\sigma^2}{2\beta} (e^{2\beta t} - 1) . \]

where (**) holds by the Itô isometry.

For \( \beta < 0 \), one sees that this model is mean reverting: when \( t \) goes to infinity, the term \( e^{\beta t} \) vanishes, so the short rate is pulled to a level \(-\frac{b}{\beta}\) at rate \(-\beta\), [20]. Until recently, a major drawback of the model was that under this model the rates can assume negative values with positive probability [5]. Now however, with the negative interest rates, this model can be very useful.

In order to determine (1.7) and (1.8), the diffusion and drift terms are written below
\[ \sigma^2(t, r) = \sigma^2 \quad \text{and} \quad b(t, r) = b + \beta r, \]
thus (1.7) and (1.8) are given by
\[
\partial_t A(t, T) = \frac{\sigma^2}{2} B^2(t, T) - bB(t, T), \quad A(T, T) = 0
\]
\[ \partial_t B(t, T) = -\beta B(t, T) - 1, \quad B(T, T) = 0. \]
Because the differential equation for $B(t, T)$ is linear and because of the condition $B(T, T) = 0$, one can write down $B(t, T)$ at once. The explicit solution of these differential equations are, as derived in [18],

$$B(t, T) = \frac{1}{\beta} \left( e^{\beta(T-t)} - 1 \right),$$

$$A(t, T) = A(T, T) - (A(t, T) - A(T, T)) = A(T, T) - \int_t^T \partial_s A(s, T) ds$$

$$= - \int_t^T \frac{\sigma^2}{2} B^2(s, T) - b B(s, T) ds$$

$$= - \int_t^T \frac{\sigma^2}{2} \left( \frac{1}{\beta} \left( e^{\beta(T-s)} - 1 \right) \right)^2 - b \left( \frac{1}{\beta} \left( e^{\beta(T-s)} - 1 \right) \right) ds$$

$$= \frac{\sigma^2}{4 \beta^3} \left( 4 e^{\beta(T-t)} - e^{2\beta(T-t)} - 2\beta(T-t) - 3 \right) + b e^{\beta(T-t)} \frac{1}{\beta^2} - 1 - \beta(T-t).$$

Thus, the Vasiček model has a closed-form solution for bond prices given by (1.4) with $A(t, T)$ and $B(t, T)$ as in (1.10).

### 1.3.2. Cox-Ingersoll-Ross (CIR) model

The next model is the Cox-Ingersoll-Ross model. The advantage of this model with respect to the Vasiček model is that this model does not allow negative rates, since it has a square root term of $r_t$ in the diffusion term. By this non-negativity and by the property that $r$ is also mean reverting in the CIR model, this model is broadly used [18]. The dynamics of $r$ are given by

$$dr_t = (b + \beta r_t)dt + \sigma \sqrt{r_t} dW_t, \quad r(0) \geq 0.$$

Since the diffusion coefficient is dependent on $r_t$, the distribution of $r_t$ is not the normal distribution as in the Vasiček model. In a similar way as done with the Vasiček model,

$$r_t = e^{\beta t} r(0) + \frac{b}{\beta} \left( e^{\beta t} - 1 \right) + \int_0^t e^{\beta(t-s)} \sigma \sqrt{r_s} dW_s.$$

Thus $r_t$ has an expectation of

$$\mathbb{E}_Q[r_t] = \mathbb{E}_Q \left[ e^{\beta t} r(0) + \frac{b}{\beta} \left( e^{\beta t} - 1 \right) + \int_0^t e^{\beta(t-s)} \sigma \sqrt{r_s} dW_s \right]$$

$$= e^{\beta t} r(0) + \int_0^t be^{\beta(t-s)} ds + \sigma e^{\beta t} \mathbb{E}_Q \left[ \int_0^t e^{-\beta s} \sqrt{r_s} dW_s \right]$$

$$= e^{\beta t} r(0) + \int_0^t be^{\beta(t-s)} ds$$

$$= e^{\beta t} r(0) + \left[ \frac{b}{\beta} e^{\beta(t-s)} \right]_0^t$$

$$= e^{\beta t} r(0) - \frac{b}{\beta} (1 - e^{\beta t}).$$
where \((\ast)\) holds, since the stochastic integral \(\int_0^t e^{\beta(t-s)} \sigma \sqrt{r_s} dW_s\) is normally distributed with mean zero. The variance satisfies

\[
\text{Var}_Q[r_t] = \text{Var}_Q\left[e^{\beta t}(0) + \frac{b}{\beta}(e^{\beta t} - 1) + \int_0^t e^{\beta(t-s)} \sigma \sqrt{r_s} dW_s\right]
\]

\[
= \sigma^2 e^{2\beta t} \text{Var}_Q\left[\int_0^t e^{-\beta s} \sqrt{r_s} dW_s\right]
\]

\[
= \sigma^2 e^{2\beta t} \mathbb{E}_Q\left[\left(\int_0^t e^{-\beta s} \sqrt{r_s} dW_s\right)^2\right]
\]

\[
= \sigma^2 e^{2\beta t} \int_0^t e^{-2\beta s} \mathbb{E}_Q[r_s] ds
\]

\[
= \sigma^2 e^{2\beta t} \int_0^t e^{-2\beta s} [e^{\beta s}(0) - \frac{b}{\beta}(1 - e^{\beta s})] ds
\]

\[
= \sigma^2 e^{2\beta t} \left(-\frac{b}{\beta} \left(-\frac{1}{2\beta} + \frac{1}{2\beta} e^{-\beta t} + \frac{1}{2\beta} e^{-\beta s} - \frac{1}{2\beta} - \frac{1}{\beta}\right) + r_0 \left(-\frac{1}{\beta} e^{-\beta t} + \frac{1}{\beta}\right)\right)
\]

where \((\ast\ast)\) holds by the Itô isometry.

The function \(A\) and \(B\) can be easily given, by just filling in the right terms in (1.7) and (1.8)

\[
\begin{align*}
\partial_t A(t, T) &= -bB(t, T), \quad A(T, T) = 0, \\
\partial_t B(t, T) &= \frac{\sigma^2}{2} B^2(t, T) - \beta B(t, T) - 1, \quad B(T, T) = 0. \\
\end{align*}
\]

(1.11)

Unlike the Vasiček model, the differential equation for \(B(t, T)\) is nonlinear, so it is harder to find the solution. However, the equation for \(\partial_t B(t, T)\) is a known equation (a Riccati equation) of which the solution is known. Filipović [18] states that for \(\gamma = \sqrt{\beta^2 + 2\sigma^2}\)

\[
B(t, T) = \frac{2(e^\gamma(T-t) - 1)}{(-\gamma - \beta)(e^\gamma(T-t) - 1) + 2\gamma},
\]

(1.12)

\[
A(t, T) = A(T, T) - \int_T^t \partial_s A(s, T) ds = \int_T^t bB(s, T) ds.
\]

Thus, also CIR model has a closed-form solution for bond prices given by (1.4) with \(A(t, T)\) and \(B(t, T)\) as above.

### 1.3.3. Hull-White model (extended Vasiček)

The last model is the Hull-White model (1990), given as an extension of the Vasiček model, where \(b\) also depends on \(t\) [18]; the diffusion term in this model, \(b_t\), is time-dependent. Hull and White introduced this model as one of the first models that can fit the currently-observed yield curve [5]; the function \(b_t\) is chosen to match the initial forward curve. As for the Vasiček model, this model implies a Gaussian distribution of the interest rate and is analytically tractable, but also allows negative interest rates. According to Brigo [5], this model is historically spoken one of the most important interest-rate models, because of the property that it can be fitted to the observed curve. Also, it is broadly used in risk-management.

The dynamics of the extended Vasiček Hull-White model are given by

\[
\begin{align*}
dr_t &= \left(b_t + \beta r_t\right) dt + \sigma dW_t,
\end{align*}
\]
where $b$ is a deterministic function of time such that it is fitted to the observed term structure, and $\beta$ and $\sigma$ are constants. Thus, following [18] gives

$$
\begin{align*}
\partial_t A(t, T) &= \frac{\sigma^2}{2} B^2(t, T) - b_t B(t, T), \quad A(T, T) = 0, \\
\partial_t B(t, T) &= -\beta B(t, T) - 1, \quad B(T, T) = 0.
\end{align*}
$$

(1.13)

So

$$
B(t, T) = \frac{1}{\beta} (e^{\beta(T-t)} - 1),
$$

$$
A(t, T) = A(T, T) - \int_t^T \frac{\sigma^2}{2} B^2(s, T) - b_s B(s, T) ds
$$

(1.14)

The initial forward curve corresponding to these expressions is then given by

$$
f_0(T) = \partial_T A(0, T) + \partial_T B(0, T)r(0)
$$

$$
= \frac{\sigma^2}{2} \int_0^T \partial_s B^2(s, T) ds + \int_0^T b_s \partial_T B(s, T) ds + \partial_T B(0, T)r(0)
$$

$$
= -\frac{\sigma^2}{2\beta^2} (e^{\beta T} - 1)^2 + \int_0^T b_s e^{\beta(T-s)} ds + e^{\beta T} r(0),
$$

where

$$
\partial_T \varphi(T) = \beta \varphi(T) + b_T, \quad \varphi(0) = r(0).
$$

Thus, it follows that

$$
b_T = \partial_T \varphi(T) - \beta \varphi(T)
$$

$$
= \partial_T \left( f_0(T) - \frac{\sigma^2}{2\beta^2} (e^{\beta T} - 1)^2 \right) - \beta \left( f_0(T) - \frac{\sigma^2}{2\beta^2} (e^{\beta T} - 1)^2 \right).
$$
2. Affine models

As already pointed out in Chapter 1 about short-rate models, the models which induce a closed-form formula of the bond price are favorable. One of them, the affine term-structures, also induce this closed-form and have nicer analytical properties, which will be discussed in this chapter.

This chapter studies the mathematical concepts of affine models in detail. This involves the definition of an affine model and the introduction of admissibility conditions and Riccati equations. After a change of measure, the affine models are considered in perspective of pricing. For this, Fourier transform techniques lead to explicit pricing formulas in the Vasicek and CIR model for derivatives of the zero-coupon bond: calls, puts and floors.

The class of affine processes is used a lot in finance; this chapter makes a start in the use of affine models in finance, but there will be elaborated on this more in a later chapter.

First, affine models will be studied by considering the characteristic function. This method is used frequently in the literature [3, 5, 15, 18].

2.1. Definition of an affine process

Let $X \subset \mathbb{R}^d$ be a closed state space with non-empty interior and fixed dimension $d \geq 1$. It is assumed that the stochastic process $X$, which has values in $X$, satisfies the following stochastic differential equation

$$dX_t = b(X_t)dt + \rho(X_t)dW_t, \quad X_0 = x,$$

(2.1)

where $W$ denotes a $d$-dimensional Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, and $b : X \rightarrow \mathbb{R}^d$ continuous and $\rho : X \rightarrow \mathbb{R}^{d \times d}$ measurable such that the diffusion matrix $a(x) = \rho(x)\rho(x)^\top$ is continuous in $x \in X$.

**Definition 2.1.1.** A process $X$ is affine if the $\mathcal{F}_t$-conditional characteristic function of $X_T$ is exponential affine in $X_t$, for all $t \leq T$. Exponential affine means that there exist $\mathbb{C}$- and $\mathbb{C}^d$-valued functions $\varphi(t, u)$ and $\psi(t, u)$, respectively, with jointly continuous $t$-derivatives such that $X = X^x$ satisfies

$$\mathbb{E}_\mathbb{P} \left[ e^{u^\top X_T} \mid \mathcal{F}_t \right] = e^{\psi(T-t, u) + \psi(T-t, u)^\top X_t}$$

(2.2)

for all $u \in \mathbb{R}^d$, $t \leq T$ and $X_0 = x$, where $x \in X$, [15, 18].

The following theorem states under what conditions a process $X$ is affine.

**Theorem 2.1.2.** Suppose $X$ is affine, then the diffusion matrix $a(x)$ and the drift $b(x)$ are affine in $x$. Thus,

$$a(x) = a + \sum_{i=1}^{d} x_i \alpha_i,$$

$$b(x) = b + \sum_{i=1}^{d} x_i \beta_i = b + Bx$$

(2.3)
Conversely, suppose the diffusion matrix $a(x)$ and drift $b(x)$ are affine of the form (2.3) and suppose there exists a solution $(\varphi, \psi)$ of the Riccati equations (2.4) such that $\varphi(t, u) \in \mathbb{R}^2$ and $x$ has a non-positive real part for all $t \geq 0$, $u \in i\mathbb{R}^d$ and $x \in \mathcal{X}$. Then $X$ is affine with conditional characteristic function (2.2), [18].

Proof. ‘$\Rightarrow$’ Suppose $X$ is affine. Define the complex-valued Itô process

$$M(t) = e^{\varphi(T-t,u)+\psi(T-t,u)^T}X_t,$$

for $T > 0$ and $u \in i\mathbb{R}^d$. Then this process is a martingale, since for all $t \leq T$

$$
\mathbb{E}[M(T) \mid F_t] = \mathbb{E}[e^{\varphi(T-T,u)+\psi(T-T,u)^T}X_T \mid F_t] = \mathbb{E}[e^{\varphi(T-t,u)+\psi(T-t,u)^T}X_T \mid F_t],
$$

where it is used that $X$ is affine, thus according to Definition 2.1.1 $(*)$ holds for since

$$\psi(0,u) = u \text{ and } (**) \text{ holds for every } u \in i\mathbb{R}^d \text{ and } t \leq T.$$

Applying Itô’s lemma (Theorem A.2.1) on the process $M(t)$ gives

$$
dM(t) = \left( \frac{\partial M(t)}{\partial X_t}b(X_t) + \frac{\partial M(t)}{\partial t} + \frac{1}{2} \frac{\partial^2 M(t)}{\partial X_t^2} \rho(X_t)^2 \right) dt + \frac{\partial M(t)}{\partial X_t} \rho(X_t) dW_t
$$

$$
= \left[ \varphi(T-t,u)^T M(t)b(X_t) + \left( \partial_t \varphi(T-t,u) + \partial_t \psi(T-t,u)^T X_t + \psi(T-t,u)^T \partial_t X_t \right) M(t) + \frac{1}{2} \varphi(T-t,u)^T \psi(T-t,u)^T \varphi(T-t,u) \right] dt
$$

$$
+ (T-t,u)^T M(t) \rho(X_t) dW_t
$$

$$
= \left[ 2\psi(T-t,u)^T b(X_t) - \partial_t \varphi(T-t,u) - \partial_t \psi(T-t,u)^T X_t + \psi(T-t,u)^T \partial_t X_t + \frac{1}{2} \psi(T-t,u)^T \psi(T-t,u) \rho(X_t)^2 \right] M(t) dt
$$

$$
+ (T-t,u)^T M(t) \rho(X_t) dW_t.
$$

(2.6)
Since $M(t)$ is a martingale, the $dt$-part of (2.6) equals zero. Hence, letting $T - t \to t$, note that this is allowed since there is continuity in $t$ of $\varphi$ and $\psi$,

$$2\psi(t,u)^\top b(x) - \partial_t \varphi(t,u) - \partial_u \psi(t,u)^\top x + \frac{1}{2} \psi(t,u)^\top a(x)\psi(t,u) = 0$$

$$\Leftrightarrow \partial_t \varphi(t,u) + \partial_u \psi(t,u)^\top x = 2\psi(t,u)^\top b(x) + \frac{1}{2} \psi(t,u)^\top a(x)\psi(t,u)$$

for all $x \in X, t \geq 0, u \in \mathbb{R}^d$, where it is used that $X(0) = x$. Then since $\psi(0,u) = u$ and $b$ are affine of the form (2.3) for parameters $a, \alpha, b, \beta$, (2.4) is obtained

$$\partial_t \varphi(t,u) + \partial_u \psi(t,u)^\top x = 2\psi(t,u)^\top b(x) + \frac{1}{2} \psi(t,u)^\top a(x)\psi(t,u)$$

$$\Leftrightarrow \partial_t \varphi(t,u) = \frac{1}{2} \psi(t,u)^\top a\psi(t,u) + b^\top \psi(t,u)$$

$$\Leftrightarrow \partial_t \psi_i(t,u)^\top = \frac{1}{2} \psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad 1 \leq i \leq d.$$ 

'\Leftrightarrow' Suppose the diffusion and drift are affine in $x$ and let $(\varphi, \psi)$ be a solution of the Riccati equations (2.4) such that $\varphi(t,u) + \psi(t,u)^\top x$ has a non-positive real part for all $t \geq 0, u \in \mathbb{R}^d$ and $x \in X$. Then it should be proved that $X$ is affine. For this, note that $M$ defined by (2.5) is uniformly bounded, since $\text{Re}(\varphi(T - t, u) + \psi(T - t, u)^\top X_t) \leq 0$. Moreover, since $M$ is also a locally bounded martingale, $M$ is a martingale. Therefore, $\mathbb{E}[M(T) \mid F_t] = M(t)$ for all $t \leq T$, hence $\mathbb{E} \left[ e^{\psi X_T} \mid F_t \right] = e^{\psi(T - t, u) + \psi(T - t, u)^\top X_t}$ and $X$ is affine.

\[\square\]

### 2.2. Canonical state space for Affine processes

From now one, consider the (canonical) state space $X = \mathbb{R}_+^m \times \mathbb{R}^n$. Also, define the index sets

$$I = \{1, \ldots, m\} \quad \text{and} \quad J = \{m + 1, \ldots, m + n\}. \quad (2.7)$$

The following theorem gives admissibility conditions for the drift and diffusion matrix in the state space $X$ as defined above. These admissibility conditions will turn out to be very useful in the theory of affine models.

**Theorem 2.2.1.** The process $X$ on the canonical space $\mathbb{R}_+^m \times \mathbb{R}^n$ is affine if and only if $a(x)$ and $b(x)$ are affine of the form (2.3) for parameters $a, \alpha, b, \beta$, which are admissible in the following sense:

- $a, \alpha_i$ are symmetric positive semi-definite,
- $a_{II} = 0$,
- $\alpha_j = 0$ for all $j \in J$,
- $\alpha_{i,kl} = \alpha_{i,lk} = 0$ for $k \in I \setminus \{i\}, 1 \leq i, l \leq d$, 
- $b \in \mathbb{R}^n_+ \times \mathbb{R}^m$,
- $B_{IJ} = 0$,
- $B_{II}$ has nonnegative off-diagonal elements.

18
So for the Riccati equations (2.4) this implies

\[
\begin{align*}
\partial_t \varphi(t, u) &= \frac{1}{2} \psi_J(t, u)^T a_J \psi_J(t, u) + b^T \psi(t, u), \\
\varphi(0, u) &= 0, \\
\partial_t \psi_i(t, u) &= \frac{1}{2} \psi(t, u)^T \alpha_i \psi(t, u) + \beta_i^T \psi(t, u), \quad i \in I, \\
\partial_t \psi_J(t, u) &= B^T J \psi_J(t, u), \\
\psi(0, u) &= u,
\end{align*}
\]  

(2.9)

and there exists a unique global solution \((\varphi(\cdot, u), \psi(\cdot, u)) : \mathbb{R}_+ \to \mathbb{C}_+ \times \mathbb{C}^m \times i\mathbb{R}^n\) for all initial values \(u \in \mathbb{C}^m \times i\mathbb{R}^n\). In particular, the equation for \(\psi_J\) forms an autonomous linear system with unique global solution \(\psi_J(t, u) = e^{B^T J t} u_J\) for all \(u_J \in \mathbb{C}^n\), [7, 18].

This theorem is proved in [18, Theorem 10.2]. In the next chapter this theorem will be extended to a theorem that can be used for pricing derivatives with a particular exponential form of payoff. However, first a shift to short-rate models is introduced below.

### 2.3. Pricing in Affine models

Now that the definition of a general affine model for \(X\) is given, the definition of an affine short-rate model for \(r\) can be considered. Therefore, let \(A_r \in \mathbb{R}\) and \(B_r \in \mathbb{R}^d\) some constant parameters, then the short-rate model \(r_t\) of the form

\[
r_t = A_r + B_r^T X_t,
\]

(2.10)

with \(X_t\) an affine process, is called an affine short-rate model [18]. It will turn out that these affine short-rate models can conveniently be used to price financial derivatives.

Consider a claim with maturity \(T > 0\) with payoff \(f(X_T)\) that satisfies

\[
\mathbb{E}_Q \left[ e^{-\int_0^T r(s) ds} | f(X_T) | \right] < \infty.
\]

When this integrability condition holds, the price at \(t \leq T\) is given by

\[
\pi(t) = \mathbb{E}_Q \left[ e^{-\int_0^T r(s) ds} f(X_T) | \mathcal{F}_t \right].
\]

(2.11)

As said above, eventually, with the affine models the aim is to price financial derivatives. Therefore, the above pricing formula should be evaluated analytically. However, for general payoff functions \(f(x)\) the distribution of \(X_T\) under the \(T\)-forward measure is not always known; if this \(\mathcal{F}_r\)-conditional distribution \(Q(t, T, dx)\) is known, the following price applies by numerical integration of \(f\) in expression (2.11). Using the Radon-Nikodym derivative \(\frac{dQ}{dQ_T}\) to switch to the \(T\)-forward measure as explained in A.3.1 and using [29, Lemma 9.2]

\[
\mathbb{E}_Q [XZ | \mathcal{F}_t] = \mathbb{E}_{Q_T} [X | \mathcal{F}_t] \mathbb{E}_Q [Z | \mathcal{F}_t]
\]

where \(Z = \frac{dQ}{dQ_T}\).  

19
Thus,

\[
\pi(t) = \mathbb{E}_Q \left[ e^{-\int_t^T r(s)ds} f(X_T) \mid \mathcal{F}_t \right] \\
= \mathbb{E}_Q \left[ \frac{B(t)}{B(T)} f(X_T) \mid \mathcal{F}_t \right] \\
= \mathbb{E}_Q \left[ \frac{dQ^T}{dQ} P(0, T) B(t) f(X_T) \mid \mathcal{F}_t \right] \\
= \mathbb{E}_{Q^r} \left[ P(0, T) B(t) f(X_T) \mid \mathcal{F}_t \right] \mathbb{E}_Q \left[ \frac{dQ^T}{dQ} \mid \mathcal{F}_t \right] \\
= P(0, T) B(t) \mathbb{E}_{Q^r} \left[ f(X_T) \mid \mathcal{F}_t \right] \frac{P(t, T)}{P(0, T) B(t)} \\
= P(t, T) \mathbb{E}_{Q^r} \left[ f(X_T) \mid \mathcal{F}_t \right] \\
= P(t, T) \int_{\mathbb{R}^d} f(x) Q(t, T, dx). \quad (2.12)
\]

Note that removing the \( e^{-\int_t^T r(s)ds} \) out of the expectation is useful, since the joint distribution of the payoff \( f(X_T) \) is often not known, but the marginal distribution of \( f(X_T) \) is. Thus, the multiplication of the price of the zero coupon bond and the integral, enables to express the price function more easily. When the distribution \( Q(t, T, dx) \) of \( X_T \) under the \( T \)-forward measure is also not known, Fourier transform techniques can be used to derive the price. Results for these general payoff functions without known distribution will be posed after the easier results for payoff functions for \( T \)-claims are stated.

The first step is to write down an analytical formula for a \( T \)-bond which has a constant payoff function of 1. This pricing formula is derived from the following theorem of [18, Theorem 10.4] that gives a formula for exponential payoff functions.

**Theorem 2.3.1.** Let \( \tau > 0 \). The following statements are equivalent:

(a) \( \mathbb{E}[e^{-\int_0^\tau r(s)ds}] < 0 \) for all \( x \in \mathbb{R}^n_+ \times \mathbb{R}^n \).

(b) There exists a unique solution \( (\Phi(t, u), \Psi(t, u)) : [0, \tau] \to \mathbb{C} \times \mathbb{C}^d \) of

\[
\begin{align*}
\partial_t \Phi(t, u) &= \frac{1}{2} \Psi_j(t, u) \alpha_{ij} \Psi_j(t, u) + b^\top \Psi(t, u) - A_r, \\
\Phi(0, u) &= 0 \\
\partial_t \Psi_i(t, u) &= \frac{1}{2} \Psi_j(t, u) \alpha_{ij} \Psi_j(t, u) + \beta_i^\top \Psi(t, u) - (B_r)_i, \\
\Psi(0, u) &= u \\
\quad \text{for } i \in \mathbb{R}^d. 
\end{align*}
\]

(2.13)

Moreover, let \( \mathcal{D}_K \) \( (K = \mathbb{R} \text{ or } \mathbb{C}) \) denote the maximal domain for the system of Riccati equations (2.13) \( (i.e. \mathcal{D}_K = \{(t, u) \in \mathbb{R}_+ \times K^d \mid t < t_+(u)\}, \text{ with } t_+(u) \in (0, \infty)\} \). If either (a) or (b) holds then \( \mathcal{D}_R(S) \), defined by \( \mathcal{D}_R(S) = \{u \in K^d \mid (S, u) \in \mathcal{D}_K\} \), is a convex open neighborhood of 0 in \( K^d \), and \( \mathcal{S}(\mathcal{D}_R(S)) \subseteq \mathcal{D}_R(S), \) for all \( S \leq \tau \). Further, the following affine representation holds

\[
\mathbb{E}[e^{-\int_t^T r(s)ds} e^{u^\top X_T} \mid \mathcal{F}_t] = e^{\Phi(T-t, u) + \Psi(T-t, u)^\top X_t}
\]

for all \( u \in \mathcal{S}(\mathcal{D}_R(S)) \), \( t \leq T \leq t + S \) and \( x \in \mathbb{R}^m_+ \times \mathbb{R}^n \) [18, Theorem 10.4].
for all \((X, Y)\). Theorem 2.2.1 now implies that there exists a unique global diffusion process. Where

\[
\frac{dX'}{dt} = \frac{b'}{A_r} + \frac{B'}{\gamma^T} X' + \sum_{i=1}^{d} x_i \alpha_i \frac{dW}{dt},
\]

so \(X'\) has admissible parameters

\[
a' = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \quad \alpha_i = \begin{pmatrix} \alpha_i & 0 \\ 0 & 0 \end{pmatrix}, \quad b' = \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} B & 0 \\ \gamma^T & 0 \end{pmatrix}.
\]

The candidate system of Riccati equations for \(i \in I\) is

\[
\begin{align*}
\partial_t \psi_i'(t, u, v) &= \frac{1}{2} \psi_i'(t, u, v) + \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \quad \psi_i'(t, u, v) + \begin{pmatrix} b \\ 0 \end{pmatrix}, \\
\partial_t \psi_i'(t, u, v) &= \frac{1}{2} \psi_i'(t, u, v) + \begin{pmatrix} \alpha_i & 0 \\ 0 & 0 \end{pmatrix}, \quad \psi_i'(t, u, v) + \begin{pmatrix} b' \\ 0 \end{pmatrix}, \\
\partial_t \psi_i'(t, u, v) &= \begin{pmatrix} B & 0 \\ \gamma^T & 0 \end{pmatrix}, \quad \psi_i'(t, u, v),
\end{align*}
\]

(2.15)

Theorem 2.2.1 now implies that there exists a unique global \(C_{-} \times C_{m} \times i \mathbb{R}^{n+1}\)-valued solution \((\psi'(-, t, u, v), \psi'(-, t, u, v))\) of (2.15) for all initial values \((u, v) \in C_{m} \times i \mathbb{R}^{n} \times i \mathbb{R}^{n+1}\). Theorem 2.1.2 asserts that \(X'\) is affine with conditional characteristic function

\[
\mathbb{E} \left[ e^{(u, v)')X'(T) \mid F_t \right] = e^{v'(T-t, u, v) + \psi'(T-t, u, v)'} X'(t)
\]

\[
\mathbb{E} \left[ e^{x'T(T)+Y(T)} \mid F_t \right] = e^{x'(T-t, u, v) + \psi'(T-t, u, v)'} X(t) + \psi'Y(t)
\]

for all \((u, v) \in C_{m} \times i \mathbb{R}^{n} \times i \mathbb{R}^{n+1}, t \leq T\) and where \(Y\) defined as in (2.14).
Let $\Phi(t,u) = \varphi'(t,u,-1)$ and $\Psi(t,u) = \psi'_1,...,d(t,u,-1)$, then with equation $(2.14)$

$$\mathbb{E}\left[e^{x^T X(t) - \int_0^t (A_x + B_x X(s))ds} \mid \mathcal{F}_t\right] = e^{\Phi(t,u) + \Psi(T-t,u)^T X(t) - \int_0^t (A_x + B_x X(s))ds}$$

$$\Leftrightarrow \mathbb{E}\left[e^{x^T X(t) - \int_t^T r_s ds} \mid \mathcal{F}_t\right] = e^{\Phi(t,u) + \Psi(T-t,u)^T X(t)}$$

By inspection it is clear that $D_K(S) = \{u \in K^d \mid (u,-1) \in D'_K(S)\}$ where $D'_K$ denotes the maximal domain for the system of Riccati equations $(2.15)$ [18, Theorem 10.4].

The following corollary links this theorem to a pricing strategy.

**Corollary 2.3.2.** For any maturity $T \leq \tau$, with $\tau$ as in Theorem 2.3.1, the $T$-bond price at $t \leq T$ is given as

$$P(t,T) = e^{-A(T-t)-B(T-t)^T X_t}$$

where

$$A(t) = -\Phi(t,0), \quad B(t) = -\Psi(t,0).$$

Moreover, for $t \leq T \leq S \leq \tau$, the $\mathcal{F}_t$-conditional characteristic function of $X_T$ under the $S$-forward measure $Q^S$ is given by

$$\mathbb{E}_{Q^S}\left[e^{u^T X_T} \mid \mathcal{F}_t\right] = \frac{e^{-A(S-T)+\Phi(T-t,u-B(S-T))+\Psi(T-t,u-B(S-T))^T X_t}}{P(t,S)} ,$$

for all $u \in S(D_R(T) + B(S-T))$, which contains $i\mathbb{R}^d$ [3, 15, 18].

Note that this corollary implies that, if the solutions $(2.13)$ of the system of Riccati equations are known, one can also derive the $T$-bond price. As will be shown, some affine short-rate models allow an easy derivation of the solutions $\Phi$ and $\Psi$ to the Riccati equations, which can be written in closed-form. For other models, it is only possible to evaluate the solutions numerically. For the models with closed-form solutions $\Phi$ and $\Psi$, the $T$-bond price is easily obtained by just substituting the solutions in the stated corollary.

### 2.3.1. Fourier transform technique

As previously mentioned, the price of a $T$-claim is given by $(2.12)$. In this section, the case where the distribution of $X_t$ under the $T$-forward measure $Q^T$ is not explicitly known is considered. In practice, this is often the case and therefore this is also used in Chapter 5. The unknown distribution of $X_t$ leads to the exploration of a Fourier transform technique. The first important theorem, which is stated below, originates from [18, Theorem 10.5]

**Theorem 2.3.3.** Suppose either (a) of (b) of Theorem 2.3.1 is met for some $\tau \geq T$, and let $D_R$ denote the maximal domain for the system of Riccati equations $(2.13)$. Assume that $f$ satisfies

$$f(x) = \int_{\mathbb{R}^d} e^{(v+iL\lambda)^T x} f(\lambda) d\lambda, \quad dx-a.s. \quad \text{(2.17)}$$

for some $v \in D_R(T)$ and $d \times q$-matrix $L$, and some integrable function $\tilde{f} : \mathbb{R}^q \rightarrow \mathbb{C}$, for a positive integer $q \leq d$. Then the $T$-bond price is well defined and given by

$$\pi(t) = \int_{\mathbb{R}^d} e^{\Phi(T-t,v+iL\lambda)+\Psi(T-t,v+iL\lambda)^T X_t} \tilde{f}(\lambda) d\lambda. \quad \text{(2.18)}$$
Thus, if a payoff function can be represented as (2.17), then the price formula is given by Theorem 2.3.3. The following example is a representation of a very similar well-known payoff function, namely those of the put and call option.

**Example 2.3.4.** Let $\mathbb{K} > 0$. For any $y \in \mathbb{R}$ the following identities hold

$$
\frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)y} K^{-(w-1+i\lambda)} (w-1+i\lambda) d\lambda = \begin{cases} (K - e^y)^+ & \text{if } w < 0, \\
(K - e^y)^+ - e^y & \text{if } 0 < w < 1, \\
(K - e^y)^+ & \text{if } w > 1. 
\end{cases}
$$

**Proof.** Suppose one wants to find the Fourier transform of the payoff function $(K - e^y)^+$, then $\hat{f}$ in

$$(K - e^y)^+ = \int_{\mathbb{R}^\mathbb{C}} e^{(v+i\lambda)y} \hat{f}(\lambda) d\lambda, \text{ dx-a.s.}
$$

should be determined. The fundamental inversion formula (see [18, Proof of Theorem 10.6]) states that for $g : \mathbb{R}^y \to \mathbb{C}$ an integrable function with integrable Fourier transform

$$
\hat{g}(\lambda) = \int_{\mathbb{R}^\mathbb{C}} e^{-i\lambda \mathbf{y}} g(y) dy,
$$

the following inversion formula holds

$$
g(y) = \frac{1}{(2\pi)^{\mathbb{C}}} \int_{\mathbb{R}^\mathbb{C}} e^{i\lambda \mathbf{y}} \hat{g}(\lambda) d\lambda \text{ dy-a.s.}
$$

Thus, for $w < 0$, function $f(y) = e^{-w y}(K - e^y)^+$ on $\mathbb{R}$ is integrable and has a Fourier transform $\hat{f}(\lambda)$ defined by

$$
\hat{f}(\lambda) = \int_{\mathbb{R}} e^{-i\lambda y} f(y) dy = \int_{\mathbb{R}} e^{-i\lambda y} e^{-w y}(K - e^y)^+ dy = \int_{\mathbb{R}} e^{-(w+i\lambda)y}(K - e^y)^+ dy.
$$

So, by the inversion formula

$$
e^{-w y}(K - e^y)^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\lambda y} \hat{f}(\lambda) d\lambda \\
\Leftrightarrow (K - e^y)^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)y} \hat{f}(\lambda) d\lambda
$$

(2.19)

In order to express this $(K - e^y)^+$ more specifically, the expression for $\hat{f}(\lambda)$ is rewritten. By substitution of $z = e^y$ with $dz = e^y dy$, it is obtained that

$$
\hat{f}(\lambda) = \int_{\mathbb{R}} e^{-(w+i\lambda)y}(K - e^y)^+ dy = \int_{\{K > e^y\}} e^{-(w+i\lambda)y}(K - e^y) dy
$$

$$
= \int_{\{0 \leq z \leq K\}} z^{-(w+i\lambda+1)} (K - z) dz.
$$

Then, apply integration by parts (i.e. $\int f'(x)g(x)dx = [f(x)g(x)] - \int f(x)g'(x)dx$ where $f'(x) = z^{-(w+i\lambda+1)}$ and $g(x) = (K - z)$, so $f(x) = z^{-(w+i\lambda+1)}/(w+i\lambda)$ and $g'(x) = -1$), to get

$$
\hat{f}(\lambda) = \int_{0}^{K} \underbrace{z^{-w+i\lambda+1}}_{-w+i\lambda} \underbrace{(K - z)}_{-1} dz = \left[ \frac{z^{-(w+i\lambda)}}{-(w+i\lambda)} (K - z) \right]_{0}^{K} - \int_{0}^{K} \frac{z^{-(w+i\lambda)-1}}{-(w+i\lambda)} -1 dz
$$

$$
= \int_{0}^{K} \frac{z^{-(w+i\lambda)}}{-(w+i\lambda)} dz = \left[ \frac{z^{-(w+i\lambda-1)}}{-(w+i\lambda)(w+i\lambda-1)} \right]_{0}^{K} = \frac{K^{-(w+i\lambda-1)}}{-(w+i\lambda)(w+i\lambda-1)}.
$$

So substituting this $\hat{f}(\lambda)$ in the expression (2.19) proves the claim for $w < 0$

$$(K - e^y)^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)y} \hat{f}(\lambda) d\lambda = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)y} \frac{K^{-(w+i\lambda-1)}}{-(w+i\lambda)(w+i\lambda-1)} d\lambda.
$$

For the proofs of the cases where $0 < w < 1$ and $w > 1$, the argumentation is similar. \(\square\)
2.3.2. Price formula of call and put options on a bond

Using the Fourier transform technique, one can now price \( T \)-claims if the payoff function \( f(X_T) \) satisfies (2.17). Two frequently considered \( T \)-claims are the call and put option on a \( S \)-bond with expiry date \( T, T < S \) and strike price \( K \). Its payoff function are very similar to the function that is discussed in Example 2.3.4. For a call option

\[
(P(T, S) - K)^+ = \left( e^{-A(S-T) - B(S-T)^\top x} - K \right)^+,
\]

where Corollary 2.3.2 is used for the expression of \( P(T, S) \). This payoff equals the expression in the example when \( y \) is substituted with \(-A(S-T) - B(S-T)^\top x\). Thus, for \( w > 1 \)

\[
\left( e^{-A(S-T) - B(S-T)^\top x} - K \right)^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(w+i\lambda)(-A(S-T)-B(S-T)^\top x)} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)} d\lambda
\]

\[
= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-(w+i\lambda)A(S-T)} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)} d\lambda
\]

\[
= \int_{\mathbb{R}} e^{-i\lambda w} \tilde{f}(w, \lambda) d\lambda,
\]

where

\[
\tilde{f}(w, \lambda) = \frac{1}{2\pi} e^{-(w+i\lambda)A(S-T)} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}.
\]

The payoff of the put option is very similar but then for \( w < 0 \). As a specific case of Theorem 2.3.3 the following corollary is stated by Filipović in [18].

**Corollary 2.3.5.** There exists some \( W_+ > 0 \) and \( W_+ > 1 \) such that \(-B(S-T)w \in \mathcal{D}_\mathbb{R}(T)\) for all \( w \in (W_-, W_+) \), where \( \mathcal{D}_\mathbb{R} \) denotes the maximal domain for the system of Riccati equations (2.13) (see Theorem 2.3.1). Define \( \tilde{f}(w, \lambda) \) as

\[
\tilde{f}(w, \lambda) = \frac{1}{2\pi} e^{-(w+i\lambda)A(S-T)} \frac{K^{-(w-1+i\lambda)}}{(w+i\lambda)(w-1+i\lambda)}.
\]

Then the line integral

\[
\Pi(w, t) = \int_{\mathbb{R}} e^{\Phi(T-t, -(w+i\lambda)B(S-T)) + \Phi(T-t, -(w+i\lambda)B(S-T))\top X_t} \tilde{f}(w, \lambda) d\lambda
\]

is well defined for all \( w \in (W_-, W_+) \setminus \{0, 1\} \) and \( t \leq T \). Moreover, the time \( t \) prices of the European call and put option on the \( S \)-bond with expiry date \( T \) and strike price \( K \) are given by any of the following identities:

\[
\pi_{\text{call}}(t) = \begin{cases} 
\Pi(w, t), & \text{if } w \in (1, W_+), \\
\Pi(w, t) + P(t, S), & \text{if } w \in (0, 1) 
\end{cases}
\]

\[
\pi_{\text{put}}(t) = \begin{cases} 
\Pi(w, t) + K P(t, T), & \text{if } w \in (0, 1), \\
\Pi(w, t), & \text{if } w \in (W_-, 0) 
\end{cases}
\]

(2.20)

where \( \mathcal{I} = (A(S-T) + \log K, \infty) \), and \( q(t, S, dy) \) and \( q(t, T, dy) \) denote the \( F_T \)-conditional distributions of the real-valued random variable \( Y = -B(S-T)^\top X_T \) under the \( S \)- and \( T \)-forward measure, respectively.
In order to determine a more explicit formula for the call and put price, consider time $t = 0$. Then the price of the European call option on the S-bond with expiry date $T < S$ and strike $K$ is given by

$$
\pi_{\text{call}}(0) = P(0,S)q(0,S,I) - KP(0,T)q(0,T,I). \tag{2.21}
$$

For $q(0,S,I)$ the following holds

$$
q(0,S,I) = q(0,S,(A(S-T) + \log K, \infty)) = Q^S[Y > A(S-T) + \log K \mid F_0] = Q^S[-B(S-T)^\top X_T > A(S-T) + \log K] = Q^S[-A(S-T) - B(S-T)^\top X_T > \log K] = Q^S[\exp(-A(S-T) - B(S-T)^\top X_T) > K] = Q^S[P(T,S) > K] = Q^S\left[\frac{P(T,T)}{P(T,S)} < \frac{1}{K}\right],
$$

and similarly for $q(0,T,I)$

$$
q(0,T,I) = Q^T\left[\frac{P(T,T)}{P(T,S)} < \frac{1}{K}\right] = Q^T\left[\frac{P(T,S)}{P(T,T)} > K\right].
$$

Hence, the distributions of $\frac{P(T,T)}{P(T,S)}$ w.r.t. $Q^S$ and $\frac{P(T,S)}{P(T,T)}$ w.r.t. $Q^T$ have to be determined. It depends on the model how these distributions are determined. Therefore, see Section 2.4 for explicit expressions of the call option price. Note that the derivation of the put option price is similar.

### 2.3.3. Price formula of cap

Another financial product that is frequently considered, is the cap (an introduction on caps is for example given in [3, Section 26.8]). A cap is a strip of caplets, so for a derivation of the price of a cap, first a caplet is explored.

Consider a caplet with $\leq T_0$ and payoff $\delta(F(T_{i-1},T_i) - \kappa)^+$ at time $T_i$, where $\delta = T_i - T_{i-1}$ and $F(T_{i-1},T_i)$ is defined as the spot rate (e.g. the LIBOR rate) given by (A.1) [3, Section 26.8]. The following holds

$$
\delta(F(T_{i-1},T_i) - \kappa)^+ = \delta \left( \frac{1}{T_i - T_{i-1}} \left( \frac{1}{P(T_{i-1},T_i)} - 1 \right) - \kappa \right)^+ = \frac{1}{\delta} \left( \frac{1}{P(T_{i-1},T_i)} - 1 \right)^+ = \left( \frac{1}{P(T_{i-1},T_i)} - (1 + \delta \kappa) \right)^+ = \frac{1 + \delta \kappa}{P(T_{i-1},T_i)} \left( \frac{1}{1 + \delta \kappa} - P(T_{i-1},T_i) \right)^+.
$$

Now, note that a payment of $(1 + \delta \kappa) \left( \frac{1}{1 + \delta \kappa} - P(T_{i-1},T_i) \right)^+$ at time $T_{i-1}$, has a time $T_i$ value of

$$
\frac{1}{P(T_{i-1},T_i)}(1 + \delta \kappa) \left( \frac{1}{1 + \delta \kappa} - P(T_{i-1},T_i) \right)^+,
$$
and the part \((1 + \delta\kappa)\left(\frac{1}{1 + \delta\kappa} - P(T_{i-1}, T_i)\right)^+\) is exactly the payoff of \((1 + \delta\kappa)\) put options on a \(T_i\)-bond with exercise date \(T_{i-1}\) and with strike price \(1/1 + \delta\kappa\). Thus, the payoff, and therefore also the price, of a caplet can be calculated by determining the payoff of a put option as shown below. With the tower property [6, Theorem 7.29(iv)] applied in the second line, the price of a caplet equals

\[
\pi_{\text{caplet}} = \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \delta(F(T_{i-1}, T_i) - \kappa)^+ | \mathcal{F}_t \right]
\]

\[
= \mathbb{E}_Q\left[ \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \delta(F(T_{i-1}, T_i) - \kappa)^+ | \mathcal{F}_{T_{i-1}} \right] | \mathcal{F}_t \right]
\]

\[
= \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \mathbb{E}_Q\left[ e^{-\int_{T_{i-1}}^{T_i} r(s)ds} \delta(F(T_{i-1}, T_i) - \kappa)^+ | \mathcal{F}_{T_{i-1}} \right] | \mathcal{F}_t \right]
\]

\[
= \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \mathbb{E}_Q\left[ P(T_{i-1}, T_i) \delta(F(T_{i-1}, T_i) - \kappa)^+ | \mathcal{F}_{T_{i-1}} \right] | \mathcal{F}_t \right]
\]

\[
= \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \left(1 + \delta\kappa\right) \left(1 +\delta\kappa - P(T_{i-1}, T_i)\right)^+ | \mathcal{F}_t \right]
\]

\[
= \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \left(1 + \delta\kappa\right) \left(1 +\delta\kappa - P(T_{i-1}, T_i)\right)^+ | \mathcal{F}_t \right]
\]

\[
= (1 + \delta\kappa) \cdot \mathbb{E}_Q\left[ e^{-\int_0^{T_i} r(s)ds} \left(1 +\delta\kappa - P(T_{i-1}, T_i)\right)^+ | \mathcal{F}_t \right]
\]

\[
= (1 + \delta\kappa) \cdot \pi_{\text{put}}, \quad (2.22)
\]

where the \(\pi_{\text{put}}\) is the time \(t\)-price of a put, where \(t < T_0\), on a \(T_i\)-bond with expiry date \(T_{i-1}\) and strike price \(1/1 + \delta\kappa\). As noted in the previous section, the price of this put can be made explicit when the interest model is chosen. Therefore, if the model and the corresponding parameters are known one can calculate the caplet price.

When the caplet prices are calculated, the cap price \(C_P(t)\) at time \(t\) can be calculated by summing the caplets over all time intervals \(T_0 < T_1 < \cdots < T_n\), where \(T_n\) is the maturity of the cap

\[
C_P(t) = \sum_{i=1}^n C_{pl}(t; T_{i-1}, T_i).
\]

### 2.4. Examples of affine short-rate models

For all examples it is supposed that \(\mathbb{E}[e^{-\int_0^T r(s)ds}] < \infty\) for all \(x \in \mathbb{R}_m^+ \times \mathbb{R}^n\), i.e. \(r_t \geq 0\), such that Theorem 2.3.1 can be used.

#### 2.4.1. Vasiček short-rate model

Recall that the Vasiček short-rate model is given by

\[
dr_t = (b + \beta r_t)dt + \sigma dW_t,
\]

with the usual notations, state space \(\mathbb{R}\), \(r_t = X_t\), short rate \(r_t\), and \(W_t\) a Brownian motion. The diffusion coefficient \(a(x)\) and the drift term \(b(x)\) as defined in (2.3) are

\[
a(x) = \sigma^2 \quad \text{and} \quad b(x) = b + \beta r_t,
\]

thus \(a = \sigma^2, \alpha_1 = 0, b = b\) and \(\beta_1 = \beta\).

Since the state space is \(\mathbb{R}\), \(m\) and \(n\) in the canonical state space \(\mathbb{R}_m^+ \times \mathbb{R}^n\) are \(m = 0\) and \(n = 1\). Thus, the index sets (2.7) are given by \(I = \emptyset\) and \(J = \{1\}\). Also, because
For the Vasicek model the price formula (2.21) can be made explicit by considering equality (2.23) in the Riccati system (2.13) gives

\[
\Phi(t, u) = \frac{1}{2} \sigma^2 \int_0^t \Psi^2(s, u) ds + b \int_0^t \Psi(s, u) ds,
\]

so \( \phi(t, u) = \frac{1}{2} \sigma^2 \int_0^t \Psi^2(s, u) ds + b \int_0^t \Psi(s, u) ds \),

\[
\phi(t, u) = e^{\beta t} u - \frac{e^{\beta t} - 1}{\beta},
\]

is the unique global solution of the system for all \( u \in \mathbb{C} \). Substituting (2.23) in the described system of equations proves this claim (see Appendix B.1.1).

**Explicit call option price formula for Vasicek**

For the Vasicek model the price formula (2.21) can be made explicit by considering equality (A.5) of Lemma A.3.1 for \( t = T \). Note that \( t = T \) can be substituted since it was stated that \( T < S \), thus \( t \leq S \wedge T \) as in the lemma. So,

\[
\frac{P(T, T)}{P(T, S)} = \frac{P(0, T)}{P(0, S)} \mathcal{E}_T(\sigma_{T, S} \cdot W^S).
\]

Hence

\[
\log \left( \frac{P(T, T)}{P(T, S)} \right) = \log \left( \frac{P(0, T)}{P(0, S)} \right) + \mathcal{E}_T(\sigma_{T, S} \cdot W^S) = \log \left( \frac{P(0, T)}{P(0, S)} \right) + \left( \int_0^T \sigma_{T, S}(s) dW^S(s) - \frac{1}{2} \int_0^T \sigma_{T, S}(s)^2 ds \right).
\]

Since the first log-term is just a constant, only the distribution of the second term must be determined. Theorem 4.4.9 from [28] implies that for the deterministic \( \sigma_{T, S}(s) = -\sigma_{S, T}(s) = -\int_s^T \sigma(s, u) du \) function of time in the Vasicek model, for each \( T \geq 0, \int_0^T \sigma_{S, T}(s) dW^S(s) \) is normally distributed with expected value zero and variance \( \int_0^T \sigma_{T, S}^2(s) ds = \int_0^T \sigma_{S, T}^2(s) ds \). Also, the second term in the second log-term is deterministic, so

\[
\int_0^T \sigma_{T, S}(s) dW^S(s) = \frac{1}{2} \int_0^T \sigma_{T, S}^2(s) ds \sim \mathcal{N} \left( -\frac{1}{2} \int_0^T \sigma_{S, T}^2(s) ds, \int_0^T \sigma_{T, S}^2(s) ds \right).
\]

Thus,

\[
\log \left( \frac{P(T, T)}{P(T, S)} \right) \sim \mathcal{N} \left( \log \left( \frac{P(0, T)}{P(0, S)} \right) - \frac{1}{2} \int_0^T \sigma_{S, T}^2(s) ds, \int_0^T \sigma_{S, T}^2(s) ds \right).
\]

For \( \frac{P(T, S)}{P(T, T)} \), there is a similar result

\[
\frac{P(T, S)}{P(T, T)} = \frac{P(0, S)}{P(0, T)} \mathcal{E}_T(\sigma_{S, T} \cdot W^T),
\]

27
thus

\[ \log \left( \frac{P(T,S)}{P(T,T)} \right) = \log \left( \frac{P(0,S)}{P(0,T)} \right) + \left( \int_0^T \sigma_{S,T}(s)dW(s) - \frac{1}{2} \int_0^T \sigma_{S,T}^2(s)ds \right), \]

and

\[ \log \left( \frac{P(T,S)}{P(T,T)} \right) \sim N \left( \log \left( \frac{P(0,S)}{P(0,T)} \right) - \frac{1}{2} \int_0^T \sigma_{S,T}^2(s)ds, \int_0^T \sigma_{S,T}^2(s)ds \right). \]

Since \( \frac{P(T,T)}{P(T,S)} \) and \( \frac{P(T,S)}{P(T,T)} \) are log-normally distributed as described above, the expressions for \( q(0,S,\mathcal{I}) \) and \( q(0,T,\mathcal{I}) \) can be written down, where it is used that if \( X \sim \mathcal{N}(\mu,\sigma^2) \), then

\[ \mathbb{P}(X < x) = \Phi \left( \frac{x - \mu}{\sigma} \right), \]

thus \( \log(q(0,S,\mathcal{I})) = \mathcal{N} \left( \log \left( \frac{P(0,S)}{P(0,T)} \right) - \frac{1}{2} \int_0^T \sigma_{S,T}^2(s)ds, \int_0^T \sigma_{S,T}^2(s)ds \right) \) implies

\[ q(0,S,\mathcal{I}) = \mathbb{Q}^S \left[ \frac{P(T,T)}{P(T,S)} < \frac{1}{K} \right] \]

\[ = \mathbb{Q}^S \left[ \log \left( \frac{P(T,T)}{P(T,S)} \right) < \log \left( \frac{1}{K} \right) \right] \]

\[ = \Phi \left( \frac{\log(1/K) - \left( \log(P(0,T)/P(0,S)) - \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds \right)}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}} \right) \]

\[ = \Phi \left( \frac{\log(P(0,S)/P(0,T)K) + \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}} \right). \]

Similarly,

\[ q(0,T,\mathcal{I}) = \mathbb{Q}^T \left[ \frac{P(T,S)}{P(T,T)} > K \right] \]

\[ = 1 - \Phi \left( \frac{\log(K) - \left( \log(P(0,S)/P(0,T)) - \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds \right)}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}} \right) \]

\[ = \Phi \left( \frac{\log(K) - \left( \log(P(0,S)/P(0,T)) - \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds \right)}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}} \right) \]

\[ = \Phi \left( \frac{\log(P(0,S)/P(0,T)K) - \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}} \right). \]

It can be concluded that the option price at time \( t = 0 \) satisfies

\[ \pi_{\text{call}}(0) = P(0,S)q(0,S,\mathcal{I}) - KP(0,T)q(0,T,\mathcal{I}) \]

\[ = P(0,S)\Phi(d_1) - KP(0,T)\Phi(d_2), \quad (2.24) \]

where

\[ d_{1,2} = \frac{\left( \log(P(0,S)/P(0,T)K) \pm \frac{1}{2} \int_0^T ||\sigma_{S,T}(s)||ds \right)}{\sqrt{\int_0^T ||\sigma_{S,T}(s)||ds}}. \]
Similarly, the price of the put option can be determined. It satisfies
\[
\pi_{\text{put}}(0) = KP(0,T)q(0,T,\mathbb{R} \setminus \mathcal{I}) - P(0,S)q(0,S,\mathbb{R} \setminus \mathcal{I})
\]
\[
= KP(0,T)\Phi(-d_2) - P(0,S)\Phi(-d_1),
\]
(2.25)
where \(d_1\) and \(d_2\) are defined as in the price of a call option.

The \(P(0,S)\) and \(P(0,T)\) in these expressions are known by Corollary 2.3.2
\[
P(0,S) = e^{-A(S)-B(S)^T r(0)} = e^{\Phi(S,0)+\Phi(S,0)^T r(0)}
\]
\[
= \exp\left\{ \frac{1}{2} \sigma^2 \left( \frac{1}{2\beta^3} (e^{2\beta S} - 4e^{\beta S} + 2\beta S + 3) \right) + b \left( -\frac{e^{\beta S} - 1 - \beta S}{\beta^2} \right) \right\}
\]
and
\[
P(0,T) = \exp\left\{ \frac{1}{2} \sigma^2 \left( \frac{1}{2\beta^3} (e^{2\beta T} - 4e^{\beta T} + 2\beta T + 3) \right) + b \left( -\frac{e^{\beta T} - 1 - \beta T}{\beta^2} \right) \right\}
\]
\[
+ \left( -\frac{e^{\beta T} - 1}{\beta} \right) r(0) \right\}.
\]

Moreover, for the Vasicek model (see Appendix B.1.2 for the definition of \(\alpha(t,T)\)) the following holds
\[
df(t,T) = \alpha(t,T)dt + \sigma e^{\beta(T-t)}dW,
\]
hence \(\sigma(t,T) = \sigma e^{\beta(T-t)}\). Thus also \(\sigma S,T(s)\) is known
\[
\sigma S,T(s) = \int_S^T \sigma(s,u)du = \int_S^T \sigma e^{\beta(u-s)}du = \left[ \frac{\sigma}{\beta} e^{\beta(u-s)} \right]_S^T = \frac{\sigma}{\beta} \left( e^{\beta(T-s)} - e^{\beta(S-s)} \right).
\]
So, when the parameters of the Vasicek model are given, the call and put option price for the Vasicek model can be calculated explicitly by using the above expressions.

**Explicit cap price formula for Vasicek**

Since the price of a put option price is known explicitly, one can also determine the cap price formula for the Vasicek model. Recall that the price of a \(T_i\)-caplet at \(t = 0\) equals (2.22), where the put has strike price \(
\frac{1}{1+\delta \kappa}\), thus
\[
\pi_{\text{caplet}} = (1 + \delta \kappa) \cdot \pi_{\text{put}}
\]
\[
= (1 + \delta \kappa) \cdot \left( \frac{1}{1 + \delta \kappa} \cdot P(0,T_{i-1}) \Phi[-d_2] - P(0,T_i) \Phi[-d_1] \right),
\]
with
\[
d_{1,2} = \log \left[ \frac{P(0,T_i)}{P(0,T_{i-1})} \right] = \frac{1}{2} \int_0^{T_{i-1}} ||\sigma_{T_{i-1},T_i}(s)||^2 ds
\]
\[
= \frac{1}{2} \log \left[ \frac{P(0,T_{i-1})}{P(0,T_{i-1})} \right] + \frac{1}{2} \int_0^{T_{i-1}} ||\sigma_{T_{i-1},T_i}(s)||^2 ds
\]
\[
\sqrt{\int_0^{T_{i-1}} ||\sigma_{T_{i-1},T_i}(s)||^2 ds}
\]
\[
= \log \left[ (1 + \delta \kappa) \cdot P(0,T_{i-1}) \right] + \frac{1}{2} \int_0^{T_{i-1}} ||\sigma_{T_{i-1},T_i}(s)||^2 ds
\]
\[
\sqrt{\int_0^{T_{i-1}} ||\sigma_{T_{i-1},T_i}(s)||^2 ds}.
\]
2.4.2. CIR short-rate model

The CIR short-rate model is given by

\[ dr = (b + \beta r)dt + \sigma \sqrt{r}dW, \]

with the usual notations, state space \( \mathbb{R}_+ \), \( r = X \), short rate \( r \), and \( W \) a Brownian motion. The diffusion matrix \( a(x) \) and the drift term \( b(x) \) as defined in (2.3) are

\[ a(x) = \sigma^2 r \text{ and } b(x) = b + \beta r, \]

thus \( a = 0 \), \( \alpha_1 = \sigma^2 \), \( b = b \) and \( \beta_1 = \beta \).

Since the state space is \( \mathbb{R}_+ \), \( m \) and \( n \) in the canonical state space \( \mathbb{R}^m \times \mathbb{R}^n \) are \( m = 1 \) and \( n = 0 \). Thus, the index sets (2.7) are given by \( I = \{1\} \) and \( J = \emptyset \). Also, because \( r = X \), \( A_r = 0 \) and \( B_r = 1 \) in (2.10). Substituting the parameters of the CIR model \( a_{JJ} = 0 \), \( b^\top = b \), \( \alpha_i = \sigma^2 \), \( \beta_i = \beta \), \( (B_r)_i = B_r \), \( B_r^\top JJ = 0 \) and \( (B_r)_J = 0 \) in the Riccati system (2.13) gives

\[ \Phi(t,u) = b \int_0^t \Psi(s,u)ds, \]
\[ \partial_t \Psi(t,u) = \frac{1}{2} \sigma^2 \Psi^2(t,u) + \beta \Psi(t,u) - 1, \]
\[ \Psi(0,u) = u. \]

The solution of this system is given by

\[ \Phi(t,u) = \frac{2b}{\sigma^2} \log \left( \frac{2\theta e^{(\theta-\beta)t}}{L_3(t) - L_4(t)u} \right), \]
\[ \Psi(t,u) = \frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u}, \]

where \( \theta = \sqrt{\beta^2 + 2\sigma^2} \) and

\[ L_1(t) = 2(e^{\theta t} - 1), \]
\[ L_2(t) = \theta(e^{\theta t} + 1) + \beta(e^{\theta t} - 1), \]
\[ L_3(t) = \theta(e^{\theta t} + 1) - \beta(e^{\theta t} - 1), \]
\[ L_4(t) = \sigma^2(e^{\theta t} - 1). \]

Explicit call and put option price formula for CIR

For the CIR model, the further derivation of the price formula (2.21) is different than for the Vasicek model, since the diffusion coefficient is not deterministic. Thus, another technique must be considered in order to price the call and put options. The \( q(0,S,I) \) of (2.21) satisfies

\[ q(0,S,I) = Q^S[Y > A(S - T) + \log K] = Q^S[-B(S - T)^\top X_T > A(S - T) + \log K] = Q^S[X_T \leq \frac{-A(S - T) - \log K}{B(S - T)^\top}] = Q^S[r_T \leq \frac{-A(S - T) - \log K}{B(S - T)^\top}] = Q^S[2r_T \leq \frac{2}{C_1(0,T,S)} \frac{-A(S - T) - \log K}{B(S - T)^\top}] = Q^S[2r_T \leq \frac{2}{C_1(0,T,S)} \frac{-A(S - T) - \log K}{B(S - T)^\top}] = Q^S(\ast). \]
It is claimed that \( \frac{2\tau r}{\left\langle \left\langle \frac{u}{s(T)} \right\rangle \right\rangle} \) is non-centrally \( \chi^2 \)-distributed under \( Q^S \) with degrees of freedom \( \delta = 4b/\sigma^2 \) and parameter of noncentrality \( \zeta = 2C_2(0, T, S)\tau(0) \) (proved in Appendix B.2.2), thus

\[
(\ast) = \text{CDF}_{\chi^2(\delta, \zeta)}\left(\frac{2}{C_1(0, T, S)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right),
\]

where \( A(t) = -\Phi(t, 0) \) and \( B(t) = -\Psi(t, 0) \) with \( \Phi \) and \( \Psi \) defined as (2.26).

Similarly, under \( Q^T \), \( \frac{2\tau r}{\left\langle \left\langle \frac{u}{s(T)} \right\rangle \right\rangle} \) is non-centrally \( \chi^2 \)-distributed with degrees of freedom \( \delta = 4b/\sigma^2 \) and parameter of noncentrality \( \zeta_T = 2C_2(0, T, T)r_0 \), thus

\[
q(0, T, T) = Q^T[Y > A(S - T) + \log K]
\]

\[
= Q^T\left[r_T \leq -\frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right]
\]

\[
= Q^T\left[2r_T \left(\frac{2}{C_1(0, T, T)} \frac{2}{C_2(0, T, S)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right)\right]
\]

\[
= \text{CDF}_{\chi^2(\delta, \zeta_T)}\left(\frac{2}{C_1(0, T, T)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right).
\]

Thus using Corollary 2.3.5, the price of a call option, \( \pi_{\text{call}} \), at time \( t = 0 \) is given by

\[
\pi_{\text{call}}(0) = P(0, S)q(0, S, T) - KP(0, T)q(0, T, T)
\]

\[
= P(0, S)\text{CDF}_{\chi^2(\delta, \zeta_S)}\left(\frac{2}{C_1(0, T, S)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right) - KP(0, T)\text{CDF}_{\chi^2(\delta, \zeta_T)}\left(\frac{2}{C_1(0, T, T)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right),
\]

and similarly

\[
\pi_{\text{put}}(0) = KP(0, T)q(0, T, R \setminus T) - P(0, S)q(0, S, R \setminus T)
\]

\[
= KP(0, T)\left(1 - \text{CDF}_{\chi^2(\delta, \zeta_T)}\left(\frac{2}{C_1(0, T, T)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right)\right) - P(0, S)\left(1 - \text{CDF}_{\chi^2(\delta, \zeta_S)}\left(\frac{2}{C_1(0, T, S)} \frac{-A(S - T) - \log K}{B(S - T)^{1/2}}\right)\right).
\]

The upper equations imply that if the parameters of the CIR model are known, the call and put option prices can be calculated explicitly, since then all the term in the above expressions are known; using Corollary 2.3.2, it is obtained that

\[
P(0, S) = e^{-A(S - T)B(S)^\tau(0)} = e^{\Phi(S(0)) + \Phi(S(0)^{-}r(0))}
\]

\[
= \exp\left\{\frac{2b}{\sigma^2} \log \left(\frac{2e^{(S - T)/2}}{L_3(s)}\right) - \frac{L_1(s)}{L_3(s)}\right\} = \left(\frac{2e^{(S - T)/2}}{L_3(S)}\right)^{\frac{2b}{\sigma^2}} \exp\left\{-\frac{L_1(S)}{L_3(S)}\right\}
\]

and

\[
P(0, T) = \left(\frac{2e^{(S - T)/2}}{L_3(T)}\right)^{\frac{2b}{\sigma^2}} \exp\left\{-\frac{L_1(T)}{L_3(T)}\right\},
\]

with the same definition of \( L_1, L_2, L_3 \) and \( L_4 \) as above. Moreover, for the CIR model (see Appendix B.2.3)

\[
d\sigma(t, T) = \alpha(t, T)dt + \partial_T B(t, T)e^{\beta(t - s)}\sigma\sqrt{\tau(t)}dW^s
\]

hence \( \sigma(t, T) = \partial_T B(t, T)e^{\beta(t - s)}\sigma\sqrt{\tau(t)} \). Thus also \( \sigma_{S,T}(s) \) is known

\[
\sigma_{S,T}(s) = \int_s^T \sigma(u, s)du = \int_s^T \partial_s B(u, s)e^{\beta(s - u)}\sigma\sqrt{\tau(u)}du.
\]
Explicit cap price formula for CIR

Since the price of a put option price is known explicitly, one can also determine the cap price formula for the CIR model. Recall that the price of a $T_i$-caplet at $t = 0$ equals (2.22), where the put has strike price $\frac{1}{1+\delta \kappa}$, thus

$$
\pi_{\text{caplet}} = (1 + \delta \kappa) \cdot \pi_{\text{put}}
$$

$$
= (1 + \delta \kappa) \cdot \left[ \frac{1}{1 + \delta \kappa} P(0, T_{i-1}) \left( 1 - CDF_{\chi^2(\delta, \zeta_{T_{i-1}})} \left( \frac{2}{C_1(0, T_{i-1}, T_{i-1})} - \frac{A(\delta) - \log \frac{1}{1+\delta \kappa}}{B(\delta)^\top} \right) \right) 
- P(0, T_i) \left( 1 - CDF_{\chi^2(\delta, \zeta_{T_i})} \left( \frac{2}{C_1(0, T_{i-1}, T_i)} - \frac{A(\delta) - \log \frac{1}{1+\delta \kappa}}{B(\delta)^\top} \right) \right) \right].
$$
3. Quadratic models

The theory about closed-form prices and analytically tractable models is continued by introducing quadratic models. As will be pointed out in this chapter, quadratic models have similar properties as affine models, but are also very different. This chapter will show that quadratic models are more sophisticated than affine models: in contrast to affine models, which assume a linear combination for the short rate, quadratic models assume a quadratic combination (compare expressions (2.10) and (3.6)). Where affine models cannot guarantee a positive nominal interest rate in a general framework without restrictions for the parameters, quadratic models allow for strictly positive interest rates without the restrictions that are sometimes required in the affine setting\(^1\) [1].

Also, in affine models nonlinearities in the data cannot be captured, while Dai and Singleton suggest that this nonlinearity may exist. They found that the pricing errors move with the magnitude of the slope of the yield curve, which might imply a nonlinearity. The extra quadratic term in quadratic models adds the possible nonlinearity to the dynamics which makes these models also more flexible [25].

Higher degree polynomials (i.e. dimension higher than two) would intuitively perform even better, because of the extra term and so extra flexibility. However, Filipović [18, Section 9.4] proves that there is no consistent polynomial term structure for degree greater than two. This consistency is necessary, since without consistency, the forward curve cannot be fully determined and also no closed form solution is guaranteed. Thus, polynomial term structures for degree greater than two cannot use theorems similar to those in the affine setting, since they assume closed form solutions. Therefore, in this thesis only affine and quadratic term structures are considered.

This chapter starts with defining quadratic processes and related concepts such as the quadratic equivalent of admissible parameters and Riccati equations. After stating theorems about the sufficient conditions for a model in the quadratic class and a quadratic process, there will be elaborated on pricing in the quadratic setting. As in the previous chapter, this pricing also involves Fourier techniques. Although, these techniques differ slightly and are therefore discussed separately in this chapter.

3.1. Definition of a quadratic process

For the definition of a quadratic process, assume the same as for the affine models. Thus, let \( \mathcal{X} \) a closed state space with non-empty interior and fixed dimension \( d \geq 1 \), and let stochastic process \( X \) have values in \( \mathcal{X} \), where, as in (2.1)

\[
dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad \text{and} \quad X_0 = x,
\]

where \( W \) denotes a \( d \)-dimensional Brownian motion on \( (\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}) \), and \( b : \mathcal{X} \to \mathbb{R}^d \) continuous and \( \rho : \mathcal{X} \to \mathbb{R}^{d \times d} \) measurable such that the diffusion matrix \( a(x) = \sigma(x)\sigma(x)^\top \) is continuous in \( x \in \mathcal{X} \).

The following definition is the quadratic version of Definition 2.1.1 for affine processes and is stated in [7, Definition 3.1].

\(^1\)Restrictions hold for example for the CIR model, which has to assume interest rates for the square root term in the diffusion. There, restrictions are necessary for the drift and diffusion terms. If \( b > \sigma^2/2 \), then \( \tau > 0 \) for all \( t \), whenever \( r(0) > 0 \) [5, 18].
Definition 3.1.1. A process $X$ is quadratic if the $\mathcal{F}_t$-conditional expectation of the continuous function $e^{u^\top X_t + X_t^\top Y X_T}$ belongs to $B(\mathbb{R}^d)$ and has an exponential quadratic form in $X_t$ for every $(t, u, V) \in \mathbb{R} \times \mathbb{C}^d \times \mathbb{C}^{d \times d}$. That is, there exist functions $\varphi(t, u, V) \in \mathbb{C}, \psi(t, u, V) \in \mathbb{C}^d$ and $\omega(t, u, V) \in \mathbb{C}^d$ such that for every $(t, u, V) \in \mathbb{R} \times \mathbb{C}^d \times \mathbb{C}^{d \times d}$ and $t \leq T$,

$$
\mathbb{E}[e^{u^\top X_t + X_t^\top Y X_T} \mid \mathcal{F}_t] = e^{\varphi(T-t, u, V) + \psi(T-t, u, V)^\top X_t + X_t^\top \omega(T-t, u, V) X_t}.
$$

(3.1)

Remark. The assumption that $e^{u^\top X_t + X_t^\top Y X_T}$ is bounded, i.e., $e^{u^\top X_t + X_t^\top Y X_T} \in B(\mathbb{R}^d)$, is essential for the expectation to exist. It is the equivalent of that $u \in i\mathbb{R}^d$ in the affine case of the definition, Definition 2.1.1). Also, for the rest of this thesis, it is assumed without loss of generality that $\omega$ is a symmetric matrix for any $(t, u, V) \in \mathbb{R} \times \mathbb{C}^d \times \mathbb{C}^{d \times d}$.

The next theorem, Theorem 3.1.3, states an important result of quadratic processes and gives an explicit system of differential equations that describes the functions $\varphi, \psi$ and $\omega$ in the definition of a quadratic process. The theorem combines the main result and a corollary of the paper of Chen, Filipović and Poor [7, Theorem 3.6 and Corollary 3.8]. For the theorem, first the definition of admissible parameters should be given.

Definition 3.1.2. A parameter set $(\alpha, b, \beta)$ is said to be admissible if

$$
a \in \text{Sem}_2^d, b \in \mathbb{R}^d, \beta \in \mathbb{R}^{d \times d}.
$$

With the admissible parameter set, the generator$^2$ acting on $f \in C_c^2(\mathbb{R}^d)$, denoted by $\mathcal{A}$, and the function $\varphi, \psi$ and $\omega$ of (3.1) can be given by using the following theorem.

Theorem 3.1.3. Suppose $X$ is a quadratic process, then there exist admissible parameters $(\alpha, b, \beta)$ such that for $f \in C_c^2(\mathbb{R}^d)$,

$$
\mathcal{A}f(x) = \frac{1}{2} \text{tr}(a \nabla^2 f(x)) + (b + \beta x)^\top (\nabla f(x)).
$$

(3.2)

Moreover, the functions $\varphi, \psi$ and $\omega$ in (3.1) satisfy the Riccati equations

$$
\begin{align*}
\partial_t \varphi(t, u, V) &= \frac{1}{2} \psi(t, u, V)^\top a^\top \psi(t, u, V) + \text{tr}(a \omega(t, u, V)) + b^\top \psi(t, u, V), \\
\varphi(0, u, V) &= 0, \\
\partial_t \psi_i(t, u, V) &= 2(\alpha \psi(t, u, V))^\top \omega_i(t, u, V) + \beta_i \psi(t, u, V) + 2b^\top \omega_i(t, u, V), \\

\psi(0, u, V) &= u, \\
\partial_t \omega_i(t, u, V) &= 2(\alpha \omega(t, u, V))^\top \omega_i(t, u, V) + \beta_i \omega^j(t, u, V) + \beta_j \omega_i(t, u, V), \\
\omega(0, u, V) &= V,
\end{align*}
$$

(3.3)

where $\omega^i$ represents the $i$-column vector of $\omega$ and $\beta_i$ the $i$-th row vector of $\beta$.

Conversely, let $(\alpha, b, \beta)$ admissible parameters, then the generator is given by (3.2) and (3.1) holds where $\varphi, \psi$ and $\omega$ are given as unique solutions of the Riccati equations (3.3).

With this theorem one can derive the pricing function for the zero-coupon bond. However, like in the affine case, the theorem has to be rephrased such that it also considers the short rate process $r_t$, since for the risk-neutral measure $\mathbb{Q}$ the formula for the price of a zero-coupon bond $P(t, T)$ depends on its corresponding short rate process $r_t$

$$
P(t, T) = \mathbb{E}_\mathbb{Q} \left[ e^{- \int_t^T r_s \, ds} \mid \mathcal{F}_t \right].
$$

(3.4)

Everything that follows is defined under the risk-neutral measure $\mathbb{Q}$.

$^2$More about generator can for example be found in [27, 2].
In the affine case, first, in context of the characteristic function, a separate system of Riccati equations was given for which $\varphi$ and $\psi$ are solutions. Then, this system is extended to a second system of Riccati equations with $\Phi$ and $\Psi$, to derive the $\mathcal{F}_t$ conditional characteristic function of $X(T)$ under the $T$-forward measure. For the quadratic model, this is done equivalently: the first system of Riccati equations, consisting of the equations for $\varphi, \psi$ and $\omega$, corresponds to the set (3.3) given in the previous Theorem 3.1.3. The extension is given in the next section.

3.2. Canonical state space for quadratic models

For the affine models the canonical state space $X = \mathbb{R}_m^+ \times \mathbb{R}_n$ was considered and for the quadratic case a similar approach is aimed for. However, in the quadratic case, the $\mathbb{R}_m^+$-part is not necessary. This follows from the to-be-proven Theorem 3.3.2 that states that the diffusion term for quadratic models should be constant and nonsingular. Thus, the canonical state space for quadratic models is just $X = \mathbb{R}_d$, where $d$ is the dimension of the model. To maintain the link to the affine models, the same notation is used in the formulation of the system of differential equations, where $m = 0$ and $n \in \mathbb{N}$: the corresponding index sets are

$I = \emptyset$ and $J = \{1, \ldots, n\}$,

where $X = \mathbb{R}_n$, so $n$ equals the dimension $d$.

3.3. Pricing with Quadratic models

As already pointed out in the earlier chapter about affine models, one needs a short-rate model $r_t$ to price derivatives. Since the definition of a quadratic model $X_t$ is now given, the quadratic form of models of $r_t$ can be stated; this quadratic class of term-structure models is defined in terms of the form of the prices of zero-coupon bonds, and is presented in the following definition.

**Definition 3.3.1.** The quadratic class of term-structure models contains models with prices for zero-coupon bonds, $P(t, T)$, that are exponential-quadratic functions of $X_t$ such that

$$P(t, T) = e^{-A(\tau) - B(\tau)^{\top} X_t - X_t^{\top} C(\tau) X_t},$$

(3.5)

where $\tau = T - t$, $A(\tau)$ is a scalar, $B(\tau)$ is a $d$-vector and $C(\tau)$ is a non-singular $d \times d$-matrix [25].

Note that, as in the affine case (Corollary 2.3.2), the fact that $P(T, T) = 1$ for all $X_t \in X$ implies boundary conditions. The boundary conditions are

$$P(T, T) = e^{-A(0) - B(0)^{\top} X_T - X_T^{\top} C(0) X_T} = 1$$

$\iff A(0) = 0, \ B(0) = 0$ and $C(0) = 0$.

The following theorem states under what necessary and sufficient conditions a term structure belongs to the quadratic class [25, Proposition 1].

**Theorem 3.3.2.** A term-structure belongs to the quadratic class if and only if

i) The instantaneous interest rate $r(X_t)$ is a quadratic function of $X_t$:

$$r(X_t) = A_r + B_r^{\top} X_t + X_t^{\top} C_r X_t,$$

(3.6)

with $A_r \in \mathbb{R}, B_r \in \mathbb{R}^n$ and $C_r \in \mathbb{R}^{n \times n}$.
ii) The drift of process $X_t$ under the risk-neutral measure, denoted by $b(X_t)$, is affine in $X_t$, thus $b(X_t) = b + \beta X_t$, with $b \in \mathbb{R}$ and $\beta \in \mathbb{R}^n$.

iii) The diffusion of the process $X_t$ denoted by $\sigma(X_t) = \sigma$ is a constant matrix with $\sigma \in \mathbb{R}^{n \times n}$.

Proof. ‘$\Rightarrow$’

Assume the term-structure to belong to the quadratic class. Then

$$P(t,T) = \mathbb{E}_Q \left[ e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right] = e^{-A(\tau) - B(\tau)^\top X_t - X_t^\top (C(\tau) X_t)}$$

where $A(\tau) \in \mathbb{R}$, $B(\tau) \in \mathbb{R}^d$ and $C(\tau) \in \mathbb{R}^d \times \mathbb{R}^d$ which is non-singular. Since the assumptions of the Feynman-Kac formula hold, this formula implies

$$r(t)P(t,T) = \frac{\partial P(t,T)}{\partial t} + AP(t,T), \quad (3.7)$$

where $A$ is the infinitesimal generator on $X_t$ under the risk-neutral measure $Q$ such that

$$AP(t,T) = \left[ \frac{\partial P(t,T)}{\partial X_i} \right]^\top b(X_t) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[ \frac{\partial^2 P(t,T)}{\partial X_i \partial X_j} \right]_{ij} [\sigma(X_t)\sigma(X_t)^\top]_{ij},$$

where the subscript $ij$ denotes the $(i,j)$th element of the matrix. The expression for $P(t,T)$ implies that $r(X_t)$ is a quadratic function of $X_t$. Hence, it is of the form as claimed in statement i). To prove the other two statements ii) and iii), the Feynman-Kac formula is worked out. The process $r(X_t)$ can be written down by taking partial derivatives of $P(t,T)$. Note that the partial derivatives to $t$ are changed to partial derivatives to $\tau$, which causes a change in sign. Substituting the partial derivatives leads to

$$r(t) = \frac{\partial A(\tau)}{\partial \tau} + \left[ \frac{\partial B(\tau)}{\partial \tau} \right]^\top X_t + X_t^\top \left[ \frac{\partial C(\tau)}{\partial \tau} \right] X_t - [B(\tau) + 2C(\tau) X_t]^\top b(X_t)$$

$$- \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[ 2C(\tau) - [2C(\tau) X_t + B(\tau)] [2C(\tau) X_t + B(\tau)]^\top \right]_{ij} [\sigma(X_t)\sigma(X_t)^\top]_{ij}.$$

Since $r(X_t)$ is quadratic in $X_t$, $\sigma(X_t)\sigma(X_t)^\top$ is constant, because if it would be dependent of $X_t$, $r(X_t)$ would have a higher degree. For the same reasons, $b(X_t)$ is of (maximal) degree one, hence it is affine in $X_t$. Note that here it is needed that $C(\tau)$ is non-singular, otherwise $\sigma(X_t)\sigma(X_t)^\top$ could be dependent of $X_t$ and $b(X_t)$ could also be quadratic in $X_t$, while it is just argued that this is not possible. For the converse argument, see [25, Appendix C].

Remark. The third condition implies that the quadratic class only consists of Ornstein-Uhlenbeck processes, so processes with a constant diffusion term. As stated in the proof, if $\sigma$ is dependent of $X_t$, the expression for $r(X_t)$ has degree higher than two, while the pricing formula implies that it has degree of at most two, so that would lead to a contradiction.

For the quadratic class of term-structures this pricing function is thus an exponential-quadratic function of $X_t$ given by (3.5). The terms $A(\tau), B(\tau)$ and $C(\tau)$ in that expression can be made explicit by using the following more general theorem which is a combination of Theorem 3.6 and Proposition 4.4 of Chen [7].

Theorem 3.3.3. For the set admissible parameters $(\alpha, \beta, b)$ and the short rate given by $r_t = A_t + B_t^\top X_t + X_t^\top C_t X_t$ the $\mathcal{F}_t$-conditional function $\mathbb{E} \left[ e^{-\int_t^T r_s ds} e^{u^\top X_T + X_T^\top V X_T} \mid \mathcal{F}_t \right]$ allows an exponential representation dependent on $\Phi(t,u,V), \Psi(t,u,V)$ and $\Omega(t,u,V)$ given by

$$\mathbb{E} \left[ e^{-\int_t^T r_s ds} e^{u^\top X_T + X_T^\top V X_T} \mid \mathcal{F}_t \right] = e^{\Phi(T-t,u,V) + \Psi(T-t,u,V)^\top X_t + X_t^\top \Omega(T-t,u,V) X_t}, \quad (3.8)$$

36
Moreover, the functions $\Phi, \Psi$ and $\Omega$ satisfy the Riccati equations
\[
\begin{align*}
\partial_t \Phi(t,u,V) &= \frac{1}{2} \Psi(t,u,V) \partial^\top \Psi(t,u,V) + \text{tr}(\sigma(t,u,V)) + b^\top \Psi(t,u,V) - A_r, \\
\Phi(0,u,V) &= 0, \\
\partial_t \Psi(t,u,V) &= 2(\sigma(t,u,V))^\top \Omega(t,u,V) + \beta(t.,V) + 2b^\top \Omega(t,u,V) - B_r, \\
\Psi(0,u,V) &= u, \\
\partial_t \Omega_{ij}(t,u,V) &= 2(\sigma(t,u,V))^\top \Omega(t,u,V) + \beta(t.,V) + \beta_j(t,u,V) + C_r, \\
\Omega(0,u,V) &= V,
\end{align*}
\]
where $\Omega^\top$ represents the $i$-column vector of $\Omega$ and $\beta_i$ the $i$-th row vector of $\beta$.

**Proof.** The proof of this theorem needs some analytical concepts that are out of scope of this thesis. However, to get some insight in the relation between this theorem and Theorem 3.1.3, we consider in detail. The rest of the proof is only considered globally. For more details, see [7, Appendix B].

The proof uses the definition of regular quadratic semigroups and their infinitesimal generators and resolvents. The generator $\mathcal{A}$ corresponds to $P_1 f_{u,V}(x) = \mathbb{E} \left[ e^{\int_0^t X_s^\top V X_s | \mathcal{F}_t} \right]$ and the generator $\mathcal{G}$ corresponds to $Q_t f_{u,V}(x) = \mathbb{E} \left[ e^{-\int_0^t r(X_s)ds} e^{\int_0^s X_s^\top V X_s | \mathcal{F}_s} \right]$, which implies the relations
\[
\partial_t P_1 f_{u,V}(x) = \mathcal{A} P_1 f_{u,V}(x) \quad \text{and} \quad \partial_t Q_t f_{u,V}(x) = \mathcal{G} Q_t f_{u,V}(x).
\]

The resolvent of $(P_1)$ and $(Q_t)$ are denoted with $R_1^P g(x)$ and $R_1^Q g(x)$. It will be used that
\[
R_1^P = (\lambda - \mathcal{A})^{-1} \quad \text{and} \quad R_1^Q g = R_1^Q g + \mathcal{R}(R_1^P g(x)).
\]

Moreover, if one writes $f = R_1^P g$, then $(\lambda - \mathcal{A}) R_1^P g = g \iff (\lambda - \mathcal{A}) f = g$. Using the given equalities, the following is obtained
\[
(\lambda - \mathcal{G}) f = (\lambda - \mathcal{G}) R_1^P g = (\lambda - \mathcal{G}) \left( R_1^Q g + \mathcal{R}(R_1^P g) \right) = (\lambda - \mathcal{G}) (\lambda - \mathcal{A})^{-1} (g + \mathcal{R}(R_1^P g)) = g + \mathcal{R} \left( (\lambda - \mathcal{A})^{-1} g \right) = (\lambda - \mathcal{A}) f + \mathcal{R} ((\lambda - \mathcal{A})^{-1} (\lambda - \mathcal{A}) f) = (\lambda - \mathcal{A}) f + \mathcal{R}(f) = (\lambda - (\lambda - \mathcal{A})) f.
\]

Thus, with this it can be concluded that $\mathcal{G} = \mathcal{A} - \mathcal{R}$. The definition of semigroups and their generators, and equation (3.10), imply that $Q_t f_{u,V}(x) = \mathbb{E} \left[ e^{-\int_0^t r(X_s)ds} e^{\int_0^s X_s^\top V X_s | \mathcal{F}_s} \right]$ is the unique solution to the initial value problem
\[
\partial_t g(t,x) = \mathcal{G} g(t,x), \quad g(0,x) = e^{\int_0^x X_s^\top X_s | \mathcal{F}_0} = e^{u^\top x + x^\top V x}.
\]

where it is easily seen that indeed $g(0,x) = Q_0 f_{u,V}(x) = \mathbb{E} \left[ e^{u^\top x + x^\top V x | \mathcal{F}_0} \right] = e^{u^\top x + x^\top V x}.

The following claim will be proved in more detail.
Claim. The function $e^{\Phi+\Psi^T x + x^T \Omega x}$ is a solution of the initial value problem, where $\Phi$, $\Psi$ and $\Omega$ are the solutions of the Riccati equations (3.3) replaced by $\Phi = \varphi - A_r$, $\Psi = \psi - B_r$ and $\Omega = \omega - C_r$.

Proof of claim. To prove the claim, it should be proved that the generator $G$ acting on $e^{\Phi+\Psi^T x + x^T \Omega x}$ equals the derivative of the function with respect to $t$. Using that $G = A - R$, and that the generator $A$ acting on the function is given by Theorem 3.1.3, one has

$$G \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right)$$

$$= A \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right) - R \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right)$$

$$= \frac{1}{2} \text{tr} \left( \frac{\partial^2}{\partial x^2} \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right) + (b + \beta x)^T \left( \nabla e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right) \right)$$

$$- r(x) \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right).$$

(3.11)

For the first part (1), the Hessian matrix is written down as

$$\nabla^2 \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right) = \begin{bmatrix} \frac{\partial^2 e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x}}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x}}{\partial x_1 \partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x}}{\partial x_d \partial x_1} & \cdots & \frac{\partial^2 e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x}}{\partial x_d \partial x_d} \end{bmatrix}.$$ 

Thus, for $f(x) = e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x}$

$$= \frac{1}{2} \text{tr} \left[ a \nabla^2 \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V) x} \right) \right]$$

$$= \frac{1}{2} \text{tr} \left[ \sum_{j=1}^d \frac{\partial^2 f(x)}{\partial x_j \partial x_1} a_{1j} + \cdots + \sum_{j=1}^d \frac{\partial^2 f(x)}{\partial x_j \partial x_d} a_{dj} \right]$$

$$= \frac{1}{2} \left( \sum_{j=1}^d \frac{\partial^2 f(x)}{\partial x_j \partial x_1} a_{1j} + \cdots + \sum_{j=1}^d \frac{\partial^2 f(x)}{\partial x_j \partial x_d} a_{dj} \right).$$

The second-order derivative in this sum is given by

$$\frac{\partial^2 f(x)}{\partial x_j \partial x_k} = \frac{\partial e^{\Phi+\sum_{i=1}^d \Psi_i x_i + \sum_{j=1}^d \Omega_{ik} x_i}}{\partial x_j \partial x_k}$$

$$= \frac{\partial}{\partial x_j} \left( e^{\Phi+\sum_{i=1}^d \Psi_i x_i + \sum_{j=1}^d \Omega_{ik} x_i} \left( \Psi_k + 2 \sum_{i=1}^d x_i \Omega_{ik} \right) \right)$$

$$= f(x) (2\Omega_{jk}) + f(x) \left( \Psi_j + 2 \sum_{i=1}^d \Omega_{ij} x_i \right) \left( \Psi_k + 2 \sum_{i=1}^d \Omega_{ik} x_i \right)$$

$$= f(x) (2\Omega_{jk}) + f(x) \left( \Psi_j \Psi_k + 2 \Psi_j \sum_{i=1}^d \Omega_{ik} x_i + 2 \Psi_k \sum_{i=1}^d \Omega_{ik} x_i + 4 \sum_{i=1}^d \Omega_{ik} x_i \sum_{i=1}^d \Omega_{ij} x_i \right).$$
So, substituting this expression in the expression for (1) gives

\[
(1) = \frac{1}{2} \sum_{j,k=1}^{d} a_{kj} \frac{\partial^2 f(x)}{\partial x_j \partial x_k}
\]

\[
= \frac{1}{2} \sum_{j,k=1}^{d} a_{kj} \left[ (2 \Omega_{jk}) + \left( \Psi_j \Psi_k + 2 \Psi_j \sum_{i=1}^{d} \Omega_{ik} x_i + 2 \Psi_k \sum_{i=1}^{d} \Omega_{ij} x_i + 4 \sum_{i=1}^{d} \Omega_{ik} x_i \sum_{i=1}^{d} \Omega_{ij} x_i \right) \right] f(x)
\]

\[
= \left[ \text{tr}(a \Omega) + \frac{1}{2} \Psi^T a \Psi + \sum_{j,k=1}^{d} a_{kj} \left( \Psi_j \sum_{i=1}^{d} \Omega_{ik} x_i + \Psi_k \sum_{i=1}^{d} \Omega_{ij} x_i \right) \right] f(x)
\]

\[
+ 2 \sum_{j,k=1}^{d} a_{kj} \left( \sum_{i=1}^{d} \Omega_{ik} x_i \sum_{i=1}^{d} \Omega_{ij} x_i \right) f(x)
\]

\[
= \left[ \text{tr}(a \Omega) + \frac{1}{2} \Psi^T a \Psi + \sum_{j,k=1}^{d} a_{kj} \left( \Psi_j \sum_{i=1}^{d} \Omega_{ik} x_i + \Psi_k \sum_{i=1}^{d} \Omega_{ij} x_i \right) \right] f(x)
\]

\[
+ 2 \sum_{l,k=1}^{d} a_{lk} \left( \sum_{j=1}^{d} \Omega_{lj} x_j \sum_{i=1}^{d} \Omega_{ik} x_i \right) f(x)
\]

The second term (2) of the expression equals

\[
(2) = (b + \beta x)^T \left( \nabla \left( e^{\phi + \Psi^T x + x^T \Omega} \right) \right)
\]

\[
= (b + \beta x)^T \left( \nabla \left( e^{\phi + \Psi^T x + x^T \Omega x} \right) \right)
\]

\[
= (b + \beta x)^T \left( \nabla \left( e^{\phi + \sum_{i=1}^{d} \Psi_i x_i + \sum_{i,j=1}^{d} x_i x_j \Omega_{ij}} \right) \right)
\]

\[
= (b + \beta x)^T \left[ f(x)(\Psi_1 + 2 \sum_{i=1}^{d} \Omega_{1i} x_i) \ldots f(x)(\Psi_d + 2 \sum_{i=1}^{d} \Omega_{id} x_i) \right]^T
\]

\[
= (b + \beta x)^T \left( \Psi + 2 \Omega^T x \right) f(x)
\]

\[
= (b^T \Psi + (\beta x)^T \Psi + 2b^T \Omega^T x + 2(\beta x)^T \Omega^T x) f(x)
\]

\[
= (b^T \Psi + (\beta x)^T \Psi + 2b^T \Omega^T x + 2x^T \beta^T \Omega x) f(x),
\]

where it is used that \( \Omega \) is symmetric. Lastly, the third term (3) of the expression equals

\[
(3) = r(x) \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V)x} \right)
\]

\[
= (A_r + B_r^T x + x^T C_r x) \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T x + x^T \Omega(t,u,V)x} \right)
\]

\[
= (A_r + B_r^T x + x^T C_r x) f(x).
\]

39
Thus, substituting (1), (2) and (3) into the expression for the generator $G$, (3.11), gives

$$G \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T}x + x^T \Omega(t,u,V)x \right) = \left[ tr(a\Omega) + \frac{1}{2} \Psi^T a^T \Psi + \sum_{j,k=1}^d a_{kj} \left( \Psi_j \sum_{i=1}^d \Omega_{ik} x_i + \Psi_k \sum_{i=1}^d \Omega_{ij} x_j \right) \right. $$

$$+ 2 \sum_{l,k=1}^d a_{lk} \left( \sum_{i,j=1}^d x_i \Omega_{ki} \Omega_{lj} x_j \right) \left[ f(x) + (\beta^T \Psi + (\beta x)^T \Psi + 2b^T \Omega^T x + 2x^T \beta^T \Omega x) f(x) \right. $$

$$\left. - (A_r + B_r^T x + x^T C_r x) \right] f(x) $$

$$= \left[ \left( tr(a\Omega) + \frac{1}{2} \Psi^T a^T \Psi + b^T \Psi - A_r \right) + \left( 2 \sum_{j,k=1}^d a_{kj} \Psi_j \sum_{i=1}^d \Omega_{ik} x_i + (\beta x)^T \Psi + 2b^T \Omega^T x - B_r^T x \right) \right. $$

$$+ \left( 2 \sum_{l,k=1}^d a_{lk} \left( \sum_{i,j=1}^d x_i \Omega_{ki} \Omega_{lj} x_j \right) + 2x^T \beta^T \Omega x - x^T C_r x \right) \right] f(x) $$

$$= \left[ \left( tr(a\Omega) + \frac{1}{2} \Psi^T a^T \Psi + b^T \Psi - A_r \right) + \left( 2 \sum_{j,k=1}^d a_{kj} \Psi_j \sum_{i=1}^d \Omega_{ik} x_i + (\beta x)^T \Psi + 2b^T \Omega^T x - B_r^T x \right) \right. $$

$$+ \left( 2 \sum_{l,k=1}^d a_{lk} \left( \sum_{i,j=1}^d x_i \Omega_{ki} \Omega_{lj} x_j \right) + 2x^T \beta^T \Omega x - x^T C_r x \right) \right] f(x).$$

On the other hand, the derivative with respect to $t$ of $e^{\Phi(t,u,V)+\Psi(t,u,V)^T}x + x^T \Omega(t,u,V)x$ is given by

$$\partial_t \left( e^{\Phi(t,u,V)+\Psi(t,u,V)^T}x + x^T \Omega(t,u,V)x \right) $$

$$= e^{\Phi(t,u,V)+\Psi(t,u,V)^T}x + x^T \Omega(t,u,V)x \partial_t (\Phi(t,u,V) + \Psi(t,u,V)^T x + x^T \Omega(t,u,V)x) $$

$$= f(x) \left[ \partial_t \Phi(t,u,V) + \partial_t \Psi(t,u,V)^T x + x^T \partial_t \Omega(t,u,V)x \right] $$

$$= f(x) \left[ \partial_t \Phi(t,u,V) + \sum_{i=1}^d \partial_t \Psi_i(t,u,V) x_i + \sum_{i,j=1}^d x_i \partial_{ij} \Omega_{ij}(t,u,V) x_j \right]. \quad (3.12)$$

Where the partial derivatives of $\Psi_i(t,u,V)$ and $\Omega_{ij}(t,u,V)$ are given below. Firstly,

$$\partial_t \Psi_i(t,u,V) = 2(a^T (t,u,V))^T \Omega_i(t,u,V) + (\beta^T)^T \Psi(t,u,V) + 2b^T \Omega(t,u,V) - (B_r)_i $$

$$= 2a^T \Omega_i + (\beta^T)^T \Psi + 2b^T \Omega - (B_r)_i $$

$$= 2 \sum_{j=1}^d a_{kj} \Psi_j \Omega_{ji} + \sum_{j=1}^d \beta_{ij} \Psi_j + 2 \sum_{j=1}^d b_j \Omega_{ji} - (B_r)_i.$$
So, summing over 1 to \( d \) gives

\[
\sum_{i=1}^{d} \frac{\partial}{\partial t} \psi_i(t, u, V)x_i = \sum_{i=1}^{d} \left( 2 \sum_{j,k} a_{kj} \psi_j \Omega_{ki} + \sum_{j=1}^{d} \beta_{ij} \psi_j + 2 \sum_{j=1}^{d} b_j \psi_j x_i - (B_r)_i \right) x_i
\]

\[
= 2 \sum_{j,k} a_{kj} \sum_{i=1}^{d} \Omega_{ki} x_i + \sum_{i,j=1}^{d} \beta_{ij} \psi_j x_i + 2 \sum_{i,j=1}^{d} b_j \psi_j x_i - \sum_{i=1}^{d} (B_r)_i x_i
\]

\[
= 2 \sum_{j,k} a_{kj} \sum_{i=1}^{d} \Omega_{ki} x_i + (\beta x)^T \psi + 2 b^T \Omega x - B_r^T x.
\]

Secondly,

\[
\partial_t \Omega_{ij}(t, u, V) = 2(a \Omega^T(t, u, V))^T \Omega^T(t, u, V) + (\beta x)^T \Omega^T(t, u, V) + (\beta x)^T \Omega^T(t, u, V) - (C_r)_{ij}
\]

\[
= 2(a \Omega^T)^T \Omega^T + (\beta x)^T \Omega^T + (\beta x)^T \Omega^T - (C_r)_{ij}
\]

\[
= 2 \sum_{k,l=1}^{d} a_{lk} \Omega_{ki} \Omega_{lj} + \sum_{k=1}^{d} \beta_{ik} \Omega_{kj} + \sum_{k=1}^{d} \beta_{jk} \Omega_{ki} - (C_r)_{ij}.
\]

Thus, the sum from 1 to \( d \) equals

\[
\sum_{i,j=1}^{d} x_i \partial_t \Omega_{ij}(t, u, V)x_j = \sum_{i,j=1}^{d} x_i \left( 2 \sum_{k,l=1}^{d} a_{lk} \Omega_{ki} \Omega_{lj} + \sum_{k=1}^{d} \beta_{ik} \Omega_{kj} + \sum_{k=1}^{d} \beta_{jk} \Omega_{ki} - (C_r)_{ij} \right) x_j
\]

\[
= 2 \sum_{k,l=1}^{d} a_{lk} \sum_{i,j=1}^{d} x_i \Omega_{ki} \Omega_{lj} x_j + \sum_{i,j=1}^{d} x_{i} \left( \sum_{k=1}^{d} \beta_{ik} \Omega_{kj} \right) x_j + \sum_{i,j=1}^{d} x_{i} \left( \sum_{k=1}^{d} \beta_{jk} \Omega_{ki} \right) x_j
\]

\[
- \sum_{i,j=1}^{d} x_i (C_r)_{ij} x_j
\]

\[
= 2 \sum_{k,l=1}^{d} a_{lk} \sum_{i,j=1}^{d} x_i \Omega_{ki} \Omega_{lj} x_j + 2x^T \beta^T \Omega x - x^T C_r x.
\]

Hence, substituting these sums of derivatives in the partial derivative to \( t \) gives

\[
\partial_t \left( \phi(t, u, V) + \psi(t, u, V)^T x + x^T \Omega(t, u, V) x \right)
\]

\[
= \left[ \frac{1}{2} \phi^T \psi + tr(a \Omega) + b^T \psi - A_r \right] + \left( 2 \sum_{j,k} a_{kj} \psi_j \sum_{i=1}^{d} \Omega_{ki} x_i + (\beta x)^T \psi + 2 b^T \Omega^T x - B_r^T x \right)
\]

\[
+ \left( 2 \sum_{k,l=1}^{d} a_{lk} \sum_{i,j=1}^{d} x_i \Omega_{ki} \Omega_{lj} x_j + 2x^T \beta^T \Omega x - x^T C_r x \right) f(x).
\]

Note that this expression exactly equals the expression for the generator (3.12)

\[
G \left( \phi(t, u, V) + \psi(t, u, V)^T x + x^T \Omega(t, u, V) x \right)
\]

\[
= \left[ tr(a \Omega) + \frac{1}{2} \phi^T a \psi + b^T \psi - A_r \right] + \left( 2 \sum_{j,k=1}^{d} a_{kj} \psi_j \sum_{i=1}^{d} \Omega_{ki} x_i + (\beta x)^T \psi + 2 b^T \Omega^T x - B_r^T x \right)
\]

\[
+ \left( 2 \sum_{l,k=1}^{d} a_{lk} \sum_{i,j=1}^{d} x_i \Omega_{ki} \Omega_{lj} x_j + 2x^T \beta^T \Omega x - x^T C_r x \right) f(x).
\]
It can be concluded that
\[ \partial_t \left( e^{\Phi(t,u,V) \cdot \Psi(t,u,V)} x + x^T \Omega(t,u,V) x \right) = \mathcal{G} \left( e^{\Phi(t,u,V) \cdot \Psi(t,u,V)} x + x^T \Omega(t,u,V) x \right), \]
where \( \Phi, \Psi \) and \( \Omega \) are the solutions of the Riccati equations with \( \varphi, \psi \) and \( \omega \) replaced by \( \Phi = \varphi - A_r, \Psi = \psi - B_r \) and \( \Omega = \omega - C_r \).

With the proved claim, the proof of the theorem can be finished. Since
\[ \mathbb{E} \left[ e^{-\int_0^s r(x) ds} e^{u^T x + x^T V x} \right] \]
was the unique solution to the stated initial value problem, it can be concluded that
\[ \mathbb{E} \left[ e^{-\int_0^t r(x) ds} e^{u^T x + x^T V x} \right] = e^{\Phi(t,u,V) \cdot \Psi(t,u,V)} x + x^T \Omega(t,u,V) x \]
holds for all \((t, u, V) \in \mathbb{R}_+ \times \mathcal{C}^d \times \mathcal{C}^{d \times d}\), where \( \Phi, \Psi \) and \( \Omega \) satisfy the system of Riccati equations (3.9).

This theorem can be used to price derivatives with exponential payoff functions. The next section will elaborate on this pricing method.

The aim of this chapter is to use quadratic models for pricing. Like in the affine case, consider the payoff of a zero-coupon bond, which equals the constant function 1. Then the pricing formula (3.4) can be expressed as an exponential function by using Theorem 3.3.3. In order to write down the formula, choose \( u = 0 \) and \( V = 0 \). Then the following corollary holds as an implication of Theorem 3.3.3.

**Corollary 3.3.4.** For any maturity \( T \), the \( T \)-bond price at \( t \leq T \) is given as
\[ P(t, T) = \mathbb{E} \left[ e^{-\int_t^T r(s) ds} \mid \mathcal{F}_t \right] = e^{-A(T-t) - B(T-t)^T X_t - X_t^T C(T-t) X_t}, \]
where \( A(t), B(t) \) and \( C(t) \) are defined by
\[ A(t) = -\Phi(t, 0, 0), \quad B(t) = -\Psi(t, 0, 0) \quad \text{and} \quad C(t) = -\Omega(t, 0, 0), \]
with \( \Phi, \Psi \) and \( \Omega \) are the solutions of the Riccati equations of Theorem 3.3.3.

For the one-dimensional case, these differential equations have closed-form solutions. The theory given by Chen, Filipović and Poor [7, Proposition 4.4 & Section 4.4] can be combined to write down the solutions for the Riccati equations of this one-dimensional quadratic model
\[ \begin{align*}
\partial_t \Phi(t, u, V) &= \frac{1}{2} a \Phi(t, u, V)^2 + a \Omega(t, u, V) + b \Psi(t, u, V) - A_r, \\
\Phi(0, u, V) &= 0, \\
\partial_t \Psi(t, u, V) &= (2a \Omega(t, u, V) + \beta) \Psi(t, u, V) + 2b \Omega(t, u, V) - B_r, \\
\Psi(0, u, V) &= u, \\
\partial_t \Omega(t, u, V) &= 2a \Omega(t, u, V)^2 + 2 \beta \Omega(t, u, V) - C_r, \\
\Omega(0, u, V) &= V. 
\end{align*} \tag{3.13} \]

It can be checked that the solutions of this system Riccati equations are given by
\[ \begin{align*}
\Phi(t, u, V) &= \int_0^t \left( \frac{1}{2} a \Psi(s, u, V)^2 + a \Omega(s, u, V) + b \Psi(s, u, V) - A_r \right) ds, \\
\Psi(t, u, V) &= \frac{L_1(t) u + L_2(t) V + L_3(t)}{L_4(t) V + L_5(t)}, \\
\Omega(t, u, V) &= \frac{L_6(t) V + L_7(t)}{L_4(t) V + L_5(t)}. \tag{3.14} \end{align*} \]
where
\[ L_1(\tau) = 2\Gamma e^{\Gamma \tau / 2} \]
\[ L_2(\tau) = -\frac{b}{a} L_1(\tau) + \frac{8a}{\Gamma} \left( \frac{b\beta}{a} + B \right) \left( e^{\Gamma \tau / 2} - 1 \right)^2 \]
\[ L_3(\tau) = \frac{b}{a} \left( 2\Gamma e^{\Gamma \tau / 2} - L_5(\tau) \right) \]
\[ - \left( \frac{b\beta}{a} + B \right) \frac{4}{\Gamma} \left( e^{\Gamma \tau / 2} - 1 \right) \left( \beta \left( 1 - e^{\Gamma \tau / 2} \right) + \frac{\Gamma}{2} \left( 1 + e^{\Gamma \tau / 2} \right) \right) \]
\[ L_4(\tau) = 4a \left( 1 - e^{\Gamma \tau} \right) \]
\[ L_5(\tau) = \Gamma \left( e^{\Gamma \tau} + 1 \right) - 2\beta \left( e^{\Gamma \tau} - 1 \right) \]
\[ L_6(\tau) = \Gamma \left( e^{\Gamma \tau} + 1 \right) + 2\beta \left( e^{\Gamma \tau} - 1 \right) \]
\[ L_7(\tau) = 2\sqrt{\beta^2 + 2\sigma^2 C_r} \]

and \( \Gamma := 2\sqrt{\beta^2 + 2\sigma^2 C_r} \).

The differential equations for quadratic models look similar to the differential equations of affine models. In fact, affine models and quadratic models overlap for some models and particular choices of parameters, but one is not a subset of the other. The relation between affine and quadratic models will be discussed in Chapter 4.

As for the affine models, for some short-rate models the system of ODE’s (4.4) has a closed-form solution. If there does not exist a closed-form solution, the solutions can be computed numerically using Fourier techniques. One of the key advantages of the affine model was that with the closed-form solutions, derivatives can be priced. In the quadratic setting also a similar pricing technique is aimed for. In the next section, the solutions of the ODE’s will be studied and it will provide a way in which the solutions can be used for pricing derivatives.

3.3.1. Price formula for exponential-quadratic payoffs

In order to apply a similar pricing method as for the affine models, first a particular form of payoffs is considered: the exponential-quadratic payoffs, i.e. a payoff function of the form
\[ \exp \left( -q_1(X_T) - \int_t^T q_2(X_s) ds \right), \] (3.15)

with \( q_j(X) = A_j + B_j^T X + X^T C_j X \). Note that this form of payoff function (3.15) is in line with the price of a zero-coupon bond (3.5) by choosing \( q_1(X_T) = \tau f(t, T) \) and \( q_2(X_T) = 0 \):

\[ \exp \left( -q_1(X_T) - \int_t^T q_2(X_s) ds \right) = \exp (-\tau f(t, T)) \]
\[ = \exp \left( -\tau \left( -\log P(t, T) \right) \right) \]
\[ = P(t, T). \]

Moreover, for assets with an exponential-quadratic payoff function one can write down the time-\( t \) price. This is stated in the following proposition which is a slightly adapted version of Proposition 3 in [25].
Proposition 3.3.5. If an asset has an exponential-quadratic payoff function, then the time-
price is given by

\[
\zeta \left( q_1 + \int_t^T q_2, \tau \right) := \mathbb{E}_Q \left[ \exp \left( - \int_t^T r_s ds \right) \exp \left( -q_1(X_T) - \int_t^T q_2(X_s) ds \right) \bigg| \mathcal{F}_t \right] \\
= \exp \left( -A(\tau) - B(\tau)X_t - X_t^T C(\tau)X_t \right), \tag{3.16}
\]

where \( A(\tau) = -\Phi(\tau, 0, 0), B(\tau) = -\Psi(\tau, 0, 0) \) and \( C(\tau) = -\Omega(\tau, 0, 0) \) with \( \Phi, \Psi \) and \( \Omega \) satisfy
the system of ODE’s given in (3.9), with boundary conditions \( \Phi(0, u, V) = -A_1, \Psi(0, u, V) = -B_1 \) and \( \Omega(0, u, V) = -C_1 \) and with \( \{A_r, B_r, C_r\} \) replaced by \( \{A_r + A_2, B_r + B_2, C_r + C_2\} \).

Proof. The proof of this proposition uses similar techniques as Theorem 3.3.2. First,

\[
\zeta \left( q_1 + \int_t^T q_2, \tau \right) = \mathbb{E}_Q \left[ \exp \left( - \int_t^T r_s ds \right) \exp \left( -q_1(X_T) - \int_t^T q_2(X_s) ds \right) \bigg| \mathcal{F}_t \right] \\
= \mathbb{E}_Q \left[ \exp \left( - \int_t^T \tilde{r}_s ds \right) \exp \left( -q_1(X_T) \right) \bigg| \mathcal{F}_t \right],
\]

where \( \tilde{r}_s = r_s + q_2(X_s) \). As in the proof of Theorem 3.3.2, applying the Feynman-Kač
formula gives

\[ \tilde{r}_s \zeta(\cdot, \tau) = \frac{\partial \zeta(\cdot, \tau)}{\partial t} + \mathcal{A} \zeta(\cdot, \tau), \]

where \( \mathcal{A} \) is the infinitesimal generator as in the proof of Theorem 3.3.2. Assuming that \( \tilde{r}_s \)
had an exponential-quadratic form as (3.16), it is obtained that

\[
\tilde{r}_s = \frac{\partial A(\tau)}{\partial \tau} + \left[ \frac{\partial B(\tau)}{\partial \tau} \right]^\top X_t + X_t^\top \left[ \frac{\partial C(\tau)}{\partial \tau} \right] X_t - [B(\tau) + 2C(\tau)X_t]^\top b(X_t) \\
- \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left[ 2C(\tau) - [2C(\tau)X_t + B(\tau)] [2C(\tau)X_t + B(\tau)]^\top \right] \sigma(X_t)^\top \sigma(X_t)_{ij}.
\]

Since \( \tilde{r}_s = r_s + q_2 \) and since the expression for \( \tilde{r}_s \) includes the same ordinary differen-
tial equations as (3.9), the corresponding ordinary differential equations for \( r_s \), and thus the \( A, B \) and \( C \) in (3.16), are those in (3.9), but with \( \{A_r, B_r, C_r\} \) substituted by
\( \{A_r + A_2, B_r + B_2, C_r + C_2\} \). The boundary conditions change from \( A(0) = 0, B(0) = 0, \) and \( C(0) = 0 \) for \( r_s \) to \( A(0) = A_1, B(0) = B_1 \) and \( C(0) = C_1 \) for \( \tilde{r}_s \). \( \Box \)

3.3.2. Fourier transform techniques for quadratic models

In combination with Proposition 3.3.5 and some Fourier techniques, the price of a call and
put option, and therefore the price of a caption, can be derived in a similar way as in the
affine framework in Section 2.3. Recall that in that section, Fourier techniques were used to
rewrite the payoff function of the options. Therefore, in this section Fourier techniques for
quadratic models are considered.

Consider a claim that has an exponential-quadratic payoff function \( \exp(-q_1(X_T)) \) which is
exercised at time \( T \), when \( q_1(X_T) \leq y \) for some fixed \( y \). Then, the time-\( t \) price of this
claim equals

\[
\pi_{q_1, q_2}^C(y, \tau) = \mathbb{E}_Q \left[ e^{-\int_t^T r_s ds} e^{-q_1(X_T)} \mathbb{I}_{q_1(X_T) \leq y} \bigg| \mathcal{F}_t \right]. \tag{3.17}
\]

Note that for \( y = \infty \), \( \pi_{C}(\infty, \tau) = \zeta(q_1, \tau) \) with \( \zeta \) as given in Proposition 3.3.5. Also, for \( q_1 = 0 \) and \( y = \infty \), this price is exactly the price of a zero-coupon bond, thus
\[ \pi_C(\infty, \tau) = P(t,T). \] Then, the Fourier transform, \( \hat{\pi}_C^{q_i,q_j} \), of this price function \( \pi_C^{q_i,q_j}(y,\tau) \) is given by

\[ \hat{\pi}_C^{q_i,q_j}(\lambda) = \int_{-\infty}^{\infty} e^{\lambda y} d\pi_C^{q_i,q_j}(y), \quad \lambda \in \mathbb{R}, \tag{3.18} \]

where the \( \tau \) is omitted. The Fourier transform is expressed explicitly in the following proposition from [25, Proposition 4].

**Proposition 3.3.6.** The Fourier transform \( \hat{\pi}_C^{q_i,q_j}(\lambda) \) of \( \pi_C^{q_i,q_j}(y,\tau) \) as defined above, is equivalent to the price of an asset with exponential-quadratic terminal payoffs,

\[ \hat{\pi}_C^{q_i,q_j}(\lambda,\tau) = \xi(\xi - i\lambda q_j, \tau) \tag{3.19} \]

\[ = E_Q \left[ e^{-\int_t^T r_s ds} e^{-q_i(X_T) + i\lambda q_j(X_T)} \mathbb{1}_{q_i(X_T) \leq y} \bigg| \mathcal{F}_t \right]. \]

**Proof.** From the definition of the given Fourier transform, Fubini’s theorem applied in equality (\( \ast \)) below and the Fourier transform of the Dirac density applied in equality (\( \ast\ast \)) below, the proposition follows directly by

\[ \hat{\pi}_C^{q_i,q_j}(\lambda,\tau) = \int_{-\infty}^{\infty} e^{\lambda y} d\left( \pi_C^{q_i,q_j}(y,\tau) \right) \]

\[ = \int_{-\infty}^{\infty} e^{\lambda y} d \left( E_Q \left[ e^{-\int_t^T r_s ds} e^{-q_i(X_T) + i\lambda q_j(X_T)} \mathbb{1}_{q_i(X_T) \leq y} \bigg| \mathcal{F}_t \right] \right) \]

\[ \overset{(*)}{=} E_Q \left[ e^{-\int_t^T r_s ds} e^{-q_i(X_T) + i\lambda q_j(X_T)} \mathbb{1}_{q_i(X_T) \leq y} \bigg| \mathcal{F}_t \right] \]

\[ \overset{(**)}{=} E_Q \left[ e^{-\int_t^T r_s ds} e^{-q_i(X_T) + i\lambda q_j(X_T)} \bigg| \mathcal{F}_t \right] \]

\[ = \xi(\xi - i\lambda q_j, \tau). \]

With this Fourier transform, the price of the \( t \)-claim (3.17) can be written down explicitly in the following proposition which also originates from [25, Proposition 5] where also the proof can be found.

**Proposition 3.3.7.** The price \( \pi_C^{q_i,q_j}(y,\tau) \) is given by the following inversion formula,

\[ \pi_C^{q_i,q_j}(y,\tau) = \frac{\hat{\pi}_C^{q_i,q_j}(0,\tau)}{2} + \frac{1}{2\pi} \int_0^{\infty} \frac{e^{i\lambda y} \hat{\pi}_C^{q_i,q_j}(-\lambda,\tau) - e^{-i\lambda y} \hat{\pi}_C^{q_i,q_j}(\lambda,\tau)}{i\lambda} d\lambda. \tag{3.20} \]

So, to derive the price for a claim with an exponential-quadratic payoff function, the upper integral (3.20) should be calculated numerically. As will be shown in the next section, the call option can be expressed in functions of the price given by (3.20).

**Remark.** The derivation of the price of claims that can be expressed in functions of the price given by (3.20) still involves a numerical calculation. However, as Leippold and Wu [25] point out this is already an improvement compared to other methods, since regardless of the dimension, one only needs one numerical integration. Moreover, compared to affine models, the computational efficiency is good and very similar. For affine models also a numerical valuation is needed; in the Vasiček model a valuation of the cumulative normal densities has to be determined.
3.3.3. Price formula of call and put options on a bond

As for affine models, one can also write down the price formula for call and put options on a bond using quadratic models. Proposition 3.3.7 in the previous section gives a general pricing formula for a claim with an exponential-quadratic payoff function. Using (3.17), the price of a call option at \( t \) is obtained

\[
\pi_{\text{call}} = E_Q \left[ e^{-\int_t^T r_s ds} (P(T, S) - K)^+ \bigg| \mathcal{F}_t \right] \\
= E_Q \left[ e^{-\int_t^T r_s ds} P(T, S) 1_{\{P(T, S) \geq K\}} \bigg| \mathcal{F}_t \right] - K E_Q \left[ e^{-\int_t^T r_s ds} 1_{\{P(T, S) \leq K\}} \bigg| \mathcal{F}_t \right] \\
= E_Q \left[ e^{-\int_t^T r_s ds} e^{-q(\tau_p)(X_T)} 1_{\{q(\tau_p)(X_T) \leq -\ln K\}} \bigg| \mathcal{F}_t \right] - K E_Q \left[ e^{-\int_t^T r_s ds} 1_{\{q(\tau_p)(X_T) \leq -\ln K\}} \bigg| \mathcal{F}_t \right]
\]

With Proposition 3.3.7 from this it is derived that

\[
\pi_{\text{call}} = \pi^C_{q(\tau_p),q(\tau_p)}(-\ln K, \tau_c) - \pi^C_{0,-q(\tau_p)}(-\ln K, \tau_c) \\
= \frac{\hat{q}^C_{q(\tau_p),q(\tau_p)}(0, \tau_c)}{2} + \frac{1}{2\pi} \int_0^\infty \frac{e^{-i\lambda\ln K} \hat{q}^C_{q(\tau_p),q(\tau_p)}(-\lambda, \tau_c) - e^{i\lambda\ln K} \hat{q}^C_{q(\tau_p),q(\tau_p)}(\lambda, \tau_c)}{i\lambda} d\lambda
\]

where

\[
\hat{q}^C_{q(\tau_p),q(\tau_p)}(\lambda, \tau_c) = E_Q \left[ e^{-\int_t^T r_s ds} e^{-i\lambda q(\tau_p)(X_T)} + i\lambda q(\tau_p)(X_T) \bigg| \mathcal{F}_t \right],
\]

in which \( q(\tau_p)(X_T) = A(S-T) + B(S-T)^\top X_T + X_T^\top C(S-T)X_T \),

with \( A(\cdot), B(\cdot) \) and \( C(\cdot) \) defined as in Corollary 3.3.4.

3.4. Example of quadratic short-rate model

To have a better insight in quadratic models in particular, an example of a quadratic short-rate model is considered. Consider a one-dimensional OU-process, for which every initial value \( x \in \mathbb{R} \) can be realized as solution of

\[
dX_t = (b + \beta X_t)dt + \sigma dW_t,
\]

where \( b, \beta, \sigma \in \mathbb{R} \) \[7\]. Assume that the short rate is quadratic, then it is of the following form

\[
r_t = A r + B r X_t + X_t^\top C r X_t.
\]

Choose the most simple quadratic model for which the short rate is not affine anymore: \( A r = 0, B r = 1 \) and \( C r = 1 \). Thus,

\[
r_t = X_t + X_t^2.
\]

The price function is given by Corollary 3.3.4 and simplified for dimension one to

\[
P(t, T) = e^{-A(\tau) - B(\tau)X_t - C(\tau)X_t^2},
\]

46
where \( A(\tau) = -\Phi(\tau, 0, 0) \), \( B(\tau) = -\Psi(\tau, 0, 0) \) and \( C(\tau) = -\Omega(\tau, 0, 0) \), with \( \Phi, \Psi \) and \( \Omega \) the solutions of the Riccati equations (3.13). Substituting \( A_\tau = 0, B_\tau = 1 \) and \( C_\tau = 1 \) gives

\[
\Omega(\tau, 0, 0) = \frac{L_\tau(\tau)}{L_5(\tau)} = \frac{2(1 - e^{\Gamma \tau})}{\Gamma (e^{\Gamma \tau} + 1) - 2\beta (e^{\Gamma \tau} - 1)},
\]

where \( \Gamma = 2\sqrt{\beta^2 + 2\sigma^2} \). The expression for \( \Psi(\tau, 0, 0) \) is given by

\[
\Psi(\tau, 0, 0) = \frac{b(2\beta(e^{\Gamma \tau/2} - 1) + 4\Gamma/2e^{\Gamma \tau/2} - \Gamma(e^{\Gamma \tau} + 1)) - 2(\beta b +1)(e^{\Gamma \tau/2} - 1)(\beta(1 - e^{\Gamma \tau/2}) + \Gamma/2(e^{\Gamma \tau} + 1))}{\beta \sigma^2}.
\]

And, lastly the expression for \( \Phi(\tau, 0, 0) \) can be derived from the expression for \( \Psi \) and \( \Omega \), but cannot be simplified further than

\[
\Phi(\tau, 0, 0) = \int_0^\tau \left( \frac{1}{2} \sigma^2 \Psi(s, 0, 0) + 2\sigma^2 \Omega(s, 0, 0) + b\Psi(s, 0, 0) - 1 \right) ds.
\]

### Explicit Call Option Price Formula

As showed in the previous section, the price formula of a call option equals

\[
\pi_{\text{call}} = \frac{\pi C(q(\tau_p), q(\tau_p))(0, \tau_e)}{2 \pi} + \frac{1}{i\lambda} \int_0^\infty e^{-i\lambda \ln K} \pi C(q(\tau_p), q(\tau_p))(-\lambda, \tau_e) - e^{i\lambda \ln \hat{K} \pi C(q(\tau_p), q(\tau_p))(\lambda, \tau_e)} d\lambda
\]

\[
- K \left( \frac{\pi C(q(\tau_p), q(\tau_p))(0, \tau_e)}{2 \pi} + \frac{1}{i\lambda} \int_0^\infty e^{-i\lambda \ln K} \pi C(q(\tau_p), q(\tau_p))(-\lambda, \tau_e) - e^{i\lambda \ln \hat{K} \pi C(q(\tau_p), q(\tau_p))(\lambda, \tau_e)} d\lambda \right),
\]

where

\[
\pi C(q(\tau_p), q(\tau_p))(\lambda, \tau_e) = E_q \left[ e^{-\int_{\tau}^{T} r_s ds} e^{-q(\tau_p)(X_T) + i\lambda q(\tau_p)(X_T)} \right].
\]

The \( q(\tau_p)(X_T) \) is defined as

\[
q(\tau_p)(X_T) = A(S - T) + B(S - T)^T X_T + X_T^T C(S - T)X_T,
\]

with \( A(\cdot), B(\cdot) \) and \( C(\cdot) \) defined as in Corollary 3.3.4. Thus, for this example

\[
q(\tau_p)(X_T) = \int_0^{S - T} \left( \frac{1}{2} \sigma^2 \Psi(s, 0, 0) + 2\sigma^2 \Omega(s, 0, 0) + b\Psi(s, 0, 0) - 1 \right) ds
\]

\[
+ \left( \frac{b(2\beta(e^{\Gamma(S - T)/2} - 1) + 4\Gamma/2e^{\Gamma(S - T)/2} - \Gamma(e^{\Gamma(S - T) + 1)) - 2(\beta b +1)(e^{\Gamma(S - T)/2} - 1)(\beta(1 - e^{\Gamma(S - T)/2}) + \Gamma/2(e^{\Gamma(S - T) + 1))}{\beta \sigma^2} \right) X_T
\]

\[
+ \frac{2(1 - e^{\Gamma(S - T)})}{\Gamma (e^{\Gamma(S - T)} + 1) - 2\beta (e^{\Gamma(S - T)} - 1)} X_T^2.
\]

Thus, although the price formula is still in closed form and can be written in analytical form, the expression is (quite) complex and extensive. Calculating an option price will for example include numerical calculations of the integral. Since the expression is so large and not really workable, pricing the call option in the quadratic case in Chapter 5 is based on a Monte Carlo simulation that uses the analytical bond price formula, but not the analytical expression for the call option.
4. Comparison of affine and quadratic models

The class of quadratic models and affine models overlap, but are not the same: the setup of the quadratic models is based on a \(d\)-dimensional Ornstein Uhlenbeck (OU) state process, where the diffusion term is constant [7], while in the affine models the diffusion term can also be dependent on \(X_t\). The quadratic models on the other hand can have an extra quadratic term in the short rate process, so the class of the quadratic term structure models is an extension of the class of affine models.

The relation between the models is visualized in the Venn diagrams below in Figure 4.1. Suppose \(r_t\) describes an affine model and \(\tilde{r}_t\) describes a quadratic model. Then, both models can be based on the same underlying process \(X_t\). For \(\tilde{r}_t\) to be quadratic, the diffusion of \(X_t\) should be constant as showed in the proof of Theorem 3.3.2; the process \(X_t\) should be an OU-process. Thus, affine models that are based on a underlying OU-process fall in the category of quadratic models; Venn diagram 4.1b visualizes this relation. However, affine models can also be based on processes with a non-constant diffusion term. These models are therefore broader than quadratic models, which is visualized in Venn diagram 4.1a.

\[
dX_t = (b + \beta X_t)dt + \sigma(X_t)dW_t
\]

(a) Venn diagram based on forms of diffusion processes \(dX_t\)

\[
r(X_t) = A_r + B_r^\top X_t + X_t^\top C_r X_t
\]

(b) Venn diagram based on forms of short-rate processes \(r(X_t)\)

Figure 4.1.: Venn diagrams showing the relation between affine, \(A\), and quadratic, \(Q\), models.

To illustrate that the quadratic class is not a subset of the affine class, consider the following example [7, Remark 4.6].

**Example 4.1** Let \(X_t\) a process satisfying

\[
dx_t = \mu(X_t)dt + \sigma(X_t)dW_t = (b + \beta X_t)dt + dW_t
\]

and define the quadratic short rate as \(r_t(X_t) = ||X_t||^2\). Using the Itô formula gives

\[
\begin{align*}
\quad dr_t & = 2X_t dX_t + \frac{1}{2} 2d(X)_t \\
& = 2X_t ((b + \beta X_t)dt + dW_t) + dt \\
& = (2X_t(b + \beta X_t) + 1) dt + 2X_t dW_t \\
& = (2bX_t + 2\beta r_t + 1) dt + 2\sqrt{r_t} d\tilde{W}_t,
\end{align*}
\]

48
where \(d\tilde{W}_t = sgn(X_t)dW_t\), which is again a Brownian motion. This interest rate has a non-affine drift term with respect to the process \(r_t\) and therefore this term structure model \(r_t\) is not affine (Theorem 2.1.2). Also, the process \((r,X)\) is not affine, which can be seen by considering the differential equation for the process and using that affine models should satisfy the admissibility conditions (2.8). The differential equation satisfies

\[
d\left(\begin{array}{c} r \\ X \end{array}\right)_t = \left(\begin{array}{c} \frac{1}{b} \\ 1 \end{array}\right)dt + \left(\begin{array}{cc} \frac{2b}{\beta} & 2b \\ 0 & \beta \end{array}\right) \left(\begin{array}{c} r \\ X \end{array}\right)_t dt + \left(\begin{array}{cc} 2\sqrt{\Gamma(T-t)} & 0 \\ 0 & 1 \end{array}\right) \left(d\tilde{W}ight)_t.
\]

Thus, using the notation of the admissibility conditions (2.8) the following holds

\[
I = \{1\}, \\
J = \{2\}, \\
\hat{B} = \left(\begin{array}{cc} 2\beta & 2b \\ 0 & \beta \end{array}\right).
\]

so \(\hat{B}_{12} = \hat{B}_{12} = 2b \neq 0\). Hence, \((r,X)\) does not satisfy all the admissibility conditions and by Theorem 2.2.1 the process is not affine.

Though, the model is a quadratic model, since it satisfies all the necessary and sufficient conditions for the quadratic class (Theorem 3.3.2): \(r_t\) is a quadratic function of \(X_t\) by definition \((r_t = X_t^2)\), the drift of \(X_t\) is affine in \(X_t\) and the diffusion of \(X_t\) is constant.

Concluding, this example shows that a quadratic model is not necessarily an affine model.

### 4.1. From quadratic models to affine models

The particular affine models with a constant diffusion term are a subset of the quadratic models. This implies that the system of ODE’s (3.9) should also hold for these particular affine models by choosing the right parameters for the quadratic model. To get a better intuition about the relation between the two kinds of models, the ODE’s of the affine models are derived from those of the quadratic models for dimension one.

Consider a one-dimensional short rate model

\[
r_t(X_t) = A_r + B_r X_t,
\]

with

\[
dX_t = (b + \beta X_t)dt + \sigma dW_t
\]

under the risk neutral measure. This \(r_t\) equals the equation for \(r_t\) in the quadratic case (3.6) for \(C_r = 0\). The short rate is an affine model and also a particular quadratic model, the one where \(C_r = 0\). Hence, the price of the zero-coupon bond could be given by the price formula of the affine model (stated in Corollary 2.3.2) and by the formula of the quadratic model, given in Definition 3.3.1. The price formulas can be equated to get the Riccati equations for the affine model by simply considering the quadratic price formula

\[
P(t,T) = e^{\Phi(t-T,0,0) + \Psi^T(t-T,0,0) X(t)} + X(t)^T \Omega(t-T,0,0) X(t).
\]

From (3.14) it is known that \(\Omega(T-t,0,0)\) should satisfy

\[
\Omega(T-t,0,0) = \frac{L_0(T-t) \cdot 0 + L_7(T-t)}{L_4(T-t) \cdot 0 + L_5(T-t)} = \frac{2C_r(1 - e^{\Gamma(T-t)})}{\Gamma(e^{\Gamma(T-t)} + 1) - 2\beta (e^{\Gamma(T-t)} - 1)} = 0,
\]

since \(C_r = 0\). Thus

\[
P(t,T) = e^{\Phi_A(T-t,0) + \Psi_A^T(T-t,0) X(t)} = e^{\Phi_Q(T-t,0,0) + \Psi_Q^T(T-t,0,0) X(t)}
\]

49
where $\Phi_A$ and $\Psi_A$ are the solutions to the Riccati equations in the affine case and $\Phi_Q$ and $\Psi_Q$ are the solutions to the Riccati equations in the quadratic case. Hence, by the equality $\Phi_A(T-t,0) = \Phi_Q(T-t,0)$ and $\Psi_A(T-t,0) = \Psi_Q(T-t,0)$. 

As stated in the beginning of this section, for the quadratic models closed-form solutions exist for dimension one, see (3.14). Since the affine models with constant diffusion term overlap with the quadratic models, the solutions for the quadratic models are also the solutions of the affine models. Here it will be shown that these equations hold for the considered short rate model from the previous chapter with a constant diffusion term, i.e. the Vasiček model.

### Solutions of the Vasiček model

Consider the Vasiček model where $r(X_t) = X_t$ and $dX_t = (\beta + b X_t) + \sigma dW_t$. This $X_t$ is the process that was considered earlier this section, and the short rate $r_t$ is a particular case of the considered model with $A_r = 0$ and $B_r = 1$, so those results can be used. Substituting these parameters in the Riccati equations for the one-dimensional case gives

$$
\begin{align*}
\partial_t \Phi(t,u,V) &= \frac{1}{2} \sigma^2 \Psi(t,u,V)^2 + \sigma^2 \Omega(t,u,V) + b \Psi(t,u,V), \\
\Phi(0,u,V) &= 0, \\
\partial_t \Psi(t,u,V) &= (2 \sigma^2 \Omega(t,u,V) + \beta) \Psi(t,u,V) + 2b \Omega(t,u,V) - 1, \\
\Psi(0,u,V) &= u, \\
\partial_t \Omega(t,u,V) &= 2 \sigma^2 \Omega(t,u,V)^2 + 2 \beta \Omega(t,u,V), \\
\Omega(0,u,V) &= V.
\end{align*}
$$

(4.1)

Moreover, the solutions of the quadratic Riccati equations (3.14) are given by

$$
\begin{align*}
\Phi(\tau,u,V) &= \int_0^\tau \left( \frac{1}{2} \sigma^2 \Psi(s,u,V)^2 + \sigma^2 \Omega(s,u,V) + b \Psi(s,u,V) \right) ds, \\
\Psi(\tau,u,V) &= \frac{L_1(\tau)u + L_2(\tau)V + L_3(\tau)}{L_4(\tau)V + L_5(\tau)}, \\
\Omega(\tau,u,V) &= \frac{L_4(\tau)V + L_2(\tau)}{L_4(\tau)V + L_5(\tau)},
\end{align*}
$$

where

$$
\begin{align*}
L_1(\tau) &= 2 \Gamma e^{\Gamma \tau/2} \\
L_2(\tau) &= -\frac{b}{\sigma^2} L_4(\tau) + \frac{8 \sigma^2}{\Gamma} \left( \frac{b \beta}{\sigma^2} + B_r \right) \left( e^{\Gamma \tau/2} - 1 \right)^2 \\
L_3(\tau) &= \frac{b}{\sigma^2} \left( 2 \Gamma e^{\Gamma \tau/2} - L_5(\tau) \right) \\
&\quad - \left( \frac{b \beta}{\sigma^2} + B_r \right) \frac{4}{\Gamma} \left( e^{\Gamma \tau/2} - 1 \right) \left( \beta \left( 1 - e^{\Gamma \tau/2} \right) + \frac{\Gamma}{2} \left( 1 + e^{\Gamma \tau/2} \right) \right) \\
L_4(\tau) &= 4 \sigma^2 \left( 1 - e^{\Gamma \tau} \right) \\
L_5(\tau) &= \Gamma \left( e^{\Gamma \tau} + 1 \right) - 2 \beta \left( e^{\Gamma \tau} - 1 \right) \\
L_6(\tau) &= \Gamma \left( e^{\Gamma \tau} + 1 \right) + 2 \beta \left( e^{\Gamma \tau} - 1 \right) \\
L_7(\tau) &= 0
\end{align*}
$$

50
and $\Gamma := 2|\beta|$. The aim is to find the (analytical) expression for the Vasiček price and to show that this is the same as in the affine case. For the price of a zero-coupon bond, the expression for $\Phi(\tau,0,0), \Psi(\tau,0,0)$ and $\Omega(\tau,0,0)$ are needed. First, consider the $\Omega(\tau,0,0)$, then since $L_1(\tau) = 0$ also $\Omega(\tau,0,0) = 0$ for all $\tau$. Secondly, the $\Psi$ and $\Phi$ are determined. The expression for $\Gamma$, which is used for $\Psi$, is dependent on the sign on $\beta$: if $\beta > 0$, then $\Gamma = 2\sqrt{\beta^2 + \sigma^2}$, $C_r = 2\beta$, but if $\beta < 0$, then $\Gamma = -2\beta$. Thus, first assume that $\beta > 0$, then it is obtained that

$$\Psi(\tau,0,0) = \frac{L_3(\tau)}{L_5(\tau)}$$

$$= \frac{b}{\sigma^2} \left( \frac{4\beta e^{\beta \tau} - L_5(t)}{\beta (1 - e^{\beta \tau} + \beta (1 + e^{\beta \tau}))} \right)$$

$$= \frac{b}{\sigma^2} \left( \frac{4\beta e^{\beta \tau} - 1}{\beta (1 - e^{\beta \tau} + \beta (1 + e^{\beta \tau}))} \right)$$

$$= \frac{b}{\sigma^2} \left( \frac{4\beta (e^{\beta \tau} - 1)}{\beta (1 - e^{\beta \tau} + \beta (1 + e^{\beta \tau}))} \right)$$

$$= \frac{b}{\sigma^2} \left( \frac{4\beta (e^{\beta \tau} - 1)}{\beta (1 - e^{\beta \tau} + \beta (1 + e^{\beta \tau}))} \right)$$

$$= \frac{e^{\beta \tau} - 1}{\beta}.$$ 

Secondly, if $\beta < 0$, then

$$\Psi(\tau,0,0) = \frac{L_3(\tau)}{L_5(\tau)}$$

$$= \frac{\beta}{\sigma^2} \left( -4\beta e^{-\beta \tau} (1 + e^{\beta \tau}) + \left( \frac{b}{\beta} + 1 \right) \left( e^{-\beta \tau} - 1 \right) \left( \beta (1 - e^{-\beta \tau}) - \beta (1 + e^{-\beta \tau}) \right) \right)$$

$$= \frac{\beta}{\sigma^2} \left( -4\beta e^{-\beta \tau} (1 + e^{\beta \tau}) - \frac{b}{\beta} (1 + e^{\beta \tau}) \left( -4\beta e^{-\beta \tau} + \frac{\beta}{2} (e^{-\beta \tau} - 1) \left( -2\beta e^{-\beta \tau} \right) \right) \right)$$

$$= \frac{e^{\beta \tau} - 1}{\beta}.$$ 

Thus for every $\beta$, $\Psi(\tau,0,0) = \frac{e^{\beta \tau} - 1}{\beta}$. With the expressions for $\Psi$ and $\Omega$, the expression for $\Phi$ can be written down

$$\Phi(\tau,0,0) = \int_0^\tau \left( \frac{1}{2} \sigma^2 \Psi(s,0,0)^2 + \sigma^2 \Omega(s,0,0) + b \Psi(s,0,0) \right) ds$$

$$= \int_0^\tau \left( \frac{1}{2} \sigma^2 \left( \frac{1 - e^{\beta s}}{\beta} \right)^2 \right) ds$$

$$= \int_0^\tau \left( \frac{1}{2} \sigma^2 \left( \frac{1 - e^{\beta s}}{\beta} \right)^2 \right) ds$$

$$= \int_0^\tau \left( \frac{1}{2} \sigma^2 \left( \frac{1 - e^{\beta s}}{\beta} \right)^2 \right) ds$$

$$= \left[ \frac{\sigma^2}{2\beta^2} \left( s - \frac{2}{\beta} e^{\beta s} + \frac{1}{2\beta} e^{2\beta s} \right) + \frac{b}{\beta} \left( s - \frac{1}{\beta} e^{\beta s} \right) \right]$$

$$= \left[ \frac{\sigma^2}{2\beta^2} \left( \tau - \frac{2}{\beta} e^{\beta \tau} + \frac{1}{2\beta} e^{2\beta \tau} \right) + \frac{b}{\beta} \left( \tau - \frac{1}{\beta} e^{\beta \tau} \right) \right]$$

$$= \left[ \frac{\sigma^2}{2\beta^2} \left( \tau - \frac{2}{\beta} e^{\beta \tau} + \frac{1}{2\beta} e^{2\beta \tau} \right) + \frac{b}{\beta} \left( \tau - \frac{1}{\beta} e^{\beta \tau} \right) \right]$$

$$= \frac{1}{2} \sigma^2 \left( \frac{1}{2\beta^2} \left( e^{2\beta \tau} - 4e^{\beta \tau} + 2\beta \tau + 3 \right) + b \left( \frac{e^{\beta \tau} - 1 - 2\beta \tau}{\beta^2} \right) \right).$$
Thus, it is shown that for the Vasiček model the solutions of $\Phi(\tau, 0, 0)$ and $\Psi(\tau, 0, 0)$ are indeed the same (compare the upper solutions with (2.23)), since it is an affine and a quadratic model.

### 4.2. Riccati equations for rescaled processes

Instead of the general process as defined above, sometimes in the literature people work with a simple form of this process under measure $\mathbb{P}$ [25]. This rescaling affects the price formula and therefore also the Riccati equations. The process is scaled such that it is mean reverting to zero and has a volatility equal to the identity matrix $I$. When assuming non-degeneracy conditions and with the possibility to rescale and rotate, the simpler process is defined by

$$dX_t = -\kappa X_t dt + d\tilde{W}_t,$$

which is a process under the $\mathbb{P}$-measure [25]. The $\kappa \in \mathbb{R}^{n \times n}$ determines the speed of reversion. Under the $Q$-measure, this process follows

$$dX_t = (b(X_t) - \rho(X_t)\lambda(X_t))dt + \rho(X_t)dW_t$$

$$= (-\kappa X_t - \lambda(X_t))dt + dW_t$$

$$= (-B\lambda - (\tilde{\kappa} + A\lambda)X_t)dt + dW_t$$

$$= (-B\lambda - \kappa X_t)dt + dW_t^*,$$

where $\kappa = \tilde{\kappa} + A\lambda$ and where $A\lambda \in \mathbb{R}^{n \times n}$ is a symmetric matrix and $B\lambda \in \mathbb{R}^n$. The $A\lambda$ and $B\lambda$ are terms in the market price of risk [25], defined by

$$\lambda(X_t) = A\lambda X_t + B\lambda. \quad (4.3)$$

The Riccati equations for the process become

$$\frac{\partial \Phi(\tau, u, V)}{\partial \tau} = \frac{1}{2} \Psi(\tau, u, V)^\top \Psi(\tau, u, V) - \Psi(\tau, u, V)^\top B\lambda + \text{tr}\Omega(\tau, u, V) - A_v,$$

$$\frac{\partial \Psi(\tau, u, V)}{\partial \tau} = -\kappa^\top \Psi(\tau, u, V) - 2\Omega(\tau, u, V)B\lambda + 2\Omega(\tau, u, V)\Psi(\tau, u, V) - B_v, \quad (4.4)$$

$$\frac{\partial \Omega(\tau, u, V)}{\partial \tau} = -\Omega(\tau, u, V)\kappa - \kappa^\top \Omega(\tau, u, V) + 2\Omega(\tau, u, V)^2 - C_v.$$

These other Riccati equations give the price formula

$$P(t, T) = e^{-A_Q(\tau) - B_Q(\tau)^\top \tilde{X}_t - \tilde{X}_t^\top C_Q(\tau) \tilde{X}_t},$$

with $A_Q(\tau) = -\Phi_Q(\tau, 0, 0), B_Q(\tau) = -\Psi_Q(\tau, 0, 0)$ and $C_Q(\tau) = -\Omega_Q(\tau, 0, 0)$, but now with $\Phi_Q, \Psi_Q$ and $\Omega_Q$ the solutions of the quadratic Riccati equations (4.4).

### 4.2.1. From rescaled quadratic processes to affine processes

To deduce the Riccati equations for the affine model from those of the quadratic model, the process $X_t$ should first be rewritten as the rescaled process. Therefore, define the rescaled process

$$\tilde{X}_t = \sigma^{-1}X_t + \frac{b}{\beta}\sigma^{-1}. \quad (4.5)$$

With the Itô formula it can be derived\(^{1}\) that the drift term then equals $\mu(\tilde{X}_t) = \sigma^{-1}\beta\sigma\tilde{X}_t$ and the diffusion term equals $\sigma(\tilde{X}_t) = I$, thus the differential equation of $\tilde{X}_t$ is given by

\[^{1}\text{Namely, }dX_t = d(\sigma^{-1}X_t + \frac{b}{\beta}\sigma^{-1}) = (\sigma^{-1}(b + \beta X_t))dt + (\sigma^{-1}\beta\sigma)\ dW_t = \left(\sigma^{-1}(b + \beta(\sigma(X_t) - \frac{b}{\beta}\sigma^{-1}))\right)dt + dW_t = \left(\sigma^{-1}b - \beta\frac{b}{\beta}\sigma^{-1} + \sigma^{-1}\beta\sigma\tilde{X}_t\right)dt + dW_t = \sigma^{-1}\beta\sigma\tilde{X}_t dt + dW_t.\]
\[ d\tilde{X}_t = \sigma^{-1}\beta\sigma\tilde{X}_t + dW_t. \] Note that the market price of risk for this process also equals zero, since there is no change of measure, i.e. \( A_\lambda = 0 \) and \( B_\lambda = 0 \).

The expression for \( r_t \) can also be expressed in \( \tilde{X}_t \) and is given by

\[
\begin{align*}
    r_t(X_t) &= A_r + B_r X_t \\
    &= A_r + B_r(\sigma\tilde{X} - \frac{b}{\beta}) \\
    &= A_r - B_r\frac{b}{\beta} + B_r\sigma\tilde{X}_t \\
    &= r(\tilde{X}_t).
\end{align*}
\]

Hence, for \( \tilde{A}_r = A_r - B_r\frac{b}{\beta} \) and \( \tilde{B}_r = B_r\sigma, r(\tilde{X}_t) = \tilde{A}_r + \tilde{B}_r\tilde{X}_t \).

Rewriting the price formula for quadratic models in terms of \( X_t \), using the expression for the rescaled process (4.5) gives

\[
P(t, T) = e^{-A_Q(t)-B_Q(t)\tau}\tilde{X}_t - \tilde{X}_t^\top C_Q(\tau)\tilde{X}_t
\]

\[
= e^{-A_Q(t)-B_Q(t)\tau}(\sigma^{-1}X_t + \frac{b}{\beta}b^{-1}) - (\sigma^{-1}X_t + \frac{b}{\beta}b^{-1})^\top C_Q(\tau)(\sigma^{-1}X_t + \frac{b}{\beta}b^{-1})
\]

\[
= e^{-A_Q(t)-B_Q(t)\tau}(\sigma^{-1}X_t + \frac{b}{\beta}b^{-1}) - C_Q(\tau)(\sigma^{-1}X_t + \frac{b}{\beta}b^{-1})^2
\]

\[
= e^{-A_Q(t)-B_Q(t)\tau}(\sigma^{-1}B_Q(\tau) - \sigma^{-1}B_Q(\tau)X_t - \sigma^{-2}C_Q(\tau)X_t^2 - 2\frac{b}{\beta}\sigma^{-2}C_Q(\tau)X_t - \frac{b^2}{\beta}\sigma^{-2}C_Q(\tau)}
\]

\[
= e^{-A_Q(t)+\frac{b}{\beta}\sigma^{-1}B_Q(\tau)+\frac{b^2}{\beta^2}\sigma^{-2}C_Q(\tau)} - (\sigma^{-1}B_Q(\tau) + 2\frac{b}{\beta}\sigma^{-2}C_Q(\tau))X_t - (\sigma^{-2}C_Q(\tau))X_t^2 \quad (4.6)
\]

As argued earlier in this section, this pricing formula equals the pricing formula for the affine models

\[
P(t, T) = e^{-A_\lambda(\tau)-B_\lambda(\tau)X_t},
\]

where \( A_\lambda(\tau) = \Phi_\lambda(\tau, 0) \) and \( B_\lambda(\tau) = \Psi_\lambda(\tau, 0) \), with \( \Phi_\lambda \) and \( \Psi_\lambda \) the solutions of the Riccati equations (2.13) (Corollary 2.3.2). Thus

\[
\begin{align*}
    A_\lambda(\tau) &= A_Q(\tau) + \frac{b}{\beta}\sigma^{-1}B_Q(\tau) + \frac{b^2}{\beta^2}\sigma^{-2}C_Q(\tau), \\
    B_\lambda(\tau) &= \sigma^{-1}B_Q(\tau) + 2\frac{b}{\beta}\sigma^{-2}C_Q(\tau), \quad (4.7)
\end{align*}
\]

where the \( A_Q, B_Q \) and \( C_Q \) are the terms of the quadratic model for \( \tilde{X}_t \) (4.6), i.e. \( A_Q(\tau) = -\Phi(\tau, 0, 0), B_Q(\tau) = -\Psi(\tau, 0, 0) \) and \( C_Q(\tau, 0, 0) = -\Omega(\tau, 0, 0) \). The equalities in (4.7) also imply

\[
\frac{\partial (-A_\lambda(\tau))}{\partial \tau} = \frac{\partial \left( -A_Q(\tau) + \frac{b}{\beta}\sigma^{-1}B_Q(\tau) \right)}{\partial \tau},
\]

\[
\frac{\partial (-B_\lambda(\tau))}{\partial \tau} = \frac{\partial \left( -\sigma^{-1}B_Q(\tau) \right)}{\partial \tau}.
\]

And therefore,

\[
\begin{align*}
    \frac{\partial (\Phi_\lambda(\tau, 0))}{\partial \tau} &= \frac{\partial \left( -A_Q(\tau) + \frac{b}{\beta}\sigma^{-1}B_Q(\tau) \right)}{\partial \tau}, \\
    \frac{\partial (\Psi_\lambda(\tau, 0))}{\partial \tau} &= \frac{\partial \left( -\sigma^{-1}B_Q(\tau) \right)}{\partial \tau},
\end{align*}
\]

where \( \Phi_\lambda \) and \( \Psi_\lambda \) are the Riccati equations for the affine model for \( X_t \), and where all the terms on the right-hand side are known. Thus, by substituting the terms on the right-hand
side, the Riccati equations of affine model for $X_t$ can be derived. So, substituting the Riccati equations, and $\tilde{B}_\lambda = 0, \tilde{A}_r = A_r - \frac{1}{2}B_r, \kappa = \hat{k} + \tilde{A}_\lambda = -\beta$ and $\tilde{B}_r = \sigma B_r$ gives

$$\frac{\partial (\Phi_A(\tau, 0))}{\partial \tau} = \frac{\partial}{\partial \tau} \left( \left(-A_Q(\tau) + \frac{\beta}{2}\sigma^{-1}B_Q(\tau)\right) \right)$$

$$= \frac{\partial (-A_Q(\tau))}{\partial \tau} + \frac{\beta}{2}\sigma^{-1} \frac{\partial (-B_Q(\tau))}{\partial \tau}$$

$$= \frac{\partial (\Phi_Q(\tau, 0, 0))}{\partial \tau} + \frac{\beta}{2}\sigma^{-1} \frac{\partial (\Psi_Q(\tau, 0, 0))}{\partial \tau}$$

$$= \left(\frac{1}{2}\psi_Q(\tau, 0, 0)^2 - \psi_Q(\tau, 0, 0)\tilde{B}_\lambda + \text{tr}Q(\tau, 0, 0) - \tilde{A}_r\right)$$

$$+ \frac{\beta}{2}\sigma^{-1} \left(-\kappa \psi_Q(\tau, 0, 0) - 2\Omega_Q(\tau, 0, 0)\tilde{B}_\lambda + 2\Omega_Q(\tau, 0, 0)\psi_Q(\tau, 0, 0) - \tilde{B}_r\right)$$

$$= \left(\frac{1}{2}\psi_Q(\tau, 0, 0)^2 - (A_r - \frac{\beta}{2}B_r)\right) + \frac{\beta}{2}\sigma^{-1} \left(\beta \psi_Q(\tau, 0, 0) - \sigma B_r\right)$$

$$= \left(\frac{1}{2}\psi_Q(\tau, 0, 0)^2 - A_r + \frac{\beta}{2}B_r + \beta\sigma^{-1} \psi_Q(\tau, 0, 0) - \frac{\beta}{2}B_r\right)$$

$$= \left(\frac{1}{2}\psi_Q(\tau, 0, 0)^2 + \beta\sigma^{-1} \psi_Q(\tau, 0, 0) - A_r\right)$$

$$= \left(\frac{1}{2}\sigma^2 \phi_A(\tau, 0)^2 + b\psi_A(\tau, 0) - A_r\right), \quad (4.8)$$

where for step (*) it is used that $\psi_Q(\tau, 0, 0) = -B_Q(\tau) \overset{(4.7)}{=} -\sigma B_A(\tau) = \sigma \psi_A(\tau, 0)$.

The derivation of the other Riccati equation goes similarly

$$\frac{\partial (\psi_A(\tau, 0))}{\partial \tau} = \frac{\partial (-\psi^{-1}B_Q(\tau))}{\partial \tau}$$

$$= \sigma^{-1} \frac{\partial (\psi_Q(\tau, 0, 0))}{\partial \tau}$$

$$= \sigma^{-1} \left(-\kappa \psi_Q(\tau, 0, 0) - 2\Omega_Q(\tau, 0, 0)B_\lambda + 2\Omega_Q(\tau, 0, 0)\psi_Q(\tau, 0, 0) - B_r\right)$$

$$= \sigma^{-1} \left(\beta \psi_Q(\tau, 0, 0) - \sigma B_r\right)$$

$$= \sigma^{-1} \beta \psi_Q(\tau, 0, 0) - B_r$$

$$= \beta \psi_A(\tau, 0) - B_r, \quad (4.9)$$

Lastly, the boundary conditions for the affine model should be chosen such that

$$e^{\psi_A(0, u) + \psi_A(0, u)^T x_0} \equiv e^{u^T X_0},$$

since that is part of the definition of an affine model. Hence, $\Phi_A(0, u) = 0$ and $\psi_A(0, u) = u$.

Together with (4.8) and (4.9), the full set of Riccati equations for the affine model is derived

$$\partial_x (\Phi_A(\tau, u)) = \frac{1}{2}\sigma^2 \psi_A(\tau, u)^2 + b\psi_A(\tau, u) - A_r,$$

$$\Phi_A(0, u) = 0,$$

$$\partial_x (\psi_A(\tau, u)) = \beta \psi_A(\tau, u) - B_r,$$

$$\psi_A(0, u) = u,$$

for $u = 0$, which is equal to the Riccati equations for the affine models (2.13).
Solutions of the Vasiček model

Consider the Vasiček model where \( r(X_t) = X_t \) and
\[
dX_t = (\beta + bX_t) + \sigma dW_t.
\]
This \( X_t \) is the process that was considered earlier this section, and the short rate \( r_t \) is a particular case of the considered model with \( A_r = 0 \) and \( B_r = 1 \), so those results can be used. In short, rescaling the process gives a process \( \tilde{X}_t = \sigma^{-1} X_t + \frac{\beta}{\sigma} \) with drift term
\[
\mu(\tilde{X}_t) = -\kappa \tilde{X}_t = -\beta \tilde{X}_t
\]
and diffusion term \( \sigma(\tilde{X}_t) = 1 \). The corresponding parameters are
\[
\tilde{A}_t = 0, \tilde{B}_t = 0, \kappa = A + \tilde{\kappa} = -\beta, \tilde{A}_r = A_r - B_r \frac{\beta}{\sigma} = -\frac{\beta}{\sigma}, \text{ and } \tilde{B}_r = B_r \sigma = \sigma \text{ and the Riccati equations of the model are given by (4.9) and (4.8). Also, since it is an affine model } \Omega_Q(\tau, u, V) = 0. \text{ Substituting these parameters in the solutions of the quadratic Riccati equations (3.14) gives }
\[
\Phi_Q(\tau, u, V) = \int_0^\tau \left( \frac{1}{2} \Psi_Q(s, u, V)^2 + \Omega_Q(s, u, V) - \tilde{B}_s \Psi_Q(s, u, V) - \tilde{A}_s \right) ds
\]
\[
\Psi_Q(\tau, u, V) = \frac{L_1(\tau) u + L_2(\tau) V + L_3(\tau)}{L_4(\tau) V + L_5(\tau)},
\]
where the boundary conditions \( \Phi_Q(0, u, V) = 0, \Psi_Q(0, u, V) = u \) and \( \Omega_Q(0, u, V) = V \) should hold. Since \( \Omega_Q(\tau, u, V) = 0 \) for every \( \tau, V = 0 \) in this model. And \( \Gamma := 2\beta, \) so
\[
L_1(\tau) = 4\beta e^{\beta \tau}
\]
\[
L_2(\tau) = \frac{4}{\beta} (\sigma) (e^{\beta \tau} - 1)^2
\]
\[
L_3(\tau) = -\left( \sigma \right) \frac{2}{\beta} \left( e^{\beta \tau} - 1 \right) \left( \beta \left( 1 - e^{\beta \tau} \right) + \beta \left( 1 + e^{\beta \tau} \right) \right) = -4\sigma (e^{\beta \tau} - 1)
\]
\[
L_4(\tau) = 4 \left( 1 - e^{2\beta \tau} \right)
\]
\[
L_5(\tau) = 2\beta \left( e^{2\beta \tau} + 1 \right) - 2\beta \left( e^{2\beta \tau} - 1 \right) = 4\beta
\]
\[
L_6(\tau) = 2\beta \left( e^{2\beta \tau} + 1 \right) + 2\beta \left( e^{2\beta \tau} - 1 \right) = 4\beta e^{2\beta \tau}
\]
\[
L_7(\tau) = 0.
\]
Thus,
\[
\Psi_Q(\tau, u, V) = \frac{L_1(\tau) u + L_3(\tau)}{L_5(\tau)} = \frac{4\beta e^{\beta \tau} u - 4\sigma (e^{\beta \tau} - 1)}{4\beta} = e^{\beta \tau} u - \frac{\sigma (e^{\beta \tau} - 1)}{\beta}.
\]
Using \( \Psi_Q(\tau, 0, 0) = \sigma \Psi_A(\tau, 0) \)
\[
\Psi_A(\tau, 0) = \sigma^{-1} \left( -\frac{\sigma (e^{\beta \tau} - 1)}{\beta} \right) = -\frac{(e^{\beta \tau} - 1)}{\beta}
\]
The additional \( u \)-term is derived by considering the boundary condition \( \Psi_A(0, u) = u \) and by using that the equation should be of the form \( -\frac{(e^{\beta \tau} - 1)}{\beta} \) with an additional \( e^{\beta \tau} \)-term, since the differential equations should still hold. It is obtained that
\[
\Psi_A(\tau, \bar{u}) = \sigma^{-1} \left( e^{\beta \tau} \bar{u} - \frac{\sigma (e^{\beta \tau} - 1)}{\beta} \right)
\]
\[
\Psi_A(0, u) = u
\]
The other solution is given by

$$\Phi_Q(\tau, u, 0) = \int_0^\tau \left( \frac{1}{2} \Psi_Q(s, u, 0)^2 + \frac{b}{\beta} \right) ds$$

$$= \int_0^\tau \left( \frac{1}{2} \left( e^{2\beta s} u^2 - \sigma \left( \frac{e^{2\beta s} - 1}{\beta} \right)^2 + \frac{b}{\beta} \right) \right) ds$$

$$= \frac{1}{2} \int_0^\tau \left( e^{2\beta s} u^2 - 2 \frac{\sigma e^{2\beta s} u - \sigma e^{\beta s} u}{\beta^2} + \frac{\sigma^2 (e^{2\beta s} - 2e^{\beta s} + 1)}{\beta^2} + \frac{b}{\beta} \right) ds$$

$$= \frac{1}{2} \int_0^\tau \left( u^2 - \frac{2\sigma}{\beta} u + \frac{\sigma^2}{\beta^2} \right) e^{2\beta s} + \left( \frac{2\sigma u}{\beta} - \frac{2\sigma^2}{\beta^2} \right) e^{\beta s} + \left( \frac{\sigma^2 + 2\beta b}{\beta^2} \right) ds$$

$$= \frac{1}{2} \left[ \frac{1}{2\beta} \left( u^2 - \frac{2\sigma}{\beta} u + \frac{\sigma^2}{\beta^2} \right) e^{2\beta s} + \left( \frac{2\sigma u}{\beta} - \frac{2\sigma^2}{\beta^2} \right) e^{\beta s} + \left( \frac{\sigma^2 + 2\beta b}{\beta^2} \right) s \right]_0^\tau$$

$$= \frac{1}{2} \left[ \left( \frac{e^{2\beta \tau} - 1}{2\beta^3} \right) (\sigma^2 e^{2\beta \tau} - 4\sigma^2 e^{\beta \tau} + 3\sigma^2 + 2\sigma^2 \beta \tau) - \frac{u}{\beta^2} (\sigma e^{2\beta \tau} - 2\sigma e^{\beta \tau} + \sigma) + \frac{2\beta \tau}{\beta^2} \right]$$

Also, this solution $\Phi_Q$ has to be rewritten to a solution of the affine diffusion for which it is used that $\Phi_Q(\tau, 0, 0) = -A_Q(\tau) = - \left( A_A(\tau) - \frac{\beta}{2\sigma} B_Q(\tau) \right) = - \left( A_A(\tau) - \frac{\beta}{2} B_A(\tau) \right) = \Phi_A(\tau, 0) - \frac{b}{\beta} \Psi_A(\tau, 0)$. Consider $\bar{u} = \sigma u$ as above, then

$$\Phi_Q(\tau, \bar{u}, 0) + \frac{b}{\beta} \Psi_A(\tau, u) = \frac{\sigma^2}{2} \left[ \frac{u}{2\beta} (e^{2\beta \tau} - 1) + \frac{1}{2\beta^3} (e^{2\beta \tau} - 4e^{\beta \tau} + 3 + 2\beta \tau) - \frac{u}{\beta^2} (e^{2\beta \tau} - 2e^{\beta \tau} + 1) \right]$$

$$+ b \left( \frac{e^{\beta \tau} - 1 - \beta \tau}{\beta^2} - \frac{u}{\beta} - e^{\beta \tau} - 1\right).$$

For $\tau = 0$ this formula equals $\frac{b}{\beta} u$, so subtracting this term gives a formula which satisfies both the boundary condition $\Phi_A(0, u) = 0$ and the differential equation, thus

$$\Phi_A(\tau, u) = \frac{\sigma^2}{2} \left[ \frac{u}{2\beta} (e^{2\beta \tau} - 1) + \frac{1}{2\beta^3} (e^{2\beta \tau} - 4e^{\beta \tau} + 3 + 2\beta \tau) - \frac{u}{\beta^2} (e^{2\beta \tau} - 2e^{\beta \tau} + 1) \right]$$

$$+ b \left( \frac{e^{\beta \tau} - 1 - \beta \tau}{\beta} - u - e^{\beta \tau} - 1\right).$$

The resulting solutions for the Riccati equations are exactly the solutions given in (2.23). Concluding, with quite some rewriting and using the boundary conditions, one can derive the Vasiček solutions from the general solutions given for dimension one in the quadratic model.
5. Financial applications of affine and quadratic models

In the financial market affine and quadratic models are used to price derivatives. The pricing of derivatives starts by choosing an underlying stochastic process $X_t$ and the dependency of the short rate on this process. Specifying the process $X_t$, and therefore also the dynamic term structure model for $r_t$, is a choice of the user of the model. As Dai and Singleton [10] point out, this choice is a trade-off between the computational efficiency and how well the state variables can replicate the econometric scenarios. Affine models are computationally more efficient than quadratic models, nevertheless they are less flexible and therefore less able to replicate more sophisticated derivative prices.

To give an intuition about the pricing, this chapter discusses applications of affine and quadratic models. First a short introduction will be given about short-rate models in practice, where also the Monte Carlo method will be briefly explained. Then the method of calibration be explained with supporting plots. Lastly, three financial products will be priced using the Vasiček and a quadratic model: zero-coupon bonds, call options and put options. All the results in this chapter are made by using Matlab.

5.1. Short-rate models in practice

In this chapter two choices for the underlying stochastic process $X_t$ are considered: the Vasiček model and the CIR model. The Vasiček model, $X_t$, and the CIR model, $\tilde{X}_t$, are described by the following diffusion processes

$$dX_t = (b + \beta X_t)dt + \sigma dW_t$$
$$d\tilde{X}_t = (b + \beta \tilde{X}_t)dt + \sigma \sqrt{\tilde{X}_t}dW_t,$$

(5.1)

where $b, \beta, \sigma \in \mathbb{R}$ and $W_t$ is a one-dimensional Brownian motion with respect to the risk-neutral measure (see also Section 1.3 and Section 2.4). Thus, the dynamics of the processes depend on the choice of parameters $b, \beta$ and $\sigma$ and on a random element that is included in the Brownian motion $W_t$.

For these stochastic processes a random path can be sampled based on the parameters. This is done by fixing the parameters and using the Monte Carlo method with a discretization scheme, called the Euler scheme [4]. It constructs a possible path of the stochastic process $X_t$ satisfying the diffusion process by repeatedly sampling from the standard normal distribution to simulate the increments of the Brownian motion $W_t$. The increments of the process $X_t$ are then described by

$$\delta X_{t+\delta t} = (b + \beta X_t)\delta t + \sigma \delta W_t, \quad \delta W_t \sim \mathcal{N}(0, \delta t).$$

5.1.1. Monte Carlo method on Vasiček and CIR

For an introduction of the Monte Carlo method, the short-rate path in the Vasiček and CIR model are simulated in Figure 5.1 and Figure 5.2. When running a Monte Carlo simulation for the stochastic process long and frequent enough, the simulated values should be distributed like the process. Therefore, the histograms of the last values at time $t = 7$ of the paths are also displayed. Note that for clarity of the plot, the Monte Carlo simulations are only plotted for ten different paths in both models. For the histogram however, 1000 samples are used.
As shown in Section 1.3 the Vasiček model is normally distributed with known mean and variance. In Figure 5.1b the blue line fits a normal distribution to the histogram. This indeed overlaps with the analytical mean of \( e^{\beta t} X_0 - \frac{\sigma^2}{2\beta} (1 - e^{\beta t}) = 0.09 \) and variance of \( \frac{\sigma^2}{2\beta} (e^{2\beta t} - 1) = 0.02 \).

Figure 5.1.: Monte Carlo simulation and corresponding distribution for the Vasiček model with parameters \( b = 0.08, \beta = -0.90, \sigma = 0.1817 \) and \( X_0 = 0.01 \).

The CIR model is not normally distributed, but the mean and variance are known. The histogram in Figure 5.2b implies a slanted normal distribution, since the normal distribution fit (the blue line) does not correspond properly to the histogram. The other fitted distribution, the gamma distribution, fits the histogram better.

Figure 5.2.: \( b = 0.08, \beta = -0.90, \sigma = 0.1817, X_0 = 0.01, T = 20, n = 1000, dt = 20/1000, m = 1000 \)
5.1.2. Sensitivity analysis

Below, the effect of changes in one of the parameters is shown by a plot of random paths of the processes. In this plot $\beta$ is adjusted, but in such way that the fraction $b/|\beta|$, the mean reversion level is kept the same. A path is simulated for these different values of the $\beta$ based for on the same seed. The latter means that previously to the construction of the path, the increments are sampled from the standard normal distribution such that for every short-rate path the randomness is used and only the effect of the parameter is shown.

![Image](https://via.placeholder.com/150)

Figure 5.3.: Plot of the sensitivity of a Vasicek model for changes in the speed of mean reversion $\beta$ with initial value $X_0 = 0.06$, where the parameter $\beta$ in (5.1) varies between -10 and -0.01, and where $b$ also varies to have a long-term mean of $b/|\beta| = 0.06$. The dashed lines represent the confidence bands of two standard deviations.

5.2. Calibration of the models

In order to use the described interest rate models of this thesis, the parameters of the models should be derived. In practice, this is done by a calibration to data from the financial market. There are various sophisticated ways to calibrate interest rate models to observed data, but this calibration is not the aim of this thesis. Therefore, a more simple calibration method is used: a nonlinear least-squares method.

5.2.1. Data set of Euribor rates

The data set used for the calibration contains daily Euribor rates from January 1999 until May 2017. These Euribor data points, denoted by $R(t, T)$ with maturity

$$T \in \{1/52, 1/12, 1/6, 1/4, 1/2, 3/4, 1\}$$

are used to calculate the price of a zero-coupon bond by applying

$$P_{EU}(t, T) = \exp(-R(t, T)(T - t)).$$
This leads to a 4691 × 7 matrix with prices where each column stands for one of the maturities and each row is a daily rate. The matrix consists of the entries as shown in Table 5.1. The

<table>
<thead>
<tr>
<th></th>
<th>( T_1 = 1/52 )</th>
<th>( T_2 = 1/12 )</th>
<th>( T_3 = 1/6 )</th>
<th>( T_4 = 1/4 )</th>
<th>( T_5 = 1/2 )</th>
<th>( T_6 = 3/4 )</th>
<th>( T_7 = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/01/'92</td>
<td>0.9394</td>
<td>0.7633</td>
<td>0.5830</td>
<td>0.4455</td>
<td>0.1996</td>
<td>0.0898</td>
<td>0.0404</td>
</tr>
<tr>
<td>05/01/'92</td>
<td>0.9396</td>
<td>0.7638</td>
<td>0.5840</td>
<td>0.4469</td>
<td>0.2015</td>
<td>0.0912</td>
<td>0.0413</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>09/05/'17</td>
<td>1.0073</td>
<td>1.0317</td>
<td>1.0583</td>
<td>1.0857</td>
<td>1.1326</td>
<td>1.1428</td>
<td>1.1309</td>
</tr>
</tbody>
</table>

Table 5.1.: Prices of a zero-coupon bond for different maturities and dates based on the observed Euribor rates.

Euribor rates and the corresponding bond prices over time are shown in Figure 5.4.

![Euribor from January 1999 until May 2017](image)

![Zero-coupon prices from January 1999 until May 2017](image)

Figure 5.4.

5.2.2. Nonlinear least-squares method for calibration

To estimate the parameters that fit the Euribor data, a calibration is performed using a nonlinear least-squares method. In general, this method searches for values of the parameters such that the squared difference between of the curve evaluated in those parameters and the observed value is minimal. In this setting of calibrating the parameters for an interest rate model, the squared error between the price function based on an interest rate model and the observed price is minimized; this involves minimizing

\[
(P(t, T) - P_{EU}(t, T))^2 = \left( e^{-A(T-t) - B(T-t)^T X_t - X_t^T C(T-t) X_t} - P_{EU}(t, T) \right)^2,
\]

where the price function is from Corollary 3.3.4 and \( A, B \) and \( C \) are dependent on the parameters of the model (i.e. \( b, \beta, \sigma, A_r, B_r \) and \( C_r \)). Note that for affine models \( C_r = 0 \) and \( C(T-t) = 0 \).

In the calibration of the Euribor rates, there are seven data point available per day: one for every maturity. Therefore, to calibrate the model without over-fitting the data, it is necessary to use the smallest number of parameters as possible. The affine cases considered
here are the Vasicek and CIR model where the short-rate model \( r_t = X_t \) is calibrated to the data. As in (2.10) and (3.6), the short-rate model is of the form

\[
r_t = A_r + B_r^T X_t + X_t^T C_r X_t
\]

For the affine and quadratic case, the short-rate models that are calibrated for this thesis are respectively given by

\[
r_t = X_t \quad \text{and} \quad r_t = X_t + C_r X_t^2.
\]

Note that the constant term, \( A_r \), is set to zero for all three models and the linear term \( B_r \) is set to 1.

The minimizing procedure to estimate the parameters is coded in a standard Matlab function called \texttt{lsqnonlin}(). For this thesis this solver uses the vector-valued function

\[
f_{b, \beta, \sigma, X_i}(t, T) = e^{-A(T-t)-B(T-t)^T X_t - X_t^T C(T-t) X_t} - P_{EU}(t, T)
\]

as input, where

\[
f_{b, \beta, \sigma, X_i}(t, T) = (f_{b, \beta, \sigma, X_i}(t, T_1) \quad f_{b, \beta, \sigma, X_i}(t, T_2) \quad \cdots \quad f_{b, \beta, \sigma, X_i}(t, T_7))^T.
\]

In combination with a vector of starting points for the parameters \texttt{lsqnonlin}() minimizes the function and gives the best estimate for the parameters, such that the error is as small as possible. This concludes in one set of parameters for one day. The Matlab function also allows more input as lower and upper bounds, and optimization option such as for example restrictions for the number of iterations or a precision boundary of the value of the error term.

For the calibration in this thesis, some restrictions had to be made for the optimization of the parameters to reduce computational time. Since the calibration is not the aim of the thesis, the \texttt{FunctionTolerance} is set at 0.01 and only a subset of the data is used. This \texttt{FunctionTolerance} terminates the optimization of the parameters, when the change in the value of the objective function during a step is below the set boundary. The subset is the set from January 2010 until the last date in the data set. This choice is substantiated by the fact that the calibration for a quadratic model took quite some time already in the current setup (approximately 9 hours for only 1000 dates).
The process of calibration using the nonlinear least-squares method is summarized in the following steps:

1. The difference between the function and the observed values is captured in the new function $f_{b,\beta,\sigma,X_t}(t,T)$ depending on $t,T,b,\beta,\sigma$ and $X_t$. This function is a seven dimensional vector, since there are seven observed prices for one day (one for every maturity $T_i$).

2. The `lsqnonlin()` function estimates the best fit for $b,\beta,\sigma,X_t$ and for quadratic models also $C_r$ for one day using:
   
   i) The seven zero-coupon bond values for the different maturities;
   
   ii) Start values of the parameters $b = 1.2, \beta = -1, \sigma = 0.01, X_t = 0$ and for quadratic models $C_r = 0$;
   
   iii) Lower and upper bounds for the parameters: $b \in [0,20], \beta \in [-10,0], \sigma \in [0,10], X_t \in [-10,10]$ and for quadratic models $C_r \in [-10,10]$;
   
   iv) In order to reduce computational time, the `FunctionTolerance` is set to 0.01.

   This step returns a matrix of parameter values of length 1877, the number of days in the subset, and width four (five for quadratic models), the number of parameters.

With the calibrated models one can make a fair comparison of the different short-rate models. In Figure 5.5 the estimated parameters per model are shown. It is remarkable that for every model the parameters are less stable after the first negative rate is observed (the black dashed line). Furthermore, the parameters of the Vasićek model and the CIR model are very similar. Also, note the difference in the estimation of the volatility sigma: for the Vasićek model it increases rapidly in the last period, while for the CIR and quadratic model it only increases a little. Lastly, the CIR model shows a dip around the start of 2016 for the beta estimate. This could be explained by the fact that for that moment the curve of the 12-month rate and price crosses the curve of the 9-month rate and price; thus the form of the curve changes which might explain the dip for the beta in the CIR model. However, it is strange that this does not seem to have an effect on the other models.
Figure 5.5.: Zero-coupon bond prices at time 01/2010 until 05/2017 and estimated parameters, $b, \beta, \sigma, X_t$, for the Vasicek, CIR and quadratic models on the same time frame. For the estimations all the available maturities are used to obtain on set of parameters for one date. The black dashed line is drawn at the moment where the first negative rate is observed.

### 5.2.3. Calibration accuracy

In order to give some insight in how well the calibrated parameters describe the observed data, a plot shown in Figure 5.6 is constructed, where the sum of squared errors of the prices of zero-coupon bond for the seven maturities is displayed. For comparison of affine and quadratic models in general, only the Vasicek model and the quadratic model are shown.

The plot shows that for the first part of the data before the end of 2012 (red dashed line) from where the curves of the prices are beginning to flatten out, the Vasicek model performs better, since it has a lower sum of squared errors. However, in the period between the end of 2012 and the moment that the Euribor goes below zero for the first time (black dashed line), both models seem to perform quite well, since they have a very low sum of squared errors. After the first moment of interest rates becoming negative, the Vasicek has a higher error term and it seems that the quadratic model performs better. To test these findings,
the total sum of errors for these three intervals is calculated and displayed in Table 5.2. The table supports the observations from the plot and even shows that the quadratic model performs better for the second interval as well.

Figure 5.6.: Zero-coupon bond prices for different maturities at time 01/2010 until 05/2017 and the sum of squared errors of zero-coupon bond prices modelled by Vasićek and a quadratic model with the calibrated parameters for different maturities on the same time frame. The red line is drawn at 19/10/2012 and the black line is drawn at 05/09/2014 at the point where the first rate was below zero.

<table>
<thead>
<tr>
<th>Period of time</th>
<th>Total sum of squared errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vasićek model</td>
</tr>
<tr>
<td>04/01/2010 - 19/10/2012</td>
<td>0.2634</td>
</tr>
<tr>
<td>22/10/2012 - 05/09/2014</td>
<td>0.0117</td>
</tr>
<tr>
<td>08/09/2014 - 09/05/2017</td>
<td>0.3613</td>
</tr>
</tbody>
</table>

Table 5.2.: Table of total sum of squared errors of the zero-coupon bond prices for three different time periods for the three used models.

To give some intuition about the calibrated values, two examples are given: one at a time that the rate curve is flat (at 30th of August, 2013) and one at the end of the set (at 9th of May, 2017), since the curve looks very different at these moments. The calibrated parameters are displayed in Table 5.3. For the first sample date the parameters for all models are very similar. For all models the $\sigma$ is very low and the mean reversion level of the underlying $X_t$ is approximately 1.3. Also, the factor of the quadratic term, $C_r$, is low, which means that the quadratic model is similar to a normal Vasićek model. The graphs in the first plot of the short rate paths of the Vasićek and Quadratic model, visualized in Figure 5.7, also show these findings. Due to the low $\sigma$, the paths are smooth and both lines tend to the mean reversion level of 1.3. Note that the reversion level of the quadratic model is slightly higher due to the addition of the squared process.
Table 5.3.: Calibrated parameters at August 30, 2013 and May 9, 2017 for the Vasiček, CIR, and Quadratic model based on the Euribor data.

For the second sample the parameters are quite different. First of all, it should be noted that this calibration calibrates the CIR model with a negative initial rate. Actually, the CIR model cannot handle negative rates, but this calibration is still possible since it is calibrated with respect to the zero-coupon bond price, where the short rate in theory can be negative. However, when sampling short-rate paths, a negative initial rate is not possible, since the CIR model has a square root of the initial rate in the diffusion term. Therefore, Matlab returns complex numbers for the short rate path of the CIR model, which are not possible in practice. Thus, for the rest of this chapter, the CIR model is left out for a comparison on the different dates.

The Vasiček model and the quadratic model have a mean reverting level of approximately 0.5, but the $\sigma$ and the coefficient $C_r$ cause very different graphs for these models. The $\sigma$ of the Vasiček model is high, and therefore it has a large deviations. The quadratic model is more flat, since it consists of a Vasiček model minus 0.77 times the square of this Vasiček model and therefore the high peaks level out.

Figure 5.7.: Sample paths based on one seed for Vasiček, CIR and Quadratic model on calibrated parameters as shown in Table 5.3, for starting date August 30th, 2013 and May 9th, 2017.
These examples show that the models are very dependent on the parameters. Also, it shows that the quadratic model based on the Vasiček model and the CIR model simulate more stable short rate paths than the Vasiček model. This could have implications for the zero-coupon bond prices and prices of derivatives based on the different models. Therefore, these samples are also discussed in the next section about pricing.

5.3. Pricing

With the calibrated models, the pricing of derivatives can be performed. As showed in both the chapters about affine models and quadratic models, the price of a zero-coupon bond can be derived analytically. These pricing formulas lead to analytical expressions for call and put options. This section shows results from pricing with an affine model, $r_t$, for which the Vasiček model is used, and pricing with a quadratic model, $\tilde{r}_t$, based on an underlying Vasiček process. So, as in the calibration,

$$r_t = X_t, \quad \tilde{r}_t = X_t + C_t X_t^2,$$

where $X_t$ is a Vasiček model.

5.3.1. Zero-coupon bonds

One of the easier product to price is the zero-coupon bond. For both the affine and the quadratic model, the analytical price can be used with the calibrated parameters to calculate the zero-coupon bond price for different maturities. This pricing is performed for the two moments in time studied in the previous section, with parameters as in Table 5.3, 30th of August in 2013 and 9th of May in 2017. The bond-prices of these models are displayed in Figure 5.8. Note that these prices are calculated with the analytical pricing formulas.

![Zero-coupon bond prices for Vasicek, CIR and Quadratic model on calibrated parameters](image)

Figure 5.8.: Zero-coupon bond prices calculated for different maturities with the analytical bond price formulas using the calibrated parameters at dates August 30, 2013 and May 9, 2017.

Figure 5.8 shows that for the first set of parameters the three models predict very much the same zero-coupon bond prices. The Vasiček and CIR prices even completely overlap, which can be explained by the small $\sigma$: the only difference between the Vasiček and CIR
model is the square-root term of $X_t$, which due to the small $\sigma$, has almost no effect. The small $C_r$ causes the prices of the quadratic model to be slightly higher than the other models.

The second set of parameters shows different results. The large $\sigma$ in the Vasiček model causes a high bond price. The curve of the quadratic model is quite different than that of the Vasiček model. This could be explained by multiple arguments, such as the addition of the quadratic term with coefficient $C_r = -0.77$ which makes the quadratic model different from a normal Vasiček model and the $\sigma$ of the quadratic model that is higher.

5.3.2. Call and Put options

The second class of products is the class of the call and put options. In Chapter 2 the analytical prices for call and put options are derived for the Vasiček and CIR model. In Chapter 3 also an analytical price is written down for a call option, but then for the quadratic model. However, as already noted, the latter price formula is complex and therefore it is decided to price the call and put options for the affine and the quadratic model with the Monte Carlo method together with the analytical bond prices.

To calculate the call and put prices on an $S$-bond with expiry date $T < S$ and strike price $K$, their payoff functions are used. A Monte Carlo simulation with a large number of samples (here 1000 samples are used per model), creates possible paths for the short rate up to expiry date $T$, according to the Vasiček model and the quadratic model based on an underlying Vasiček process. For all 1000 values of the short rates in the simulation, the analytical bond price and discount factor are calculated. Then, after calculating the payoffs using the bond and strike prices, the mean of these discounted payoffs gives the estimated call and put prices. The results on the two sample dates for the Vasiček and CIR model are shown in Figure 5.9 and Figure 5.10.

![Call option price for Vasiček model](image1)

![Call option price for Quadratic model](image2)

![Put option price for Vasiček model](image3)

![Put option price for Quadratic model](image4)

Figure 5.9.: Prices of call and put options for a Vasiček and a quadratic model for the second sample date: the 30th of August in 2013.
The graphs of the call and put prices show that on the first sample date, the 30th of August in 2013, the prices for the Vasiček model and the quadratic model are almost the same. Since the example of the sampled short-rates paths and the prices of the zero-coupon bonds also were very similar for these models, this is what is expected. The graphs of the second sample date, the 9th of May in 2017, are quite different. The short-rate paths showed in Figure 5.7 already implied this pricing difference: since the volatility of the Vasiček model is high compared to that of the quadratic model, the option prices for this model are also high. Since call and put options have an asymmetric payoff pattern without the risk of negative payoff, a higher volatility implies a higher price for both options.
6. Discussion

This thesis studied affine and quadratic interest rate models in mostly a theoretical and partially empirical way. Affine models have already proven to be very delicate in modeling short-rates because of their analytical forms and computational tractability. However, quadratic short-rate models are not widely studied yet. In this thesis it is shown that quadratic models have similar analytical forms, although they are somewhat more sophisticated. For affine and quadratic models, it has been proven that zero-coupon bond prices can be written down in closed form, assuming the so-called admissible parameters and using the Riccati equations, specified by the model. For one-dimensional quadratic models, even closed-form solutions are written down for those models. Moreover, also analytical forms for call and put options can be written down for both the affine and quadratic models.

On the computational side however, quadratic models do need more time. The quadratic term in these models cause a longer computation, but it is empirically shown that the quadratic models do model different short-rates, and therefore also different prices of derivatives. In times with negative, or very low, interest rates, the calibrated quadratic model had a lower calibration error than the Vasićek and the CIR model. Notably, the calibration of the quadratic model also resulted in a lower volatility term than the other models, which implies a different modeling of the market.

This thesis mostly aimed at the theoretical aspect of affine and quadratic models and pointed out the differences and similarities. This involved quite sophisticated mathematics, which caused that there was less time for the computational part. Therefore, to conclude whether quadratic models are a valuable addition to the short-rate modeling (in practice), more research in the computational sense should be done. However, there is already some evidence that quadratic models could improve modeling short-rates (for example see [1, 8, 26]). Also, because of the (just) observed negative interest rates and the performing of the quadratic models for these rates, further research could indicate if quadratic models are computational feasible.

6.1. Models in practice

That the affine and quadratic models are mathematically elegant and are (probably) both computational tractable, is not sufficient for the models to be used in the financial industry. During the TopQuants event in Autumn Event 2016 on November 24th, Peter den Iseger discussed 'New developments in risk modelling: affine models', which showed that affine models are just being introduced in financial risk modeling. For transparency in the financial market, it is important that every model is explainable. However, the affine models turned out to be already quite hard to explain, thus introducing even more complicated models such as quadratic models, could cause issues.

Another point of concern of using sophisticated models in general, is that these models are less transparent and therefore more prone to errors. Also, complicated models could imply false accuracy, while there are still only a model of the real world. Lastly, the question remains whether it is worth all the effort, if there already exist models that explain the real world pretty well.
Layman’s summary

In the financial world people try to speculate about the financial market. There are many variables that are unknown and that one wants to describe by, for example, stochastic models. These models help to get insight in the financial variables and are sometimes even used to predict the future development of the variable in order to do proper investments or protect themselves against risk. The latter, in form of interest rate risk modeling, is studied in this thesis.

An interest rate is defined as the rate that is charged for borrowing money and is expressed as a percentage of the loaned amount. If you put money on your savings-account, you will be rewarded by a positive interest rate, since the bank is actually borrowing, or using, your money. However, lately this interest rate developed differently than the financial experts expected and dropped below zero. If the bank would directly charge this change of interest rates to their customers, this would imply that you have to pay money for saving money. Luckily, the banks take the losses. If the experts would have considered that the interest rates could drop below zero, they could have hedged (protected themselves) against this risk, such that the bank not looses money unexpectedly. In general, there could be gained a lot in modeling and understanding the interest rates.

The interest rate risk models play are very important in the financial industry. Banks and insurance companies, for example, heavily rely on the models for managing their risk [23]. The level of the interest rate is not only important for Dutch banks and our personal bank account, it is also very important globally: according to The Bank for International Settlements (BIS), the worldwide amount of debt instruments in which interest plays a big role, is no less than $21,288 billion [11]. Moreover, in The Netherlands alone the total amount of mortgage loan outstanding is €650 billion, which equals 95% of the gross domestic product [12]. These examples show that the interest rate level has a high impact, even on a global scale, which proves that it is important to model them properly.

One of the popular models of interest rates is the class of affine models. They are becoming increasingly popular due to their analytical and computational tractability. Affine processes have a nice pricing formula for multiple financial products. Quadratic processes are, to some extent, an extension of affine models and have similar properties as affine models. This thesis compares these affine and quadratic models on a theoretical and an empirical level. For the theoretical level, the mathematics of affine and quadratic interest rate models is explained. For both affine and quadratic models analytical (‘nice’) formulas for some financial products are provided using admissible parameters and Riccati equations. Also, using the analytical bond prices, a small empirical comparison is performed where some computational examples are discussed.
Bibliography


A. Basics

A.1. Interest rates

Definition A.1.1. The simple spot rate for \([t, T]\) is given by

\[ F(t, T) = \frac{1}{T-t} \left( \frac{1}{P(t, T)} - 1 \right), \]  

(A.1)

where \(P(t, T)\) is the value of a zero-coupon bond with maturity \(T\).

Definition A.1.2. The continuously compounded forward rate for \([T, S]\) is given by

\[ R(t; T, S) = \frac{-\log P(t, S) - \log P(t, T)}{S - T}, \]  

(A.2)

where \(P(t, T)\) is the value of a zero-coupon bond with maturity \(T\).

Definition A.1.3. The expression of the instantaneous forward rate with maturity \(T\) at time \(t\) follows from the continuously compounded forward rate. It is given by

\[ f(t, T) = \lim_{S \downarrow T} R(t; T, S) = -\frac{\partial \log P(t, T)}{\partial T}. \]  

(A.3)

Following from the expression of the forward rate, \(P(t, T)\) can be given in terms of the forward rate

\[ P(t, T) = e^{-\int_{t}^{T} f(t, u) \, du}. \]

Note that for this expression \(P(T, T) = 1\) is required.

The last rate which is often used is the instantaneous short rate.

Definition A.1.4. The short rate also follows from continuously compounded forward rate \(R(t; T, S)\), but via the continuously compounded sport rate on \([t, T]\), \(R(t, T) = R(t; t, T)\). The rate is defined as

\[ r_t = f(t, t) = \lim_{T \downarrow t} R(t, T). \]  

(A.4)

A.2. Itô’s lemma

Theorem A.2.1. (Itô’s lemma) Suppose \(X\) follows the Itô process

\[ dx_t = a(x_t, t) dt + b(x_t, t) dz_t, \]

where \(dz\) is a Brownian Motion and \(a\) and \(b\) are functions of \(x\) and \(t\). Then the function \(G\) of \(x\) and \(t\) follows the process

\[ dG = \left( \partial_{x} a + \partial_{t} + \frac{1}{2} \partial_{x}^{2} b^{2} \right) dt + \partial_{x} b dz_t \quad [20]. \]
A.3. Forward measures

In the theory of interest rate models, it is sometimes useful to consider an other asset as numeraire than the risk-free one. For example with option pricing this technique is used where the traded asset is a T-bond. The following definition and lemma follow [3, Chapter 26.4] and [18, Chapter 7].

A.3.1. Forward as numeraire

Assume that there exists an equivalent martingale measure \( Q \) such that all price processes of the T-bond discounted by \( B(t) \) are \( Q \)-martingales and let \( W \) be a \( Q \)-Brownian motion. Define for fixed \( T > 0 \) an equivalent probability measure \( Q_T \sim Q \) on \( \mathcal{F}_T \) by

\[
\frac{dQ_T}{dQ} = \frac{P(T,T)}{P(0,T)B(T)} = \frac{1}{P(0,T)B(T)}.
\]

The \( T \)-forward measure for \( t \leq T \) is defined by

\[
\frac{dQ_T}{dQ} \bigg|_{\mathcal{F}_t} = \mathbb{E}_Q \left[ \frac{dQ_T}{dQ} \bigg| \mathcal{F}_t \right] = \frac{P(t,T)}{P(0,T)B(t)} = \mathcal{E}_t(v(\cdot, T) \bullet W),
\]

where \( v(\cdot, T) = \int_T^T \sigma(\cdot, u)du \). Define the process \( W_T \) as

\[
W_T(t) = W_t - \int_0^t v(s, T)^T ds, \quad t \leq T,
\]

which is a \( Q_T \)-Brownian motion using Girsanov’s Theorem [18, Theorem 4.6]. This leads to the following lemma for forward measures which is very useful in financial modeling. For the proof, see [18, Lemma 7.1].

**Lemma A.3.1.** For any \( S < 0 \), the \( T \)-bond discounted S-bond price process

\[
\frac{P(t,S)}{P(t,T)} = \frac{P(t,S)}{P(0,T)} \mathcal{E}_t(\sigma_{S,T} \bullet W^T), \quad t \leq S \wedge T
\]

is a \( Q^T \)-martingale, where \( \sigma_{S,T} \) is defined as

\[
\sigma_{S,T}(t) = -\sigma_{T,S}(t) = v(t, S) - v(t, T) = \int_T^S \sigma(t, u) du.
\]

Moreover, the \( T \)- and \( S \)-forward measures are related by

\[
\frac{dQ^S}{dQ^T} \bigg|_{\mathcal{F}_t} = \frac{P(t,S)P(0,T)}{P(t,T)P(0,S)} = \mathcal{E}_t \left( \sigma_{S,T} \bullet W^T \right), \quad t \leq S \wedge T.
\]

---

\( ^1 \)The latter equality follows from the HJM Drift Condition defined in [18, Theorem 6.1] which is not studied in this thesis. The \( \mathcal{E}(X) \) is defined as the stochastic exponential of an Itô process \( X \), i.e. \( \mathcal{E}_t(X) = e^{X_t - \frac{1}{2} \langle X, X \rangle_t} \).
B. Calculations on popular short-rate models

In this chapter some calculations are given for popular short-rate models. It serves as addendum to what is written in the rest of the thesis about the models, so for every model the necessary formulas are recapped at the start, but no context is given.

B.1. Vasiček short-rate model

The Vasiček model is given by
\[ dr_t = (b + \beta r_t)dt + \sigma dW_t, \]
with Riccati equations
\[ \Phi(t, u) = \frac{1}{2} \sigma^2 \int_0^t \Psi^2(s, u)ds + b \int_0^t \Psi(s, u)ds, \]
\[ \partial_t \Psi(t, u) = \beta \Psi(t, u) - 1, \]
\[ \Psi(0, u) = u. \]

B.1.1. Unique global solution of the Vasiček short-rate model

Note that for \( \Psi(t, u) = e^{\beta t}u - \frac{e^{\beta t} - 1}{\beta} \)
\[ \Psi(0, u) = e^{\beta 0}u - \frac{e^{\beta 0} - 1}{\beta} = u, \]
and
\[ \partial_t \Psi(t, u) = \partial_t \left( e^{\beta t}u - \frac{e^{\beta t} - 1}{\beta} \right) \]
\[ = \beta e^{\beta t}u - e^{\beta t} \]
\[ = \beta \left( e^{\beta t}u - \frac{e^{\beta t} - 1}{\beta} \right) - 1 \]
\[ = \beta \Psi(t, u) - 1. \]
Thus, for this $\Psi(t, u)$, the solution $\Phi(t, u)$ equals

$$
\Phi(t, u) = \frac{1}{2} \sigma^2 \int_0^t \Psi^2(s, u) ds + b \int_0^t \Psi(s, u) ds
$$

$$
= \frac{1}{2} \sigma^2 \int_0^t \left( e^{\beta s} u - \frac{e^{\beta s} - 1}{\beta} \right)^2 ds + b \int_0^t \left( e^{\beta s} u - \frac{e^{\beta s} - 1}{\beta} \right) ds,
$$

$$
= \frac{1}{2} \sigma^2 \int_0^t \left( e^{2\beta s} u^2 - 2 e^{\beta s} u \frac{e^{\beta s} - 1}{\beta} + \frac{e^{2\beta s} - 2 e^{\beta s} + 1}{\beta^2} \right) ds + b \int_0^t \left( e^{\beta s} u - \frac{e^{\beta s} - 1}{\beta} \right) ds,
$$

$$
= \frac{1}{2} \sigma^2 \left( \frac{u^2}{2\beta} \int_0^t e^{2\beta s} ds - \frac{u}{\beta} \int_0^t 2 \left( e^{2\beta s} - e^{\beta s} \right) ds + \frac{1}{\beta^2} \int_0^t \left( e^{2\beta t} - 2 e^{\beta t} + 1 \right) ds \right)
$$

$$
+ b \left( \frac{u}{\beta} \int_0^t e^{\beta s} ds + \frac{1}{\beta} \int_0^t e^{\beta s} ds \right)
$$

$$
= \frac{1}{2} \sigma^2 \left( \frac{u^2}{2\beta} \left( e^{2\beta t} - 1 \right) - \frac{u}{\beta^2} \left( e^{2\beta t} - 1 \right) - \left( e^{2\beta t} - 1 \right) \right) + \frac{4}{2\beta^3} \left( e^{2\beta t} - 1 \right) - 4 \left( e^{\beta t} - 1 \right)
$$

$$
+ 2\beta t \right)
$$

$$
= \frac{1}{2} \sigma^2 \left( \frac{u^2}{2\beta} \left( e^{2\beta t} - 1 \right) + \frac{1}{2\beta^3} \left( e^{2\beta t} - 4 \beta e^{\beta t} + 2 \beta t + 3 \right) - \frac{u}{\beta^2} \left( e^{2\beta t} - 2 e^{\beta t} + 1 \right) \right)
$$

$$
+ b \left( \frac{e^{\beta t} - 1}{\beta} u - \frac{e^{\beta t} - 1 - \beta t}{\beta^2} \right)
$$

$$
= \frac{1}{2} \sigma^2 \left( \frac{u^2}{2\beta} \left( e^{2\beta t} - 1 \right) + \frac{1}{2\beta^3} \left( e^{2\beta t} - 4 \beta e^{\beta t} + 2 \beta t + 3 \right) - \frac{u}{\beta^2} \left( e^{2\beta t} - 2 e^{\beta t} + 1 \right) \right)
$$

$$
+ b \left( \frac{e^{\beta t} - 1}{\beta} u - \frac{e^{\beta t} - 1 - \beta t}{\beta^2} \right).
$$

### B.1.2. Explicit expression for the forward rate

Consider the Hull-White extended Vasiček short-rate dynamics under the EMM $Q \sim \mathbb{P}$

$$
dr_t = (b(t) + \beta r_t) dt + \sigma dW_t^*,
$$

where $W^*$ is a standard real-valued $Q$-Brownian motion, $\beta$ and $\sigma > 0$ are constants, and $b(t)$ is a deterministic continuous function. The corresponding HJM forward rate dynamics are of the form

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

hence

$$
df(t, T) = \alpha(t, T) dt + \sigma(t, T) dW^*(t).
$$

Consider the forward rate $f(t, T)$, then

$$
f(t, T) = \partial_T A(t, T) + \partial_T B(t, T) r_t.
$$
Where, for the Hull-White model

\[ A(t, T) = -\frac{\sigma^2}{2} \int_t^T B^2(s, T) ds + \int_t^T b(s)B(s, T) ds \quad \text{and} \quad B(t, T) = \frac{1}{\beta} (e^{\beta(T-t)} - 1). \]

Thus,

\[ f(t, T) = \partial_T A(t, T) + \partial_T B(t, T) \]

\[ = \partial_T \left\{ -\frac{\sigma^2}{2} \int_t^T B^2(s, T) ds + \int_t^T b(s)B(s, T) ds \right\} + \partial_T B(t, T) r_t \]

\[ = \frac{\sigma^2}{2} \int_t^T \partial_s B^2(s, T) ds + \int_t^T b(s)\partial_s B(s, T) ds + \partial_T B(t, T) r_t \]

\[ = \frac{\sigma^2}{2} \left[ \frac{1}{\beta^2} (e^{\beta(T-s)} - 1)^2 \right]_t^T + \int_t^T b(s)e^{\beta(T-s)} ds + e^{\beta(T-t)} r_t \]

\[ = -\frac{\sigma^2}{2\beta^2} (e^{\beta(T-t)} - 1)^2 + \int_t^T b(s)e^{\beta(T-s)} ds + e^{\beta(T-t)} r_t, \]

where Fubini is used and also that \( \partial_T B(s, T) = -\partial_s B(s, T). \) So, in particular

\[ f(0, T) = -\frac{\sigma^2}{2\beta^2} (e^{\beta T} - 1)^2 + \int_0^T b(s)e^{\beta(T-s)} ds + e^{\beta T} r(0). \]

In order to write the forward rate in the form with the integrals as described above, the short rate \( r_t \) must be specified. The short rate dynamics are

\[ dr_t = (b(t) + \beta r_t) dt + \sigma dW_t^* . \]

Consider the process \( Y_t = e^{-\beta t} r_t, \) then applying Itô’s lemma gives

\[ dY_t = e^{-\beta t} dr_t - \beta e^{-\beta t} r_t dt \]

\[ = e^{-\beta t} ((b(t) + \beta r_t) dt + \sigma dW_t^*) - \beta e^{-\beta t} r_t dt \]

\[ = b(t)e^{-\beta t} dt + e^{-\beta t} \sigma dW_t^*, \]

thus

\[ r_t = e^{\beta t} Y_t = e^{\beta t} r(0) + \int_0^t b(s)e^{\beta(t-s)} ds + \int_0^t e^{\beta(t-s)} \sigma dW_s^*. \]
Substituting this expression for $r_t$ into the forward rate $f(t, T)$ above gives

$$f(t, T) = -\frac{\sigma^2}{2\beta^2}(e^{\beta(T-t)} - 1)^2 + \int_1^T b(s)e^{\beta(T-s)}ds + e^{\beta(T-t)}r_t$$

$$= -\frac{\sigma^2}{2\beta^2}(e^{\beta(T-t)} - 1)^2 + \int_1^T b(s)e^{\beta(T-s)}ds$$

$$+ e^{\beta(T-t)} \left( e^{\beta T}(0) + \int_0^t b(s)e^{\beta(t-s)}ds + \int_0^t e^{\beta(t-s)}\sigma dW_s^* \right)$$

$$= -\frac{\sigma^2}{2\beta^2}(e^{\beta(T-t)} - 1)^2 + \int_1^T b(s)e^{\beta(T-s)}ds$$

$$+ e^{\beta T}(0) + \int_0^t b(s)e^{\beta(T-s)}ds + \int_0^t e^{\beta(T-s)}\sigma dW_s^*$$

$$= -\frac{\sigma^2}{2\beta^2}(e^{\beta(T-t)} - 1)^2 + f(0, T) + \frac{\sigma^2}{2\beta^2}(e^{\beta T} - 1)^2 + \int_0^t e^{\beta(T-s)}\sigma dW_s^*$$

$$= f(0, T) + \int_0^t \frac{\sigma^2}{\beta} \left( e^{\beta(T-s)}(e^{\beta(T-s)} - 1) \right) ds + \int_0^t e^{\beta(T-s)}\sigma dW_s^*$$

$$= f(0, T) + \int_0^t \frac{\sigma^2}{\beta} \left( e^{2\beta(T-s)} - e^{\beta(T-s)} \right) ds + \int_0^t e^{\beta(T-s)}\sigma dW_s^*.$$  

Hence

$$\alpha(s, T) = \frac{\sigma^2}{\beta} \left( e^{2\beta(T-s)} - e^{\beta(T-s)} \right) \quad \text{and} \quad \sigma(s, T) = e^{\beta(T-s)}\sigma.$$  

It is checked that this satisfies the HJM drift condition

$$\sigma(s, T) \int_s^T \sigma(s, u)du = e^{\beta(T-s)}\sigma \int_s^u e^{\beta(u-s)}\sigma du = e^{\beta(T-s)}\sigma^2 \left[ \frac{1}{\beta}e^{\beta(u-s)} \right]_s^T = \frac{\sigma^2}{\beta} \left( e^{2\beta(T-s)} - e^{\beta(T-s)} \right).$$  

This equals the $ds$-part of the integration, thus it indeed satisfies the HJM drift condition.
B.2. The CIR short-rate model

The CIR model is given by

\[ dr_t = (b + \beta r_t)dt + \sigma \sqrt{r_t}dW_t, \]

with Riccati equations

\[ \Phi(t, u) = b \int_0^t \Psi(s, u)ds, \]
\[ \partial_t \Psi(t, u) = \frac{1}{2} \sigma^2 \Psi^2(t, u) + \beta \Psi(t, u) - 1, \]
\[ \Psi(0, u) = u. \]

B.2.1. Unique global solution of the CIR short-rate model

\[ \Phi(t, u) = \frac{2b}{\sigma^2} \log \left( \frac{2 \theta e^{(\theta - \beta) t}}{L_3(t) - L_4(t)u} \right), \]
\[ \Psi(t, u) = -\frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u}, \]

where \( \theta = \sqrt{\beta^2 + 2\sigma^2} \) and

\[
L_1(t) = 2(e^{\theta t} - 1), \\
L_2(t) = \theta(e^{\theta t} + 1) + \beta(e^{\theta t} - 1), \\
L_3(t) = \theta(e^{\theta t} + 1) - \beta(e^{\theta t} - 1), \\
L_4(t) = \sigma(e^{\theta t} - 1).
\]

First, note that for \( \Psi(t, u) = -\frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u} \)

\[
\Psi(0, u) = -\frac{L_1(0) - L_2(0)u}{L_3(0) - L_4(0)u} = -\frac{2\theta u}{2\theta} = u,
\]

and

\[
\partial_t \Psi(t, u) = \partial_t \left( -\frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u} \right) \\
= -\frac{\partial_t(L_1(t) - L_2(t)u)(L_3(t) - L_4(t)u) - (L_1(t) - L_2(t)u)\partial_t(L_3(t) - L_4(t)u)}{(L_3(t) - L_4(t)u)^2} \\
= -\frac{(2\theta e^{\theta t} - \theta^2 e^{2\theta t} + \theta e^{\theta t})u(L_3(t) - L_4(t)u) - (L_1(t) - L_2(t)u)(\theta^2 e^{\theta t} - \beta e^{\theta t} - \sigma e^{\theta t}u)}{(L_3(t) - L_4(t)u)^2} \\
= \ldots \\
= \frac{1}{2} \sigma^2 \left( -\frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u} \right)^2 - \beta \frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u} - 1 \\
= \frac{1}{2} \sigma^2 \Psi^2(t, u) + \beta \Psi(t, u) - 1.
\]
B.2.2. Proof of claim

Claim. $\chi^2_{2T(t,S),r^0}$ is non-centrally $\chi^2$-distributed with degrees of freedom $\delta = 4b/\sigma^2$ and parameter of non-centrality $\zeta = 2C_2(t,S)r(0)$.

Proof. According to (2.16) the $F_t$-conditional characteristic function of $r_T$ under the $S$-forward measure $Q^S$ with $t \leq T \leq S$ is given by

$$
\mathbb{E}_{Q^S}\left[e^{u^\top r_T} \mid F_t\right] = \frac{e^{-A(S-T) + \Psi(T-t,u + B(S-T))} + \Psi(T-t,u + B(S-T))^\top r_t}{P(t,S)},
$$

so substituting $-A(S-T) = \Phi(S-T,0)$ and $-B(S-T) = \Psi(S-T,0)$, where $\Phi$ and $\Psi$ are the solutions for the Riccati system for CIR (2.26), gives

$$
\mathbb{E}_{Q^S}\left[e^{u^\top r_T} \mid F_t\right] = \exp\left\{\Phi(S-T,0) + \Phi(T-t,u + \Psi(S-T,0)) + \Psi(T-t,u + \Psi(S-T,0))^\top r_t\right\}
$$

$$
= \exp\left\{\Phi(S-T,0) + \Phi(T-t,u + \Psi(S-T,0)) - \Phi(S-t,0)ight\}
$$

$$
+ \left(\Psi(T-t,u + \Psi(S-T,0))^\top - \Psi(S-t,0)^\top\right) r_t
$$

$$
= \exp\{1\} \exp\{2r_t\}.
$$

The $\exp\{1\}$ and $\exp\{2r_t\}$ are determined separately. First, $\exp\{1\}$:

$$
\exp\{1\} = \exp\left\{\Phi(S-T,0) + \Phi(T-t,u + \Psi(S-T,0)) - \Phi(S-t,0)\right\}
$$

$$
= \exp\left\{\frac{2b}{\sigma^2} \log\left(\frac{2\theta e^{(\frac{\theta}{\sigma^2})(S-T)}}{L_3(S-T)}\right) + \frac{2b}{\sigma^2} \log\left(\frac{2\theta e^{(\frac{\theta}{\sigma^2})(T-t)}}{L_3(T-t) - L_4(T-t) \cdot (u + \Psi(S-T,0))}\right)\right\}
$$

$$
= \exp\left\{\frac{2b}{\sigma^2} \log\left(\frac{(2\theta)^2 e^{(\frac{\theta}{\sigma^2})(S-t)}}{L_3(S-t) \cdot (L_3(T-t) - L_4(T-t) \cdot (u + \Psi(S-T,0)))}\right)\right\}
$$

$$
= \exp\left\{\frac{2b}{\sigma^2} \log\left(\frac{2\theta L_3(S-t)}{L_3(S-T) \cdot (L_3(T-t) - L_4(T-t) \cdot (u - \frac{L_1(S-T)}{L_3(S-T)}))}\right)\right\}
$$

$$
= \left(\frac{2\theta L_3(S-t)}{L_3(S-T) \cdot (L_3(T-t) - L_4(T-t) \cdot (u - \frac{L_1(S-T)}{L_3(S-T)}))}\right)^{2\delta/\sigma^2}
$$

$$
= \left(\frac{L_3(S-T) \cdot (L_3(T-t) - L_4(T-t) \cdot (u - \frac{L_1(S-T)}{L_3(S-T)}))}{2\theta L_3(S-t)}\right)^{-2\delta/\sigma^2}
$$

$$
= \left(\frac{L_3(S-T) L_4(T-t)}{2\theta L_3(S-t)} - \frac{L_3(S-T) L_4(T-t) \cdot (u - \frac{L_1(S-T)}{L_3(S-T)})}{2\theta L_3(S-t)}\right)^{-2\delta/\sigma^2}
$$

$$
= \left(\frac{L_3(S-T) L_4(T-t) - L_3(S-T) L_4(T-t) u + L_4(T-t) L_1(S-T)}{2\theta L_3(S-t)}\right)^{-2\delta/\sigma^2}
$$

80
Thus hence

\[
L_3(S-T)L_3(T-t) + L_4(T-t)L_1(S-T) - C_1(t, T, S)u
\]

\[
\frac{2\theta L_3(S-t)}{2\theta L_3(S-t)}
\]

\[
\left( 1 - C_1(t, T, S)u \right)^{-2
\}
\]

where \((**\) holds, since the enumerator of the first part satisfies

\[
L_3(S-T)L_3(T-t) + L_4(T-t)L_1(S-T)
\]

\[
= \left( \theta(e^{\theta(S-T)} + 1) - \beta(e^{\theta(S-T)} - 1) \right) \left( \theta(e^{\theta(T-t)} + 1) - \beta(e^{\theta(T-t)} - 1) \right)
\]

\[
+ 2\sigma^2(e^{\theta(T-t)} - 1)e^{\theta t} - 1
\]

\[
= \theta^2(e^{\theta(S-T)} + 1)(e^{\theta(T-t)} + 1) - \theta \beta(e^{\theta(S-T)} + 1)(e^{\theta(T-t)} - 1)
\]

\[
- \theta \beta(e^{\theta(S-T)} - 1)(e^{\theta(T-t)} + 1) + \beta^2(e^{\theta(S-T)} - 1)(e^{\theta(T-t)} - 1)
\]

\[
+ 2\sigma^2(e^{\theta(T-t)} - 1)(e^{\theta t} - 1)
\]

\[
= \theta^2(e^{\theta(S-T)} + 1)(e^{\theta(T-t)} + 1) - \theta \beta(e^{\theta(S-T)} + 1)(e^{\theta(T-t)} - 1)
\]

\[
- \theta \beta(e^{\theta(S-T)} - 1)(e^{\theta(T-t)} + 1) + \left( \beta^2 + 2\sigma^2 \right)(e^{\theta(T-t)} - 1)(e^{\theta t} - 1)
\]

\[
= \theta^2 \left( e^{\theta(S-T)} + 1 \right)(e^{\theta(T-t)} + 1) - \theta \beta(e^{\theta(S-T)} + 1)(e^{\theta(T-t)} - 1)
\]

\[
= \theta^2 \left( e^{\theta(S-T)} + e^{\theta(S-T)} + e^{\theta(T-t)} + 1 \right) + \left( e^{\theta(S-T)} - e^{\theta(S-T)} - e^{\theta(T-t)} + 1 \right)
\]

\[
- \theta \beta(e^{\theta(S-T)} - e^{\theta(S-T)} - e^{\theta(T-t)} + e^{\theta(S-T)} - 1)
\]

\[
= \theta^2 \left( 2e^{\theta(S-T)} + 2 \right) - \theta \beta \left( 2e^{\theta(S-t)} - 2 \right)
\]

\[
= 2\theta \left( e^{\theta(S-t)} + 1 - \beta(e^{\theta(S-t)} - 1) \right)
\]

\[
= 2\theta L_3(S - t).
\]

Thus

\[
\frac{L_3(S - T)L_3(T - t) + L_4(T - t)L_1(S - T)}{2\theta L_3(S - t)} = \frac{2\theta L_3(S - t)}{2\theta L_3(S - t)} = 1,
\]

hence

\[
\exp\{1\} = \left( 1 - C_1(t, T, S)u \right)^{-2
\}
\]

Next, consider (2) in \(\exp\{2\}\)

\[
(2) = \Psi(T - t, u + \Psi(S - T, 0))^\top - \Psi(S - t, 0)^\top
\]

\[
= \frac{L_1(T - t) - L_2(T - t) \cdot \left( u + \Psi(S - T, 0) \right)}{L_3(T - t) - L_4(T - t) \cdot \left( u + \Psi(S - T, 0) \right)} + \frac{L_1(S - t)}{L_3(S - t)}
\]

\[
= \frac{L_1(T - t) - L_2(T - t)u + L_2(T - t)\frac{L_3(S - T)}{L_3(S - T)}}{L_3(T - t) - L_4(T - t)(u - \frac{L_3(S - T)}{L_3(S - T)})} + \frac{L_1(S - t)}{L_3(S - t)}
\]

81
Consider the inner part in this equation (i.e. the part between the big braces)

\[
\frac{L_2(T-t)u - L_1(T-t) - L_2(T-t)L_3(S-T)}{L_3(T-t) - L_2(T-t)(u - \frac{L_1(S-T)}{L_3(S-T)\theta})} + \frac{L_1(S-t)}{L_3(S-t)}
\]

\[
= \frac{L_3(S-T)(L_2(T-t)u - L_1(T-t)) - L_2(T-t)L_1(S-T)}{L_3(S-T)(L_3(T-t) - L_4(T-t)(u - \frac{L_1(S-T)}{L_3(S-T)\theta}))} + \frac{L_1(S-t)}{L_3(S-t)}
\]

\[
= \frac{2\theta L_3(S-t) - L_3(S-T)(L_2(T-t)u - L_1(T-t)) - L_2(T-t)L_1(S-T)}{2\theta L_3(S-t)(L_3(T-t) - L_4(T-t)(u - \frac{L_1(S-T)}{L_3(S-T)\theta})))} + \frac{L_1(S-t)}{L_3(S-t)}
\]

From step (*) to the final expression in the determination of (1), it is known that

\[
\frac{2\theta L_3(S-t)}{L_3(S-T) \cdot (L_3(T-t) - L_4(T-t) \cdot (u - \frac{L_1(S-T)}{L_3(S-T)\theta}))} = \left(1 - C_1(t, T, S)u\right)^{-1}.
\]

Thus

\[
(2) = \left(1 - C_1(t, T, S)u\right)^{-1} \left(\frac{L_3(S-T)(L_2(T-t)u - L_1(T-t)) - L_2(T-t)L_1(S-T)}{2\theta L_3(S-t)}\right) + \frac{L_1(S-t)}{L_3(S-t)}
\]

\[
= \left(1 - C_1(t, T, S)u\right)^{-1} \left(\frac{L_3(S-T)(L_2(T-t)u - L_1(T-t)) - L_2(T-t)L_1(S-T)}{2\theta L_3(S-t)}\right) + \frac{L_1(S-t)}{L_3(S-t)} \left(1 - C_1(t, T, S)u\right).
\]

Consider the inner part in this equation (i.e. the part between the big braces)

\[
\frac{L_3(S-T)(L_2(T-t)u - L_1(T-t)) - L_2(T-t)L_1(S-T)}{2\theta L_3(S-t)} = \frac{L_1(S-t)}{L_3(S-t)} \left(1 - C_1(t, T, S)u\right)
\]

\[
= L_2(T-t) \frac{L_3(S-T) - L_2(T-t)L_3(S-T)u}{L_4(T-t) \frac{2\theta L_3(S-t)}{L_3(S-t)}} - L_3(S-T)L_1(T-t) + L_2(T-t)L_1(S-T) - L_1(S-t)2\theta
\]

\[
= \frac{L_2(T-t) L_1(S-t) C_1(t, T, S)u}{L_4(T-t) \frac{L_3(S-t)}{L_3(S-t)}} - \frac{L_3(S-T)L_1(T-t) + L_2(T-t)L_1(S-T) - L_1(S-t)2\theta}{2\theta L_3(S-t)}
\]

\[
= C_2(t, T, S)C_1(t, T, S)u - \frac{L_3(S-T)L_1(T-t) + L_2(T-t)L_1(S-T) - L_1(S-t)2\theta}{2\theta L_3(S-t)}.
\]

And the enumerator of the right term equals

\[
L_3(S-T)L_1(T-t) + L_2(T-t)L_1(S-T) - L_1(S-t)2\theta
\]

\[
= \left(\theta(e^{\theta(S-T)} + 1) - \theta(e^{\theta(S-T)} - 1)\right) \cdot 2\left(e^{\theta(T-t)} - 1\right)
\]

\[
+ \left(\theta(e^{\theta(T-t)} + 1) + \theta(e^{\theta(T-t)} - 1)\right) \cdot 2\left(e^{\theta(S-T)} - 1\right)
\]

\[
- 2\left(e^{(S-t)\theta} - 1\right) \cdot 2\theta
\]
Thus, hence:

\[
(2) = \left(1 - C_1(t, T, S)u\right)^{-1} \left[\frac{L_3(S - T)\left(L_2(T - t)u - L_1(T - t)\right) - L_2(T - t)L_1(S - T)}{2θL_3(S - t)}\right] + \frac{L_1(S - t)}{L_3(S - t)}(1 - C_1(t, T, S)u) \\
= \left(1 - C_1(t, T, S)u\right)^{-1} \left[\frac{C_2(t, T, S)C_1(t, T, S)u - \frac{L_3(S - T)L_1(T - t) + L_2(T - t)L_1(S - T) - L_1(S - t)2θ}{2θL_3(S - t)}}{1 - C_1(t, T, S)u}\right] \\
= \frac{C_2(t, T, S)C_1(t, T, S)u}{1 - C_1(t, T, S)u}.
\]

Hence:

\[
E_Q^S\left[e^{\left(u^Trt\right)|F_t}\right] = \exp\{(1)\}\exp\{(2)rt\} \\
= \frac{1}{(1 - C_1(t, T, S)u)^{2θ/\sigma^2}} \cdot e^{\frac{C_2(t, T, S)C_1(t, T, S)u}{1 - C_1(t, T, S)u}rt},
\]

Take \(\bar{u} = \frac{C_1(t, T, S)u}{2}\) and let \(δ = 4b/σ^2\) and \(ζ = C2(t, T, S)rt \cdot 2\), then:

\[
(*) = \left(\frac{e^{\frac{C_2(t, T, S)rt2δ}{1 - 2\bar{u}^2/σ^2}}}{(1 - 2\bar{u})^{δ/2}}\right)
\]

In order to write down this expression in an integral form, such that the distribution, thus the expectation and variation are known, the following lemma [18, Lemma 10.4] is posed about the non-central \(χ^2\)-distribution.

**Lemma B.2.1.** The non-central \(χ^2\)-distribution with \(δ > 0\) degrees of freedom and non-centrality parameter \(ζ > 0\) has density function

\[
f_{χ^2(δ, ζ)}(x) = \frac{1}{2}e^{-\frac{x+ζ}{2}}\left(\frac{x}{ζ}\right)^{\frac{δ}{2} - \frac{1}{2}}I_{\frac{δ}{2} - 1}(\sqrt{ζ}x), \quad x ≥ 0
\]

and characteristic function

\[
\int_{R_+} e^{ux} f_{χ^2(δ, ζ)}(x) dx = \frac{e^{\frac{ux}{2}}}{(1 - 2u)^{δ/2}}, \quad u \in \mathbb{C}_-.
\]

Here \(I_ν(x) = \sum_{j≥0} \frac{1}{j!(j+ν+1)} \left(\frac{x}{2}\right)^{2j+ν}\) denotes the modified Bessel function of the first kind of order \(ν > -1\).
Thus, where $A$.

B.2.3. Explicit expression of the forward rate

Consider the expression of the forward rate

$$f(t, T) = \partial_T A(t, T) + \partial_T B(t, T)r_t,$$

where $A(t, T)$ can be rewritten as $A(t, T) = A(T, T) - \int_t^T \partial_s A(s, T) = \int_b^T bB(s, T)ds$, using (1.7), and

$$\partial_T B(s, T) = \frac{2(\varepsilon^{(T-s)} - 1)}{(\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma}$$

$$= \frac{(\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma}{((\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma)^2}$$

$$= \frac{2\gamma e^{(T-s)} - 2\gamma - 2\beta e^{(T-s)} + 2\beta + 4\gamma}{((\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma)^2}$$

$$= \frac{\gamma e^{(T-s)}}{((\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma)^2}.$$

Thus,

$$f(t, T) = \partial_T \left\{ \int_t^T bB(s, T)ds \right\} + \partial_T B(t, T)r_t$$

$$= \int_t^T b\partial_T B(s, T)ds + \partial_T B(t, T)r_t$$

$$= \int_t^T b\partial_T B(s, T)ds + \partial_T B(t, T)r_t$$

$$= \int_t^T \frac{\gamma e^{(T-s)}(2(\gamma - \beta)e^{(T-s)} + \gamma + 3\beta)}{((\gamma - \beta)(\varepsilon^{(T-s)} - 1) + 2\gamma)^2}ds$$

$$+ \frac{\gamma e^{(T-t)}(2(\gamma - \beta)e^{(T-t)} + \gamma + 3\beta)}{((\gamma - \beta)(\varepsilon^{(T-t)} - 1) + 2\gamma)^2}r_t.$$

The short rate $r_t$ for the CIR model satisfies

$$dr_t = (b + \beta r_t)dt + \sigma \sqrt{r_t}dW_t^*,$$

hence according to Itô’s lemma $Y_t = e^{-\beta t}r_t$ satisfies

$$dY_t = e^{-\beta t}dr_t - \beta e^{-\beta t}r_t dt$$

$$= e^{-\beta t}((b + \beta r_t)dt + \sigma \sqrt{r_t}dW_t^*) - \beta e^{-\beta t}r_t dt$$

$$= be^{-\beta t}dt + e^{-\beta t}\sigma \sqrt{r_t}dW_t^*.$$

Thus

$$r_t = e^{\beta t}Y_t = e^{\beta t} \left( r(0) + \int_0^t be^{-\beta s}ds + \int_0^t e^{-\beta s}\sigma \sqrt{r_t}dW_s^* \right)$$

$$= e^{\beta t}r(0) + \int_0^t be^{\beta(t-s)}ds + \int_0^t e^{\beta(t-s)}\sigma \sqrt{r_t}dW_s^*.$$
Substituting the short rate into the forward rate gives

\[
f(t, T) = \int_t^T b \gamma e^{\gamma(t-s)} \left(2(\gamma - \beta)e^{\gamma(t-s)} + \gamma + 3\beta\right) ds + \frac{\gamma e^{\gamma(T-t)} (2(\gamma - \beta)e^{\gamma(T-t)} + \gamma + 3\beta)}{((\gamma - \beta)(e^{\gamma(T-t)} - 1) + 2\gamma)^2} r_t
\]

\[
= \int_t^T b \gamma e^{\gamma(t-s)} \left(2(\gamma - \beta)e^{\gamma(t-s)} + \gamma + 3\beta\right) ds
\]

\[
+ \frac{\gamma e^{\gamma(T-t)} (2(\gamma - \beta)e^{\gamma(T-t)} + \gamma + 3\beta)}{((\gamma - \beta)(e^{\gamma(T-t)} - 1) + 2\gamma)^2} \left(e^{\beta t} r(0) + \int_0^t b e^{\beta(s-t)} ds + \int_0^t e^{\beta(t-s)} \sigma \sqrt{r} dW_s^*\right).
\]

In order to specify the \(\sigma(t, T)\), the forward rate has to be rewritten in the form \(df(t, T) = \alpha(t, T)dt + \sigma(t, T)dW_t^*\), where only the \(dW_t^*\)-part is needed. Since in the first terms of \(f(t, T)\) there is no \(dW_t^*\), only the last term is considered. It can be concluded that

\[
\sigma(t, T) = \partial_T B(t, T)e^{\beta(t-s)} \sigma \sqrt{r(t)}.
\]
Index

Affine
canonical state space, 18
admissible parameters, 18
process, 16
term-structure (ATS), 10

Call option
analytical price, 24
CIR, 30
quadratic, 45
Vasiček, 27

Cap
analytical price, 25
CIR, 32
Vasiček, 29

Exponential-quadratic payoff, 43

Feynman-Kač, 9
Forward measure, 74
Forward rate
instantaneous, 73

Fourier
transform technique, 22, 44
example, 23

Interest rate, 73
Forward rate
continuously compounded, 73
Spot rate
simple, 73

Money-market account, 9
Monte Carlo method, 57

Put option
analytical price, 24
CIR, 30
quadratic, 45
Vasiček, 29

Quadratic
admissible parameters, 34
canonical state space, 35
class, 35
process, 34

Riccati equations
affine, 17, 19
adjusted, 20
quadratic, 34
adjusted, 37
one-dimensional, 42

Short rate, 73
Short-rate model
affine, 19
CIR, 13, 30
Hull-White, 14
Quadratic, 46
Vasiček, 11, 26