Bounded short-rate models with Ehrenfest and Jacobi processes

DISSERTATION

zur Erlangung des Grades
eines Doktors der Naturwissenschaften

der Fakultät für Mathematik
der Technischen Universität Dortmund
vorgelegt von

Dipl.-Math. Alexander Kaplun

Dortmund, Juli 2010

Betreuer: Prof. Dr. M. Voit

To my parents
Natalia and Iakov

Acknowledgments

First, I would like to thank my thesis adviser Prof. Dr. Michael Voit for his invaluable support and for suggesting this interesting topic to me. He gave me guidance throughout the development of this thesis, as well as during my graduate and undergraduate studies, starting with the very first course in calculus.

I am very thankful to Prof. Dr. Jeanette Woerner for the advice that she gave to me, both of a mathematical and of a professional nature.

I wholeheartedly thank Silvetta for her love and inspiration. I am deeply indebted and grateful to my parents Natalia and Iakov for giving me unconditional support throughout my life and encouragement to pursue my goals. Родители, я от всего сердца благодарю вас за ваше терпение и поддержку.

Contents

Introduction

Over the last 40 years the field of financial mathematics dedicated to the theory of stochastic interest rates has been constantly growing. One of the fundamental approaches to term structure modelling is based on the specification of the short-term interest rate – the short-rate. A typical way of modelling the short-rate is to describe the underlying short-rate process as a diffusion process, or, more generally, in terms of the solution of a stochastic differential equation. Vasicek [62] first adopted the principles of arbitrage-free valuation of contingent claims from the seminal work of Merton and Black and Scholes. In his pioneering work Vasicek derived a closed-form representation for the zero-coupon bond (ZCB) price under the assumption of a mean-reverting short-rate model with Gaussian distribution. Since then a variety of short-rate models have become established. Some of the most prominent models are the Black-Karasinski model [5], the Cox-Ingersoll-Ross model [11] and the Hull-White model [25]. Each of these models has its advantages and disadvantages. A less common approach to modelling the short-rate is based on assuming a Markov chain model in discrete or continuous time: see e.g. [10], [16] or [46]. In this work, we will consider two short-rate models, one for each approach, i.e. a diffusion model and a continuous-time Markov process model.

Albeit the earliest, the Vasicek's model and its generalizations are very popular among practitioners, which can be ascribed to its analytical tractability in regard to ZCB prices and the European options thereof. Unfortunately, there are some shortcomings, the most prominent of which is the possibility of interest rates becoming negative – a fact concerning all models with Gaussian distribution. Even though the probability of negative rates is rather small, not only does the realism of the model come into question, but problems may also appear while valuing ZCBs with a long time to maturity and a low interest rate level.

This work examines two mean-reverting models for the short-rate whose characteristic feature is the possibility of choosing arbitrary lower and upper bounds for the interest rate, thus preventing negative interest rates. The first model is a finite-state model based on the continuous time Ehrenfest process. The idea of using both the discrete and the continuous time versions of the Ehrenfest process in finance is well known. The discrete time approach was used, e.g. by Okunev and Tippett [48] in modelling accumulated cashflows, by Takahashi [61] in exploring changes in stock prices and exchange rates for currencies, and by Buehlmann [10] in modelling interest rates. Sumita, Gotoh and Jin [58] studied the passage times and the historical maximum of the Ornstein-Uhlenbeck process via an approximation by a special case of the continuous time Ehrenfest process. With regard to the modelling of interest rates, it seems that the discrete time approach leads in general only to a recursively computable term structure.

The second short-rate model this work looks at is a linearly transformed Jacobi diffusion. The Jacobi diffusion and related processes have well-known applications in finance. Larsen and Sørensen [40] proposed an analytically tractable model for an exchange rate in a target zone based on the Jacobi diffusion, and provided estimators for the model parameters. Delbaen and Shirakawa [13] studied an interest rate model with lower and upper bounds based on the Jacobi diffusion. In [63] Veraart A. and Veraart L. introduced a stochastic volatility model, where the correlation parameter between the stock and the volatility is modelled by a linearly transformed Jacobi diffusion.

In the first part of this thesis, we propose a finite-state mean-reverting model for the short-rate related to the continuous time Ehrenfest process. By choosing arbitrary lower and upper boundaries for the interest rate, we can treat the respective short-rate process as a suitably linearly-transformed birth-and-death process on $\{0, 1, \ldots, N\}$, $N \in \mathbb{N}$. By choosing the lower boundary as non-negative, the problem of negative interest rates can be avoided. Furthermore, the model allows for the explicit evaluation of ZCB prices. In this way, the model aims at realism and analytical tractability.

The main outcome of this work is the derivation of pricing formulae for ZCBs in the general and the special symmetric cases of the model. In both cases the arbitrage-free ZCB price at time t and maturity T is given as follows:

$$P(t,T) = C \cdot P_1(t,T)^k \cdot P_0(t,T)^{N-k},$$

where C is a constant, $k \in \{0, 1, \ldots N\}$, and P_1 and P_0 can be expressed in terms of ${}_1F_1$ hypergeometric functions of the matrix argument given in Section 1 (see also [21]). In the general case the model is governed by five parameters – a valuable fact considering the fitting of the model to the market data. The special case provides four parameters and is characterised by the symmetry of the underlying distribution with respect to the mean-reverting value. The advantage here is that we have more tractable expressions of P_1

and P_0 from the computational point of view. Moreover, after a suitable transformation, the model yields the Vasicek model in the limit as N tends to infinity.

The second short-rate model that we examine, is based on the Jacobi diffusion, and is given as follows:

$$dr_t = k[\theta - r_t]dt + \sigma\sqrt{(r_t - r_m)(r_M - r_t)}dW_t,$$

where $r_0, \theta, k > 0, \sigma > 0$ are the constants denoting the starting state of the process, its mean-reverting value, the speed of mean-reversion and the volatility parameter, respectively. The constants $r_m < \theta < r_M$ denote the lower and upper bounds of the process. This model was first introduced by Delbaen and Shirakawa in [13], where they calculated the transition density of the underlying process and derived a semi-closed expression for the ZCB prices. We will extend their results concerning the pricing of ZCBs and show how under suitable transformations the model converges to the Vasicek model as well as to the Cox-Ingersoll-Ross model.

This dissertation is organized as follows: the following chapter gives a short review of the special functions we shall encounter throughout this work, i.e. the ${}_1F_1$ hypergeometric functions of a matrix argument, the Krawtchouk polynomials and the Jacobi polynomials. Chapter 2 summarises basic ideas of stochastic finance and some standard concepts of short-rate modelling. Chapter 3 is divided into three sections. Section 3.1 deals with the Ehrenfest process in discrete and continuous time. The main results of this work are presented in Section 3.2, where the *Ehrenfest short-rate model* is defined and the ZCB pricing formulae are derived. In Section 3.3 we discuss the advantages of the Ehrenfest short-rate model over the Vasicek model in a case study valuing ZCB bonds. Chapter 4 examines the Jacobi diffusion and its application as the *Jacobi short-rate model.* We present the main properties of this model and derive the respective ZCB pricing formula. The last, fifth chapter recapitulates the results of the work and points out some problems that remain open for further research.

Part of this thesis has been accepted by the Journal of Applied Probability and is due to appear in September 2010, in Vol. 47, Number 3, under the title "The continuous time Ehrenfest process in term structure modelling".

Chapter 1

Special functions and orthogonal polynomials

Throughout this work we will need to be aware of some well-known facts concerning the *Krawtchouk polynomials* (see [35], [60] and [64]), the *Jacobi polynomials* (see [2] and [60]) and the ${}_1F_1$ *hypergeometric functions of a matrix argument* (see [20] and [21]) as well as some of their practical implications. For the sake of clarity we give in this section an overview of these special functions.

1.1 Hypergeometric functions of a matrix argument

Definition 1.1 (Hypergeometric functions of a matrix argument).

(a) A *partition* $m = (m_1, m_2, \ldots, m_n)$ is an n-tuple ($n \in \mathbb{N}$) of non-negative integers such that $m_1 \geq m_2 \geq \cdots \geq m_n$.

(b) For a partition m the *generalized Pochhammer symbol* is defined by

$$[a]_m = \prod_{j=1}^{n} \prod_{i=1}^{m_j} (a - j + i).$$

(c) For a partition m the *normalized Schur function of index* m is defined as a real-analytic function on the space S_n ($n \in \mathbb{N}$) of $n \times n$ Hermitian matrices with eigenvalues $z = (z_1, \ldots, z_n)^T \in \mathbb{R}^n$ as follows:

$$Z_m(z) = |m|! \cdot \frac{\prod_{1 \leq j < k \leq n}(m_j - m_k - j + k)}{\prod_{j=1}^{n}(m_j + n - j)!} \cdot \frac{\det\left(z_i^{m_j+n-j}\right)}{\prod_{1 \leq i < j \leq n}(z_i - z_j)}, \quad (1.1)$$

where $|m| := m_1 + \cdots + m_n$.

(d) The *hypergeometric function* ${}_pF_q$ *of a matrix argument* is given by the series

$$ {}_pF_q(a_1, \ldots, a_p; b_1, \ldots, b_q; z) = \sum_{j=0}^{\infty} \frac{1}{j!} \sum_{|m|=j} \frac{[a_1]_m \ldots [a_p]_m}{[b_1]_m \ldots [b_q]_m} \cdot Z_m(z), \quad (1.2) $$

where, for $1 \leq i \leq p$ and $1 \leq j \leq q$, $a_i, \; b_j \in \mathbb{C}$ and none of the numbers $-b_j + k - 1$ $(k = 1, \ldots, n)$ is a non-negative integer.

Remark 1.2. From definition (1.1) we immediately see that ${}_pF_q(.,.,z)$ is invariant under the permutations of z, where $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ is an eigenvalue vector of some hermitian matrix $H \in S_n$.

The question of convergence of the ${}_pF_q$ function is answered by the following theorem adopted from [20].

Theorem 1.3. *Let $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ be an eigenvalue vector of some hermitian matrix $H \in S_n$.*

(a) *If $p \leq q$, then the hypergeometric series (1.2) converges absolutely for all z.*

(b) *If $p = q + 1$, then the series (1.2) converges absolutely for $\|z\| < 1$ and diverges for $\|z\| > 1$.*

(c) *If $p > q$, then the series (1.2) diverges unless it terminates.*

We will be particularly concerned with the ${}_1F_1$ function, which is also known as the *confluent hypergeometric function of a matrix argument.* With Theorem 1.3 we see that it converges absolutely for all $z \in S_n$.

An important result, which will be crucial later on, is given without proof in the following remark (see [21], p. 25).

Remark 1.4. Let Δ_n denote the *standard simplex* in $\mathbb{R}^n$ $(n \in \mathbb{N})$, defined by

$$ \Delta_n = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0, i = 1, \ldots n, \sum_{i=1}^{n} x_i \leq 1 \right\}. \quad (1.3) $$

Then for $a > 0$ the following equation holds

$$ {}_1F_1(1; a+n; z) = (a)_n \int_{\Delta_n} \left(1 - \sum_{i=1}^{n} x_i \right)^{a-1} \cdot \exp\left(\sum_{i=1}^{n} z_i x_i \right) dx_1 \ldots dx_n. \quad (1.4) $$

In order to compute the ${}_pF_q$ function numerically we truncate the series (1.2) by $|m| \leq H$ as follows:

$$ {}_pF_q^H(a_1,\ldots,a_p;b_1,\ldots,b_q;z) = \sum_{j=0}^{H} \frac{1}{j!} \sum_{|m|=j} \frac{[a_1]_m \cdots [a_p]_m}{[b_1]_m \cdots [b_q]_m} \cdot Z_m(z). \tag{1.5} $$

Koev and Edelman [36] provide an effective algorithm for computing the ${}_pF_q^H$ function. For $z \in S_n$ the complexity of their algorithm is linear in n and subexponential in H, which will in the following chapters turn out to be acceptable, as we will compute the ${}_1F_1^H(1; n+1; .)$ for growing $n \in \mathbb{N}_0$.

1.2 Krawtchouk polynomials

For given $N \in \mathbb{N}$, $l \in \{0,\ldots,N\}$ and $0 < p = 1-q < 1$ the Krawtchouk polynomials $K_l(x) = K_l(x;N;p)$ are the orthogonal polynomials with respect to the binomial distribution $B_{N,p}$ with the probability mass function $\omega(x) = \binom{N}{x} p^x q^{N-x}$ at the points $x = 0, 1, \ldots, N$. They can be defined in two different, but equivalent ways.

Definition 1.5 (Krawtchouk polynomials). For $x, l \in \{0, 1, \ldots, N\}$ we set

$$ K_l(x) = {}_2\mathcal{F}_1(-l,-x;-N;1/p) = \sum_{k=0}^{N} \frac{(-l)_k(-x)_k}{(-N)_k k!} \left(\frac{1}{p}\right)^k, \tag{1.6} $$

or

$$ K_l(x) = \binom{N}{l}^{-1} \sum_{k=0}^{N} (-1)^k \binom{N-x}{l-k} \binom{x}{k} \left(\frac{q}{p}\right)^k, \tag{1.7} $$

where ${}_2\mathcal{F}_1$ is the classical Gauss hypergeometric function (see [60], §4.21).

Definition 1.5 leads to the following well-known basic properties.

Proposition 1.6 (Properties of the Krawtchouk polynomials).

(a) *Symmetry:*

$$ K_l(x) = K_x(l) \tag{1.8} $$

for all $x, l \in \{0, 1, \ldots, N\}$.

(b) $K_0(x) = K_l(0) = 1$ *for all* $x, l \in \{0, 1, \ldots, N\}$.

(c) $K_1(x) = 1 - \frac{x}{Np}$, $K_N(x) = (-1)^x \left(\frac{q}{p}\right)^x$ *for all* $x \in \{0, 1, \ldots, N\}$.

(*d*) *Generating function:*

$$\big(1 - \frac{q}{p} \cdot s\big)^i \cdot \big(1 + s\big)^{N-i} = \sum_{l=0}^{N} \binom{N}{l} K_l(i) s^l \tag{1.9}$$

for all $x \in \{0, 1, \ldots, N\}$ *and* $s \in \mathbb{C}$.

(*e*) *Recurrence relation:*

$$-xK_l(x) = (N-l)pK_{l+1}(x) - [(N-l)p + lq]\, K_l(x) + lqK_{l-1}(x) \tag{1.10}$$

for all $x, l \in \{0, 1, \ldots, N\}$, *where we set* $K_{-1}(x) = 0$ *and* $K_{N+1}(x) = 0$.

(*f*) *Orthogonality relation:*

$$\sum_{x=0}^{N} K_l(x) K_m(x) \omega(x) = \frac{\delta_{l,m}}{\pi_l}, \tag{1.11}$$

where

$$\pi_l = \binom{N}{l} \left(\frac{p}{q}\right)^l = \omega(l) q^{-N} \tag{1.12}$$

for all $l, m \in \{0, 1, \ldots, N\}$.

(*g*) *For* $l, m \in \{0, 1, \ldots, N\}$,

$$\begin{aligned} B_{m,l} &= \sum_{x=0}^{N} x K_l(x) K_m(x) \omega(x) \\ &= \begin{cases} 0 & \text{if } |l - m| \geq 2, \\ -kq/\pi_{k-1} & \text{if } |l - m| = 1, \\ ((N-l)p + lq)/\pi_l & \text{if } m = l, \end{cases} \end{aligned} \tag{1.13}$$

where $k = \max(l, m)$. *Note that* $B_{l,m} = B_{m,l}$.

Proof. $(a) - (e)$ follow from (1.6), (f) follows from (1.7) and (g) is a direct application of (1.10) and (1.11) (see e.g. [60], §2.82). □

1.3 Jacobi polynomials

For given $n \in \mathbb{N}_0$ and $\alpha, \beta > -1$ the *Jacobi polynomials* $P_n^{(\alpha,\beta)}$ are the orthogonal polynomials on $[-1, +1]$ with respect to the weight function $\omega(x) = (1-x)^\alpha (1+x)^\beta$. They can be defined by the *Rodrigues formula* (see [60], §4.3.) as given in the following definition.

Definition 1.7 (Jacobi polynomials). Let $\alpha, \beta > -1$. For $x \in [-1,1]$ and $n \in \mathbb{N}_0$ the Jacobi polynomials are given as follows:

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n! \cdot \omega(x)} \cdot \left(\frac{d}{dx}\right)^n (1-x)^{n+\alpha}(1+x)^{n+\beta}.$$

Definition 1.7 leads to the following well-known properties.

Proposition 1.8 (Properties of the Jacobi polynomials). *Let $\alpha, \beta > -1$, $x \in [-1,1]$ and $n \in \mathbb{N}_0$.*

(*a*) *Series representation:*

$$P_n^{(\alpha,\beta)}(x) = \frac{\Gamma(n+\alpha+1)}{n!\Gamma(n+\alpha+\beta+1)} \cdot \sum_{k=0}^{n} \binom{n}{k} \frac{\Gamma(n+\alpha+\beta+k+1)}{\Gamma(\alpha+k+1)} \left(\frac{x-1}{2}\right)^k, \tag{1.14}$$

where Γ denotes the Gamma function. In particular, we have

$$P_0^{(\alpha,\beta)}(x) = 1 \quad \text{and} \quad P_1^{(\alpha,\beta)}(x) = \alpha + 1 + (\alpha+\beta+2)\frac{x-1}{2}. \tag{1.15}$$

(*b*) *Orthogonality relation:*

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(x) P_m^{(\alpha,\beta)}(x) \omega(x)\, dx = \delta_{n,m} \cdot c_n^{(\alpha,\beta)}, \tag{1.16}$$

where

$$c_n^{(\alpha,\beta)} = \frac{2^{\alpha+\beta+1}}{2n+\alpha+\beta+1} \cdot \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!\Gamma(n+\alpha+\beta+1)} \tag{1.17}$$

for all $m \in \mathbb{N}_0$.

(*c*) *Differential equation:*

$$(1-z^2)\frac{d^2y}{dz^2} + [\beta - \alpha - (\alpha+\beta+2)z]\frac{dy}{dz} + n(n+\alpha+\beta+1)y = 0 \tag{1.18}$$

for $y = P_n^{(\alpha,\beta)}$.

(*d*) *Recurrence relation:*

$$\begin{aligned} 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha,\beta)}(x) = \\ \{(2n+\alpha+\beta+1)(\alpha^2-\beta^2) + (2n+\alpha+\beta)_3 \cdot x\}P_n^{(\alpha,\beta)}(x) \\ -2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha,\beta)}(x) \end{aligned} \tag{1.19}$$

for $n \in \mathbb{N}_0$, where we set $P_{-1}^{(\alpha,\beta)}(x) = 0$ for all $x \in [-1,1]$.

In the following we will need the *modified Jacobi polynomials* given by the following definition.

Definition 1.9 (Modified Jacobi polynomials). Let $\alpha, \beta > -1$. For $n \in \mathbb{N}_0$ and $x \in [0,1]$ the modified Jacobi polynomials are given as follows:

$$J_n^{(\alpha,\beta)}(x) = P_n^{(\alpha,\beta)}(1-2x), \quad x \in [0,1]. \tag{1.20}$$

The corresponding weight function is given by

$$\pi(x) = \omega(1-2x) = x^\alpha(1-x)^\beta, \quad x \in [0,1].$$

This definition leads to the following properties.

Proposition 1.10 (Properties of the modified Jacobi polynomials). *Let $\alpha, \beta > -1,\ x \in [0,1]$ and $n \in \mathbb{N}_0$.*

(a) *Series representation:*

$$J_n^{(\alpha,\beta)}(x) = \frac{\Gamma(n+\alpha+1)}{n!\Gamma(n+\alpha+\beta+1)} \sum_{k=0}^{n} \binom{n}{k} \frac{\Gamma(n+\alpha+\beta+k+1)}{\Gamma(\alpha+k+1)}(-1)^k x^k. \tag{1.21}$$

where Γ denotes the Gamma function. In particular, we have

$$J_0^{(\alpha,\beta)}(x) = 1 \quad \text{and} \quad J_1^{(\alpha,\beta)}(x) = \alpha+1-(\alpha+\beta+2)x. \tag{1.22}$$

(b) *Orthogonality relation:*

$$\int_0^1 x^\alpha(1-x)^\beta J_n^{(\alpha,\beta)}(x) J_m^{(\alpha,\beta)}(x)\, dx = \delta_{n,m} \cdot h_n^{(\alpha,\beta)}, \tag{1.23}$$

where

$$h_n^{(\alpha,\beta)} = \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)} \tag{1.24}$$

for all $n \in \mathbb{N}_0$.

(c) *Differential equation:*

$$x(1-x)\frac{d^2y}{dx^2} + (\alpha+1-(\alpha+\beta+2)x)\frac{dy}{dx} + n(n+\alpha+\beta+1)y = 0 \tag{1.25}$$

for $y = J_n^{(\alpha,\beta)}$.

(*d*) *Recurrence relation:*

$$\begin{aligned} x \cdot J_n^{(\alpha,\beta)}(x) &= \frac{(\alpha^2-\beta^2)+(2n+\alpha+\beta)(2n+\alpha+\beta+2)}{2(2n+\alpha+\beta)(2n+\alpha+\beta+2)} J_n^{(\alpha,\beta)}(x) \\ &\quad - \frac{(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} J_{n+1}^{(\alpha,\beta)}(x) \\ &\quad - \frac{(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} J_{n-1}^{(\alpha,\beta)}(x), \end{aligned} \tag{1.26}$$

for $n \in \mathbb{N}_0$, *where we set* $J_{-1}^{(\alpha,\beta)}(x) = 0$ *for all* $x \in [0,1]$.

(*e*) *For* $m, n \in \mathbb{N}_0$,

$$B_{m,n} = \int_0^1 x \cdot J_n^{(\alpha,\beta)}(x) J_m^{(\alpha,\beta)}(x) \pi(x)\, dx \tag{1.27}$$

$$= \begin{cases} 0 & |n-m| \geq 2, \\ -\frac{\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{(k-1)! \cdot (2k+\alpha+\beta-1)_3 \Gamma(k+\alpha+\beta)} & |n-m| = 1, \\ \frac{\{(\alpha^2-\beta^2)+(2n+\alpha+\beta)(2n+\alpha+\beta+2)\}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{2 \cdot n! \cdot (2n+\alpha+\beta)_3 \Gamma(n+\alpha+\beta+1)} & m = n, \end{cases}$$

where $k = \max(n,m)$. *Note that* $B_{m,n} = B_{n,m}$.

Proof. The proof is given in the Appendix B.

Chapter 2

Introduction to interest rate modelling

In this chapter we present the main definitions and concepts of the continuous-time short-rate modelling that we will need throughout this work. There are a variety of monographs on this topic (see e.g [9], [28], [50], [51] and [65]). Here, we will follow the works of Brigo and Mercurio [9] and Zagst [65].

2.1 Basics of financial modelling

In this section we will give a brief overview of the continuous time approach to the financial modelling. Some excellent books that go into depth on this topic are e.g. [7], [15], [45] and [55]. In the following, we provide the basic theory that comprises the fundamental results of Harrison and Kreps [22] and Harrison and Pliska [23], [24]. We start with the following assumption.

Assumption 2.1. *Let* $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T^*]}, \mathbb{P})$ *be a filtered probability space with filtration* $(\mathcal{F}_t)_{t\in[0,T^*]}$ *satisfying the usual conditions, i.e. it is complete and right-continuous. Let all stochastic processes considered in the following be defined on this probability space and be adopted to the filtration* $(\mathcal{F}_t)_{t\in[0,T^*]}$*, where* T^* *is the time horizon.*

An intuitive interpretation of our probability space is that Ω describes the set of possible environmental conditions that influence the market, while the filtration $(\mathcal{F}_t)_{t\in[0,T^*]}$ determines the market information available at time $t \in [0, T^*]$. Next we need the definition of a *bank account*:

Definition 2.2 (**Bank account**). Let $B(t),\ t \in [0, T^*]$, be the value of a bank account or bank process at time t. We assume $B(0) = 1$ and that the bank account evolves according to the following differential equation:

$$dB(t) = r(t)B(t)dt,$$

where $r(t)$ is a positive measurable function of time. Clearly, we have:

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

Hence, the bank account represents a riskless investment, where profits are accrued continuously at the risk-free rate $r(t)$ prevailing in the market at any time $t \in [0, T^*]$. Thus, we can use the bank account to discount future payments in terms of a *discount factor*. In the case that the interest rate $r(t)$ is not deterministic but a stochastic process $(r_t)_{t\in[0,T^*]}$, we obtain the notion of the **stochastic** discount factor:

Definition 2.3 (**(Stochastic) discount factor**)**.** The (stochastic) discount factor $D(t, T)$ between two instants $t \in [0, T^*]$ and $T \in [t, T^*]$ is the amount of money at time t that is equivalent to one monetary unit payable at time T, and is given by

$$D(t, T) = \frac{B(t)}{B(T)} = \exp\left(-\int_t^T r_s\ ds\right).$$

Clearly, the stochastic discount factor differs from the deterministic one in terms of the evolution of the underlying process $(r_t)_{t\geq 0}$. A definition and details of the modelling of this interest rate – later we will denote it as the *short-rate* – will be given in the next section and in Sections 2.3, 3.2 and 4.2.

Assumption 2.4. *Consider $N+1$ non-dividend-paying securities, which are traded continuously from time 0 until time $T \leq T^*$. Their prices are modelled by $N + 1$ Itô-processes*[1] $(S_t^i)_{t\in[0,T]}$*, $i \in \{0, \dots, N\}$. We assume $(S_t^0)_{t\in[0,T]}$ to be the bank process $(B(t))_{t\in[0,T]}$ as given in Definition 2.2. Furthermore, we assume*

$$S_t^i(\omega) > 0 \quad \textit{for all } i, t \textit{ and } \omega.$$

A central concept of financial modelling is the notion of a *self-financing portfolio strategy.*

Definition 2.5 (**Portfolio strategy**)**.** Let $\varphi = (\varphi_t)_{t\in[0,T]} = (\varphi_t^0, \dots, \varphi_t^k)_{t\in[0,T]}$ be an $(N+1)$-dimensional $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process, whose components are

[1]See Definition 4.1.1 in [47] and Appendix A.4 for a definition and details of Itô-processes.

locally bounded[2]. Then, φ is called a portfolio strategy and

$$V_t(\varphi) = \varphi_t S_t = \sum_{i=0}^{N} \varphi_t^i S_t^i, \quad t \in [0,T],$$

is called the corresponding *value process*. The following Itô-Integral

$$G_t(\varphi) = \int_0^t \varphi_s \, dS_s = \sum_{i=0}^{N} \int_0^t \varphi_s^i \, dS_s^i, \quad t \in [0,T],$$

is called the corresponding *gains process*.

Thus, a portfolio strategy is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted process φ, specifying at each state $\omega \in \Omega$ and time $t \in [0,T]$ the number $\varphi_t(\omega)$ of units of the security to hold. $V_t(\varphi)$ and $G_t(\varphi)$ are respectively interpreted as the market value of the portfolio φ_t and the cumulative gains (losses) realized by the investor up to time $t \in [0,T]$ by adopting the portfolio strategy φ.

Definition 2.6. A portfolio strategy φ is called self-financing if

$$V_t(\varphi) = V_0(\varphi) + G_t(\varphi), \quad t \in [0,T], \tag{2.1}$$

or equally

$$dV_t(\varphi) = dG_t(\varphi) \quad \left(= \sum_{i=0}^{N} \varphi_t^i dS_t^i \right), \quad t \in [0,T].$$

Intuitively, a strategy is self-financing if its value after an initial investment alters only due to changes in the security prices, i.e. no additional cash inflows or outflows occur after the initial input.

In the following chapters we will focus on pricing interest rate derivatives or, more generally, *contingent claims*.

Definition 2.7 (Contingent claim). A contingent claim is a positive and square-integrable random variable X on $(\Omega, \mathcal{A}, \mathbb{P})$. A contingent claim X is called *attainable* if a self-financing portfolio strategy φ exists, such that $V_T(\varphi) = X,\ T \in [0,T^*]$.

One of the central concepts in financial modelling is that of the absence of arbitrage. The mathematical definition of an *arbitrage portfolio strategy* is given as follows.

[2]If we consider the security prices $(S_t^i)_{t\in[0,T]}$, $i \in \{0,\dots,N\}$, to be Brownian semimartingales rather than Itô-processes that have $\mathbb{P}$-a.e. continuous paths, we also need the portfolio strategy to be predictable (for a definition and details see e.g. [32], p. 131, and [9], p. 25)

Definition 2.8 (**Arbitrage portfolio strategy**). An arbitrage portfolio strategy is a self-financing portfolio strategy φ such that corresponding value process $V_t(\varphi)$ satisfies the following properties:

(a) $P(V_0(\varphi) = 0) = 1$ and

(b) $P(V_T(\varphi) \geq 0) = 1 \quad \text{and} \quad P(V_T(\varphi) > 0) > 0$

for some $T \in (0, T^*]$. If no arbitrage portfolio strategy exist for any $T \in (0, T^*]$, we say that the market is *arbitrage free.*

The concept of no-arbitrage is closely connected with the notion of an *equivalent martingale measure* or *risk-neutral measure*, which goes back to the works of Harrison and Kreps [22] and Harrison and Pliska [23], [24]. Some preliminary definitions and the main results without proofs will be given in the following.

Definition 2.9 (**Equivalent martingale measure**). An equivalent martingale measure $\mathbb{Q}$ is a probability measure on the space $(\Omega, \mathcal{A})$ such that

(a) $\mathbb{P}$ and $\mathbb{Q}$ are equivalent measures, i.e. $\mathbb{P}(A) = 0$ if and only if $\mathbb{Q}(A) = 0$ for every $A \in \mathcal{A}$;

(b) the Radon-Nikodym derivative $d\mathbb{Q}/d\mathbb{P}$ belongs to $L^2(\Omega, \mathcal{A}, \mathbb{P})$;

(c) the discounted security price process $(D(0,t)S_t^i)_{t\in[0,T]}$, $i = 0, \cdots, N$, is a $\mathbb{Q}$-martingale, i.e.

$$\mathbb{E}^{\mathbb{Q}}[D(0,t)S_t^i|\mathcal{F}_s] = D(0,s)S_s^i,$$

for all $i = 0, \cdots, N$ and $0 \leq s \leq t \leq T$, where $D(0, .)$ is the (stochastic) discount factor as given in Definition 2.3.

The following proposition establishes the connection between the existence of a martingale measure and the uniqueness of a no-arbitrage price associated with any attainable contingent claim.

Proposition 2.10 (**Risk-neutral valuation**). *Let X be an attainable contingent claim. If there is an equivalent martingale measure $\mathbb{Q}$ on $(\Omega, \mathcal{A})$, then, for every $t \in [0,T]$, there will be a unique*[3] *no-arbitrage price π_t associated with X, i.e.*

$$\pi_t = \mathbb{E}^{\mathbb{Q}}[D(t,T)X|\mathcal{F}_t], \tag{2.2}$$

[3] In an arbitrage-free market the price of any attainable contingent claim is uniquely given, either by the value of the associated replicating strategy, or by the risk neutral expectation of the discounted claim payoff under any of the equivalent (risk-neutral) measures (see [9], p. 26).

or equivalently with Definition 2.3 and Assumption 2.4,

$$\pi_t = \mathbb{E}^{\mathbb{Q}} \left[\exp \left(- \int_t^T r_s \, ds \right) X \,\middle|\, \mathcal{F}_t \right] = S_t^0 \cdot \mathbb{E}^{\mathbb{Q}} \left[\frac{X}{S_T^0} \,\middle|\, \mathcal{F}_t \right].$$

A key definition in the theory of valuation derivative securities is in the notion of the completeness of the financial market.

Definition 2.11. A financial market is *complete* if, and only if, every contingent claim is attainable.

Harrison and Pliska [24] proved the following fundamental result.

Theorem 2.12. *A financial market is arbitrage-free and complete if, and only if, a unique equivalent martingale measure exists.*

2.2 Bonds and interest rates

In this section we describe some basic interest rate contracts and the concept of the short-rate, which will be the main object of research in this work.

A fundamental and primary asset in the interest rate world is the so-called *zero-coupon bond (ZCB).*

Definition 2.13 (Zero-coupon bond). A T-maturity zero-coupon bond (pure discount bond or T-bond) is a contract that guarantees its holder the payment of one monetary unit at time T, with no intermediate payments. The contract value at time $t \leq T$ is denoted by $P(t, T)$. Clearly, $P(T, T) = 1$ for all T. The time $T - t$ is called *the time to maturity* of a ZCB.

While a ZCB has no payments until the date of maturity T, a *coupon-bearing bond* or *coupon bond* (CB) is characterized by periodic payments - coupons - during the life of the CB.

Definition 2.14 (Coupon-bearing bond). A coupon-bearing bond is a contract that guarantees its holder the payment at future times $T_i, \ i = 1, \ldots, n$ with $t \leq T_1 < T_2 < \cdots < T_n = T$ of deterministic amounts of monetary units $C_i, \ i = 1, \ldots, n$. The last payment includes the reimbursement of the notional value of the bond, i.e. one monetary unit. The contract value at time $t \leq T$ is denoted by $CB(t, T)$ and is given by

$$CB(t, T) = \sum_{i=1}^{n} C_i P(t, T_i).$$

There is a variety of interest rate contracts and derivatives thereof, such as (see [9], Chapter 1) *forward rate agreements, swaps, caps, floors, collars, swaptions*, etc., which we will not examine in this work. However, all the listed contracts can be expressed in terms of ZCBs and/or the European options thereof (see [9], §3.3.2).

In the market model considered in the previous section, we have $N+1$ traded securities. In the interest rate market (the bond market) we see that ZCBs and CBs are typically issued with maturities in $1/4, 1/2, 1, 2, \ldots, 10, 15$ and 30 years. However, the ZCB market has theoretically an infinite number of traded securities, namely ZCBs for every maturity $T \in [0, T^*]$. The no-arbitrage theory in such a situation is more complicated (see [7], [45] or [65]) and we will not consider this topic. We rather make the following simplifying assumption.

Assumption 2.15. (*a*) *There is an equivalent martingale measure* $\mathbb{Q}$ *on* $(\Omega, \mathcal{A})$.

(*b*) *There is a (frictionless) market*[4] *for ZCBs with maturity* T *for every* $T \in [0, T^*]$.

(*c*) *For every contingent claim there is a self-financing portfolio strategy involving a finite number of securities at any trading time.*

Remark 2.16. Under Assumptions 2.15 (*b*) and (*c*), every contingent claim can be replicated via a self-financing portfolio strategy and hence is attainable (see Definition 2.7). From Assumption (*a*) and Proposition 2.10, we immediately obtain the no-arbitrage price at time $t \geq 0$ of a ZCB with maturity $T \in [t, T^*]$, as follows:

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s\, ds\right) \Bigm| \mathcal{F}_t\right]. \tag{2.3}$$

In practice, the ZCB prices $P(t,T)$ are usually translated into an implicit rate of return $R(t,T)$, the so-called (continuous) *zero-bond rate* at time t for the maturity T.

Definition 2.17 (Zero-bond rate). The (continuous) zero-bond rate or the spot-rate at time t for the maturity $T \in [t, T^*]$ is given by

$$R(t,T) = -\frac{\ln P(t,T)}{T-t}, \tag{2.4}$$

[4] An ideal trading environment that imposes no costs or restraints on transactions.

or equivalently,

$$P(t,T) = e^{-R(t,T)(T-t)}. \tag{2.5}$$

The graph of the mapping

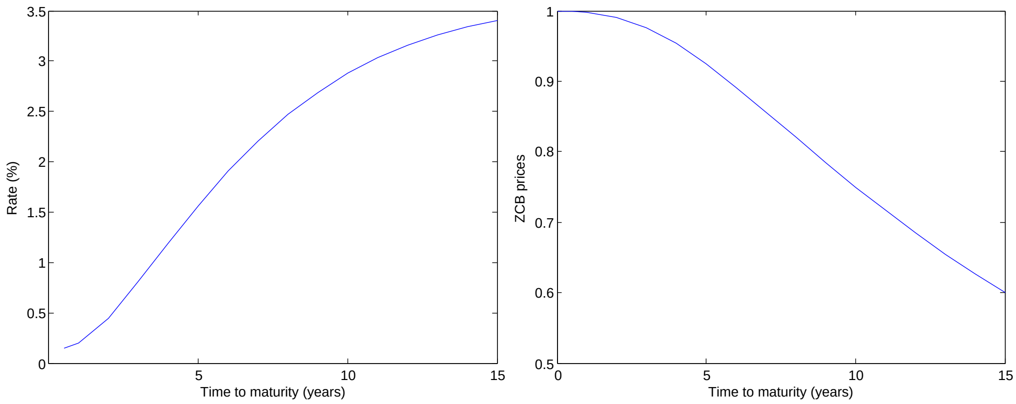
$$T \longmapsto R(t,T), \quad T \geq t,$$

is called the *yield curve* or the *term structure of the market* at time t.

An important notion we will use later in this work is the so-called *zero-bond curve*, which is the graph of the mapping

$$T \longmapsto P(t,T), \quad T \geq t,$$

for a given t. Examples of a yield curve and a zero-bond curve are given in Figure 2.1.



(a) Yield curve

(b) Zero-bond curve

Figure 2.1: Term structure of the German debt securities market on 11 June 2010, at 11:07, see [14].

If the maturity $T \geq t$ converges to $t \geq 0$, the corresponding limit of the zero-bond rate yields the notion of the *instantaneous spot rate,* which is of special interest to us.

Definition 2.18 (Short-rate). The instantaneous spot rate prevailing at time $t \geq 0$ is denoted by $r(t)$, and is defined by

$$r(t) = \lim_{T \to t} R(t,T). \tag{2.6}$$

From Definition 2.17 of the zero-bond rate we immediately obtain

$$r(t) = -\lim_{T \to t} \frac{\ln P(t,T)}{T-t} = -\frac{\partial \ln P(t,t)}{\partial T},$$

where $\partial \ln P(t,t)/\partial T$ is defined as the partial derivative $\partial \ln P(t,T)/\partial T$ evaluated at $T = t$. Here and in the following we assume that all derivatives, integrals and limits exist as they appear.

If we consider the interest rate $r(t)$ to be not a deterministic function but a stochastic process $(r_t)_{t\in[0,T^*]}$, we denote it as the *short-rate process*.

2.3 Short-rate models

The most widely used method of stochastic modelling of the short-rate (2.6) in continuous time is to describe the short-rate process in terms of its dynamics as a solution of a stochastic differential equation (SDE) (see [4], [32] and [47] for a detailed study of SDEs; see [9],[28], [50] and [51] for applications to short-rate modelling). There are basically two different approaches. The first deals with modelling the short-rate dynamics as a one-dimensional stochastic process $(r_t)_{t\in[0,T]}$ under the *objective measure* $\mathbb{P}$. The second suggests considering the dynamics of the short-rate process under an equivalent martingale measure $\mathbb{Q}$ which is assumed to exist due to Assumption 2.15.

On the one hand, a modelling under the objective measure $\mathbb{P}$ comes with a major difficulty. Since we do not have the risk-neutral valuation formula (2.2), we cannot a priori obtain unique arbitrage-free prices of contingent claims. On the contrary, we immediately have the unique arbitrage-free prices of all attainable contingent claims if we consider the dynamics of the short-rate process under an equivalent martingale measure $\mathbb{Q}$. On the other hand, under $\mathbb{Q}$ we pay the price of having to *calibrate* the short-rate model to the market data. This procedure, also called *fitting*, is rather numerically unstable and often turns out to be an ill-posed problem. A thorough analysis of calibration procedure can be found for example in [9] or [28].

In this section we will give a brief overview of some classical short-rate models. Here, we will restrict out attention to the modelling of the underlying processes under an equivalent martingale measure $\mathbb{Q}$. Let the short-rate process $(r_t)_{t\in[0,T]}$, $T \leq T^*$, under $\mathbb{Q}$ be given by the SDE

$$dr_t = \mu(r_t,t)dt + \sigma(r_t,t)dW_t, \tag{2.7}$$

where r_0 is a positive constant, $(W_t)_{t\in[0,T]}$ is the standard Brownian motion on the probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$ with the natural filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and the functions μ and σ are some sufficiently regular functions that will be specified later.

Consider the price $\pi(r,t)$ of an attainable contingent claim depending on time $t \in [0,T]$ and on the value $r = r_t$ at time t of the short-rate process $(r_t)_{t\in[0,T]}$. Let $\phi = \phi(r_T)$ be the payoff function of this contingent claim at maturity $T \le T^*$. Thus, we have from Proposition 2.10 the risk-neutral valuation formula

$$\pi(r,t) = E^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s\,ds\right)\cdot\phi(r_T)\,\middle|\, r_t = r\right], \quad 0 \le t \le T. \tag{2.8}$$

If ϕ satisfies some conditions of smoothness we can apply the *Feynman-Kac Theorem* (see e.g. Theorem 8.2.1 in [47]) and obtain the following fundamental partial differential equation (PDE) for the pricing function π:

$$\pi_t + \mu(r,t)\pi_r + \frac{1}{2}\sigma^2(r,t)\pi_{rr} - r\pi = 0 \quad ((r,t) \in \mathbb{R} \times (0,T)), \tag{2.9}$$

with the terminal condition $\pi(r,T) = \phi(r)$ for all $r \in \mathbb{R}$ and with μ and σ as in (2.7).

Under an absence of arbitrage, the PDE (2.9) for the pricing function of a contingent claim can also be derived with a hedging argument, similarly to the original derivation of the *Black-Scholes PDE*. A thorough study of this approach can be found, for example, in [7].

In the case of a ZCB we obtain the arbitrage-free price $P(r,t,T) = P(t,T)$ at time $t \in (0,T)$ with maturity $T \le T^*$ and an underlying short-rate process $(r_t)_{t\in[0,T]}$ given due to (2.7) with $r_t = r \in \mathbb{R}$ as the solution of the following *term structure equation*:

$$\begin{cases} P_t + \mu(r,t)P_r + \frac{1}{2}\sigma^2(r,t)P_{rr} - rP = 0 \\ P(r,T,T) = 1. \end{cases} \tag{2.10}$$

In the following we present some of the most widely used short-rate models providing their dynamics under $\mathbb{Q}$:

- **Vasicek model** (1977):

$$dr_t = k[\theta - r_t]dt + \sigma dW_t, \tag{2.11}$$

- **Cox-Ingersoll-Ross model (CIR)** (1985):

$$dr_t = k[\theta - r_t]dt + \sigma\sqrt{r_t}dW_t, \tag{2.12}$$

- **Hull-White model (HW)** (1990):

$$dr_t = k(t)[\theta(t) - r_t]dt + \sigma(t)dW_t, \tag{2.13}$$

- **Black-Karasinski model (BK)** (1991):

$$d \ln r_t = k[\theta(t) - \ln r_t]dt + \sigma dW_t, \tag{2.14}$$

where r_0, k, θ and σ are positive constants and $\theta(.), k(.)$ and $\sigma(.)$ are continuous non-negative functions. A comprehensive study of these (and other models) can be found in [9] and in the original works of Vasicek [62], Cox, Ingersoll and Ross [11], Hull and White [25], and Black and Karasinski [5].

In the practice, the following questions play a key role in the choice of a specific model:

- Does the dynamic imply non-negative interest rates, i.e. $r_t \geq 0 \quad \forall t \in [0, T]$?
- What distribution does the dynamic imply for the short-rate $(r_t)_{t\in[0,T]}$? Is this distribution realistic?
- Is the model *mean reverting*? This means that the expected value of the short-rate tends to a constant value as time t converges to infinity, while its variance remains bounded. This property is consistent with the empirical observation that interest rates fluctuate around a long-term average.
- Are ZCB prices explicitly computable? Are European option prices on ZCBs explicitly computable?

These properties are summarized in Table 2.1 below.

Model	$\mathbf{r_t \geq 0}$	**Distribution**	**Mean reversion**	**Explicit ZCB prices / option prices**
Vasicek	no	normal	yes	yes/yes
CIR	yes	non-central χ^2	yes	yes/yes
HW	no	normal	yes	yes/yes
BK	yes	lognormal	yes	no/no

Table 2.1: Properties of short-rate models (2.11) - (2.14).

The choice of the specific short-rate process should ensure that the calculations required for the evaluation of derivatives remain manageable. In this respect, the so-called *affine term structure models*, which we will consider in more detail in the following section, are advantageous. All short-rate models except the BK model presented above belong to this class.

2.4 Affine term structure models

Affine term structure models are short-rate models where the ZCB price $P(t,T)$ at time $t \in [0,T]$ can be written in the form

$$P(t,T) = A(t,T) \cdot e^{-B(t,T) \cdot r_t} \quad , t \in [0,T], \tag{2.15}$$

where $A(t,T)$ and $B(t,T)$ are some deterministic functions. This relation is always satisfied when the spot interest rate $R(t,T)$ given as in Definition 2.17 can be written in the form

$$R(t,T) = \alpha(t,T) + \beta(t,T) r_t,$$

where $\alpha(t,T)$ and $\beta(t,T)$ are some deterministic functions. This relation follows immediately from the definition (2.4)

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

with $\alpha(t,T) = -\frac{\ln A(t,T)}{T-t}$ and $\beta(t,T) = \frac{B(t,T)}{T-t}$.

It can be shown (see Duffie [15], pp. 136-138) that the equation (2.15) is true if μ and σ^2 from (2.7) are affine themselves, i.e.

$$\begin{cases} \mu(r,t) = \alpha_0(t) r + \alpha_1(t), \\ \sigma^2(r,t) = \beta_0(t) r + \beta_1(t), \end{cases} \tag{2.16}$$

where $\alpha_0, \alpha_1, \beta_0$ and β_1 are continuous functions. And in case μ and σ do not depend on t, i.e. if a short-rate model is defined in terms of the SDE

$$dr_t = \mu(r_t) dt + \sigma(r_t) dW_t,$$

where r_0 is a positive constant and the functions μ and σ are some sufficiently regular functions, and if the model belongs to the class of affine term structure models, then the following is true

$$\begin{cases} \mu(r) = \alpha_0 r + \alpha_1, \\ \sigma^2(r) = \beta_0 r + \beta_1, \end{cases} \tag{2.17}$$

where $\alpha_0, \alpha_1, \beta_0$ and β_1 are constant.

In the following we exploit the special form of the affine term structure models. We consider the general term structure equation (2.10) for a ZCB price given due to the representation (2.15). First, we make some preliminary calculations and compute the partial derivatives of $P(r,t,T) = P(t,T)$ for a given short-rate process $(r_t)_{t\in[0,T]}$ and the state of the process $r = r_t$ at time $t \in [0,T]$:

$$\begin{aligned} P_t(r,t,T) &= A_t(t,t)e^{-B(t,T)r} - A(t,T)B_t(t,T)re^{-B(t,T)r} \\ &= \left[\frac{\partial}{\partial t}\ln A(t,T) - B_t(t,T)r\right]P(r,t,T), \\ P_r(r,t,T) &= -A(t,T)B(t,T)e^{-B(t,T)r} = -B(t,T)P(r,t,T), \\ P_{rr}(r,t,T) &= B(t,T)^2P(r,t,T). \end{aligned}$$

Hence, we can rewrite the term structure equation (2.10) as follows:

$$\begin{aligned} &\left[\frac{\partial}{\partial t}\ln A(t,T) - B_t(t,T)r\right]P(r,t,T) - \mu(r,t)B(t,T)P(r,t,T) + \\ &+\frac{1}{2}\sigma^2(r,t)B(t,T)^2P(r,t,T) - rP(r,t,T) = 0. \end{aligned} \tag{2.18}$$

We use the affine structure of μ and σ^2 given by (2.16) and obtain

$$\begin{aligned} &\frac{\partial}{\partial t}\ln A(t,T) - \alpha_1(t)B(t,T) + \frac{1}{2}\beta_2(t)B(t,T)^2 - \\ &-r\cdot\left[B_t(t,T) + \alpha_0(t)B(t,T) - \frac{1}{2}\beta_1(t)B(t,T)^2 + 1\right] = 0. \end{aligned}$$

This equation must hold for all $r \in \mathbb{R}$ and a fixed $t \in (0,T),\ T \in (0,T^*]$. Hence, $A(t,T)$ and $B(t,T)$ are solutions of the *ordinary differential equations* (ODEs)

$$\begin{cases} B_t(t,T) + \alpha_0(t)B(t,T) - \frac{1}{2}\beta_1(t)B(t,T)^2 + 1 = 0 \\ \frac{\partial}{\partial t}\ln A(t,T) - \alpha_1(t)B(t,T) + \frac{1}{2}\beta_2(t)B(t,T)^2 = 0. \end{cases} \tag{2.19}$$

From the terminal condition $P(r,T,T) = 1$ in (2.10) we immediately obtain the terminal conditions $A(T,T) = 1$ and $B(T,T) = 0$ for (2.19).

The first equation in (2.19) is called the *Riccati differential equation* (see [26] for details). In general this ODE can only be solved numerically. However, for the Vasicek model (2.11) and the CIR model (2.12) solutions can be derived explicitly.

In the following two sections we give a short review of the Vasicek model (2.11) and the CIR model (2.12). We also provide the derivation of the ZCB prices in the CIR model via the PDE approach described above.

2.4.1 Vasicek model

The Vasicek model [62] is one of the earliest and still one of the most popular short-rate models. Closed-form expressions of ZCB prices and European options thereof make the model highly appealing to practitioners. However, it also has some shortcomings.

In this section we give a short description of the Vasicek model. We examine its main properties, point out its advantages and disadvantages, and derive the spectral decomposition of its transition probability density.

The short-rate process proposed by Vasicek is a mean-reverting version of the well-known *Ornstein-Uhlenbeck* process. The formulation of the model under a risk-neutral measure $\mathbb{Q}$ is given by the SDE

$$dr_t = k[\theta - r_t]dt + \sigma dW_t, \tag{2.20}$$

where k, θ, σ, r_0 are positive constants, and $(W_t)_{t\geq 0}$ is the standard Brownian motion on the probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ with the natural filtration $(\mathcal{F}_t)_{t\geq 0}$. With the characterization (2.16) it is clear that the Vasicek model is an affine term structure model.

Integration of the equation (2.20) yields for $t \geq s$ (see e.g. [37] for details on solving SDEs)

$$r_t = r_s e^{-k(t-s)} + \theta\left(1 - e^{-k(t-s)}\right) + \sigma \int_s^t e^{-k(t-u)}\, dW_u. \tag{2.21}$$

Thus, conditional on $\mathcal{F}_s$ $(t \geq s)$, r_t is normally distributed with mean and variance

$$\mathbb{E}[r_t|\mathcal{F}_s] = r_s e^{-k(t-s)} + \theta\left(1 - e^{-k(t-s)}\right), \tag{2.22}$$

$$\mathbb{V}ar[r_t|\mathcal{F}_s] = \frac{\sigma^2}{2k}\left(1 - e^{-2k(t-s)}\right). \tag{2.23}$$

Hence, the short-rate process $(r_t)_{t\geq 0}$ tends to the mean-reverting value θ for $t \to \infty$.

In Section 3.2 we will need the spectral representation of the transition probability density of the Vasicek process (2.20), or, more precisely, of the shifted process

$$z_t = r_t - \theta, \quad t \geq 0. \tag{2.24}$$

With *Itô's lemma*, we immediately see that $(z_t)_{t\geq 0}$ satisfies the following SDE:

$$dz_t = -kz_t dt + \sigma dW_t, \quad z_0 = r_0 - \theta, \tag{2.25}$$

i.e. $(z_t)_{t\geq 0}$ is an Ornstein-Uhlenbeck process. From the distribution of the Vasicek process (2.21) we immediately obtain the distribution of z_t, conditional on $z_s = z$ $(s \leq t)$, by setting $\theta = 0$ in (2.24), (2.22) and (2.23) as follows:

$$z_t | z_s \sim \mathcal{N}\left(z e^{-k(t-s)}, \frac{\sigma^2}{2k}\left(1 - e^{-2k(t-s)}\right)\right),$$

i.e. the density function of the transition probability of $(z_t)_{t\geq 0}$ is given for $x, y \in (-\infty, \infty)$ by

$$p(t; x, y) = \frac{1}{\sqrt{2\pi\left(\frac{1-e^{-2kt}}{2C}\right)}} \cdot \exp\left(-\frac{C(y - xe^{-kt})^2}{1 - e^{-2kt}}\right), \tag{2.26}$$

with $C = k/\sigma^2$. Consider the *Mehler formula* (see [18], p. 194)

$$\frac{1}{\sqrt{1-\xi^2}} \exp\left(\frac{2xy\xi - (x^2+y^2)\xi^2}{1-\xi^2}\right) = \sum_{n=0}^{\infty} \frac{1}{2^n n!} H_n(x) H_n(y) \xi^n,$$

where $H_n(.)$, $n \in \mathbb{N}_0$, are the *Hermite polynomials* given by (see [60], §5.5)

$$H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \frac{n!}{(n-2k)!k!} (2x)^{n-2k}.$$

With $\xi = e^{-kt}$ we obtain for $x, y \in (-\infty, \infty)$ the spectral decomposition of the density function (2.26) as follows:

$$p(t; x, y) = \frac{\sqrt{C}}{\sqrt{\pi}} \cdot e^{-Cy^2} \sum_{n=0}^{\infty} \frac{1}{2^n n!} H_n(\sqrt{C}x) H_n(\sqrt{C}y) e^{-knt}. \tag{2.27}$$

It is possible to obtain this result directly from the SDE (2.25). We will use such approach later on in Section 4.1 for a more general stochastic process, where an explicit formula for the transition density function is not known.

The closed-form representation of the transition density (2.26) of the Ornstein-Uhlenbeck process, and hence of the Vasicek process (2.21) (and Gaussian processes in general) is one of the main advantages of the Vasicek model, since it thereby allows for the explicit pricing of ZCBs and the European options thereof (see e.g. [9], §3.2.1 and §3.3.2 for details).

The drawbacks of the Vasicek model are, however, the possible negativity of the interest rates, implied by the Gaussian distribution, and the fact that it is driven by only three parameters, which renders the calibration an ill-posed problem and often yields poor results.

The price at time $t \geq 0$ of a ZCB with maturity T, conditional on $r_t = r$, is given by (see [9], §3.2.1)

$$P(t,T) = A(t,T)e^{-B(t,T)\cdot r}, \tag{2.28}$$

where

$$\begin{aligned} A(t,T) &= \exp\left\{\left(\theta - \frac{\sigma^2}{2k^2}\right)[B(t,T) - T + t] - \frac{\sigma^2}{4k}B(t,T)^2\right\}, \\ B(t,T) &= \frac{1}{k}\left(1 - e^{-k(T-t)}\right). \end{aligned}$$

2.4.2 Cox-Ingersoll-Ross model

The Cox-Ingersoll-Ross (CIR) model [11] is, beside the Vasicek model, a very popular short-rate model. The factors that make it highly appealing to practitioners are the positivity of the short rate and the closed-form expressions of the ZCB prices and European options thereof.

In this section we give a short overview of the CIR model. We describe the distribution of the short-rate process, derive its transition probability density in terms of its spectral decomposition, and use the PDE approach described in Section 2.4 for the explicit pricing of ZCBs.

The general equilibrium approach to term structure modelling developed in [11] leads to the modification of the Vasicek model (2.20), also known as the *square-root process*. Its formulation under a risk-neutral measure $\mathbb{Q}$ is given by the SDE

$$dr_t = k[\theta - r_t]dt + \sigma\sqrt{r_t}dW_t, \tag{2.29}$$

where k, θ, σ, r_0 are positive constants and $(W_t)_{t\geq 0}$ is the standard Wiener process on the probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ with the natural filtration $(\mathcal{F}_t)_{t\geq 0}$. Similarly to the Vasicek model, the CIR model belongs to the class of affine term structure models, which is easily verified via (2.16).

The stability condition

$$\frac{\sigma^2}{2k} \leq \theta \tag{2.30}$$

has to be imposed to ensure that the origin is inaccessible to the process (2.29). We will show the explicit derivation of a similar stability condition for a more general process later on in Section 4.1.

We will see that the distribution of the process $(r_t)_{t\geq 0}$ is given by a *non-central* χ^2 *distribution* whose definition and some properties are given in the following (see [30], Chapter 29, for details of this distribution).

Definition 2.19 (**Noncentral χ^2 distribution**). The density function of the noncentral χ^2 distribution with $v > 0$ degrees of freedom and the non-centrality parameter $\lambda > 0$ is given for $x > 0$ by

$$\begin{aligned} f_{\chi^2(v,\lambda)}(x) &= \sum_{i=0}^{\infty} \frac{(\lambda/2)^i}{i!} e^{-\lambda/2} \cdot f_{\Gamma(i+\frac{v}{2},2)}(x) \\ &= \frac{1}{2}\left(\frac{x}{\lambda}\right)^{(v-2)/4} e^{-(\lambda+x)/2} I_{v/2-1}(\sqrt{\lambda x}), \end{aligned} \tag{2.31}$$

where $f_{\Gamma(k,\eta)}(.)$ is the density function of the *Gamma distribution* $\Gamma(k,\eta)$ given by (see [29], p. 343)

$$f_{\Gamma(k,\eta)}(x) = \frac{x^{k-1}e^{-x/\eta}}{\eta^k \Gamma(k)}, \quad x, k, \eta > 0, \tag{2.32}$$

and $I_\nu(.)$ is the *modified Bessel function of the first kind of order* ν given by (see [1], §9.6)

$$I_\nu(y) = \left(\frac{y}{2}\right)^\nu \sum_{j=1}^{\infty} \frac{(y^2/4)^j}{j!\Gamma(\nu+j+1)}, \quad y > 0, \nu > -1. \tag{2.33}$$

The mean and the variance of a noncentral χ^2-distributed random variable Y are given by

$$\mathbb{E}[Y] = v + \lambda, \tag{2.34}$$

$$\mathbb{V}ar[Y] = 2(v + 2\lambda). \tag{2.35}$$

Theorem 2.20. *Let $(r_t)_{t\geq 0}$ be the CIR process given by (2.29).*

(*a*) *The density function of its transition probability is given by*

$$p(t; x, y) = c_t f_{\chi^2(v,\lambda_t)}(c_t y), \quad x, y > 0, \tag{2.36}$$

where

$$\begin{aligned} c_t &= \frac{4k}{\sigma^2(1-\exp(-kt))}, \\ v &= \frac{4k\theta}{\sigma^2}, \\ \lambda_t &= c_t x \exp(-kt), \end{aligned}$$

and f_{χ^2} is the probability density function of the noncentral χ^2-distribution given as in Theorem 2.19.

(*b*) *The spectral decomposition of* $p(t;x,y)$ *is given for* $x,y>0$ *by*

$$p(t;x,y) = e^{-hy}h^{\alpha+1}y^{\alpha}\sum_{n=0}^{\infty}\frac{n!}{\Gamma(n+\alpha+1)}L_n^{(\alpha)}(hx)L_n^{(\alpha)}(hy)e^{-knt}, \tag{2.37}$$

where $h=\frac{\sigma^2}{2k}$, $\alpha=\frac{2k\theta}{\sigma^2}$ *and* $L_n^{\alpha}(.)$, $n\in\mathbb{N}_0$, *are the Laguerre polynomials given by (see [60], Chapter 5)*

$$L_n^{(\alpha)}(x) = \sum_{m=0}^{n}\binom{n+\alpha}{n-m}\frac{(-x)^m}{m!}, \quad \alpha>-1, x\geq 0.$$

(*c*) *The mean and the variance of* r_t, *conditional on* $\mathcal{F}_s$ $(s\leq t)$, *are given as follows:*

$$\begin{aligned}\mathbb{E}[r_t|\mathcal{F}_s] &= r_s e^{-k(t-s)} + \theta\left(1-e^{-k(t-s)}\right),\\ \mathbb{V}ar[r_t|\mathcal{F}_s] &= r_s\frac{\sigma^2}{k}\left(e^{-k(t-s)}-e^{-2k(t-s)}\right) + \theta\frac{\sigma^2}{2k}\left(1-e^{-k(t-s)}\right)^2.\end{aligned}$$

(*d*) *The stationary distribution is the limiting distribution, and is given by the Gamma distribution* $\Gamma(\frac{2k\theta}{\sigma^2},\frac{2k}{\sigma^2})$, *where the Gamma distribution is given as in (2.32).*

Proof. The proof is given in Appendix B.2 □

Thus, the CIR process $(r_t)_{t\geq 0}$ tends to the mean-reverting value θ as $t\to\infty$. The main advantage of the model compared to a Gaussian model such as the Vasicek model is the non-negativity of the short-rate. Moreover, under the stability condition (2.30), the short-rate will never reach zero, but remain positive all the time.

In the following, we give the arbitrage-free ZCB price in the CIR model.

Theorem 2.21. *The price at time* $t\in[0,T]$ *of a ZCB with maturity* $T\leq T^*$, *conditional on* $r_t=r$, *is given by*

$$P(t,T) = A(t,T)e^{-B(t,T)\cdot r}, \tag{2.38}$$

where

$$\begin{aligned}A(t,T) &= \left[\frac{2h\cdot e^{(k+h)(T-t)/2}}{2h+(k+h)(e^{(T-t)h}-1)}\right]^{\frac{2k\theta}{\sigma^2}},\\ B(t,T) &= \frac{2(e^{(T-t)h}-1)}{2h+(k+h)(e^{(T-t)h}-1)},\\ h &= \sqrt{k^2+2\sigma^2}.\end{aligned}$$

Proof. The proof is given in Appendix B.3.

Chapter 3

Ehrenfest short-rate model

One of the approaches to the stochastic modelling of interest rates is based on the assumption that the short-rate is given by a Markov chain model in discrete or continuous time. Having a predefined finite state space, such a model allows for a descriptive interpretation of interest rate development and is used in insurance mathematics (see [46] for details and other references) and finance (see e.g. [10] and [16]). Contrary to some diffusion short-rate models, introduced in the previous chapter, Markov chain models often fail to provide explicit pricing formulae even for basic interest rate contracts like ZCBs.

In this chapter, we propose a finite-state mean-reverting model for the short-rate related to the continuous time Ehrenfest process. By choosing the lower boundary as non-negative, the problem of negative interest rates can be prevented. Furthermore, the model allows for the explicit evaluation of ZCB prices. In this way, the model aims at realism and analytical tractability.

In the following, we will examine the main properties of the model and derive explicit pricing formulae for ZCBs in its general and special symmetric cases. We will show that after a suitable linear transformation this model converges to the Vasicek model (see Section 2.4.1), and we will provide some numerical results.

3.1 Original Ehrenfest model

The original Ehrenfest model describes the heat exchange between two isolated bodies, each of arbitrary temperature. The temperatures are symbolized by the number of fluctuating balls in two urns with a total of $N \in \mathbb{N}$ balls. For details of the continuous and discrete time versions of the model we refer to [6], [35], [38] and [57].

First, we shall discuss the discrete and continuous time versions of the Ehrenfest process. Primarily, the representation of the continuous time Ehrenfest process as a sum of simple independent processes will allow us to show the main result of this thesis, which comes up in Section 3.2. Further, we will examine the transition semigroup of the Ehrenfest process and explore some of its basic properties.

3.1.1 Ehrenfest chain

Initially, N balls are distributed between two urns, labelled I and II. A ball is selected at random, each having the probability $\frac{1}{N}$. If the selected ball is in urn I, it changes the urn with a probability of $0 < \alpha \leq 1$ and stays in urn I with the probability $1-\alpha$. If the selected ball is in urn II, it changes the urn with a probability of $0 < \beta \leq 1$ and stays in urn II with the probability $1-\beta$. The process is repeated any number of times. Let $(\hat{X}_n)_{n\in\mathbb{N}}$ be the number of balls in urn I after n selections. Clearly, $(\hat{X}_n)_{n\in\mathbb{N}}$ is a Markov chain with a transition probability matrix P given by

$$p_{ij} = \begin{cases} \alpha \cdot \frac{N-i}{N} & \text{if} \quad j = i+1, \\ \beta \cdot \frac{i}{N} & \text{if} \quad j = i-1, \\ 1 - \alpha \cdot \frac{N-i}{N} - \beta \cdot \frac{i}{N} & \text{if} \quad j = i, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Without going into detail we mention only that the transition probability matrix P has the spectral representation of the form

$$P = \sum_{i=1}^{N} \mu_i v_i u_i^T,$$

where $\mu_i,\ i \in \{1, \dots, N\}$, denote the eigenvalues of P and v_i and $u_i,\ i \in \{1, \dots, N\}$, are the respective right and left eigenvectors of P (see e.g. [27], §1.9 or [38], pp. 965-966). Hence, the n-step transition probabilities can be calculated as follows:

$$P^n = \sum_{i=1}^{N} \mu_i^n v_i u_i^T.$$

Kraft and Schaefer [38] examined $(\hat{X}_n)_{n\in\mathbb{N}}$, denoting it by the *two-parameter Ehrenfest chain* and provided the following theorem that is a modification of the well-known Kac's theorem [31], where the special case of $\alpha = \beta = 1$ is studied.

Theorem 3.1. *Let $(\hat{X}_n)_{n\in\mathbb{N}}$ be the Ehrenfest chain with a transition probability matrix given by (3.1).*

(a) *The n-step transition probabilities $p_{ij}^{(n)}$ are given by*

$$p_{ij}^{(n)} = \binom{N}{j}\left(\frac{p}{q}\right)^j \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) \left(1 - \frac{(\alpha+\beta)x}{N}\right)^n ,$$

where $p = \frac{\alpha}{\alpha+\beta}$, $q = 1 - p$ and $K_i(.) = K_i(.;p;N)$, $i = 0, \ldots, N$, are the Krawtchouk polynomials as given in Definition 1.5.

(b) *The (unique) stationary distribution of $(\hat{X}_n)_{n\in\mathbb{N}}$ is given by the binomial distribution $B_{N,p}$ on $\{0, \ldots, N\}$ with the parameter p.*

Proof. The proof is given in [38], pp. 965-966 and p. 970. □

3.1.2 Ehrenfest process in continuous time

Let N balls, initially distributed between urns I and II in such a way that k balls are in urn I and $N - k$ balls are in urn II, fluctuate independently in continuous time between the two urns. We fix a fluctuation parameter $\lambda > 0$ and independent Poisson processes $(N_t^1)_{t\geq 0}, \ldots, (N_t^N)_{t\geq 0}$ with intensity λ. Let $(\hat{Y}_n)_{n\in\mathbb{N}}$ be a Markov chain with the state space $\{0, 1\}$ and transition probability matrix

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \quad \alpha, \beta \in (0, 1]. \tag{3.2}$$

Then, the subordinated Markov chain $(Y_t^l = \hat{Y}_{N_t^l})_{t\geq 0}$ describes the state of the l-th ball at time t, where $Y_t^l = 1$ or 0 when the l-th ball is in urn I or II respectively. Let k balls be initially in urn I and $N - k$ balls in urn II. Then,

$$\left(X_t = \sum_{l=1}^{N} Y_t^l\right)_{t\geq 0} \tag{3.3}$$

is a Markov process with the state space $E = \{0, 1, \ldots, N\}$, denoting the number of balls in urn I at time t. We call $(X_t)_{t\geq 0}$ the *(continuous time) Ehrenfest process.*

Remark 3.2. A special case of (3.3) where $\alpha = \beta = 1$, first suggested by Siegert [57] and also studied by Bingham [6], where the transitions become "deterministic" in the sense of switching between the states 0 and 1, will be important for us later on in Section 3.2. (Its discrete time analogue leads to the original Ehrenfest chain).

Before computing the transition semigroup of $(X_t)_{t\geq 0}$ we need the following result.

Lemma 3.3. (*a*) *The transition semigroup of* $(Y_t^l)_{t\geq 0}$ *is given by*

$$P(t) = \begin{pmatrix} q + pe^{-\lambda(\alpha+\beta)t} & p - pe^{-\lambda(\alpha+\beta)t} \\ q - qe^{-\lambda(\alpha+\beta)t} & p + qe^{-\lambda(\alpha+\beta)t} \end{pmatrix}, \tag{3.4}$$

where $p = \frac{\alpha}{\alpha+\beta}$ *and* $q = 1 - p$.

(*b*) $(Y_t^l)_{t\geq 0}$ *has a stationary distribution, which is the binomial distribution* $B_{1,p}$.

Proof. Ad (*a*) : Since $(Y_t^l)_{t\geq 0}$ is a Markov chain subordinated by a Poisson process with index λ, the associated transition semigroup can be written as (see [8], p.333)

$$P(t) = e^{-\lambda t} \cdot e^{\lambda t P} = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} P^n. \tag{3.5}$$

We can avoid the computation of P^n by writing

$$P = \mu S + (1-\mu) I, \tag{3.6}$$

where $S = \begin{pmatrix} q & p \\ q & p \end{pmatrix}$ is a stochastic matrix with $S^2 = S$, I is the identity matrix, and $\mu = \alpha + \beta$. Then, (3.5) becomes

$$\begin{aligned} P(t) &= e^{-\lambda t} \cdot e^{\lambda t(\mu S + (1-\mu) I)} = e^{-\lambda t} \cdot e^{\lambda \mu t S} \cdot e^{\lambda (1-\mu) t I} \\ &= e^{-\lambda \mu t} \cdot e^{\lambda \mu t S} = e^{-\lambda(\alpha+\beta)t} \sum_{n=0}^{\infty} \frac{(\lambda(\alpha+\beta)t)^n}{n!} S^n. \end{aligned}$$

Since $S^n = S$ for all $n \in \mathbb{N}$, we can easily compute the above series, which completes the proof.
Ad (*b*) : Let π denote a stationary distribution of $(\hat{Y}_n)_{n\in\mathbb{N}}$ defined by (3.2). It is immediately clear that π is given uniquely by the binomial distribution $B_{1,p}$. Since $(Y_t^l)_{t\geq 0}$ is the subordinated Markov chain, we use (3.5), obtaining

$$\begin{aligned} \pi^T P(t) &= \pi^T e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} P^n = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \pi^T P^n \\ &= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \pi^T = \pi^T, \end{aligned}$$

which completes the proof. □

Remark 3.4. Consider the transition semigroup of $(Y_t^l)_{t\geq 0}$ given by (3.4). We see that any variation of λ is equal to a variation of $\alpha+\beta$, as long as α and β are in an admissible range (α and β may not exceed 1), while $p=\frac{\alpha}{\alpha+\beta}$ remains constant. Heuristically it is clear, since in order to intensify the jumps of the underlying process $(Y_t^l)_{t\geq 0}$ with stationary distribution $B_{1,p}$ and a fixed p, we can either change the intensity λ of the subordinating Poisson process or we can vary the jump probabilities α and β. To put it roughly, the distribution of $(Y_t^l)_{t\geq 0}$ depends either on p and λ or on α and β, where λ is set equal 1.

Theorem 3.5. *(Properties of the Ehrenfest process) Let $(X_t)_{t\geq 0}$ be the Ehrenfest process given by (3.3).*

(*a*) *The transition probabilities $p_{ij}(t)=\mathbb{P}(X_{t+s}=j|X_s=i)$ are given by*

$$p_{ij}(t)=\binom{N}{j}\left(\frac{p}{q}\right)^j\sum_{x=0}^{N}\binom{N}{x}p^xq^{N-x}K_i(x)K_j(x)e^{-\lambda(\alpha+\beta)xt}, \qquad (3.7)$$

where $p=\frac{\alpha}{\alpha+\beta}$, $q=1-p$, and $K_i(.)=K_i(.;p;N)$ $(i\in E)$ are the Krawtchouk polynomials as given in Definition 1.5 .

(*b*) *The conditional mean and variance of $(X_t)_{t\geq 0}$ are given by*

$$\begin{aligned}\mathbb{E}[X_t|X_0=i] &= Np-(Np-i)e^{-\lambda(\alpha+\beta)t}, &(3.8)\\ \mathbb{V}ar[X_t|X_0=i] &= Np(1-p)+(Np-i)(2p-1)e^{-\lambda(\alpha+\beta)t} &(3.9)\\ &\quad -(Np-i)^2(2p-1)e^{-2\lambda(\alpha+\beta)t}.\end{aligned}$$

(*c*) *The limiting distribution of $(X_t)_{t\geq 0}$ is a stationary distribution and is given by the binomial distribution $B_{N,p}$ on E with parameter p. The rate of convergence to the limiting distribution is exponential of order $e^{-\lambda(\alpha+\beta)t}$. More precisely,*

$$p_{ij}(t)=\binom{N}{j}p^jq^{N-j}\left[1+\frac{(Np-i)(Np-j)}{Npq}e^{-\lambda(\alpha+\beta)t}+O\left(e^{-2\lambda(\alpha+\beta)t}\right)\right].$$

Proof. Ad (a) : The proof here is similar to that in [6]. First, we compute the moment-generating function of $(Y_t^l)_{t\geq 0}$, conditional on $Y_0^l\in\{0,1\}$. In the following we suppress the dependence of $(Y_t^l)_{t\geq 0}$ on l when it is clear from the context. From Lemma 3.3 we have

$$\begin{aligned}\mathbb{E}\left[z^{Y_t}\middle|Y_0=1\right] &= p_{10}(t)+p_{11}(t)z=q+pz-q(1-z)e^{-\lambda(\alpha+\beta)t},\\ \mathbb{E}\left[z^{Y_t}\middle|Y_0=0\right] &= p_{00}(t)+p_{01}(t)z=q+pz+p(1-z)e^{-\lambda(\alpha+\beta)t}.\end{aligned}$$

Since $(Y_t^l)_{t\geq 0}$ are independent for all $l = 1, \ldots, N$, we obtain from (3.3) the moment-generating function of $(X_t)_{t\geq 0}$, given $X_0 = i$, as follows:

$$\begin{aligned}
\sum_{j=0}^{N} p_{ij}(t) z^j &= \mathbb{E}\left[z^{\sum_{l=1}^{N} Y_t^l} \big| X_0 = i\right] = \mathbb{E}\left[\prod_{l=1}^{N} z^{Y_t^l} \big| X_0 = i\right] \\
&= \mathbb{E}\left[z^{Y_t} \big| Y_0 = 1\right]^i \cdot \mathbb{E}\left[z^{Y_t} \big| Y_0 = 0\right]^{N-i} \\
&= \quad \left[q + pz - q(1-z)e^{-\lambda(\alpha+\beta)t}\right]^i \\
&\quad \cdot \left[q + pz + p(1-z)e^{-\lambda(\alpha+\beta)t}\right]^{N-i} \\
&= (q+pz)^N \cdot \left[1 - \frac{q}{p} \cdot \frac{p(1-z)}{q+pz} e^{-\lambda(\alpha+\beta)t}\right]^i \\
&\qquad \cdot \left[1 + \frac{p(1-z)}{q+pz} e^{-\lambda(\alpha+\beta)t}\right]^{N-i}.
\end{aligned}$$

Applying (a) and (d) of Proposition 1.6, we obtain

$$\begin{aligned}
\sum_{j=0}^{N} p_{ij}(t) z^j &= (q+pz)^N \sum_{x=0}^{N} \binom{N}{x} K_i(x) \left(\frac{p(1-z)}{q+pz} e^{-\lambda(\alpha+\beta)t}\right)^x \\
&= \sum_{x=0}^{N} \binom{N}{x} K_i(x) (q+pz)^{N-x} p^x (1-z)^x e^{-\lambda(\alpha+\beta)xt} \\
&= \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) \\
&\quad \cdot \left(1 - \frac{q}{p} \cdot \frac{pz}{q}\right)^x \left(1 + \frac{pz}{q}\right)^{N-x} e^{-\lambda(\alpha+\beta)xt}.
\end{aligned}$$

Applying (1.9) once again, we get

$$\begin{aligned}
\sum_{j=0}^{N} p_{ij}(t) z^j &= \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) \left[\sum_{j=0}^{N} \binom{N}{j} K_j(x) \left(\frac{pz}{q}\right)^j\right] e^{-\lambda(\alpha+\beta)xt} \\
&= \sum_{j=0}^{N} \left[\binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(k) K_j(x) e^{-\lambda(\alpha+\beta)xt}\right] z^j.
\end{aligned}$$

Equating coefficients of z^j completes the proof of the claim.
Ad (b) : Combining (3.7) with results $(a), (b)$ and (g) of Proposition 1.6, we

obtain

$$\begin{aligned}
\mathbb{E}[X_t|X_0=i] &= \sum_{j=0}^{N} j\cdot p_{ij}(t) = \sum_{j=0}^{N} j\cdot\omega(j)\sum_{x=0}^{N}\pi_x K_i(x)K_j(x)e^{-\lambda(\alpha+\beta)xt} \\
&= \sum_{x=0}^{N}\pi_x K_i(x)e^{-\lambda(\alpha+\beta)xt}\sum_{j=0}^{N} j\cdot K_x(j)\omega(j) \\
&= \sum_{x=0}^{N}\pi_x B_{0,x}K_i(x)e^{-\lambda(\alpha+\beta)xt}, \qquad (3.10)
\end{aligned}$$

where $B_{0,x}\neq 0$ for $x\in\{0,1\}$ as given in (1.13), which implies (3.8). Moreover,

$$\begin{aligned}
\mathbb{E}[X_t^2|X_0=i] &= \sum_{j=0}^{N} j^2\cdot p_{ij}(t) \\
&= \sum_{j=0}^{N} j^2\cdot\omega(j)\sum_{x=0}^{N}\pi_x K_i(x)K_j(x)e^{-\lambda(\alpha+\beta)xt} \\
&= \sum_{x=0}^{N}\pi_x K_i(x)e^{-\lambda(\alpha+\beta)xt}\sum_{j=0}^{N} j^2\cdot K_x(j)\omega(j) \\
&= \sum_{x=0}^{N}\pi_x K_i(x)e^{-\lambda(\alpha+\beta)xt}\sum_{j=0}^{N} j\cdot\omega(j)\Big[-(N-x)pK_{x+1}(j) \\
&\qquad + \left[(N-x)p+xq\right]K_x(j) - xqK_{x-1}(j)\Big] \\
&= \sum_{x=0}^{N}\pi_x K_i(x)e^{-\lambda(\alpha+\beta)xt}\Big[-(N-x)pB_{0,x+1} \\
&\qquad + \left[(N-x)p+xq\right]B_{0,x} - xqB_{0,x-1}\Big],
\end{aligned}$$

where $B_{0,x}$ is given by (1.13). A straightforward computation of

$$\mathbb{V}ar[X_t|X_0=i] = \mathbb{E}[X_t^2|X_0=i] - \mathbb{E}[X_t|X_0=i]^2$$

leads immediately to (3.9).
Ad (c) : From the representation (3.7) of the transition probability and

Proposition 1.6 $(a), (b)$ and (c) we obtain

$$\begin{aligned} p_{ij}(t) &= \binom{N}{j} p^j q^{N-j} \left[1 + N\frac{p}{q} K_i(1) K_j(1) e^{-\lambda(\alpha+\beta)t} \right. \\ &\qquad \left. + \sum_{x=2}^{N} \binom{N}{x} \left(\frac{p}{q}\right)^x K_i(x) K_j(x) e^{-\lambda(\alpha+\beta)xt} \right] \\ &= \binom{N}{j} p^j q^{N-j} \left[1 + N\frac{p}{q} \left(1 - \frac{i}{Np}\right) \left(1 - \frac{j}{Np}\right) e^{-\lambda(\alpha+\beta)t} \right. \\ &\qquad \left. + O\left(e^{-2\lambda(\alpha+\beta)t}\right) \right] \\ &= \binom{N}{j} p^j q^{N-j} \left[1 + \frac{(Np-i)(Np-j)}{Npq} e^{-\lambda(\alpha+\beta)t} + O\left(e^{-2\lambda(\alpha+\beta)t}\right) \right]. \end{aligned}$$

Hence, the limiting distribution is obtained as follows:

$$\lim_{t\to\infty} p_{ij}(t) = \binom{N}{j} p^j q^{N-j},$$

which is the probability mass function of the binomial distribution $B_{N,p}$ on E with the parameter p. Since the limiting distribution is always a stationary distribution, the claim follows. □

Remark 3.6. Karlin and McGregor [35] provided an alternative but equivalent definition of (3.3) as a *birth-and-death process* with the state space E. Here, the time intervals between events are independently exponentially distributed with intensity γ, and for $i \in E$ the birth and death rates are $\lambda_i = \gamma\alpha\frac{(N-i)}{N}$ and $\mu_i = \gamma\beta\frac{i}{N}$ respectively, where α and β are given as above. To see this we derive the transition probabilities of the Ehrenfest chain defined by (3.1), subordinated by a Poisson process $(N_t)_{t\geq 0}$ with intensity $\gamma = \lambda N$, where $\lambda > 0$ is given as above. From Theorem 3.1 we know the transition probabilities $p_{ij}^{(n)}$ and with the formula (3.5) we obtain

$$\begin{aligned} p_{ij}(t) &= \sum_{n=0}^{\infty} e^{-\gamma t} \frac{(\gamma t)^n}{n!} p_{ij}^{(n)} \\ &= \sum_{n=0}^{\infty} e^{-\gamma t} \frac{(\gamma t)^n}{n!} \binom{N}{j} \left(\frac{p}{q}\right)^j \\ &\quad \cdot \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) \left(1 - \frac{(\alpha+\beta)x}{N}\right)^n \\ &= \binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) e^{-\gamma t} \\ &\quad \cdot \sum_{n=0}^{\infty} \frac{\left(\gamma t \left(1 - \frac{(\alpha+\beta)x}{N}\right)\right)^n}{n!} \\ &= \binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) \exp\left(-\frac{\gamma t(\alpha+\beta)x}{N}\right)^n \\ &= \binom{N}{j} \left(\frac{p}{q}\right)^j \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} K_i(x) K_j(x) \exp\left(-\lambda t(\alpha+\beta)x\right)^n, \end{aligned}$$

which is the transition probability (3.7) of the Ehrenfest process.

3.2 Ehrenfest short-rate model

In this section we introduce a finite-state mean-reverting short-rate model associated with the continuous time Ehrenfest process (3.3) and give its basic properties. As a main result, we exploit the algebraic-combinatorial roots of the Ehrenfest process and derive explicit pricing formulae for ZCBs given as in Definition 2.13 in the general and the special cases of the process, both of which have their advantages.

3.2.1 Definition and properties

Let $[r_m, r_M] \subseteq \mathbb{R}$ be an interval on the real line. We decompose it into N equal pieces of length $h = \frac{r_M - r_m}{N}$ and consider the process

$$(R_t^{(N)} = hX_t^{(N)} + r_m)_{t \geq 0} \tag{3.11}$$

as a short-rate process (see Definition 2.18) with the state space $E = \{r_k = hk + r_m,\ k = 0, \ldots, N\}$, where $(X_t^{(N)})_{t\geq 0}$ is the Ehrenfest process given by (3.3) where $\alpha, \beta \in (0, 1]$. From Remark 3.4 we know that the pairs α, β and p, λ equally determine the distribution of the Ehrenfest process and hence the distribution of (3.11). In the following we use both terminologies. Bearing this in mind, we denote this short-rate model as the $\mathcal{E}(p, \lambda)$ *model*.

Considering Remark 3.6, we notice that $(R_t^{(N)})_{t\geq 0}$ can be seen as an affine linearly transformed birth-and-death process on $\{0, 1, \cdots, N\}$. In the case at hand, N can be interpreted as the state space discretization parameter. Clearly, $(R_t = R_t^{(N)})_{t\geq 0}$ also depends on N. We will suppress this dependence when it is clear from the context.

From Theorem 3.5 we immediately obtain the mean and variance of $(R_t)_{t\geq 0}$, conditional on $R_0 = r_k \in E$, as follows:

$$\begin{aligned}
\mathbb{E}[R_t|R_0 = r_k] &= h \cdot \mathbb{E}[X_t|X_0 = k] + r_m \\
&= (r_M - r_m)\Big(p - \big(p - \frac{i}{N}\big)e^{-\lambda(\alpha+\beta)t}\Big) + r_m, \qquad (3.12)\\
\mathbb{V}ar[R_t|R_0 = r_k] &= h^2 \cdot \mathbb{V}ar[X_t|X_0 = k] \\
&= \frac{(r_M - r_m)^2}{N^2}\Big(Np(1-p) + (Np - i)(2p-1)e^{-\lambda(\alpha+\beta)t} \\
&\qquad -(Np - i)^2(2p-1)e^{-2\lambda(\alpha+\beta)t}\Big), \qquad (3.13)
\end{aligned}$$

where $p = \frac{\alpha}{\alpha+\beta}$. We also obtain the *mean reversion* of $(R_t)_{t\geq 0}$:

$$\lim_{t\to\infty} \mathbb{E}[R_t|R_0] = pr_M + (1-p)r_m, \qquad (3.14)$$

$$\lim_{t\to\infty} \mathbb{V}ar[R_t|R_0] = \frac{(r_M - r_m)^2}{N} p(1-p) < \infty. \qquad (3.15)$$

Thus, we have a total of five parameters, r_m, r_M, p, λ and N, to fit the model to the market data. Here, p governs the skewness of the underlying distribution, $r_M - r_m, N$ and p have an impact on its kurtosis, and λ influences the speed of reversion to the mean reverting value $pr_M + (1-p)r_m$.

3.2.2 Zero-coupon bond

In Chapter 2, Definition 2.13, we introduced the notion of a ZCB that stands for a stochastic discount factor and is essential for the world of *Fixed Income*. In this section we derive the arbitrage-free price $P(t, T)$ of a ZCB at time t with maturity $T \leq T^*$ within the Ehrenfest short-rate model.

Let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration of $(R_t)_{t\geq 0}$. Then, due to Assumption 2.15 there exists an equivalent martingale measure $\mathbb{Q}$ such that the price

$P(t,T)$ is given due to the risk-neutral valuation formula (2.3) as follows:

$$P^{(N)}(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T R_s\,ds\right)\middle|\mathcal{F}_t\right]. \tag{3.16}$$

In the following we omit explicitly writing out the dependence of $P(t,T)$ on N when it is clear from the context.

The calculation of (3.16) within the $\mathcal{E}(p,\lambda)$ model with arbitrary $\alpha, \beta \in (0,1]$ is inspired by the proof of Theorem 3.1 in [13]. There, Delbaen and Shirakawa represent the transition probabilities of the underlying short-rate process as a weighted series of the Jacobi polynomials. Using orthogonality relations of the Jacobi polynomials, they obtain a pricing formula for ZCBs in the associated model. However, this formula is only semi-explicit, since it contains multiple integrals that have to be calculated iteratively. We will avoid this problem by representing such integrals in terms of the ${}_1F_1$ hypergeometric functions of a matrix argument as given in Definition 1.1.

Theorem 3.7 (**ZCB price in $\mathcal{E}(p,\lambda)$ model**)**.** *Let $(R_t)_{t\geq 0}$ be given by the definition (3.11) where $\alpha, \beta \in (0,1]$, $p = \frac{\alpha}{\alpha+\beta}$ and $\lambda > 0$. The price at time $t \geq 0$ of a ZCB with maturity at T is given by*

$$P(t,T) = e^{-r_m(T-t)} \cdot P_1(t,T)^k \cdot P_0(t,T)^{N-k}, \tag{3.17}$$

where $k = \frac{R_t - r_m}{h} \in \{0,\ldots,N\}$ and for $m \in \{0,1\}$

$$\begin{aligned} P_m(t,T) \;&=\; 1 + \sum_{n=1}^{\infty}(-1)^n \frac{(h(T-t))^n}{n!} \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} K_m(i_n) \cdot \\ &\quad (-1)^{C_n} p^{n-\left\lfloor \frac{C_n}{2}\right\rfloor}(1-p)^{\left\lfloor \frac{C_n}{2}\right\rfloor} \cdot {}_1F_1\left(1; n+1; z^{(n)}\right). \end{aligned} \tag{3.18}$$

Here, K_m is given according to (b) and (c) of Proposition 1.6, ${}_1F_1$ is defined by (1.2), $z^{(n)} = -\lambda(\alpha+\beta)(T-t)(i_1,\ldots,i_n)^T \in \mathbb{R}^n$, $i_0 = 0$, and

$$C_n = \sum_{j=1}^{n} |i_j - i_{j-1}|.$$

Proof. Let $r_k = hk + r_m$, $k \in \{0,\ldots,N\}$, be the state of $(R_t)_{t\geq 0}$ at time t. Bearing in mind that $(R_t)_{t\geq 0}$ is a Markov process, and using the definitions

(3.3) and (3.11), we get from (3.16)

$$\begin{aligned} P(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \left(h\sum_{l=1}^N Y_s^l + r_m\right) ds\right) \middle| X_t = k\right] \\ &= e^{-r_m(T-t)} \cdot \mathbb{E}^{\mathbb{Q}}\left[\prod_{l=1}^N \exp\left(-\int_t^T hY_s^l\, ds\right) \middle| X_t = k\right] \\ &= e^{-r_m(T-t)} \cdot \mathbb{E}_{1,t}\left[\exp\left(-\int_t^T hY_s\, ds\right)\right]^k \qquad (3.19) \\ &\quad \cdot \mathbb{E}_{0,t}\left[\exp\left(-\int_t^T hY_s\, ds\right)\right]^{N-k}, \end{aligned}$$

where we set $\mathbb{E}_{m,t}[\ .\ ] = \mathbb{E}^{\mathbb{Q}}[\ .\ |Y_t = m]$ for $m \in \{0,1\}$. The equality (3.19) holds because of the independence of $(Y_t^l)_{t\geq 0}$ for all $l = 1,\ldots,N$. In the following we omit writing out the dependence on a particular l, and set

$$P_m(t,T) = \mathbb{E}_{m,t}\left[\exp\left(-\int_t^T hY_s\, ds\right)\right], \quad m \in \{0,1\}. \qquad (3.20)$$

Using (3.20), we rewrite (3.19) as follows:

$$P(t,T) = e^{-r_m(T-t)} \cdot P_1(t,T)^k \cdot P_0(t,T)^{N-k}. \qquad (3.21)$$

From the power series representation of the exponential function, we obtain

$$\begin{aligned} P_m(t,T) &= \mathbb{E}_{m,t}\left[1 + \sum_{n=1}^{\infty}\frac{1}{n!}\left(-\int_t^T hY_s\, ds\right)^n\right] \qquad (3.22) \\ &= 1 + \sum_{n=1}^{\infty}(-1)^n h^n \int_t^T \int_{s_1}^T \ldots \int_{s_{n-1}}^T \mathbb{E}_{m,t}[Y_{s_n}\cdots Y_{s_1}]\, ds_n \ldots ds_2 ds_1, \end{aligned}$$

where $t = s_0 < s_1 < \cdots < s_n < T$. The last equality follows from

$$\left(\int_t^T Y_s\, ds\right)^n = n! \int_t^T \int_{s_1}^T \ldots \int_{s_{n-1}}^T Y_{s_n}\cdots Y_{s_1}\, ds_n \ldots ds_2 ds_1$$

and the dominated convergence theorem.

For the given $t < s_1 < \cdots < s_n < T$ and $t_j = s_j - s_{j-1}$ we have

$$\mathbb{E}_{m,t}[Y_{s_n}\cdots Y_{s_1}] = \sum_{m_1=0}^{1}\cdots\sum_{m_n=0}^{1}\prod_{j=1}^{n} m_j p_{m_{j-1},m_j}(t_j),$$

where $m_0 = m$. Using the symmetry relation (1.8), we write the transition probabilities $p_{m_{j-1},m_j}(t_j)$ given in Theorem 3.5 as follows:

$$p_{m_{j-1},m_j}(t_j) = w(m_j)\sum_{i=0}^{1} \pi_i K_{m_{j-1}}(i) K_{m_j}(i) e^{-\lambda_i t_j},$$

where $\lambda_i = \lambda(\alpha+\beta)i,\ \pi_i = \left(\frac{p}{1-p}\right)^i,\ w(i) = \pi_i(1-p)^i,\ i \in \{0,1\}$. Analogously to the calculation of the expected value (3.10) in the proof of Theorem 3.5, we obtain

$$\mathbb{E}_{m_{n-1},s_{n-1}}[Y_{s_n}] = \sum_{i_1=0}^{1} \pi_{i_1} B_{i_0,i_1} K_{m_{n-1}}(i_1) e^{-\lambda_{i_1} t_n},$$

where B_{i_0,i_1} is defined by (1.13) and $i_0 = 0$. Iteratively, we get

$$\mathbb{E}_{m,t}[Y_{s_n} \cdots Y_{s_1}] = \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} K_m(i_n) \left[\prod_{j=1}^{n} \pi_{i_j} B_{i_{j-1},i_j} e^{-\lambda_{i_j} t_{n-j+1}} \right].$$

Hence, (3.22) becomes

$$\begin{aligned} P_m(t,T) &= 1 + \sum_{n=1}^{\infty} (-1)^n h^n \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} K_m(i_n) \left[\prod_{j=1}^{n} \pi_{i_j} B_{i_{j-1},i_j} \right] \\ &\quad \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T \exp\left(-\sum_{k=1}^{n} \lambda_{i_k}(s_{n-k+1} - s_{n-k}) \right) ds_n \ldots ds_1 \\ &= 1 + \sum_{n=1}^{\infty} (-1)^n h^n \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} K_m(i_n) \left[\prod_{j=1}^{n} \pi_{i_j} B_{i_{j-1},i_j} \right] \qquad (3.23) \\ &\quad \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T \exp\left(-\sum_{k=1}^{n} \lambda_{i_{n-k+1}}(s_k - s_{k-1}) \right) ds_n \ldots ds_1. \end{aligned}$$

In order to evaluate the multiple integrals above, we transform the integration domain to the standard simplex Δ_n defined by (1.3) via the mapping

$$J : \mathbb{R}^n \longrightarrow \mathbb{R}^n,\ \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \longmapsto \begin{pmatrix} (T-t)s_1 + t \\ (T-t)(s_1+s_2) + t \\ \vdots \\ (T-t)(s_1 + \cdots + s_n) + t \end{pmatrix}. \qquad (3.24)$$

Using (3.24), we rewrite (3.23) as follows:

$$P_m(t,T) = 1 + \sum_{n=1}^{\infty}(-1)^n h^n \sum_{i_1=0}^{1}\cdots\sum_{i_n=0}^{1} K_m(i_n)\left[\prod_{j=1}^{n}\pi_{i_j}B_{i_{j-1},i_j}\right] \cdot (T-t)^n \int_{\Delta_n} e^{\langle z^{(n)},x\rangle}\ dx, \tag{3.25}$$

where $z^{(n)} = -(T-t)(\lambda_{i_n},\cdots,\lambda_{i_1})^T \in \mathbb{R}^n$ and $\langle\cdot,\cdot\rangle$ denotes the standard inner product. Applying (1.4) with $a=1$, we express the integrals in (3.25) as ${}_1F_1$ functions as follows:

$$n!\int_{\Delta_n} e^{\langle z^{(n)},x\rangle}\ dx = {}_1F_1\left(1;n+1;z^{(n)}\right) \tag{3.26}$$

for all $z^{(n)} \in \mathbb{R}^n$ ($n \in \mathbb{N}_0$), setting ${}_1F_1(1;1;z^{(0)}) = 1$. From Remark 1.2 we know that ${}_1F_1(.,.,z^{(n)})$ is invariant under the permutations of $z^{(n)}$ and hence we consider in the following $z^{(n)} = -(T-t)(\lambda_{i_1},\cdots,\lambda_{i_n})^T \in \mathbb{R}^n$. If we combine (3.25) and (3.26), we obtain

$$P_m(t,T) = 1 + \sum_{n=1}^{\infty}\frac{(h(t-T))^n}{n!}\cdot\sum_{i_1=0}^{1}\cdots\sum_{i_n=0}^{1}K_m(i_n)\cdot \left[\prod_{j=1}^{n}\left(\frac{p}{q}\right)^{i_j}B_{i_{j-1},i_j}\right]\cdot {}_1F_1\left(1;n+1;z^{(n)}\right), \tag{3.27}$$

where

$$B_{i_{j-1},i_j} = \begin{cases}(p-1) & \text{if} \quad |i_{j-1}-i_j| = 1,\\ i_j(1-2p)+p & \text{if} \quad i_{j-1} = i_j.\end{cases} \tag{3.28}$$

We denote the number of unequal adjacent elements of vector $(i_0,i_1,\ldots,i_n) \in \{0,1\}^{n+1}$ as C_n. Hence,

$$C_n = \sum_{j=1}^{n}|i_j - i_{j-1}|.$$

We also denote the number of ones in $(i_0,i_1,\ldots,i_n)$ by η_n and the number of zeros by ζ_n. Thus, bearing in mind that $i_0 = 0$, the number of adjacent ones in $(i_0,i_1,\ldots,i_n)$ equals $\eta_n - \lceil C_n/2\rceil$ and the number of adjacent zeros equals $\zeta_n - 1 - \lfloor C_n/2\rfloor$, where $\lfloor.\rfloor$ and $\lceil.\rceil$ denote the floor and the ceiling functions. Hence, with $\pi_{i_j} = \left(\frac{p}{1-p}\right)^{i_j}$, $i_j \in \{0,1\}$, and (3.28) we obtain

$$\begin{aligned}\prod_{j=1}^{n}\pi_{i_j}B_{i_{j-1},i_j} &= \left(\frac{p}{1-p}\right)^{\eta_n}(p-1)^{C_n}(1-p)^{\eta_n-\lceil C_n/2\rceil}p^{\zeta_n-1-\lfloor C_n/2\rfloor}\\ &= (-1)^{C_n}(1-p)^{C_n-\lceil C_n/2\rceil}p^{\zeta_n-1-\lfloor C_n/2\rfloor+\eta_n}.\end{aligned}$$

If we use the facts that $\zeta_n = n+1-\eta_n$ and $C_n = \lceil C_n/2 \rceil + \lfloor C_n/2 \rfloor$, $C_n \in \mathbb{N}_0$, we get

$$\prod_{j=1}^{n} \pi_{i_j} B_{i_{j-1},i_j} = (-1)^{C_n} p^{n-\left\lfloor \frac{C_n}{2} \right\rfloor} (1-p)^{\left\lfloor \frac{C_n}{2} \right\rfloor}.$$

If we combine this result with (3.21) and (3.27), the theorem follows. □

Now we consider the $\mathcal{E}(1/2, \lambda)$ model where $\alpha = \beta = 1$. Hence, instead of two variable parameters p and λ, determining the distribution of the Ehrenfest process, we now have only one variable parameter λ. This means that, on the one hand, we lose one of the fitting parameters for the short-rate model, although, the model is still well-suited to model the term structure, and it yields the famous *Vasicek model* in the limit (see Section 2.4.1). On the other hand, we obtain a more tractable pricing formula for ZCBs, where, in contrast to the general case, no multiple sums need calculation, which improves the computational speed.

The calculation of the arbitrage-free ZCB price (3.16) in this setting is very intuitive and requires no knowledge of the transition probabilities of $(R_t)_{t\geq 0}$, since the only stochastic parameters are the arrival times of the underlying Poisson process.

Theorem 3.8 (ZCB price in $\mathcal{E}(1/2, \lambda)$ model). *Let $(R_t)_{t\geq 0}$ be given by the definition (3.11) where $\lambda > 0$ and $\alpha = \beta = 1$. The price at time $t \geq 0$ of a ZCB with maturity at T is given by*

$$P(t,T) = e^{-(r_m + \lambda N)(T-t)} \cdot P_1(t,T)^k \cdot P_0(t,T)^{N-k}, \tag{3.29}$$

where $k = \frac{R_t - r_m}{h} \in \{0, \dots, N\}$,

$$P_1(t,T) = \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \cdot \Big\{ e^{-h(T-t)} \cdot {}_1F_1\left(1; 2n+1; z^{(2n)}\right) + \frac{\lambda(T-t)}{2n+1} \cdot {}_1F_1\left(1; 2n+2; -z^{(2n+1)}\right) \Big\}, \tag{3.30}$$

$$P_0(t,T) = \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \cdot \Big\{ {}_1F_1\left(1; 2n+1; -z^{(2n)}\right) + \frac{\lambda(T-t)}{2n+1} e^{-h(T-t)} \cdot {}_1F_1\left(1; 2n+2; z^{(2n+1)}\right) \Big\}, \tag{3.31}$$

where ${}_1F_1$ is defined by (1.2), $z^{(2n)} = h(T-t)(0,1,\dots,0,1)^T \in \mathbb{R}^{2n}$ and $z^{(2n+1)} = h(T-t)(1,0,1,\dots,0,1)^T \in \mathbb{R}^{2n+1}$.

Proof. Let $r_k = hk + r_m,\ k \in \{0, \dots, N\}$, be the state of $(R_t)_{t\geq 0}$ at time t. Analogously to the derivation of the expression (3.21) in the proof of Theorem 3.7, we obtain

$$P(t,T) = e^{-r_m(T-t)} \cdot \tilde{P}_1(t,T)^k \cdot \tilde{P}_0(t,T)^{N-k}, \tag{3.32}$$

where

$$\tilde{P}_y(t,T) = \mathbb{E}_{y,t}\left[\exp\left(-\int_t^T hY_s\, ds\right)\right], \quad y \in \{0,1\}. \tag{3.33}$$

In order to evaluate $\tilde{P}_y(t,T)$, we count the number of jumps in the underlying Poisson process $(N_t)_{t\geq 0}$ within the time interval $(t,T]$, and, denoting the jump times by $(\tau_i)_{i\in\mathbb{N}}$ and setting $\tau_0 = t$, we split the integral on the right-hand side of (3.33), obtaining

$$\tilde{P}_y(t,T) = \mathbb{E}_{y,t}\left[\sum_{n=0}^{\infty} \mathbb{1}_{\{N_{T-t}=n\}} \cdot \exp\left(-\sum_{i=0}^{n-1}\int_{\tau_i}^{\tau_{i+1}} h\hat{Y}_i\, ds - \int_{\tau_n}^{T} h\hat{Y}_n\, ds\right)\right].$$

At this point, we have to distinguish between even and odd numbers of jumps, since $(\hat{Y}_n)_{n\in\mathbb{N}}$ switches between 0 and 1 $\mathbb{Q}$−a.s. according to its transition probability matrix (3.2). Thus, conditional on $\{\hat{Y}_0 = 1\}$, the Markov chain $(\hat{Y}_n)_{n\in\mathbb{N}}$ stays in 1 after an even number of jumps, whereas it stays in 0 after an odd number of jumps. This consideration yields

$$\begin{aligned}
\tilde{P}_1(t,T) &= \sum_{n=0}^{\infty}\Bigg\{\mathbb{E}\left[\mathbb{1}_{\{N_{T-t}=2n\}} \cdot \exp\left(-\sum_{i=0}^{n-1}\int_{\tau_{2i}}^{\tau_{2i+1}} h\, ds - \int_{\tau_{2n}}^{T} h\, ds\right)\right] \\
&\quad + \mathbb{E}\left[\mathbb{1}_{\{N_{T-t}=2n+1\}} \cdot \exp\left(-\sum_{i=0}^{n}\int_{\tau_{2i}}^{\tau_{2i+1}} h\, ds\right)\right]\Bigg\} \\
&= \sum_{n=0}^{\infty}\Bigg\{\mathbb{E}\left[\exp\left(h\left(\sum_{i=1}^{2n}(-1)^i\tau_i + t - T\right)\right)\Bigg| N_{T-t} = 2n\right] \\
&\qquad \cdot \mathbb{P}\left(N_{T-t} = 2n\right) \\
&\quad + \mathbb{E}\left[\exp\left(h\left(\sum_{i=1}^{2n+1}(-1)^i\tau_i + t\right)\right)\Bigg| N_{T-t} = 2n+1\right] \\
&\qquad \cdot \mathbb{P}\left(N_{T-t} = 2n+1\right)\Bigg\}.
\end{aligned}$$

Furthermore, from the order statistics property of the Poisson process (see, for example, [34], pp. 101-102), we know that the joint density of the arrival times $\tau_1, \dots, \tau_k$ ($k \in \mathbb{N}$) of $(N_t)_{t\geq 0}$ in $(t,T]$, conditional on $\{N_{T-t} = k\}$, is given by

$$\begin{aligned}&\mathbb{P}(t < \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq T | N_{T-t} = k)\\ &\qquad = \frac{n!}{(T-t)^k} \int_t^T \int_{t_1}^T \dots \int_{t_{k-1}}^T \, dt_k \dots dt_2 dt_1.\end{aligned}$$

Hence, for $\tilde{P}_1(t,T)$ we obtain

$$\begin{aligned}\tilde{P}_1(t,T) \;=\; \sum_{n=0}^{\infty} \Bigg\{ &\Bigg[e^{-h(T-t)} \cdot \frac{(2n)!}{(T-t)^{2n}} \\ &\cdot \int_t^T \dots \int_{t_{2n-1}}^T \exp\left(h \cdot \sum_{i=1}^{2n} (-1)^i t_i \right) \, dt_{2n} \dots dt_1 \Bigg] \\ &\cdot e^{-\lambda(T-t)} \frac{(\lambda(T-t))^{2n}}{(2n)!} \\ + &\Bigg[e^{-ht} \cdot \frac{(2n+1)!}{(T-t)^{2n+1}} \\ &\cdot \int_t^T \dots \int_{t_{2n}}^T \exp\left(h \cdot \sum_{i=1}^{2n+1} (-1)^i t_i \right) \, dt_{2n+1} \dots dt_1 \Bigg] \\ &\cdot e^{-\lambda(T-t)} \frac{(\lambda(T-t))^{2n+1}}{(2n+1)!} \Bigg\}.\end{aligned} \tag{3.34}$$

Analogously to the proof of Theorem 3.7, we consider the mapping J given by (3.24) and rewrite (3.34) as follows:

$$\begin{aligned}&\tilde{P}_1^{(N)}(t,T) = e^{-\lambda(T-t)} \sum_{n=0}^{\infty} \frac{(\lambda(T-t))^{2n}}{(2n)!} \cdot \\ &\left\{ e^{-h(T-t)} \cdot (2n)! \int_{\Delta_{2n}} e^{\langle z^{(2n)}, x\rangle} \, dx + \; (2n+1)! \int_{\Delta_{2n+1}} e^{\langle -z^{(2n+1)}, x\rangle} \, dx \right\},\end{aligned} \tag{3.35}$$

where

$$\begin{aligned} z^{(2n)} &= h(T-t)(0,1,\dots,0,1)^T \in \mathbb{R}^{2n}, \\ z^{(2n+1)} &= h(T-t)(1,0,1,\dots,0,1)^T \in \mathbb{R}^{2n+1}, \end{aligned}$$

and $\langle \cdot\,, \cdot \rangle$ denotes the standard inner product. Applying the relation (3.26), we rewrite (3.35) as follows:

$$\begin{aligned}\tilde{P}_1(t,T) &= e^{-\lambda(T-t)}\sum_{n=0}^{\infty}\frac{(\lambda(T-t))^{2n}}{(2n)!}\Big\{e^{-h(T-t)}\cdot {}_1F_1\left(1;2n+1;z^{(2n)}\right)\\ &\qquad + \frac{\lambda(T-t)}{2n+1}\cdot {}_1F_1\left(1;2n+2;-z^{(2n+1)}\right)\Big\}\\ &=: e^{-\lambda(T-t)}P_1(t,T). \qquad (3.36)\end{aligned}$$

In similar fashion, we obtain

$$\begin{aligned}\tilde{P}_0(t,T) &= e^{-\lambda(T-t)}\sum_{n=0}^{\infty}\frac{(\lambda(T-t))^{2n}}{(2n)!}\Big\{{}_1F_1\left(1;2n+1;-z^{(2n)}\right)\\ &\qquad + \frac{\lambda(T-t)}{2n+1}e^{-h(T-t)}\,{}_1F_1\left(1;2n+2;z^{(2n+1)}\right)\Big\}\\ &=: e^{-\lambda(T-t)}P_0(t,T). \qquad (3.37)\end{aligned}$$

If we combine (3.36) and (3.37) with (3.32), the theorem follows. □

Remark 3.9. We notice that the proofs of Theorems 3.7 and 3.8 are based on two completely different approaches and yield different representations of the ZCB prices. However, both formulae involve the confluent hypergeometric function of a matrix argument ${}_1F_1$ defined by (1.2).

3.2.3 Practical implementation

From Theorems 3.7 and 3.8, the ZCB prices can be computed approximately by truncating the series in the respective formulae. We also use the truncated ${}_1F_1^H$ function defined by (1.5) as an approximation for the ${}_1F_1$ function.

Thus, in the setting of Theorem 3.7, we truncate the sum of series (3.18), obtaining

$$P(t,T;M,H) = e^{-r_m(T-t)}\cdot P_1(t,T;M,H)^k\cdot P_0(t,T;M,H)^{N-k}, \qquad (3.38)$$

$$\begin{aligned}P_y(t,T) = 1 + \sum_{n=1}^{\infty}(-1)^n\frac{(h(T-t))^n}{n!}\sum_{i_1=0}^{1}\cdots\sum_{i_n=0}^{1}K_m(i_n)\cdot \qquad (3.39)\\ (-1)^{C_n}p^{n-\left\lfloor\frac{C_n}{2}\right\rfloor}(1-p)^{\left\lfloor\frac{C_n}{2}\right\rfloor}\cdot {}_1F_1\left(1;n+1;z^{(n)}\right).\end{aligned}$$

for $y\in\{0,1\}$.

In the setting of Theorem 3.8, we truncate the series (3.30) and (3.31), obtaining

$$
\begin{aligned}
P(t,T;M,H) &= e^{-(r_m+a)(T-t)} P_1(t,T,M,H)^k P_0(t,T,M,H)^{N-k}, \qquad (3.40)\\
P_1(t,T;M,H) &= \sum_{n=0}^{M} \frac{(\lambda(T-t))^{2n}}{(2n)!}\\
&\cdot\Big\{ e^{-h(T-t)} \cdot {}_1F_1^H \left(1; 2n+1; h(T-t)z^{(2n)}\right)\\
&+ \frac{\lambda(T-t)}{2n+1} \cdot {}_1F_1^H \left(1; 2n+2; -h(T-t)z^{(2n+1)}\right) \Big\},\\
P_0(t,T;M,H) &= \sum_{n=0}^{M} \frac{(\lambda(T-t))^{2n}}{(2n)!} \cdot \Big\{ {}_1F_1^H \left(1; 2n+1; -h(T-t)z^{(2n)}\right)\\
&+ \frac{\lambda(T-t)}{2n+1} e^{-h(T-t)} \cdot {}_1F_1^H \left(1; 2n+2; h(T-t)z^{(2n+1)}\right) \Big\}.
\end{aligned}
$$

The choice of the truncation parameters M and H is left to the practitioner and should be made in the way of maintaining a balance between the accuracy of the results and reasonable computation speed. Some numerical examples, which provide numerical accuracy and computation speed for the formulae (3.38) and (3.40), will be given at the end of the next section.

3.2.4 Connection to the Vasicek model

In this section we provide a convergence result, which shows that after a suitable linear rescaling the $\mathcal{E}(p,\lambda)$ model converges weakly to the Vasicek model introduced in Section 2.4.1. We show the convergence of the respective ZCB prices and provide some numerical examples.

It is well known that the Ehrenfest process converges weakly to the Ornstein-Uhlenbeck (OU) process (see e.g [34], pp. 168-173, or [58]). The following theorem shows that after a suitable adaptation of the model parameters, the $\mathcal{E}(p,\lambda)$ model also converges weakly to the Vasicek model.

Theorem 3.10. *Let $(r_t)_{t\geq 0}$ be given as in (2.20). Consider $(R_t^{(N)})_{t\geq 0}$ as defined in (3.11) where $\alpha,\ \beta \in (0,1],\ \lambda = k/(\alpha+\beta),\ r_m = \theta - \frac{1}{\sqrt{C}} \cdot \sqrt{\frac{Np}{2q}},\ r_M = \theta + \frac{1}{\sqrt{C}} \cdot \sqrt{\frac{Nq}{2p}}$ and $C = k/\sigma^2$ Then,*

$$(R_t^{(N)})_{t\in[0,T]} \Rightarrow (r_t)_{t\in[0,T]} \quad as \quad N \to \infty,$$

where " $\Rightarrow$ " denotes the weak convergence.

Proof. We show the claim for the case of centered processes $(r_t)_{t\geq 0}$ and $(R_t)_{t\geq 0}$. For this purpose, we consider the OU process $(z_t)_{t\geq}$ given by (2.25) and the respective spectral decomposition of the transition probability density (2.27). Recall that the OU process is the Vasicek process $(r_t)_{t\geq 0}$ under the mapping $x \mapsto x - \theta$, i.e. the mean-reverting value of $(r_t)_{t\geq 0}$ is being shifted to 0. In order to centre the Ehrenfest short-rate process, we consider $\tilde{R}_t = R_t - \tilde{\theta},\ t \geq 0$, where

$$\tilde{\theta} = pr_M - qr_m = \theta + p\frac{1}{\sqrt{C}} \cdot \sqrt{\frac{Nq}{2p}} - q\frac{1}{\sqrt{C}} \cdot \sqrt{\frac{Np}{2q}}$$

is the mean-reverting value of $(R_t)_{t\geq 0}$ due to (3.12). Notice that for the symmetric case of $\alpha = \beta$, we have $\tilde{\theta} = \theta$. Summing up some well-known results, we will show that the transition density of the process $(\tilde{R}_t)_{t\geq 0}$ converges for $N \to \infty$ to the transition density (2.27) of $(z_t)_{t\geq 0}$, which yields the general case for an arbitrary mean-reverting value after transforming $(z_t)_{t\geq 0}$ via the mapping $x \mapsto x + \theta$ and $(\tilde{R}_t)_{t\geq 0}$ via the mapping $x \mapsto x + \tilde{\theta}$.

Consider

$$\vartheta(y) = Np + \sqrt{C}y\sqrt{2Npq},\ y \in \left[-\frac{1}{\sqrt{C}}\sqrt{\frac{Np}{2q}}, \frac{1}{\sqrt{C}}\sqrt{\frac{Nq}{2p}}\right].$$

Let $r_1, r_2 \in \{hk+r_m, k = 0, \ldots, N\}$. Thus, with (3.11) we have $X_t = \vartheta(\tilde{R}_t) \in \{0, \ldots, N\}$ for $t \geq 0$. From Lemma 1 in [43], we obtain the transition density $p_{\tilde{R}}(t, r_1, r_2)$ of $(\tilde{R}_t)_{t\geq 0}$ from the transition probability $p_{\vartheta(r_1),\vartheta(r_2)}(t)$ given by (3.7) of the Ehrenfest process $(X_t)_{t\geq 0}$ as follows:

$$p_{\tilde{R}}(t, r_1, r_2) = p_{\vartheta(r_1),\vartheta(r_2)}(t) \cdot \vartheta'(r_2) = \sqrt{C}\sqrt{2Npq} \cdot p_{\vartheta(r_1),\vartheta(r_2)}(t),$$

where $\vartheta'(.)$ is the derivative of $\vartheta(.)$. Hence, with (3.7), $k = \lambda(\alpha+\beta),\ i = \vartheta(r_1)$ and $j = \vartheta(r_2)$, we obtain:

$$p_{\tilde{R}}(t, r_1, r_2) = \sqrt{C}\sqrt{2Npq}\binom{N}{j}p^j q^{N-j} \sum_{x=0}^{N} \binom{N}{x} \left(\frac{p}{q}\right)^x K_i(x)K_j(x)e^{-kxt}$$

Karlin and McGregor showed in a rigorous way (see [35], pp. 369, 371-373) the following result:

$$\sum_{x=0}^{N} \binom{N}{x} \left(\frac{p}{q}\right)^x K_i(x)K_j(x)e^{-kxt} \overset{N\to\infty}{\longrightarrow} \sum_{n=0}^{\infty} \frac{1}{2^n n!} H_n(\sqrt{C}r_1)H_n(\sqrt{C}r_2)e^{-knt} \tag{3.41}$$

uniformly for r_1, r_2 in finite intervals. Furthermore, the *Stirling's approximation* $n! \cong \sqrt{2\pi n} n^n e^{-n}$ yields (see [35], p.359)

$$\sqrt{2Npq}\binom{N}{j}p^j q^{N-j} \overset{N\to\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}}e^{-\sqrt{C}r_2^2}. \tag{3.42}$$

If we use (3.41) and (3.42), we obtain:

$$p_{\tilde{R}}(t, r_1, r_2) \overset{N\to\infty}{\longrightarrow} p(t, r_1, r_2),$$

where $p(t, ., .)$ is the transition density (2.27) of the Ornstein-Uhlenbeck process $(z_t)_{t\geq 0}$, which completes the proof. □

A direct consequence of Theorem 3.10 is the convergence of the respective ZCB prices.

Corollary 3.11. *Consider $(r_t)_{t\in[0,T]}$ and $(R_t^{(N)})_{t\in[0,T]}$ as in Theorem 3.10 and let $P(t,T)$ and $P^{(N)}(t,T)$ denote the associated ZCB prices at time t with maturity at T given by (2.28) and (3.29) respectively. Then,*

$$P^{(N)}(t,T) \overset{N\to\infty}{\longrightarrow} P(t,T).$$

Proof. W.l.o.g. let $t = 0$. We denote by $\mathbb{D} = \mathbb{D}[0,T]$ the space of the real-valued functions on $[0,T]$ that are *right continuous and have left-hand limits (RCLL)*. From [3] (see p. 123), we know that a metric exists that makes $\mathbb{D}$ a *Polish space*, i.e. a metric, separable and complete space. Clearly, $R^{(N)} = (R_t^{(N)})_{t\in[0,T]}$ and $r = (r_t)_{t\in[0,T]}$ both lie in $\mathbb{D}$. With Theorem 3.10, it follows that $R^{(N)} \Rightarrow r$ in $\mathbb{D}$ as N tends to infinity.

Consider the *Volterra operator* $\tilde{S} : \mathbb{D} \to \mathbb{L}^2[0,T]$ defined by

$$(\tilde{S}f)(t) = \int_0^t f(s)\, ds \quad \text{for } f \in \mathbb{D} \text{ and } t \in [0,T].$$

It is well known (see e.g. [12], p. 143) that $\tilde{S}$ is a *compact* operator on $\mathbb{L}^2[0,T] \supseteq \mathbb{D}$. Hence, $\tilde{S}$ is a continuous operator on $\mathbb{D}$ (see [12] for details on compact operators). Then, for $f \in \mathbb{D}$ and $t \in [0,T]$, operator S, defined by

$$(Sf)(t) = \exp\left(-(\tilde{S}R_N)(t)\right) = \exp\left(-\int_0^t f(s)\, ds\right),$$

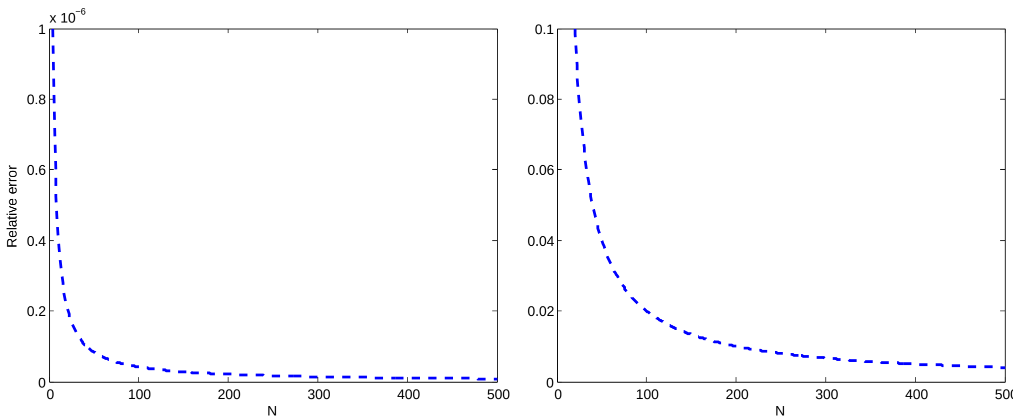
is also a continuous operator on $\mathbb{D}$. Let $Y_N = (SR_t^{(N)})_{t\in[0,T]}$ and $Y = (Sr_t)_{t\in[0,T]}$. Then, Theorem 5.1 in [3] yields $Y_N \Rightarrow Y$. Since Y_N is *uniformly integrable*, it follows from Theorem 5.4 in [3] that

$$\mathbb{E}\left[Y_N\right] \overset{N\to\infty}{\longrightarrow} \mathbb{E}\left[Y\right],$$

which completes the proof. □

In the following, we compare for growing N the approximative ZCB prices $P(0,T;M,H)$ obtained for the Ehrenfest short-rate process $(R_t^{(N)})_{t\in[0,T]}$ with ZCB prices $P(0,T)$ explicitly given in the Vasicek model by (2.28). We compute the ZCB prices $P(0,T;M,H)$ with $M=10$ and $H=30$ according to the convergence Theorem 3.10 and Corollary 3.11. For the computation of the ZCB prices $P(0,T;M,H)$ we can use either formula (3.38) or simplified formula (3.40). If we use the first formula, the computation time of one ZCB price is 2.67 seconds; in the latter case it is less than 0.1 seconds. All computations were made on an INTEL Core2Duo 2400MHz machine. We consider two sets of parameters for the Vasicek model, where $k=0.2$, $\theta=0.08$ and $r_0=0.05$ are fixed for both cases. In the first case (a), we choose volatility $\sigma=0.05$ and a time to maturity $T=1$ year. In the second case (b), we choose an unrealistically high for the interest rate market volatility $\sigma=0.2$ and a time to maturity of 10 years, which is unfavourable for the numerical computation of the ZCB prices.

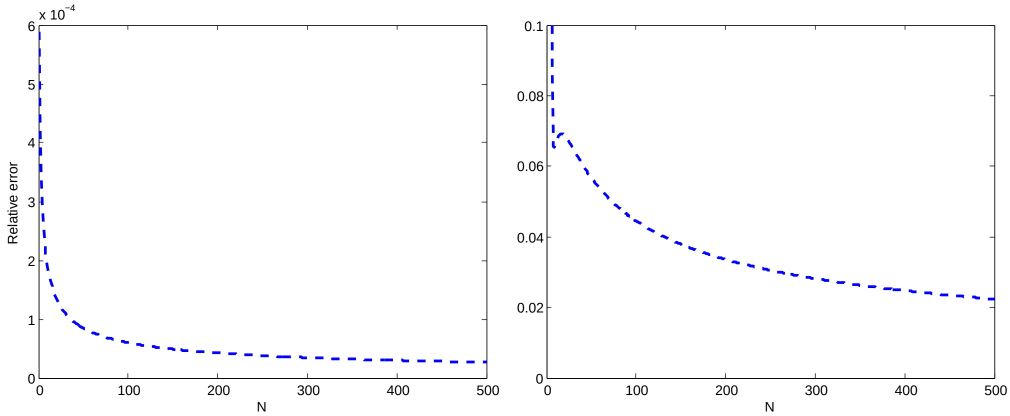
Figure 3.1 illustrates the convergence results for the symmetric case, i.e. the $\mathcal{E}(1/2,\lambda)$ model where $\alpha=\beta=1$ and the simplified formula (3.40). Figure 3.2 demonstrates the convergence results for the general $\mathcal{E}(p,\lambda)$ model where $\alpha=2/3$ and $\beta=1/3$ and the pricing formula (3.38). In case (a) we



(a) Numerically favourable case (b) Numerically unfavourable case

Figure 3.1: Relative price errors against N by an approximation of ZCB prices $P(0,T)$ in the Vasicek model by $P^{(N)}(0,T;10,30)$ in the $\mathcal{E}(1/2,\lambda)$ model.

observe for both formulae a fast convergence of the respective ZCB prices and see that the choice of the truncating parameters $M=10$ and $H=30$ is satisfactory for our purpose. In case (b) we see that at least the simplified formula (3.40) provides satisfactory results, whereas the ZCB prices provided by a convergence of the general formula (3.38) converge relatively slowly.



(a) Numerically favourable case   (b) Numerically unfavourable case

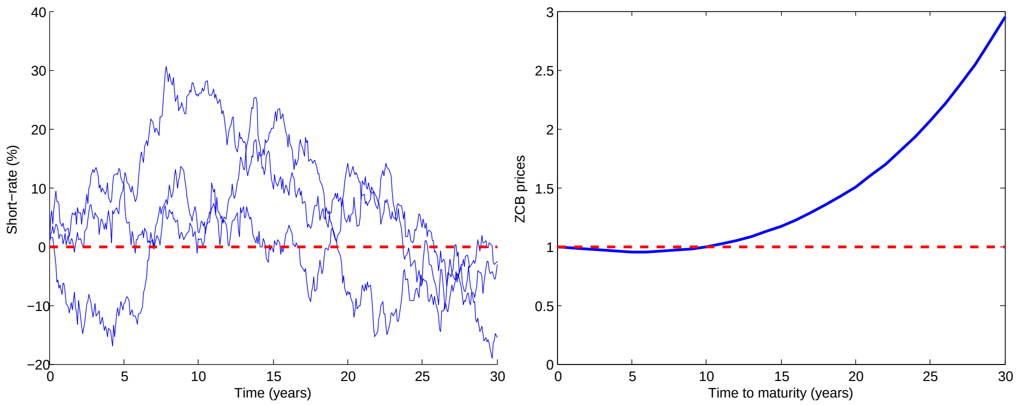
Figure 3.2: Relative price errors against N by an approximation of ZCB prices $P(0,T)$ in the Vasicek model by $P^{(N)}(0,T;10,30)$ in the $\mathcal{E}(p,\lambda)$ model where $\alpha = 2/3$ and $\beta = 1/3$.

3.3 Discussion

In this section we discuss the advantages of the $\mathcal{E}(p,\lambda)$ model with respect to the positivity of the interest rates. We use the case study of a ZCB valuation, showing that the $\mathcal{E}(p,\lambda)$ model can still be used when the Vasicek model reaches its limits.

The main shortcoming of all models with Gaussian distribution, including the Vasicek model, is the positive probability of the interest rates becoming negative. Although this probability is rather small, some problems may appear while valuing ZCBs with long residual maturity. For example, Rogers [54] illustrates how an attempt to keep the probability of negative interest rates negligible by choosing suitable parameters of the Vasicek model in the limiting case $t \to \infty$ leads to an exponential growth in t of the ZCB prices. Conversely, the $\mathcal{E}(p,\lambda)$ model admits the choice of the lower and upper boundaries r_m and r_M for the interest rate, and excludes the possibility of negative as well as unrealistically high positive interest rates.

Times of financial crisis are often followed by interest rates near 0, as we see at present. The following example of pricing ZCBs in a respective scenario illustrates the advantage of the $\mathcal{E}(p,\lambda)$ model over the Vasicek model examined in Section 2.4.1. First, we assume the Vasicek model given according to (2.20) with mean-reverting value $\theta = 0.04$, mean-reversion speed $k = 0.1$, volatility $\sigma = 0.05$ and initial interest rate $r_0 = 0.01$. Figure 3.3 (a) shows three sample paths of the underlying process $(r_t)_{t \geq 0}$ over a period of 30 years simulated according to the *Euler–Maruyama method* (see [19], pp. 7 - 9). We see that every path of the simulated process spends some time below the zero mark. Figure 3.3 (b) demonstrates the weakness of the model in the case at hand, as we observe that the ZCB prices do not monotonically decrease in the time to maturity and even exceed the upper bound of 1 monetary unit, which is contradictory to no-arbitrage principles (see Figure 2.1 and the current market zero-bond curve).



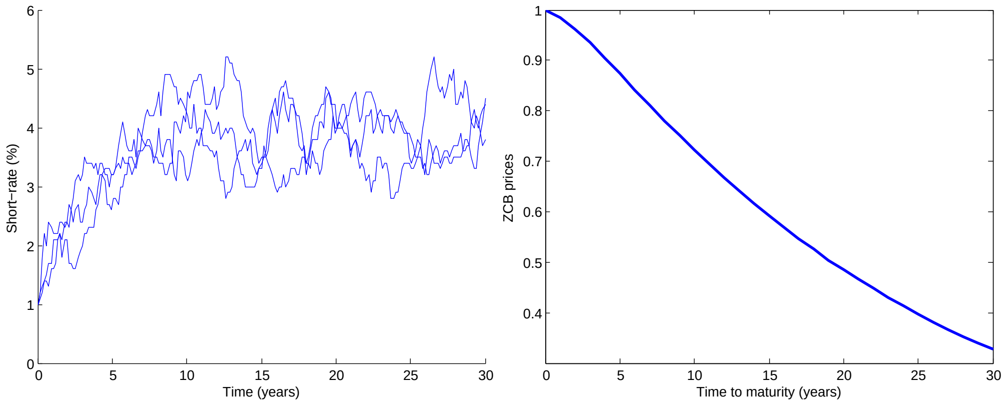
(a) Short-rate

(b) ZCB prices

Figure 3.3: (a) Sample paths of the short-rate process (2.20) in the Vasicek model with $k = 0.1$, $\theta = 0.04$, $\sigma = 0.05$ and $r_0 = 0.01$. (b) ZCB prices in the Vasicek model with the given parameters and residual maturities from 1 to 30 years.

Now we consider the $\mathcal{E}(p, \lambda)$ model in a corresponding hypothetical setting. We set the lower and upper boundaries at $r_m = 0$ and $r_M = 0.16$, and the state space discretization parameter $N = 160$. We choose $\lambda = 1$, $\alpha = 0.1$ and $\beta = 0.3$, in that we have with (3.14) a mean-reverting value of 0.04 as in the case above. Here, we set $R_0 = 0.01$ as well. Figure 3.4 (a) gives three possible trajectories of the short-rate process $(R_t)_{t \geq 0}$ over 30 years, simulated on the basis of the underlying distribution. In Figure 3.4 (b) we see the strictly monotonically decreasing character of the respective ZCB prices as a function of the time to maturity, which is highly plausible.

In summary, we have seen that in some cases, for example, when interest rates are low and/or the volatility of the market is high – as is usually the



(a) Short-rate (b) ZCB prices

Figure 3.4: (a) Sample paths of the short-rate process (3.11) in the $\mathcal{E}(p, \lambda)$ model where $\alpha = 0.1$, $\beta = 0.3$, $\lambda = 1$, $r_m = 0$, $r_M = 0.16$, $N = 160$ and $R_0 = 0.01$. (b) ZCB prices in the $\mathcal{E}(p, \lambda)$ model with the given parameters and residual maturities from 1 to 30 years.

case in times of financial crises – the Vasicek model reaches its limits and fails to provide reasonable ZCB prices due to the possible negativity of the interest rates, which is the main drawback of the model. In contrast, the Ehrenfest short-rate model performs well under similar conditions. Moreover, the existence of explicit pricing formulae for ZCB prices, makes the model an interesting alternative to other short-rate models (see Table 2.1) that provide only positive interest rates.

Chapter 4

Jacobi short-rate model

Diffusion processes with boundaries make up an important part of the theory of stochastic processes. They have also found various applications in different areas of finance. For example, they have been used for modelling exchange rates in a target zone (see [40]), stochastic volatility modelling with a stochastic correlation parameter ([63]), or as an interest rate model with boundaries (see [13]). Other references can be found in [41].

In this chapter we discuss a short-rate model based upon the well-known Jacobi diffusion. This model was first introduced by Delbaen and Shirakawa [13]. The main feature of the model is that it admits lower and upper boundaries for the interest rate, hence preventing negative interest rates, and is mean-reverting. In the following we will take a closer look at the model, derive an explicit pricing formula for ZCBs, show the limiting relations to the CIR [11] and Vasicek [62] models, and provide some numerical results.

4.1 Jacobi diffusion

In this section we examine a diffusion process whose transition probability can be represented as a weighted series of the Jacobi polynomials given in Section 1.3.

Let a time-homogeneous stochastic process $(z_t)_{t\geq 0}$ on the state space $E = [0,1]$ be given by the following stochastic differential equation (SDE):

$$dz_t = k[\gamma - z_t]dt + \sigma\sqrt{z_t(1-z_t)}dW_t, \tag{4.1}$$

where $z_0, k > 0, \sigma > 0$ and $0 < \gamma < 1$ are constants and $(W_t)_{t\geq 0}$ is the standard Brownian motion on the probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ equipped with the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions. We also impose the following stability condition (similar to the stability condition for the CIR process 2.30):

$$\frac{\sigma^2}{2k} \leq \gamma \leq 1 - \frac{\sigma^2}{2k}, \tag{4.2}$$

which ensures that the boundaries $\{0\}$ and $\{1\}$ are inaccessible to the process (4.1). The derivation of this condition is given in the following Proposition and Corollary.

Proposition 4.1. *Let $(z_t)_{t\geq 0}$ be given as in (4.1) starting in $z_0 \in (0,1)$. The hitting probabilities $\rho_{z_0,0}$ and $\rho_{z_0,1}$ of $(z_t)_{t\geq 0}$ hitting the boundaries $\{0\}$ and $\{1\}$ respectively are given – if they exist – as follows:*

$$\rho_{z_0,0} = \lim_{y\to 0, z\to 1} \frac{B_{z_0,z}(p,q)}{B_{y,z}(p,q)} \quad \text{and} \quad \rho_{z_0,1} = \lim_{y\to 0, z\to 1} \frac{B_{y,z_0}(p,q)}{B_{y,z}(p,q)},$$

where $p = 1 - \frac{2k\gamma}{\sigma^2}$, $q = 1 - \frac{2k(1-\gamma)}{\sigma^2}$ and $B_{x,y}(p,q) = \int_x^y z^{p-1}(1-z)^{q-1}$.

Proof. We consider $\tilde{x}, z \in (0,1)$ and $I(\tilde{x}, z)$ given as in (A.5) and make some preliminary calculations

$$\begin{aligned}
I(\tilde{x}, z) &= \int_{\tilde{x}}^{z} \frac{2\mu(y)}{\sigma^2(y)}\,dy = \int_{\tilde{x}}^{z} \frac{2k(\gamma - y)}{\sigma^2 y(1-y)}\,dy \\
&= \frac{2k}{\sigma^2} \int_{\tilde{x}}^{z} \frac{\gamma}{y} + \frac{\gamma - 1}{1-y}\,dy \\
&= \frac{2k\gamma}{\sigma^2} \ln y \Big|_{y=\tilde{x}}^{z} - \frac{2k(\gamma-1)}{\sigma^2} \ln(1-y)\Big|_{y=\tilde{x}}^{z} \\
&= (1-p)\ln\left(\frac{z}{\tilde{x}}\right) + (1-q)\ln\left(\frac{1-z}{1-\tilde{x}}\right) \\
&= \ln\left(\frac{z}{\tilde{x}}\right)^{1-p} + \ln\left(\frac{1-z}{1-\tilde{x}}\right)^{1-q} \\
&= \ln\left(\frac{z^{1-p}(1-z)^{1-q}}{\tilde{x}^{1-p}(1-\tilde{x})^{1-q}}\right).
\end{aligned}$$

Hence,

$$\exp(-I(\tilde{x}, z)) = \left(\frac{z}{\tilde{x}}\right)^{p-1} \left(\frac{1-z}{1-x_0}\right)^{q-1} = K z^{p-1}(1-z)^{q-1},$$

where K is a suitable constant. From Corollary A.11 (b) we obtain the probability of $(z_t)_{t\geq 0}$ hitting $\{0\}$ as follows:

$$\rho_{z_0,0} = \frac{s(1)-s(z_0)}{s(1)-s(0)} = \frac{\int_{z_0}^1 \exp(-I(\tilde{x},z))\,dz}{\int_0^1 \exp(-I(\tilde{x},z))\,dz}.$$

Using the previous calculations we obtain

$$\rho_{z_0,0} = \frac{\int_{z_0}^1 Kz^{p-1}(1-z)^{q-1}\,dz}{\int_0^1 Kz^{p-1}(1-z)^{q-1}\,dz} = \frac{\int_{z_0}^1 z^{p-1}(1-z)^{q-1}\,dz}{\int_0^1 z^{p-1}(1-z)^{q-1}\,dz}.$$

In a similar fashion we obtain from Corollary A.11 (a) the probability of $(z_t)_{t\geq 0}$ hitting $\{1\}$:

$$\rho_{z_0,1} = \frac{\int_0^{z_0} z^{p-1}(1-z)^{q-1}\,dz}{\int_0^1 z^{p-1}(1-z)^{q-1}\,dz}.$$

□

Corollary 4.2. *The process $(z_t)_{t\geq 0}$ defined by (4.1) starting in $z_0 \in (0,1)$ will not access the boundaries $\{0\}$ and $\{1\}$ if*

$$\frac{\sigma^2}{2k} \leq \gamma \leq 1 - \frac{\sigma^2}{2k}. \tag{4.3}$$

Proof. With the notations of Proposition 4.1 we have for $z_0 \in (0,1)$

$$\rho_{z_0,0} = 0 \quad \text{if} \quad p \leq 0 \quad \text{and}$$
$$\rho_{z_0,1} = 0 \quad \text{if} \quad q \leq 0,$$

which is equivalent to (4.3). □

Proposition 4.3. *The stochastic process $(z_t)_{t\geq 0}$ given by (4.1) is a diffusion which we denote as the Jacobi diffusion.*

Proof. First, we have to show that $(z_t)_{t\geq 0}$ is *pathwise unique*. Therefore, we prove that the coefficients of (4.1) satisfy the following conditions of the Yamata-Watanabe theorem (see [53], Theorem 40.1):

(i) there exists increasing $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\lim_{s\to\infty} \int_{0+}^{s} \rho(u)^{-1}\,du = \infty,$$

and for all $x, y \in \mathbb{R}$,

$$(\sigma(x)-\sigma(y))^2 \leq \rho(|x-y|),$$

(ii) $\mu(.)$ is Lipschitz.

Ad (i): Consider the modified process $(\tilde{z}_t)_{t\geq 0}$ which coincides with $(z_t)_{t\geq 0}$ on $[0,1]$, and where the drift and diffusion coefficients of the process are set to equal 0 outside $[0,1]$. Consider also the function $\rho(x)=\sqrt{x}$. Then,

$$\lim_{s\to\infty}\int_{0+}^{s}\rho(u)^{-1}\,du=\lim_{s\to\infty}\sqrt{s}=\infty,$$

and the second inequality holds with Theorem 11 of [44]:

$$\sqrt{|x|}-\sqrt{|y|}\leq\sqrt{|x-y|},\quad x,y\in\mathbb{R}.$$

Ad (ii): Clear with Lipschitz constant k.
Since the coefficients of the modified process $(\tilde{z}_t)_{t\geq 0}$ are continuous, bounded functions, we know from Theorem 23.5 of [53] and the uniqueness that there is a *(weak) solution* to $(z_t)_{t\geq 0}$. Hence, from Theorem 21.1 of [53] we conclude that $(z_t)_{t\geq 0}$ is a *strong Markov process*, and therefore a diffusion. □

In order to use the representation of the probability transition density (A.10) we do in the following some preliminary calculations. From Proposition A.6 we obtain the infinitesimal generator $\mathcal{A}$ of (4.1)

$$(\mathcal{A}f)(z)=k(\gamma-z)f'(z)+\frac{1}{2}\sigma^2 z(1-z)f''(z),\quad f\in\mathcal{D}_{\mathcal{A}}.\tag{4.4}$$

Lemma 4.4. *Let $\mathcal{A}$ be the infinitesimal operator as given in (4.4). Its associated eigenfunctions are the modified Jacobi polynomials $\big(\varphi_n=J_n^{(\alpha,\beta)}\big)_{n\in\mathbb{N}_0}$ as given in Definition 1.9, where*

$$\alpha=\frac{2k\gamma}{\sigma^2}-1,\quad\beta=\frac{2k(1-\gamma)}{\sigma^2}-1,$$

and the corresponding eigenvalues

$$\lambda_n=-kn-\frac{\sigma^2}{2}n(n-1),\quad n\in\mathbb{N}_0.$$

Proof. Let φ_n and λ_n $(n\in\mathbb{N}_0)$ be the eigenfunctions and the corresponding eigenvalues of $\mathcal{A}$. Then, from the definition (A.7) and the representation (4.4) we obtain

$$k(\gamma-z)\varphi_n'(z)+\frac{1}{2}\sigma^2 z(1-z)f''(z)-\lambda_n\varphi_n=0,$$

which is equivalent to the differential equation (1.25) for the modified Jacobi polynomials, where α, β and λ_n are obtained by a comparison of coefficients as follows:

$$\begin{aligned} \frac{2k}{\sigma^2}\gamma = \alpha + 1 \quad &\Leftrightarrow \quad \alpha = \frac{2k\gamma}{\sigma^2} - 1, \\ \frac{2k}{\sigma^2} = \alpha + \beta + 2 \quad &\Leftrightarrow \quad \beta = \frac{2k(1-\gamma)}{\sigma^2} - 1, \\ -\frac{2}{\sigma^2}\lambda_n = n(n+\alpha+\beta+1) \quad &\Leftrightarrow \quad \lambda_n = -kn - \frac{\sigma^2}{2}n(n-1). \end{aligned}$$

□

Theorem 4.5 (Properties of the Jacobi diffusion). *Let $(z_t)_{t\geq 0}$ be the Jacobi diffusion given by (4.1).*

(*a*) *The density function of its transition probability is given by*

$$p(t; x, y) = \sum_{n=0}^{\infty} h_n^{-1} J_n^{(\alpha,\beta)}(x) J_n^{(\alpha,\beta)}(y)\pi(y)e^{\lambda_n t}, \quad x, y \in [0, 1], \qquad (4.5)$$

where

$$\begin{aligned} \alpha &= \frac{2k\gamma}{\sigma^2} - 1 > -1, \\ \beta &= \frac{2k(1-\gamma)}{\sigma^2} - 1 > -1, \\ \lambda_n &= -kn - \frac{\sigma^2}{2}n(n-1) < 0, \\ h_n^{-1} &= \frac{n!(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}, \\ J_n^{(\alpha,\beta)}(y) &= \frac{\Gamma(n+\alpha+1)}{n!\Gamma(n+\alpha+\beta+1)} \\ &\quad \cdot \sum_{k=0}^{n} \binom{n}{k} \frac{\Gamma(n+\alpha+\beta+k+1)}{\Gamma(\alpha+k+1)}(-1)^k x^k, \\ \pi(y) &= y^\alpha(1-y)^\beta. \end{aligned}$$

(*b*) *The conditional mean and variance of $(z_t)_{t\geq 0}$ are given by*

$$\mathbb{E}[z_t|z_0] = \gamma + \frac{\sigma^2}{2k} J_1^{(\alpha,\beta)}(z_0) \cdot e^{-kt}, \qquad (4.6)$$

$$\begin{aligned} \mathbb{V}ar[z_t|z_0] &= \frac{\sigma^2\gamma(1-\gamma)}{\sigma^2+2k} + V_1 J_1^{(\alpha,\beta)}(z_0) \cdot e^{-kt} \qquad (4.7) \\ &\quad + V_2\left(J_1^{(\alpha,\beta)}(z_0) \cdot e^{-kt}\right)^2 + V_3 J_2^{(\alpha,\beta)}(z_0) \cdot e^{-(\sigma^2+2k)t}, \end{aligned}$$

where

$$\begin{aligned} V_1 &= \frac{(-\gamma\sigma^2+\sigma^2+k\gamma)\sigma^2}{2k(\sigma^2+k)} - \frac{\gamma\sigma^2}{2k} + \frac{4k\gamma^2(1-\gamma)}{\sigma^2+2k}, \\ V_2 &= -\frac{4k^2\gamma^2(1-\gamma)^2}{(\sigma^2+2k)^2}, \\ V_3 &= -\frac{\sigma^4}{(\sigma^2+k)(\sigma^2+2k)}. \end{aligned}$$

(c) *The stationary distribution is the limiting distribution, and is given by the Beta distribution on* $[0,1]$ *with parameters* $\alpha+1$ *and* $\beta+1$. *The rate of convergence to the limiting distribution is exponential of order* e^{-kt}. *More precisely,*

$$p(t;x,y) = \frac{y^\alpha(1-y)^\beta}{B(\alpha+1,\beta+1)} + \frac{(\alpha+\beta+2)y^\alpha(1-y)^\beta}{B(\alpha+2,\beta+2)} \cdot$$
$$\cdot \left(\frac{\alpha+1}{\alpha+\beta+2} - x\right)\left(\frac{\alpha+1}{\alpha+\beta+2} - y\right) e^{-kt} + O\left(e^{-(2k+\sigma^2)t}\right),$$

where B *denotes the Beta function given as follows*

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad \text{for } a,b>0.$$

Proof. Ad (a) : Since $(z_t)_{t\geq 0}$ is a diffusion process, the density function can be represented in the form (A.10), where the calculation of π is straightforward, since with (A.4) and (A.5) we have for arbitrary $x, x_0 \in E$

$$\begin{aligned} I(x_0,x) &= \int_{x_0}^{x} \frac{2\mu(y)}{\sigma^2(y)}\,dy = \int_{x_0}^{x} \frac{2k(\gamma-y)}{\sigma^2 y(1-y)}\,dy \\ &= \frac{2k}{\sigma^2}\int_{x_0}^{x} \frac{\gamma}{y} + \frac{\gamma-1}{1-y}\,dy \\ &= \frac{2k\gamma}{\sigma^2}\ln y\Big|_{y=x_0}^{x} - \frac{2k(\gamma-1)}{\sigma^2}\ln(1-y)\Big|_{y=x_0}^{x} \\ &= \ln\left(\frac{x}{x_0}\right)^{\frac{2k\gamma}{\sigma^2}} + \ln\left(\frac{1-x}{1-x_0}\right)^{\frac{2k(1-\gamma)}{\sigma^2}} \end{aligned}$$

and hence, with $\alpha = \frac{2k\gamma}{\sigma^2} - 1$ and $\beta = \frac{2k(1-\gamma)}{\sigma^2} - 1$ we obtain

$$\begin{aligned} \pi(x) &= \frac{2K}{\sigma^2(x)}\exp(I(x_0,x)) = \frac{2K}{\sigma^2 x(1-x)} \cdot \frac{x^{\alpha+1}(1-x)^{\beta+1}}{x_0^{\alpha+1}(1-x_0)^{\beta+1}} \\ &= x^\alpha(1-x)^\beta, \end{aligned}$$

where K is a suitable constant. We notice that π is the weight function of the modified Jacobi polynomials given in Definition 1.9. If we use Lemma 4.4, the claim follows.
Ad (b) : If we use the representation (4.5) of the density function and Proposition 1.10 (e), we obtain

$$\begin{aligned}\mathbb{E}[z_t|z_0] &= \int_0^1 z \cdot p(t; z_0, z)\; dz \\ &= \sum_{n=0}^{\infty} h_n^{-1} J_n^{(\alpha,\beta)}(z_0) e^{\lambda_n t} \int_0^1 z \cdot J_n^{(\alpha,\beta)}(z)\pi(z)\; dz \\ &= \sum_{n=0}^{\infty} h_n^{-1} B_{0,n} J_n^{(\alpha,\beta)}(z_0) e^{\lambda_n t},\end{aligned}$$

where $B_{0,n} \neq 0$ for $n \in \{0, 1\}$ as given in (1.27), which implies (4.6).
After a similar elementary but tedious calculation of $\mathbb{E}[z_t^2|z_0]$ and the application of the identity $\mathbb{V}ar[z_t|z_0] = \mathbb{E}[z_t^2|z_0] - \mathbb{E}[z_t|z_0]^2$, we obtain the asserted formula (4.7) for the conditional variance.
Ad (c) : We consider the density function (4.5) of $(z_t)_{t\geq 0}$ and use the calculation (1.22) to obtain

$$\begin{aligned}p(t; x, y) &= \frac{(\alpha+\beta+1)\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta+1)} J_0^{(\alpha,\beta)}(x) J_0^{(\alpha,\beta)}(y) y^\alpha (1-y)^\beta \\ &\quad + \frac{(2+\alpha+\beta+2)\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+2)\Gamma(\beta+2)} J_1^{(\alpha,\beta)}(x) J_1^{(\alpha,\beta)}(y) y^\alpha (1-y)^\beta e^{-kt} \\ &\quad + \sum_{n=2}^{\infty} h_n^{-1} J_n^{(\alpha,\beta)}(x) J_n^{(\alpha,\beta)}(y)\pi(y) e^{\left(-kn - \frac{\sigma^2}{2} n(n-1)\right)t} \\ &= \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} y^\alpha (1-y)^\beta + \frac{\Gamma(\alpha+\beta+4)}{(\alpha+\beta+2)\Gamma(\alpha+2)\Gamma(\beta+2)} \cdot \\ &\quad \cdot (\alpha+1-(\alpha+\beta+2)x)(\alpha+1-(\alpha+\beta+2)y) y^\alpha (1-y)^\beta e^{-kt} \\ &\quad + O\left(e^{-(2k+\sigma^2)t}\right) \\ &= \frac{y^\alpha (1-y)^\beta}{B(\alpha+1, \beta+1)} + \frac{y^\alpha (1-y)^\beta}{(\alpha+\beta+2) B(\alpha+2, \beta+2)} \cdot \\ &\quad \cdot (\alpha+\beta+2)^2 \left(\frac{\alpha+1}{\alpha+\beta+2} - x\right)\left(\frac{\alpha+1}{\alpha+\beta+2} - x\right) e^{-kt} \\ &\quad + O\left(e^{-(2k+\sigma^2)t}\right).\end{aligned}$$

Thus, the limiting distribution, which is always a stationary distribution as well, is obtained by

$$\lim_{t\to\infty} p(t;x,y) = \frac{y^\alpha(1-y)^\beta}{B(\alpha+1,\beta+1)},$$

which is the density function of the Beta distribution on $[0,1]$ with the parameters $\alpha+1$ and $\beta+1$. □

4.2 Jacobi short-rate model

In this section we examine a mean-reverting short-rate model associated with the Jacobi diffusion given in the previous section. Here, we follow the idea of Delbaen and Shirakawa [13], who proposed a short-rate model with upper and lower bounds. We will extend their results concerning the pricing of ZCBs and show that under suitable transformations the model converges to the famous Vasicek model [62] and to the CIR model [11]. We provide some numerical results showing the accuracy of the ZCB pricing formula.

4.2.1 Definition and properties

Let $(r_t)_{t\geq 0}$ be the stochastic time-homogeneous process given by the SDE

$$dr_t = k[\theta - r_t]dt + \sigma\sqrt{(r_t - r_m)(r_M - r_t)}dW_t, \tag{4.8}$$

where k, θ, σ are positive constants and $r_m < r_0 < r_M$ and $(W_t)_{t\geq 0}$ is the standard Brownian motion on the probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ equipped with the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual conditions.

We consider the transformed process

$$z_t = \frac{r_t - r_m}{r_M - r_m}, \quad t \geq 0. \tag{4.9}$$

Hence, the state space of the process $(z_t)_{t\geq 0}$ is $[0,1]$. The inverse transformation is given by

$$r_t = z_t(r_M - r_m) + r_m, \quad t \geq 0. \tag{4.10}$$

From Itô's lemma we obtain, with $\gamma = \frac{\theta - r_m}{r_M - r_m}$ for $t \geq 0$,

$$\begin{aligned}
dz_t &= \frac{1}{r_M - r_m}k[\theta - r_t]dt + \frac{1}{r_M - r_m}\sigma\sqrt{(r_t - r_m)(r_M - r_t)}dW_t \\
&= \frac{k}{r_M - r_m}[\theta - z_t(r_M - r_m) - r_m]dt \\
&\quad + \frac{\sigma}{r_M - r_m}\sqrt{(z_t(r_M - r_m))(r_M - z_t(r_M - r_m) - r_m)}dW_t \\
&= k[\gamma - z_t]dt + \sigma\sqrt{z_t(1 - z_t)}dW_t, \qquad (4.11)
\end{aligned}$$

which is the Jacobi diffusion examined in Section 4.1. We denote the short-rate model (4.8) as the *Jacobi short-rate model*.

From Theorem 4.5 and the identity (4.9) we immediately obtain the conditional mean and variance of r_t, conditional on $r_0 = z_0(r_M - r_m) + r_m \in [r_m, r_M]$ with a suitable $z_0 \in [0,1]$, as follows:

$$\begin{aligned}
\mathbb{E}[r_t|r_0] &= (r_M - r_m)\cdot \mathbb{E}[z_t|z_0] + r_m \\
&= \theta + \frac{\sigma^2(r_M - r_m)}{2k} J_1^{(\alpha,\beta)}(z_0)\cdot e^{-kt}, \\
\mathbb{V}ar[r_t|r_0] &= (r_M - r_m)^2 \cdot \mathbb{V}ar[z_t|z_0] = \frac{\sigma^2(r_M-\theta)(\theta - r_m)}{\sigma^2 + 2k} \\
&\quad + (r_M - r_m)^2 \Big(V_1 J_1^{(\alpha,\beta)}(z_0)\cdot e^{-kt} + V_2 \big(J_1^{(\alpha,\beta)}(z_0)\cdot e^{-kt}\big)^2 \\
&\quad + V_3 J_2^{(\alpha,\beta)}(z_0)\cdot e^{-(\sigma^2+2k)t} \Big),
\end{aligned}$$

where V_1, V_2 and V_3 are given as in Theorem 4.5. We also obtain the *mean reversion* of $(r_t)_{t\geq 0}$:

$$\lim_{t\to\infty} \mathbb{E}[r_t|r_0] = \theta, \tag{4.12}$$

$$\lim_{t\to\infty} \mathbb{V}ar[r_t|r_0] = \frac{\sigma^2(r_M-\theta)(\theta - r_m)}{\sigma^2 + 2k} < \infty. \tag{4.13}$$

Overall, we have five parameters, r_m, r_M, k, θ and σ for the calibration of the model to the market data. Here, θ is the mean-reverting value, σ and $r_M - r_m$ govern the volatility of the short-rate process, and k has an impact on the speed of reversion to θ.

Figure 4.1 shows three sample paths of the Jacobi short-rate process $(r_t)_{t\geq 0}$ over a period of 30 years simulated according to the *Euler–Maruyama method* (see [19], pp. 7 - 9). The parameters of the model are chosen as follows: mean-reverting value $\theta = 0.04$, mean-revertion speed $k = 0.1$, volatility $\sigma = 0.05$, initial interest rate $r_0 = 0.01$, and lower and upper boundaries $r_m = 0$ and $r_M = 0.1$.

4.2.2 Zero-coupon bond

The following theorem gives the arbitrage-free price of a ZCB in the Jacobi short-rate model, slightly extending the results of Delbaen and Shirakawa [13] by expressing certain multiple integrals in terms of the ${}_1F_1$ hypergeometric functions of a matrix argument given in Section 1.

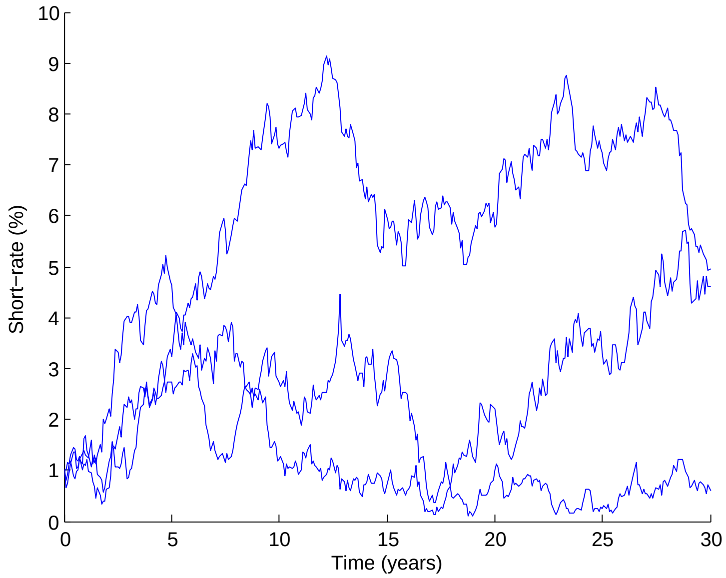


Figure 4.1: Sample paths of the Jacobi short-rate process (4.8) with $k = 0.1$, $\theta = 0.04$, $\sigma = 0.2$, $r_m = 0$, $r_M = 0.1$ and $r_0 = 0.01$.

Theorem 4.6 (**ZCB price in the Jacobi short-rate model**). *Let $(r_t)_{t\geq 0}$ be the Jacobi short-rate process defined by (4.8) with the state $r_t = r$ at time $t \geq 0$. The price at time $t \geq 0$ of a ZCB with maturity T is given by*

$$P(t,T) = e^{-r_m(T-t)} \cdot \Bigg[1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \cdot (r_M - r_m)^n (T-t)^n \cdot \sum_{(i_1,\ldots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \cdot \prod_{j=1}^{n} k_{i_{j-1},i_j} \cdot {}_1F_1\left(1; n+1; v^{(n)}\right)\Bigg], \qquad (4.14)$$

where ${}_1F_1$ is defined by (1.2), $v^{(n)} = (T-t)(\lambda_{i_1}, \ldots, \lambda_{i_n})^T \in \mathbb{R}^n$, $J_n^{(\alpha,\beta)}, \alpha, \beta, h_n$, and λ_n are given due to Theorem 4.5, $z = \frac{r-r_m}{r_M-r_m}$,

$$I^n = \{(i_1, \ldots, i_n) \in \mathbb{N}_0 \,:\, |i_j - i_{j-1}| \leq 1,\ 1 \leq j \leq n,\ i_0 = 0\},$$

and

$$k_{m,n} = \begin{cases} -\frac{n(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)} & \text{if} \quad m = n-1, \\ \frac{\alpha^2-\beta^2}{2(2n+\alpha+\beta)(2n+\alpha+\beta+2)} + \frac{1}{2} & \text{if} \quad m = n, \\ -\frac{(n+\alpha+1)(n+\beta+1)}{(2n+\alpha+\beta+2)(2n+\alpha+\beta+3)} & \text{if} \quad m = n+1, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $r \in (r_m, r_M)$ be the state of $(r_t)_{t\geq 0}$ at time t. From (3.16), (4.10) and the Markov property of the underlying process we obtain

$$\begin{aligned} P(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_u\,du\right)\Big|\, r_t = r\right] \\ &= \mathbb{E}_{z,t}\left[\exp\left(-r_m(T-t) - (r_M - r_m)\cdot\int_t^T z_u\,du\right)\right], \end{aligned}$$

where we set $\mathbb{E}_{z,t}[\,\cdot\,] = \mathbb{E}^{\mathbb{Q}}[\,\cdot\,|z_t = z]$ and $z = \frac{r-r_m}{r_M-r_m}$. From the power series representation of the exponential function, we obtain

$$\begin{aligned} P(t,T) &= e^{-r_m(T-t)}\cdot\mathbb{E}_{z,t}\left[\exp\left(-(r_M-r_m)\cdot\int_t^T z_u\,du\right)\right] \\ &= e^{-r_m(T-t)}\cdot\mathbb{E}_{z,t}\left[\sum_{n=0}^{\infty}\frac{(-1)^n(r_M-r_m)^n}{n!}\left(\int_t^T z_u\,du\right)^n\right] \\ &= e^{-r_m(T-t)}\cdot\left[1+\sum_{n=1}^{\infty}\frac{(-1)^n(r_M-r_m)^n}{n!}\cdot\mathbb{E}_{z,t}\left[\left(\int_t^T z_u\,du\right)^n\right]\right] \\ &= e^{-r_m(T-t)}\cdot\Bigg[1+\sum_{n=1}^{\infty}(-1)^n(r_M-r_m)^n\cdot \qquad (4.15) \\ &\qquad \int_t^T\int_{s_1}^T\cdots\int_{s_{n-1}}^T \mathbb{E}_{z,t}[z_{s_n}\cdots z_{s_1}]\,ds_n\ldots ds_2ds_1\Bigg], \end{aligned}$$

where $t = s_0 < s_1 < \cdots < s_n < T$. The equality (4.15) follows from

$$\left(\int_t^T z_u\,du\right)^n = n!\int_t^T\int_{s_1}^T\cdots\int_{s_{n-1}}^T z_{s_n}\cdots z_{s_1}\,ds_n\ldots ds_2ds_1$$

and the dominated convergence theorem.

For the given $t < s_1 < \cdots < s_n < T$ and $t_j = s_j - s_{j-1}$ we have

$$\mathbb{E}_{z,t}[z_{s_n}\cdots z_{s_1}] = \int_0^1\cdots\int_0^1\prod_{j=0}^{n-1} z_j p(t_j; z_{j-1}, z_j)\,dz_1\ldots dz_n,$$

where $z_0 = z$. From Theorem 4.5 we know the density of the transition probability of $(z_t)_{t\geq 0}$

$$p(t_j; z_{j-1}, z_j) = \sum_{i=0}^{\infty} h_i^{-1} J_i^{(\alpha,\beta)}(z_{j-1}) J_i^{(\alpha,\beta)}(z_j)\pi(z_j)e^{\lambda_i t_j}.$$

Hence, we calculate the expected value $\mathbb{E}_{z_{n-1},s_{n-1}}[z_{s_n}]$ as follows:

$$\begin{aligned}
\mathbb{E}_{z_{n-1},s_{n-1}}[z_{s_n}] &= \int_0^1 z_n \cdot p(t_n; z_{n-1}, z_n)\, dz_n \\
&= \int_0^1 z_n \cdot \sum_{i_1=0}^{\infty} h_{i_1}^{-1} J_{i_1}^{(\alpha,\beta)}(z_{n-1}) J_{i_1}^{(\alpha,\beta)}(z_n)\pi(z_n) e^{\lambda_{i_1} t_n}\, dz_n \\
&= \sum_{i_1=0}^{\infty} h_{i_1}^{-1} J_{i_1}^{(\alpha,\beta)}(z_{n-1}) e^{\lambda_{i_1} t_n} \underbrace{\int_0^1 z_n \cdot J_{i_1}^{(\alpha,\beta)}(z_n)\pi(z_n)\, dz_n}_{B_{i_0,i_1}} \\
&= \sum_{i_1=0}^{1} h_{i_1}^{-1} J_{i_1}^{(\alpha,\beta)}(z_{n-1}) B_{i_0,i_1} e^{\lambda_{i_1} t_n},
\end{aligned}$$

where $B_{n,m}$ $(n, m \in \mathbb{N})$ is given as in Proposition 1.10 (e) and $i_0 = 0$. We calculate the following expected value in a similar fashion:

$$\begin{aligned}
&\mathbb{E}_{z_{n-2},s_{n-2}}[z_{s_n} z_{s_{n-1}}] \\
&= \int_0^1 z_{n-1} \left(\int_0^1 z_n \cdot p(t_n; z_{n-1}, z_n)\, dz_n \right) p(t_{n-1}; z_{n-2}, z_{n-1})\, dz_{n-1} \\
&= \int_0^1 z_{n-1} \left(\sum_{i_1=0}^{1} h_{i_1}^{-1} J_{i_1}^{(\alpha,\beta)}(z_{n-1}) B_{i_0,i_1} e^{\lambda_{i_1} t_n} \right) \cdot \\
&\qquad \left(\sum_{i_2=0}^{\infty} h_{i_2}^{-1} J_{i_2}^{(\alpha,\beta)}(z_{n-2}) J_{i_2}^{(\alpha,\beta)}(z_{n-1}) e^{\lambda_{i_2} t_{n-1}} \pi(z_{n-1}) \right) dz_{n-1} \\
&= \sum_{i_1=0}^{1} \sum_{i_2=0}^{\infty} h_{i_1}^{-1} h_{i_2}^{-1} J_{i_2}^{(\alpha,\beta)}(z_{n-2}) e^{\lambda_{i_1} t_n} e^{\lambda_{i_2} t_{n-1}} B_{i_0,i_1} \cdot \\
&\qquad \underbrace{\int_0^1 z_{n-1} \cdot J_{i_1}^{(\alpha,\beta)}(z_{n-1}) J_{i_2}^{(\alpha,\beta)}(z_{n-1}) \pi(z_{n-1})\, dz_{n-1}}_{B_{i_1,i_2}} \\
&= \sum_{i_1=0}^{1} \sum_{i_2=i_1-1}^{i_1+1} h_{i_1}^{-1} h_{i_2}^{-1} J_{i_2}^{(\alpha,\beta)}(z_{n-2}) B_{i_0,i_1} B_{i_1,i_2} e^{\lambda_{i_1} t_n} e^{\lambda_{i_2} t_{n-1}}.
\end{aligned}$$

Iteratively, we obtain

$$\begin{aligned}
\mathbb{E}_{z,t}[z_{s_1} \dots z_{s_n}] &= \sum_{i_1=0}^{1} \sum_{i_2=i_1-1}^{i_1+1} \cdots \sum_{i_n=i_{n-1}-1}^{i_{n-1}+1} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j} e^{\lambda_{i_j} t_{n-j+1}} \\
&= \sum_{(i_1,\dots,i_n)\in I^n} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j} e^{\lambda_{i_j} t_{n-j+1}},
\end{aligned}$$

where

$$I^n = \{(i_1, \dots, i_n) \in \mathbb{N}_0 \,:\, |i_j - i_{j-1}| \le 1,\, 1 \le j \le n,\, i_0 = 0\}.$$

Hence, the expression (4.15) for the ZCB becomes

$$\begin{aligned} P(t,T) &= e^{-r_m(T-t)} \cdot \Bigg[1 + \sum_{n=1}^{\infty}(-1)^n (r_M - r_m)^n \cdot \\ &\quad \sum_{(i_1,\dots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j} \cdot \\ &\quad \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T \prod_{j=1}^{n} \exp\left(\lambda_{i_j}(s_{n-j+1} - s_{n-j})\right)\, ds_n \dots ds_2 ds_1\Bigg] \\ &= e^{-r_m(T-t)} \cdot \Bigg[1 + \sum_{n=1}^{\infty}(-1)^n (r_M - r_m)^n \cdot \\ &\quad \sum_{(i_1,\dots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j} \cdot \\ &\quad \int_t^T \int_{s_1}^T \cdots \int_{s_{n-1}}^T \prod_{k=1}^{n} \exp\left(\lambda_{i_{n-k+1}}(s_k - s_{k-1})\right)\, ds_n \dots ds_2 ds_1\Bigg]. \end{aligned}$$

Analogously to the proof of Theorem 3.7, we consider the mapping J given in (3.24), which yields for the above equation

$$\begin{aligned} P(t,T) &= e^{-r_m(T-t)} \cdot \Bigg[1 + \sum_{n=1}^{\infty}(-1)^n (r_M - r_m)^n \cdot \\ &\quad \sum_{(i_1,\dots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j} (T-t)^n \int_{\Delta_n} e^{\langle v^{(n)},x\rangle}\, dx\Bigg], \end{aligned} \tag{4.16}$$

where $v^{(n)} = (T-t)(\lambda_{i_1}, \dots,\ \lambda_{i_n})^T \in \mathbb{R}^n$, λ_{i_j} are given due to Theorem 4.5, Δ_n is the standard simplex in $\mathbb{R}^n$ defined by (1.3), and $\langle \cdot\,,\cdot\rangle$ denotes the standard inner product. Expressing the integrals in (4.16) as the ${}_1F_1$ functions according to (1.4) or (3.26) respectively, we obtain

$$\begin{aligned} P(t,T) &= e^{-r_m(T-t)} \cdot \Bigg[1 + \sum_{n=1}^{\infty}\frac{(-1)^n (r_M - r_m)^n (T-t)^n}{n!} \cdot \\ &\quad \sum_{(i_1,\dots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \prod_{j=1}^{n} h_{i_j}^{-1} B_{i_{j-1},i_j}\, {}_1F_1\left(1; n+1; v^{(n)}\right)\Bigg]. \end{aligned}$$

A straightforward calculation of $k_{i_{j-1},i_j} = h_{i_j}^{-1} B_{i_{j-1},i_j}$ according to Proposition 1.10 (b) and (e) completes the proof. $\square$

Zero-coupon bond: the PDE approach Another canonical method of computing the ZCB prices would be to solve the term structure equation (2.10) associated with the SDE of the underlying short-rate $(r_t)_{t\geq 0}$. In the case of the Jacobi short-rate model given by (4.8) it is preferable to consider the SDE (4.1) of the Jacobi diffusion, which is connected to the Jacobi short-rate model via the affine-linear transformation (4.9). The associated PDE in this case is given as follows:

$$\begin{cases} P_t(r,t,T) = -\frac{\sigma^2}{2}(z-z^2)P_{zz}(z,t,T) - k(\gamma - z)P_z(z,t,T) + zP(z,t,T), \\ P(z,T,T) = 1. \end{cases} \tag{4.17}$$

Considering the complicated formula for the ZCB prices derived in Theorem 4.6, it seems that the above PDE does have no simple solution, although we can simplify this PDE if we impose some restrictions on the underlying model.

Consider the following mapping for an arbitrary $\tilde{z}$:

$$z \longmapsto \int_{\tilde{z}}^{z} \frac{1}{\sqrt{s(1-s)}}\, ds = C \cdot \arccos(2z-1) =: x(z),$$

where C is a suitable constant. We have

$$\begin{aligned} P_z &= P_x \cdot \frac{C}{\sqrt{z(1-z)}}, \\ P_{zz} &= \frac{d}{dz}(P_z) = \frac{d}{dx}\frac{dx}{dz}\left(P_x \cdot \frac{C}{\sqrt{z(1-z)}}\right) \\ &= \frac{dx}{dz} \cdot \left(P_{xx} \cdot \frac{C}{\sqrt{z(1-z)}} + P_x \cdot \frac{d}{dx}\left(\frac{C}{\sqrt{z(1-z)}}\right)\right) \\ &= \frac{C}{\sqrt{z(1-z)}}\left(P_{xx} \cdot \frac{C}{\sqrt{z(1-z)}} - P_x \cdot \frac{1}{2} \cdot \frac{1-2z}{z(1-z)}\right) \\ &= P_{xx} \cdot \frac{C^2}{z(1-z)} - P_x \cdot \frac{C}{2} \cdot \frac{1-2z}{(z(1-z))^{3/2}}. \end{aligned}$$

Then (4.17) becomes

$$P_t = -\frac{\sigma^2 C^2}{2} \cdot P_{xx} + \left(\frac{\sigma^2 C(1-2z)}{4\sqrt{z(1-z)}} - \frac{Ck(\gamma - z)}{\sqrt{z(1-z)}}\right) \cdot P_x + \frac{\cos(x/C)+1}{2} \cdot P.$$

If we now choose $\gamma = \frac{1}{2}$, which is the case if the mean-reverting point lies in the middle of the state space, and $k = \frac{\sigma^2}{2}$, the middle term equals zero. By setting $C = \frac{\sqrt{2}}{\sigma}$, we obtain

$$P_t = -P_{xx} + \frac{\cos\left(x \cdot \frac{\sigma}{\sqrt{2}}\right) + 1}{2} \cdot P,$$

which is known as the *Kolmogorov-Petrovskii-Piskunov equation* (see [49], p. 71).

4.2.3 Practical implementation

From Theorem 4.6, the ZCB prices can be computed approximately by truncating the series in the relevant formula. We also use the truncated ${}_1F_1^H$ function defined by (1.5) as an approximation for the ${}_1F_1$ function.

Thus, with the notations of Theorem 4.6, we truncate the sum of series (4.14), obtaining

$$\begin{aligned} P(t, T; M, H) \;=\; & e^{-r_m(T-t)} \cdot \left[1 + \sum_{n=1}^{M} \frac{(-1)^n}{n!} \cdot (r_M - r_m)^n (T-t)^n \cdot \right. \\ & \left. \sum_{(i_1,\ldots,i_n \in I^n)} J_{i_n}^{(\alpha,\beta)}(z) \cdot \left(\prod_{j=1}^{n} k_{i_{j-1},i_j}\right) \cdot {}_1F_1^H\left(1; n+1; v^{(n)}\right)\right]. \end{aligned} \tag{4.18}$$

The choice of the truncation parameters M and H is left to the practitioner and should be made in the way of maintaining a balance between the accuracy of the results and reasonable computational speed. Some numerical examples, which provide numerical accuracy and computational speed for the formula (4.18), will be given at the end of the next section.

4.2.4 Connection to the CIR model

In this section we point out that after a suitable adaptation of the diffusion coefficient the Jacobi model converges to the CIR model examined in Section 2.4.2. Thus, we can use exact values of the ZCB prices in the CIR model as reference values to test how the ZCB pricing formula for the Jacobi model given in Section 4.2.2 performs.

The following convergence theorem shows that after a careful adaptation of the diffusion coefficient, the Jacobi model converges weakly to the CIR model.

Theorem 4.7. *Let* $(r_t^C)_{t\in[0,T]}$ *be the CIR short-rate process given by (2.29) with the initial value* r_0 *and the volatility parameter* $\sigma > 0$*, and let* $(r_t^J)_{t\in[0,T]}$ *be the Jacobi short-rate process given by (4.8) with the initial value* r_0 *and the volatility parameter* $\sigma_J > 0$*. If we set* $\sigma_J = \sigma/\sqrt{r_M}$ *and* $r_m = 0$*, then*

$$(r_t^J)_{t\in[0,T]} \Rightarrow (r_t^C)_{t\in[0,T]} \quad as \quad r_M \to \infty,$$

where " $\Rightarrow$ " denotes the weak convergence.

Proof. We use the convergence Theorem A.12. We know that the CIR process is the unique weak solution of (2.29) with the locally bounded diffusion coefficient $\sigma(x) = \sigma\sqrt{x}$. Moreover, we know from Proposition 4.3 that the Jacobi short-rate model $(r_t^J)_{t\in[0,T]}$ is the unique weak solution of the SDE (4.1). With $\sigma_J = \sigma/\sqrt{r_M}$ and $r_m = 0$, we obtain the following SDE:

$$dr_t^J = k[\theta - r_t^J]dt + \sigma\sqrt{r_t^J \cdot \frac{r_M - r_t^J}{r_M}}dW_t.$$

Now we need to verify the assumptions A.15 and A.16 of Theorem A.12. Let $T > 0$ and $R > 0$, then:

$$\begin{aligned}\sup_{r_M>0}\sup_{|x|\le R}|\sigma^{(n)}(x)| + |\mu^{(n)}(x)| &= \sup_{r_M>0}\sup_{|x|\le R}\sigma\sqrt{x\cdot\frac{r_M - x}{r_M}} + |k(\theta - x)| \\ &= \sup_{r_M>0}\sigma\sqrt{R} + k(R-\theta) < \infty\end{aligned}$$

and

$$\begin{aligned}&\lim_{r_M\to\infty}\int_0^T \sup_{|x|\le R}\left(|\sigma(x) - \sigma_n(x)| + |\mu(x) - \mu_n(x)|\right)ds = \\ &T\cdot\lim_{r_M\to\infty}\sup_{|x|\le R}\left|\sigma\sqrt{x\cdot\frac{r_M - x}{r_M}} - \sigma\sqrt{x}\right| = \\ &T\cdot\lim_{r_M\to\infty}\sup_{|x|\le R}\sigma\sqrt{x}\left(\sqrt{1 - \frac{x}{r_M}} - 1\right) = 0.\end{aligned}$$

The stability condition (4.2) of the Jacobi diffusion converges after the transformation of the model parameters to the stability condition (2.30) of the CIR process, which completes the proof. □

Theorem 4.8. *Let* $(r_t^V)_{t\in[0,T]}$ *be the Vasicek short-rate process given by (2.20) with the initial value* r_0 *and the volatility parameter* $\sigma > 0$*, and let* $(r_t^J)_{t\in[0,T]}$ *be the Jacobi short-rate process given by (4.8) with the initial value* r_0 *and the volatility parameter* $\sigma_J > 0$*. If we set* $\sigma_J = \sigma/\sqrt{-r_m \cdot r_M}$*, then*

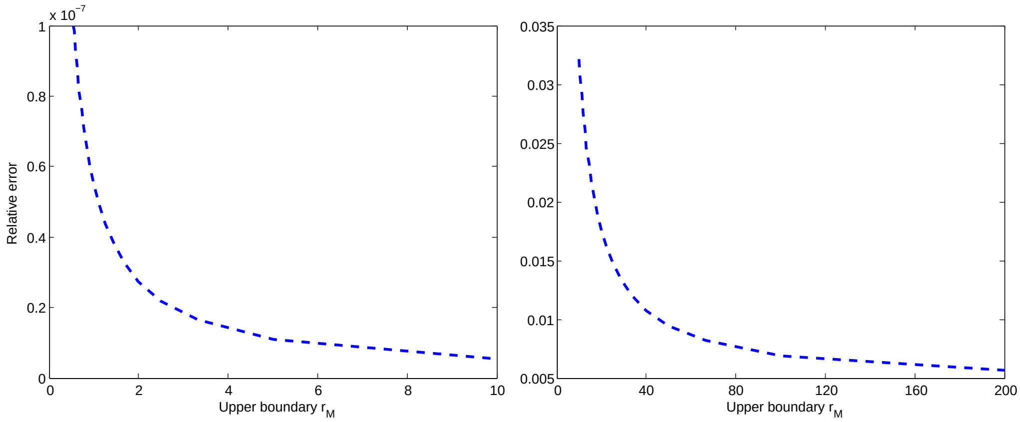
$$(r_t^J)_{t\in[0,T]} \Rightarrow (r_t^V)_{t\in[0,T]} \quad as \quad r_M \to \infty,\ r_m \to -\infty.$$

Proof. The proof is similar to that of Theorem 4.7. □

Corollary 4.9. *In the situations of Theorems 4.7 and 4.8 let $P^V(t,T)$ and $P^C(t,T)$ be the respective prices at time $t \in [0,T]$ of a ZCB with maturity $T \leq T^*$ in the Vasicek and CIR short-rate models given by (2.28) and (2.38). Let $P^J(t,T)$ be given as in (4.14) denoting the ZCB price in the Jacobi short-rate model at time t with maturity T. Then, $P^J(t,T)$ converges to $P^C(t,T)$ for $r_M \to \infty$ as well as to $P^V(t,T)$ for $r_M \to \infty, r_m \to -\infty$ if we adjust the diffusion coefficient of the Jacobi short-rate model according to Theorems 4.7 and 4.8.*

Proof. The proof is similar that of Corollary 3.10, where the convergence of the ZCB prices in the Ehrenfest short-rate model against those in the Vasicek model is shown. In fact, the proof for the Jacobi model is somewhat easier, since the Jacobi diffusion has $\mathbb{Q}$-a.s. continuous paths, in contrast to the Ehrenfest process, which makes the proof less technical. □

In the following, we compare for $r_M \to \infty$ the approximative ZCB prices $P(0,T;M,H)$ given in the Jacobi short-rate model by (4.18) with the ZCB prices $P(0,T)$ explicitly given in the CIR model by Theorem 2.21. We compute the the ZCB prices $P(0,T;M,H)$ with $M = 5$ and $H = 10$, according to the convergence Theorem 4.7 and Corollary 4.9. The computation time of one ZCB price in the Jacobi short-rate model due to (4.18) is 0.09 seconds. All computations were made on an INTEL Core2Duo 2400MHz machine.



(a) Numerically favourable case (b) Numerically unfavourable case

Figure 4.2: Relative price errors against r_M by an approximation of ZCB prices $P(0,T)$ in the CIR model by $P(0,T;5,10)$ in the Jacobi short-rate model.

Similarly to the discussion of the Ehrenfest short-rate model in Section 3.3, we consider two sets of parameters for the CIR model, where $k = 0.2$, $\theta = 0.05$ and $r_0 = 0.01$ are fixed for both cases. In the first case (a), we choose volatility $\sigma = 0.05$ and a time to maturity $T = 1$ year. In the second case (b), we choose a high volatility $\sigma = 0.2$ and a time to maturity of 10 years, which is unfavourable for the numerical computation of the ZCB prices. Figure 4.2 illustrates the convergence results.

In both cases we observe fast convergence of the respective ZCB prices and see that the choice of the truncating parameters $M = 5$ and $H = 10$ is satisfactory for our purpose.

Chapter 5

Conclusion and outlook

This work has explored two models for the short-rate whose characteristic feature lies in the possibility of choosing lower and upper boundaries for the interest rate. The first model is a finite-state mean-reverting short-rate model based on the Ehrenfest process. The respective short-rate process can be seen as an affine linearly transformed birth-and-death process on $\{0, 1, \cdots, N\}$, $N \in \mathbb{N}$. The second short-rate model is also mean-reverting and is based on the Jacobi diffusion. Both models provide a certain degree of analytical tractability, since they allow for the explicit pricing of ZCBs and also solve the problem of the negative interest rates characteristic of Gaussian models.

The pricing formulae for ZCBs in the Ehrenfest short-rate model have been derived for the general case and the special case of the model, in which the underlying distribution is symmetric with respect to the mean-reverting value. The special case benefits from a more tractable pricing formula for ZCBs. The key to both approaches has turned out to be the representation of the underlying Ehrenfest process as a sum of independent binary processes, which is possible only in continuous time. For the general case of the Ehrenfest short-rate model as well as for the Jacobi short-rate model we used the spectral representations of the underlying transition probability densities and derived the arbitrage-free ZCB prices in the risk-neutral valuation framework. For both models we used the hypergeometric functions of a matrix argument, which extends the known results onto the Jacobi short-rate model provided in [13]. We also used the Krawtchouk and Jacobi orthogonal polynomials, whose properties turned out to be the main tool for exploiting the spectral representations of the transition probabilities in both short-rate models.

We have seen that the Ehrenfest short-rate model is a good approximation to the Vasicek model under normal conditions and a better alternative to it in extreme cases, where the interest rates are low and the volatility high, providing solely positive interest rates. A further advantage of the model is the availability of five fitting parameters in the general case. The Jacobi short-rate model can furthermore be used to approximate the Vasicek model but it converges to the CIR model after suitably adapting the parameters as well. Here we also have five determining parameters to incorporate different structures of interest rates.

Our conclusion is that both the Ehrenfest and the Jacobi short-rate models are interesting enrichments in the field of term structure modelling, combining analytical tractability with the desired property of the interest rates remaining positive. Moreover, the possibility in the Ehrenfest short-rate model of choosing the states for the interest rate by means of varying the lower and upper bounds as well as the number of states between them makes the model very illustrative. The Jacobi short-rate model fits in the widely-used framework of modelling the short-rates via diffusion processes and profits from the strongly developed theory of diffusions.

Problems that remain open for both of the short-rate models that we have examined here are the derivation of explicit pricing formulae for the European options on ZCBs and the parameter estimates for the models under the objective measure. Another open problem is that of providing error estimates for the approximative ZCB prices as truncated series and determining the best way to truncate these.

An interesting generalization of the Ehrenfest short-rate model would consist in allowing the interest rate to jump more than one unit. It is also feasible to consider the limiting case of the Ehrenfest short-rate model, where the stationary distribution parameter p converges to 0 as N converges to infinity. This would yield the birth-and-death queuing process with infinitely many servers (see [35], p. 306) involving the so-called Poisson-Charlier polynomials (see [60], §2.81).

Appendix A

Introduction to Diffusion Theory

In this section we give a short overview of some of the basic concepts and results in the theory of diffusion processes that we need throughout this work. There is a vast amount of literature on this topic. We will provide some results from the works of Karatzas and Shreve [32], Karlin and Taylor [34], Øksendal [47], and Stroock and Varadhan [59].

A.1 One-dimensional diffusions

Throughout this work we encounter only *time-homogeneous* stochastic processes. In the following we will restrict our attention to this case. Let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space with filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions, i.e. it is complete and right-continuous. Let all stochastic processes considered in the following be defined on this probability space and be adopted to the filtration $(\mathcal{F}_t)_{t\geq 0}$.

Definition A.1 (**Diffusion**). A time-homogeneous Markov process $(X_t)_{t\geq 0}$ on the state space $E = (a, b)$ is said to be a (one-dimensional) diffusion with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.)$, if

(*i*) it has continuous paths, and

(*ii*) the conditions

$$\begin{aligned} \mathbb{E}[X_{s+t} - X_s \mid X_s = x] &= t\mu(x) + o(t), \\ \mathbb{E}[(X_{s+t} - X_s)^2 \mid X_s = x] &= t\sigma^2(x) + o(t), \\ \mathbb{E}[|X_{s+t} - X_s|^3 \mid X_s = x] &= o(t), \end{aligned} \tag{A.1}$$

hold, as $t \downarrow 0$, for $s \geq 0$ and every $x \in E$.

The condition (A.1) can be formulated in a more general way which does not require the existence of finite moments (see [4], p. 368).

The following assumption is essential for the further discussion.

Assumption A.2. *Let $(X_t)_{t\geq 0}$ be a diffusion on $E = (-\infty, \infty)$ with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$. The functions $\mu(.)$ and $\sigma(.)$ are continuously differentiable, with bounded derivatives on E. Also, $\sigma''(.)$ exists and is continuous.*

Remark A.3. (a) It is known that under Assumption A.2 a strictly positive and continuous density $p(t; x, y)$ of the transition probability distribution of $(X_t)_{t\geq 0}$ exists (see [4], pp. 368 and 497).

(b) The result in (a) can be stated for diffusions on $E = (a, b)$ if we modify Assumption A.2 accordingly (see Proposition V.3.1 in [4]).

For a diffusion $(X_t)_{t\geq 0}$ on the state space E consider the set $\mathcal{B}(E)$ of all real-valued, bounded and Borel measurable functions f on E, and define the *transition operator*

$$(T_t f)(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int_E f(y)p(t; x, y)\, dy, \quad x \in E,\ t \geq 0. \quad \text{(A.2)}$$

Lemma A.4. *In the setting above, the family of transition operators $(T_t)_{t\geq 0}$ is a commutative, contractive semigroup on $\mathcal{B}(E)$, i.e. $(T_t)_{t\geq 0}$ is a family of linear operators*

$$\begin{aligned} T_t : \mathcal{B}(E) &\longrightarrow \mathcal{B}(E) \\ f &\longmapsto T_t f \end{aligned}$$

satisfying the following properties:

(i) $T_0 = id$,

(ii) $T_s \cdot T_t = T_{s+t},\ \forall s, t \geq 0$,

(iii) $||T_t f|| \leq ||f||\ \forall f \in \mathcal{B}(E),\ t \geq 0$,

where $||\,.\,||$ denotes the uniform norm $||f|| = \sup\{|f(x)| \ : \ x \in E\}$.

Proof. See [4], p. 372. □

Definition A.5 (Infinitesimal generator). The *(infinitesimal) generator* $\mathcal{A}$ of the Markov process $(X_t)_{t\geq 0}$ on E, or of $(T_t)_{t>0}$, is a linear operator A defined by

$$(\mathcal{A}f)(x) = \lim_{s\downarrow 0} \frac{(T_s f)(x) - f(x)}{s}, \quad \text{(A.3)}$$

for all $f \in \mathcal{B}(E)$ for which the right side of (A.3) uniformly converges to some function in x. The class of all such f comprises the *domain* $\mathcal{D}_{\mathcal{A}}$ of $\mathcal{A}$.

The following result, showing the explicit relationship between the diffusion and the respective generator, is taken from [4], p. 374, without a proof.

Proposition A.6. *Let $(X_t)_{t\geq 0}$ be a diffusion on E with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$. Then, all $f \in \mathcal{C}_c^2(E)$, i.e. twice continuously differentiable f, vanishing outside a closed bounded subinterval of E, belong to $\mathcal{D}_{\mathcal{A}}$, and for such f*

$$(\mathcal{A}f)(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x).$$

Proposition A.7. *Let $(T_t)_{t>0}$ be the family of transition operators for a diffusion on E. If $f \in \mathcal{D}_{\mathcal{A}}$, then the following backward equation holds*

$$\frac{\partial}{\partial t}(T_t f)(x) = \mathcal{A}(T_t f)(x), \quad x \in E, \ t > 0.$$

Furthermore, T_t and $\mathcal{A}$ commute on $\mathcal{D}_{\mathcal{A}}$, i.e.

$$\mathcal{A}(T_t f) = T_t(\mathcal{A}f), \quad f \in \mathcal{D}_{\mathcal{A}}.$$

Proof. See [4], p. 375. □

A.2 Spectral decomposition of transition probability densities

In this section we give a short review of the calculation of the transition probabilities of diffusions using spectral methods. We follow Section V.8 of [4].

Consider an arbitrary diffusion $(X_t)_{t\geq 0}$ on E with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$. Now consider the function

$$\pi(x) = \frac{2K}{\sigma^2(x)} \exp\left(I(\tilde{x}, x)\right), \quad x \in E, \tag{A.4}$$

where K is an arbitrary positive constant, $\tilde{x}$ an arbitrary chosen state and

$$I(\tilde{x}, x) = \int_{\tilde{x}}^{x} \frac{2\mu(z)}{\sigma^2(z)}\, dz. \tag{A.5}$$

Consider the space $L^2(E, \pi)$ of real-valued functions on E that are square integrable with respect to the function π and the inner product $\langle . \, , . \rangle_\pi$ defined by

$$\langle f, h \rangle_\pi = \int_E f(x)h(x)\pi(x) \, dx. \tag{A.6}$$

Lemma A.8. *Let $(X_t)_{t\geq 0}$ be a diffusion on E with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$, and let $\mathcal{A}$ be the corresponding generator. Then,*

$$\langle \mathcal{A}f, h \rangle_\pi = \langle f, \mathcal{A}h \rangle_\pi \quad f, h \in \mathcal{D}_\mathcal{A}.$$

Proof. See [4], p. 408. □

Now let φ be an eigenfunction of $\mathcal{A}$ with eigenvalue λ, i.e.

$$\mathcal{A}\varphi = \lambda\varphi, \tag{A.7}$$

then $u(t, x) = e^{\lambda t}\lambda\varphi(x)$ solves the backward differential equation

$$\frac{\partial u}{\partial t}(t, x) = e^{\lambda t}\lambda\varphi(x) = e^{\lambda t}\mathcal{A}\varphi(x) = \mathcal{A}u(t, x). \tag{A.8}$$

Similarly, if $u(., .)$ is a linear combination of such functions, then the same will be true.

Let $(\lambda_n)_{n\in\mathbb{N}_0}$ be the set of the eigenvalues and $(\varphi_n)_{n\in\mathbb{N}_0}$ the set of the corresponding eigenfunctions with $\langle \varphi_m, \varphi_m \rangle_\pi = 1$. In the situation at hand we obtain the following result.

Lemma A.9. *Let $(\lambda_n)_{n\in\mathbb{N}_0}$ be the distinct eigenvalues and $(\varphi_n)_{n\in\mathbb{N}_0}$ be the corresponding eigenfunctions of a generator $\mathcal{A}$. Then, $(\varphi_n)_{n\in\mathbb{N}_0}$ is an orthogonal system, i.e.,*

$$\langle \varphi_n, \varphi_m \rangle_\pi = \delta_{n,m}, \quad n, m \in \mathbb{N}_0.$$

Proof. Let λ_n and λ_m be two distinct eigenvalues with eigenfunctions φ_n and φ_m. Then, with definition (A.7) we obtain on the one hand

$$\begin{aligned} \langle \varphi_n, \varphi_m \rangle_\pi &= \int_E \varphi_n(x)\varphi_m(x)\pi(x) \, dx = \frac{1}{\lambda_n}\int_E \lambda_n\varphi_n(x)\varphi_m(x)\pi(x) \, dx \\ &= \frac{1}{\lambda_n}\int_E \mathcal{A}\varphi_n(x)\varphi_m(x)\pi(x) \, dx = \frac{1}{\lambda_n}\langle \mathcal{A}\varphi_n, \varphi_m \rangle_\pi. \end{aligned}$$

An analogous calculation yields on the other hand

$$\langle \varphi_n, \varphi_m \rangle_\pi = \frac{1}{\lambda_m}\langle \varphi_n, \mathcal{A}\varphi_m \rangle_\pi.$$

With Lemma A.8 we obtain

$$\frac{1}{\lambda_n}\langle \mathcal{A}\varphi_n, \varphi_m\rangle_\pi = \frac{1}{\lambda_m}\langle \mathcal{A}\varphi_n, \varphi_m\rangle_\pi,$$

which immediately yields

$$\langle \varphi_n, \varphi_m\rangle_\pi = \frac{1}{\lambda_n}\langle \mathcal{A}\varphi_n, \varphi_m\rangle_\pi = 0.$$

□

It is also easy to show that if there is more than one linearly independent eigenfunction for a single eigenvalue, then these eigenfunctions can be orthogonalized by the Gram-Schmidt procedure. So, $(\varphi_n)_{n\in\mathbb{N}_0}$ can be taken to be orthonormal.

If the set of finite linear combinations of eigenfunctions is complete in $L^2(E,\pi)$, then each $f \in L^2(E,\pi)$ has a Fourier expansion of the form (see [60], p. 24)

$$f = \sum_{n=0}^{\infty} \langle f, \varphi_n\rangle_\pi \cdot \varphi_n.$$

Consider the linear combination defined by

$$\tilde{u}(t,x) = \sum_{n=0}^{\infty} e^{\lambda_n t}\langle f, \varphi_n\rangle_\pi \cdot \varphi_n. \tag{A.9}$$

Then $\tilde{u}$ satisfies the backward equation (A.8) with the initial condition

$$\tilde{u}(0,x) = f(x).$$

From Proposition A.7 we know that $T_t f$ defined by (A.2) also satisfies the same backward equation and initial condition. So if there is uniqueness for a sufficiently large class of initial functions f, then we obtain

$$T_t f(x) = \tilde{u}(t,x), \quad x \in E,$$

which is, due to definition (A.9), equivalent to

$$\int_E f(y)p(t;,x,y)\ dy = \int_E f(y)\left(\sum_{n=0}^{\infty} e^{\lambda_n t}\varphi_n(x)\varphi_n(y)\right)\pi(y)\ dy.$$

In such cases, therefore, the transition probability density $p(t;x,y)$ is given by

$$p(t;x,y) = \sum_{n=0}^{\infty} e^{\lambda_n t}\varphi_n(x)\varphi_n(y)\pi(y), \quad x,y \in E. \tag{A.10}$$

A.3 Hitting probability

For this section we consider an arbitrary diffusion $(X_t)_{t\geq 0}$ on $E = [a, b]$ with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$. Let τ_x be the hitting time defined by

$$\tau_x = \inf\{t \geq 0 : X_t = x\},\ x \in E. \tag{A.11}$$

Since $(X_t)_{t\geq 0}$ has $\mathbb{P}$-a.s. continuous paths and $\{x\}$ is a closed set, we know that τ_x is a stopping time (see e.g. [52], Proposition 4.5). Let $\rho_{x_0,x}$ denote the probability that $(X_t)_{t\geq 0}$, starting in x_0, hits x in finite time, i.e.

$$\rho_{x_0,x} = P(\tau_x < \infty | X_0 = x_0), \tag{A.12}$$

and let the *scale function* be defined as follows:

$$s(x) = s(\tilde{x}, x) = \int_{\tilde{x}}^{x} \exp\left(-I(\tilde{x}, z)\right)\, dz, \tag{A.13}$$

where $\tilde{x} \in E$ is arbitrary and I is defined as in (A.5). The following results can be found in Chapter V.9 of [4] and are given here without proof.

Theorem A.10. *Let $(X_t)_{t\geq 0}$ be an arbitrary diffusion on $E = [a, b]$ with drift coefficient $\mu(.)$ and diffusion coefficient $\sigma^2(.) > 0$ and a start in $x_0 \in E$. Consider $[c, d] \subseteq E, c < d$ and let*

$$\psi(x) = P((X_t)_{t\geq 0}\, \textit{hits } d \textit{ before } c\,), \quad c \leq x_0 \leq d.$$

Then,

$$\psi(x) = \frac{s(x_0) - s(c)}{s(d) - s(c)}.$$

Corollary A.11. *Let $E = (a, b)$ and $x_0 \in S$ be arbitrary.*

(a) If $s(a) = -\infty$, then $\rho_{x_0,x} = 1\ \forall x > x_0$. If $s(a) < \infty$, then

$$\rho_{x_0,x} = \frac{s(x_0) - s(a)}{s(x) - s(a)}, \quad (x > x_0).$$

(b) If $s(b) = -\infty$, then $\rho_{x_0,x} = 1\ \forall x < x_0$. If $s(b) < \infty$, then

$$\rho_{x_0,x} = \frac{s(b) - s(x_0)}{s(b) - s(x)}, \quad (x < x_0).$$

□

A.4 Weak convergence of solutions of SDEs

In this section we use some well-known results from the theory of diffusions and provide a convergence result for solutions of SDEs. In the following we omit constructing the underlying Itô-Integral, a comprehensive study of which is given in e.g. [32] or [47].

Let $(X_t)_{t\geq 0}$ be an Itô-process in the sense of Definition 4.1.1 in [47], satisfying the following SDE:

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R} \tag{A.14}$$

where μ and σ are measurable functions denoting the drift and diffusion coefficients of the process.

In the following we provide a convergence result for the solutions of SDEs which is a slight modification of the convergence theorem 11.1.4 in [59]. There, the theorem provides the weak convergence of distributions associated with solutions of the so-called *martingale problem*. A deeper discussion of this topic would be beyond the scope of this work (see Section 5.4 in [32] or Chapter 6 in [59] for more details).

Proposition A.12. *Let $X^{(n)} = (X_t^{(n)})_{t\geq 0}$ be a sequence of unique weak solutions of SDEs in the sense of Definition 3.1 and 3.2 of [32]:*

$$dX_t^{(n)} = \mu_n(t, X_t^{(n)})dt + \sigma_n(t, X_t^{(n)})dX_t^{(n)}, \quad X_0^{(n)} = x_0^{(n)} \in \mathbb{R},$$

where $\mu_n : [0,\infty) \times \mathbb{R} \longrightarrow \mathbb{R}$ and $\sigma_n : [0,\infty) \times \mathbb{R} \longrightarrow [0,\infty)$ are measurable functions such that for all $T > 0$ and $R > 0$:

$$\sup_{n\geq 1} \sup_{0\leq s\leq T} \sup_{|x|\leq R} |\sigma^{(n)}(s,x)| + |\mu^{(n)}(s,x)| < \infty. \tag{A.15}$$

Let $X = (X_t)_{t\geq 0}$ be the unique weak solution of the SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dX_t, \quad X_0 = x_0 \in \mathbb{R},$$

where $\mu : [0,\infty) \times \mathbb{R} \longrightarrow \mathbb{R}$ and $\sigma : [0,\infty) \times \mathbb{R} \longrightarrow [0,\infty)$ are locally bounded measurable functions which are continuous in x for each $t \geq 0$ satisfying for all $T > 0$ and $R > 0$:

$$\lim_{n\to\infty} \int_0^T \sup_{|x|\leq R} \left(|\sigma(s,x) - \sigma_n(s,x)| + |\mu(s,x) - \mu_n(s,x)|\right) ds = 0. \tag{A.16}$$

If $x_0^{(n)} \overset{n\to\infty}{\longrightarrow} x_0$, then $X^{(n)} \Rightarrow X$ as $n \to \infty$, where " $\Rightarrow$ " denotes the weak convergence.

Proof. We consider Proposition 4.11 in [32], where, under the condition of local boundedness of σ, the equivalence between the existence of a weak solution of an SDE and the existence of a solution to the martingale problem is shown. Hence, this proposition is equivalent to the convergence theorem 11.1.4 in [59]. □

Appendix B

Proofs

B.1 Proof of Proposition 1.10

Ad. (a) : The claim follows immediately from the series representation (1.14) of the Jacobi polynomials $P_n^{(\alpha,\beta)}$.
Ad. (b) : We use Definition 1.9 and the orthogonality relation (1.16) of the Jacobi polynomials $P_n^{(\alpha,\beta)}$ as follows:

$$\begin{aligned}
&\int_0^1 x^\alpha (1-x)^\beta J_n^{(\alpha,\beta)}(x) J_m^{(\alpha,\beta)}(x)\,dx \\
&\quad = \int_0^1 x^\alpha (1-x)^\beta P_n^{(\alpha,\beta)}(1-2x) P_m^{(\alpha,\beta)}(1-2x)\,dx \\
&\quad = -\frac{1}{2}\int_1^{-1} \left(\frac{1-t}{2}\right)^\alpha \left(1-\frac{1-t}{2}\right)^\beta P_n^{(\alpha,\beta)}(t) P_m^{(\alpha,\beta)}(t)\,dt \\
&\quad = \frac{1}{2}\int_{-1}^1 \frac{1}{2^\alpha}(1-t)^\alpha \frac{1}{2^\beta}(1+t)^\beta P_n^{(\alpha,\beta)}(t) P_m^{(\alpha,\beta)}(t)\,dt \\
&\quad = \frac{1}{2^{\alpha+\beta+1}}\int_{-1}^1 (1-t)^\alpha (1+t)^\beta P_n^{(\alpha,\beta)}(t) P_m^{(\alpha,\beta)}(t)\,dt \\
&\quad = \delta_{n,m}\cdot\frac{1}{2^{\alpha+\beta+1}}\cdot\frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)} \\
&\quad = \delta_{n,m}\cdot\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)} = \delta_{n,m}\cdot h_n^{(\alpha,\beta)}.
\end{aligned}$$

Ad. (c) : We consider the following affine linear transformation:

$$\begin{aligned}
z : [0,1] &\longrightarrow [-1,1] \\
x &\longmapsto 1-2x.
\end{aligned}$$

An auxiliary calculation yields

$$\begin{aligned}
\frac{dz}{dx} &= -2,\\
\frac{dy}{dz} &= \frac{dy}{dx}\frac{dx}{dz} = -\frac{1}{2}\frac{dy}{dx},\\
\frac{d^2y}{dz^2} &= \frac{d}{dx}\frac{dx}{dz}\left(\frac{dy}{dz}\right) = \frac{dx}{dz}\frac{d}{dx}\left(-\frac{1}{2}\frac{dy}{dx}\right) = \frac{1}{4}\frac{d^2y}{dx^2}.
\end{aligned}$$

We rewrite the differential equation (1.18) for the Jacobi polynomials $P_n^{(\alpha,\beta)}$ in the variable $x \in [0,1]$ as follows:

$$4x(1-x)\frac{1}{4}\frac{d^2y}{dx^2}+\underbrace{(\beta-\alpha-(\alpha+\beta+2)(1-2x))}_{(-2\alpha-2+2(\alpha+\beta+2)x)}\left(-\frac{1}{2}\frac{dy}{dx}\right)+n(n+\alpha+\beta+1)y=0,$$

which is equivalent to (1.25).
Ad. (d) : We use the recurrence relation (1.19) for the Jacobi polynomials $P_n^{(\alpha,\beta)}$ at $1-2x$ for $x \in [0,1]$:

$$\begin{aligned}
&2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha,\beta)}(1-2x) =\\
&\{(2n+\alpha+\beta+1)(\alpha^2-\beta^2)+(2n+\alpha+\beta)_3(1-2x)\}P_n^{(\alpha,\beta)}(1-2x)\\
&-2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha,\beta)}(1-2x),
\end{aligned}$$

An equivalent reformulation of the above equation yields

$$\begin{aligned}
xP_n^{(\alpha,\beta)}(1-2x) &= \frac{(2n+\alpha+\beta+1)(\alpha^2-\beta^2)+(2n+\alpha+\beta)_3}{2(2n+\alpha+\beta)_3}P_n^{(\alpha,\beta)}(1-2x)\\
&\quad -\frac{2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)}{2(2n+\alpha+\beta)_3}P_{n+1}^{(\alpha,\beta)}(1-2x)\\
&\quad -\frac{2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)}{2(2n+\alpha+\beta)_3}P_{n-1}^{(\alpha,\beta)}(1-2x).
\end{aligned}$$

If we use $J_n^{(\alpha,\beta)}(x) = P_n^{(\alpha,\beta)}(1-2x)$, the claim follows immediately.
Ad. (e) : The claim follows from the orthogonality relation (b) and the recurrence relation (d). □

B.2 Proof of Theorem 2.20

First, we show the spectral representation given in (b), from which we will follow the other claims.

Ad. (b) : Consider the linearly transformed process $z_t = \vartheta(r_t),\ t \geq 0$, with

$$\vartheta(x) = \frac{2k}{\sigma^2} x. \tag{B.1}$$

With Itô's lemma, we immediately see that $(z_t)_{t\geq 0}$ satisfies the SDE

$$dz_z = \left(\frac{2k^2\theta}{\sigma^2} - kz_t \right) dt + \sqrt{2kz_t} dW_t, \quad z_0 = \vartheta(r_0). \tag{B.2}$$

It is well known that under the stability condition (2.30), the CIR process $(r_t)_{t\geq 0}$, and hence the transformed process $(z_t)_{t\geq 0}$, are diffusions in terms of Definition A.1. Hence, we can apply the theory of spectral representation of transition probabilities provided in Appendix A.2, and thus obtain from Proposition A.6 and formula (A.10) the transition probability density function $p_z(t; x, y)$ of $(z_t)_{t\geq 0}$ as follows:

$$p_z(t; x, y) = \sum_{n=0}^{\infty} \pi(y)\varphi_n(x)\varphi_n(y)e^{\lambda_n t}, \quad x, y > 0,$$

where λ_n and φ_n solve the differential equation

$$kz\varphi''_n(z) + \left(\frac{2k^2\theta}{\sigma^2} - kz \right) \varphi_n^{'}(z) - \lambda_n \varphi_n(z) = 0, \quad n \in \mathbb{N}_0, \tag{B.3}$$

and the weight function π is given such that $(\varphi_n)_{n\in\mathbb{N}_0}$ is the orthonormal basis of the space $L((0,\infty),\pi)$.

It is well known that the Laguerre polynomials $h_n L_n^{(\alpha)},\ \alpha > -1, n \in \mathbb{N}_0$, with $h_n = \sqrt{\frac{n!}{\Gamma(n+\alpha+1)}}$ solve the differential equation (see [60], p.100)

$$zy'' + (\alpha + 1 - z)y^{'} + ny = 0, \quad y = L_n^{(\alpha)}(z) \tag{B.4}$$

and build the orthonormal basis of $L((0,\infty),\pi)$ with the uniquely determined weight function $\pi(z) = e^{-z}z^{\alpha},\ z > 0$. If we divide equation (B.3) by k and compare the coefficients of (B.3) and (B.4), we obtain

$$p_z(t; x, y) = \sum_{n=0}^{\infty} e^{-y}y^{\alpha} \frac{n!}{\Gamma(n+\alpha+1)} L_n^{(\alpha)}(x)L_n^{(\alpha)}(y)e^{-knt}, \quad x, y > 0,$$

where $\alpha = \frac{2k\theta}{\sigma^2} - 1$. From Lemma 1 in [43], we obtain the transition density function $p(t, x, y)$ of the CIR process $(r_t)_{t\geq 0}$ from the transition density function $p_z(t; x, y)$ as follows:

$$p(t, x, y) = p_z(t; \vartheta(x), \vartheta(y)) \cdot \vartheta'(y),$$

where $\vartheta(.)$ is defined by (B.1). Hence, with $h = \frac{2k}{\sigma^2}$ the claim follows.

Ad. (a) : Consider the *Hille-Hardy formula* (see [18], p.189)

$$\begin{aligned}&\sum_{n=0}^{\infty} \frac{n!}{\Gamma(n+\alpha+1)} L_n^{(\alpha)}(x) L_n^{(\alpha)}(y) z^n \\ &= (1-z)^{-1} \exp\left(-z\frac{x+y}{1-z}\right)(xyz)^{-\frac{\alpha}{2}} I_\alpha\left(2\frac{(xyz)^{1/2}}{1-z}\right), \quad |z|<1,\end{aligned}$$

where I_α is the modified Bessel function of the first kind of order α given by (2.33). We apply the Hille-Hardy formula with $z = e^{-kt}$ to the spectral representation (2.37) of the transition density and obtain

$$p(t;x,y) = \frac{e^{-hy} h(hy)^\alpha}{1-e^{-kt}} \exp\left(-\frac{h(x+y)e^{-kt}}{1-e^{-kt}}\right)\left(h^2 xy e^{-kt}\right)^{-\frac{\alpha}{2}} I_\alpha\left(\frac{2\sqrt{h^2 xy e^{-kt}}}{1-e^{-kt}}\right).$$

With

$$\begin{aligned} c_t &= \frac{4k}{\sigma^2(1-\exp(-kt))} = \frac{2h}{(1-e^{-kt})} \quad \text{and} \\ \lambda_t &= c_t x \exp(-kt) \end{aligned}$$

we have

$$\begin{aligned} p(t;x,y) &= \frac{c_t}{2} \exp\left(-hy - \frac{h(x+y)e^{-kt}}{1-e^{-kt}}\right)\left(\frac{(hy)^2}{h^2 xy e^{-kt}}\right)^{\frac{\alpha}{2}} I_\alpha\left(\sqrt{\lambda_t c_t y}\right) \\ &= \frac{c_t}{2} \exp\left(-\frac{\lambda_t + c_t y}{2}\right)\left(\frac{c_t y}{\lambda_t}\right)^{\frac{\alpha}{2}} I_\alpha\left(\sqrt{\lambda_t c_t y}\right) \end{aligned}$$

With

$$v = \frac{4k\theta}{\sigma^2} = 2\alpha - 2 \quad \text{or equivalently} \quad \alpha = \frac{v}{2} - 1$$

and the density formula (2.31) the claim follows.

Ad. (c) : If we use (2.36) and substitute $z = c_t y$, we obtain for the moments of r_t:

$$\mathbb{E}[r_t^n] = \int_0^\infty y^n c_t f_{\chi^2(v,\lambda_t)}(c_t y)\, dy = \frac{1}{c_t^n}\int_0^\infty z^n c_t f_{\chi^2(v,\lambda_t)}(z)\, dz.$$

Hence, we have with (2.34)

$$
\begin{aligned}
\mathbb{E}[r_t] &= \frac{v+\lambda_t}{c_t} = \frac{4k\theta}{\sigma^2 c_t} + r_0 \cdot e^{-kt} = r_0 \cdot e^{-kt} + \theta\left(1-e^{-kt}\right) \\
\mathbb{V}ar[r_t] &= \frac{2(v+2\lambda_t)}{c_t^2} \\
&= \frac{2\cdot 4k\theta\sigma^4\left(1-e^{-kt}\right)^2}{\sigma^2 \cdot 16k} + \frac{4r_0\cdot e^{-kt}\sigma^2\left(1-e^{-kt}\right)}{4k} \\
&= r_0\frac{\sigma^2}{k}\left(e^{-kt}-e^{-2kt}\right) + \theta\frac{\sigma^2}{2k}\left(1-e^{-kt}\right)^2.
\end{aligned}
$$

The claim follows with the time-homogeneity of the process $(r_t)_{t\geq 0}$.

Ad (d) : The spectral decomposition (2.37) and the fact that $L_0^{(\alpha)} \equiv 1$, yields the limiting – and hence a stationary – distribution of the CIR process:

$$
\lim_{t\to\infty} p(t;x,y) = e^{-hy}h(hy)^{\alpha}\frac{1}{\Gamma(\alpha+1)},
$$

which is the density function of the Gamma distribution with parameters $\alpha+1$ and h.

B.3 Proof of Theorem 2.21

In Section 2.3 we have seen that the price $P(r,t,T) = P(t,T)$ at time $t \in [0,T]$ of a ZCB with maturity $T \leq T^*$ and the state of the short-rate $r = r_t$ safisfies the term structure equation (2.10). Since the CIR model belongs with $\mu(r,t) = k[\theta - r]$ and $\sigma(r,t) = \sigma\sqrt{r}$ to the class of affine term structure models, we can rewrite the term structure equation (2.18) as follows:

$$
\begin{aligned}
&\left[\frac{\partial}{\partial t}\ln A(t,T) - B_t(t,T)\cdot r\right]P(r,t,T) - k[\theta-r]B(t,T)P(r,t,T) + \\
&+\frac{1}{2}\sigma^2\cdot r\cdot B(t,T)^2\cdot P(r,t,T) - r\cdot P(r,t,T) = 0,
\end{aligned}
$$

which is equal to

$$
\begin{aligned}
&\frac{\partial}{\partial t}\ln A(t,T) - k\theta B(t,T) - \\
&r\cdot\left(B_t(t,T) - kB(t,T) - \frac{1}{2}\sigma^2\cdot B(t,T)^2 + 1\right) = 0
\end{aligned}
$$

for all $t \in (0,T)$ and $r \in \mathbb{R}$. This equation is satisfied for all $r \in \mathbb{R}$ if $A(t,T)$ and $B(t,T)$ solve the ordinary differential equations

$$B_t(t,T) - kB(t,T) - \frac{1}{2}\sigma^2 \cdot B(t,T)^2 + 1 = 0, \quad B(T,T) = 0, \tag{B.5}$$

$$\frac{\partial}{\partial t} \ln A(t,T) - k\theta B(t,T) = 0, \qquad A(T,T) = 1. \tag{B.6}$$

We first solve the Riccati-equation (B.5)

$$B(t,T)^2 + \frac{2k}{\sigma^2} B(t,T) - \frac{2}{\sigma^2} = \frac{2}{\sigma^2} B_t(t,T),$$

which can be rewritten as

$$(B(t,T) - c_1)(B(t,T) + c_2) = \frac{2}{\sigma^2} B_t(t,T),$$

where

$$c_1 = \frac{1}{\sigma^2}(-k+h), \; c_2 = \frac{1}{\sigma^2}(k+h) \text{ and } h = \sqrt{k^2 + 2\sigma^2}. \tag{B.7}$$

By a separation of the variables, we have

$$\int_{B(t,T)}^{0} \frac{1}{(y-c_1)(y+c_2)} \, dy = \int_t^T \frac{\sigma^2}{2} \, ds,$$

which is equal to

$$\int_{B(t,T)}^{0} \frac{\frac{1}{c_1+c_2}}{y-c_1} - \frac{\frac{1}{c_1+c_2}}{y+c_2} \, dy = \frac{\sigma^2}{2}(T-t). \tag{B.8}$$

For the left hand side of (B.8), we have

$$\begin{aligned}
&\int_{B(t,T)}^{0} \frac{\frac{1}{c_1+c_2}}{y-c_1} - \frac{\frac{1}{c_1+c_2}}{y+c_2} = \frac{1}{c_1+c_2} \left[\ln(y-c_1) - \ln(y+c_2)\right] \Big|_{B(t,T)}^{0} \\
&= \frac{\sigma^2}{2h} \ln\left(\frac{y-c_1}{y+c_2}\right) \Big|_{B(t,T)}^{0} = \frac{\sigma^2}{2h}\left[\ln\left(\frac{-c_1}{c_2}\right) - \ln\left(\frac{B(t,T)-c_1}{B(t,T)+c_2}\right)\right] \\
&= \frac{\sigma^2}{2h} \ln\left(\frac{-c_1}{c_2} \cdot \frac{B(t,T)+c_2}{B(t,T)-c_1}\right).
\end{aligned}$$

If we set this equal to the right hand side of (B.8) and apply the exponential function, we obtain

$$-\frac{c_1}{c_2} \cdot \frac{B(t,T)+c_2}{B(t,T)-c_1} = e^{h(T-t)},$$

which is equal to

$$-c_1 B(t,T) - c_2 B(t,T) \cdot e^{h(T-t)} = -c_1 c_2 \cdot e^{h(T-t)} + c_1 c_2.$$

Hence, we use (B.7) and obtain for $B(t,T)$

$$\begin{aligned} B(t,T) &= \frac{c_1 c_2 (e^{h(T-t)} - 1)}{c_1 + c_2 \cdot e^{h(T-t)}} = \frac{\frac{2}{\sigma^2}(e^{h(T-t)} - 1)}{\frac{1}{\sigma^2}(-k+h) + \frac{1}{\sigma^2}(k+h) \cdot e^{h(T-t)}} \\ &= \frac{2(e^{h(T-t)} - 1)}{2h - (k+h) + (k+h) \cdot e^{h(T-t)}} \\ &= \frac{2(e^{h(T-t)} - 1)}{2h + (k+h)(e^{h(T-t)} - 1)}. \end{aligned} \tag{B.9}$$

Now, when we have $B(t,T)$, we can derive the representation for $A(t,T)$ by integrating the equation (B.6) on both sides over the interval $[t,T]$ as follows:

$$\underbrace{\ln A(T,T)}_{=0} - \ln A(t,T) = \int_t^T k\theta B(u,T)\ du.$$

Hence, we obtain

$$A(t,T) = \exp\left(-k\theta \int_t^T B(u,T)\ du\right). \tag{B.10}$$

From (B.9), we have

$$\begin{aligned} &\int_t^T B(u,T)\,\mathrm{d}u = \int_t^T \frac{2(e^{h(T-u)} - 1)}{2h + (k+h)(e^{h(T-u)} - 1)}\ du \\ &= 2\underbrace{\int_t^T \frac{e^{h(T-u)}}{2h + (k+h)(e^{h(T-u)} - 1)}\ du}_{=:\mathbf{I}} \\ &-2\underbrace{\int_t^T \frac{1}{2h + (k+h)(e^{h(T-u)} - 1)}\ du}_{=:\mathbf{II}}. \end{aligned}$$

We compute the integrals **I** and **II** as follows:

$$
\begin{aligned}
\mathbf{I} &= \int_t^T \frac{e^{h(T-u)}}{2h+(k+h)(e^{h(T-u)}-1)}\,\mathrm{d}u \\
&= -\frac{1}{h(k+h)} \cdot \ln\left(2h+(k+h)(e^{h(T-u)}-1)\right)\Big|_t^T \\
&= -\frac{1}{h(k+h)} \cdot \ln\left(\frac{2h}{2h+(k+h)(e^{h(T-u)}-1)}\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{II} &= \int_t^T \frac{1}{2h+(k+h)(e^{h(T-u)}-1)}\,\mathrm{d}u \\
&= \int_t^T \frac{1}{e^{h(T-u)}\cdot(k+h)+(h-k)}\,\mathrm{d}u \\
&= \frac{1}{h-k}\int_t^T 1-\frac{(k+h)e^{h(T-u)}}{e^{h(T-u)}\cdot(k+h)+(h-k)} \\
&= \frac{1}{h-k}\left[u+\frac{1}{h}\ln\left(2h+(k+h)(e^{h(T-u)}-1)\right)\right]_t^T \\
&= \frac{1}{h-k}\left[(T-t)+\ln\left(\frac{2h}{2h+(k+h)(e^{h(T-t)}-1)}\right)\right].
\end{aligned}
$$

Hence, we have:

$$
\begin{aligned}
\int_t^T B(u,T)\,\mathrm{d}u &= 2\Big\{\frac{-1}{h(k+h)}\big(\ln 2h - \ln(e^{h(T-t)}\cdot(k+h)+2h)\big) \\
&- \frac{T-t}{h-k} - \frac{1}{h(h-k)}\big(\ln 2h - \ln(e^{h(T-t)}\cdot(k+h)+2h)\big)\Big\} \\
&= 2\Bigg\{-\frac{\ln 2h}{h(k+h)} - \frac{\ln 2h}{h(h-k)} - \frac{T-t}{h-k} \\
&\quad + \ln\big(e^{h(T-t)}\cdot(k+h)+2h\big)\left(\frac{1}{h(k+h)}+\frac{1}{h(h-k)}\right)\Bigg\} \\
&= 2\Bigg\{\ln 2h\cdot\left(\frac{-(h-k)-(k+h)}{h(h^2-k^2)}\right) - \frac{T-t}{h-k} \\
&\quad + \ln\big(e^{h(T-t)}\cdot(k+h)+2h\big)\cdot\frac{h-k+k+h}{h(h^2-k^2)}\Bigg\} \\
&= -\frac{2}{\sigma^2}\left(\sigma^2\cdot\frac{T-t}{h-k} + \ln\left(\frac{2h}{2h+(k+h)(e^{h(T-t)}-1)}\right)\right) \\
&= -\frac{2}{\sigma^2}\ln\left(\exp\left(\frac{\sigma^2(T-t)}{h-k}\right)\cdot\frac{2h}{2h+(k+h)(e^{h(T-t)}-1)}\right).
\end{aligned}
$$

Putting the above calculation and equation (B.10) together, we obtain:

$$
\begin{aligned}
A(t,T) &= \exp\left(\frac{2k\theta}{\sigma^2}\ln\left(\frac{2h\cdot e^{\frac{\sigma^2(T-t)}{h-k}}}{2h+(k+h)(e^{h(T-t)}-1)}\right)\right) \\
&= \frac{2h\cdot e^{\frac{\sigma^2(T-t)}{h-k}}}{2h+(k+h)(e^{h(T-t)}-1)}.
\end{aligned}
$$

With

$$
\frac{\sigma^2(T-t)}{h-k} = \frac{2\sigma^2(T-t)}{2(h-k)} = \frac{(h^2-k^2)(T-t)}{2(h-k)} = \frac{(k+h)(T-t)}{2}
$$

we complete the proof. □

Bibliography

[1] Abramowitz, M., Stegun I. A. (1965). *Handbook of mathematical functions.* Dover Publications, New York.

[2] Askey, R. (1974). *Orthogonal Polynomials and Special Functions.* CBMS Regional Conference Series, Volume 21. Society for Industrial and Applied Mathematics, Philadelphia.

[3] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

[4] Bhattacharya, R. N. and Waymire, E. C. (1990). *Stochastic Processes with Applications.* John Wiley, New York.

[5] Black, F. and Karasinski (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal* **47,** 52–59.

[6] Bingham, N. H. (1991). Fluctuation theory for the Ehrenfest urn. *Adv. Appl. Prob.* **23,** 598–611.

[7] Björk, T. (2004). *Arbitrage Theory in Continuous Time*, 2nd edn. Oxford University Press, New York.

[8] Brémaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, New York.

[9] Brigo, D. and Mercurio, F. (2005). *Interest Rate Models - Theory and Practice.* Springer, Berlin.

[10] Buehlmann, H. (1994). Continuous and discrete models in finance, in particular for stochastic interest rates. *Riv. Mat. Sci. Econom. Social.* **17,** 3–20.

[11] Cox, J. C., Ingersoll, J. E. and Ross S.A. (1985). A Theory of the Term Structure of Interest Rates. *Math. Finance* **64,** 379–406.

[12] CONWAY, J. B. (2000). *A Course in Operator Theory.* Graduate Studies in Mathematics, Volume 21. Oxford University Press, New York.

[13] DELBAEN, F. AND SHIRAKAWA, H. (2002). An interest rate model with upper and lower bounds. *Asia-Pacific Finan. Markets* **9,** 191–209.

[14] Deutsche Bundesbank - Term structure of interest rates on listed Federal securities (method by Svensson), by residual maturity (daily data). Last update: 11 June 2010, at 11:07. `http://www.bundesbank.de/statistik/statistik_zinsen.en.php`.

[15] DUFFIE, D. (1996). *Dynamic Asset Pricing Theory*, 2nd edn. Princeton University Press, New Jersey.

[16] ELLIOTT, R. J., HUNTER, W.C. AND JAMIERSON, B.M. (1998). Drift and Volatility estimation in discrete time. *J. Econ. Dynamics Control* **22,** 209–218.

[17] ENGELBERT H.J. AND SCHMIDT W. (1984). On one-dimensional stochastic differential with generalized drift. *Lecture Notes in Control and Information Sciences* **69**, 143–155, Springer, Berlin.

[18] ERDÉLYI, A. (1953). *Higher Transcendental Functions*, Vol. 2. McGraw-Hill, New York.

[19] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer, Berlin.

[20] GROSS, K. I. AND RICHARDS, D. ST. P. (1987). Special functions of matrix argument. I. Algebraic induction, zonal polynomials and hypergeometric functions, *Trans. Amer. Math. Soc.* **301**, 781–811.

[21] GROSS, K. I. AND RICHARDS, D. ST. P. (1989). Total positivity, spherical series, and hypergeometric functions of matrix argument. *J. Approx. Theory* **59**, no. 2, 224–246.

[22] HARRISON, J.M. AND KREPS, D.M. (1979). Martingales and arbitrage in multiperiod securities markets. *J. Econ. Theory* **20**, no. 3, 381–408.

[23] HARRISON, J.M. AND PLISKA, S.R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes Appl.* **11**, 215–260.

[24] HARRISON, J.M. AND PLISKA, S.R. (1983). A stochastic calculus model of continuous trading: complete markets. *Stochastic Processes Appl.* **15**, 313–316.

[25] HULL, J. AND WHITE, A. (1990). Pricing interest rate derivative securities *The Review of Financial Studies* **3,** 573–592.

[26] INCE, E.L. (1956) *Ordinary Differential Equations*, Dover Publications, New York.

[27] IOSEFESCU, M. (1980) *Finite Markov Processes and Their Applications*, John Willey & Sons, New York.

[28] JAMES, J. AND WEBBER, N. (2004). *Interest Rate Modelling*, Chichester, Wiley.

[29] JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions*, Vol. 1. Wiley, New York.

[30] JOHNSON, N. L., KOTZ, S. AND BALAKRISHNAN, N. (1995). *Continuous Univariate Distributions*, Vol. 2. Wiley, New York.

[31] KAC, M. (1947). Random walk and the theory of Brownian motion. *Amer. Math. Monthly* **54,** 369–391.

[32] KARATZAS, I. AND SHREVE, S. E. (1988). *Brownian Motion and Stochastic Calculus*, Springer, New York.

[33] KARLIN, S. AND TAYLOR, H. M. (1975). *A First Course in Stochastic Processes*, 2nd edn. Academic Press, New York.

[34] KARLIN, S. AND TAYLOR, H. M. (1981). *A Second Course in Stochastic Processes*. Academic Press, New York.

[35] KARLIN, S. AND MCGREGOR, J. (1965). Ehrenfest urn models. *J. Appl. Prob.* **2,** 352–376.

[36] KOEV, P. AND EDELMAN, A. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Math. Comput.* **75,** 833–846.

[37] KORN, R. UND KORN, E. (2001) *Optionsbewertung und Portfolio-Optimierung*. Vieweg, Braunschweig.

[38] Kraft, O. and Schaefer, M. (1993). Mean passage times for tridiagonal transition matrices and a two-parameter Ehrenfest urn model. *J. Appl. Prob.* **30,** 964–970.

[39] Lakner, P. (2006). Martingale measures for a class of right-continuous processes. *Math. Finance* **3,** no. 1, 43–53.

[40] Larsen, K. S., Sørensen, M. (2007). Diffusion models for exchange rates in a target zone. *Math. Finance* **17,** no. 2, 285–306.

[41] Linetsky, V. (2005). On the transition densities for reflected diffusions. *Adv. Appl. Prob.* **37,** 435–460.

[42] Linetsky, V. (2007). Spectral methods in derivatives pricing. *Handbook of Financial Engineering, Handbooks in Operations Research and Management Sciences* Vol. 15. Elsevier, Amsterdam.

[43] Milian, A. (1992). Transition probability density of a certain diffusion process concentrated on a finite spatial interval. *J. Appl. Prob.* **29,** 334–342.

[44] Mitrinovic, D. S. (1970). *Analytic Inequalities*. Springer, Berlin.

[45] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*. Springer, Berlin.

[46] Norberg, R. (1995). A time-continuous Markov chain interest model with applications to insurance. *J. Appl. Stoch. Models and Data Analysis* **11,** 245–256.

[47] Øksendal B. (1998). *Stochastic Differential Equations: An Introduction with Applications*. Springer, Berlin.

[48] Okunev, J. and Tippett, M. (1990). Continuous-time stochastic calculus: A survey of applications in finance and accountancy. *IMA J. Math. Appl. Bus. Ind.* **2,** no. 2, 157–171.

[49] Polyanin, A. D. and Zaitsev, V. F. (2004). *Handbook of Nonlinear Partial Differential Equations*. Chapman & Hall/CRC, New York.

[50] Rebonato, R. (1996). *Interest-rate Option Models*, 2nd edn. John Wiley, New York.

[51] Reitz, S., Schwarz W. and Martin, M.R.W. (2004). *Zinsderivate.* Vieweg, Wiesbaden.

[52] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin.

[53] Rogers, L. C. G. and Williams, D. (1987). *Diffusions, Markov Processes, and Martingales, Vol. 2*. John Wiley, New York.

[54] Rogers, L. C. G. (1995). Which Model for Term Structure of Interest Rates Should One Use? In M. Davis, D. Duffie, W. Fleming and S. Shreve (eds.), Mathematical Finance. IMA Vol. Math. Appl. Vol. 65. Springer, New York.

[55] Shreve, S. E. (2000). *Stochastic calculus for Finance II. Continuous-Time Models*. Springer, Berlin.

[56] Schoutens, W. (2000). *Stochastic Processes and Orthogonal Polynomials*, Lecture Notes in Statistics, **146**. Springer, New York.

[57] Siegert, A. J. F. (1949). On the approach to the statistical equilibrium. *Phys. Rev.* **71,** 1708–1714.

[58] Sumita, U., Gotoh, J. and Jin, H. (2006). Numerical exploration of dynamic behavior of Ornstein-Uhlenbeck process via Ehrenfest process approximation. *J. Oper. Res. Japan* **49,** 256–278.

[59] Stroock, D. W., Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, Berlin.

[60] Szegö, G. (1983). *Orthogonal Polynomials*. Amer. Math. Soc. Colloquium Publ. 23, AMS. Providence, RI.

[61] Takahashi, H. (2004). Ehrenfest model with large jumps in finance. *Physica D* **189,** no. 1-2, 61–69.

[62] Vasicek, O. (1977). An equilibrium characterization of the term structure. *J. Financ. Econ.* **5,** 177–188.

[63] Veraart, A. E. D. and Veraart, L. A. M. (2009) Stochastic volatility and stochastic leverage. CREATES Research Papers 2009-20.

[64] Voit, M. (1996). Asymptotic distributions for the Ehrenfest urn and related random walks. *J. Appl. Prob.* **33,** no. 3, 340–356.

[65] Zagst, R. (2002). *Interest-Rate Management*. Springer, New York.