YIELD CURVE SHAPES AND THE ASYMPTOTIC SHORT RATE DISTRIBUTION IN AFFINE ONE-FACTOR MODELS

MARTIN KELLER-RESSEL AND THOMAS STEINER

Abstract. We consider a model for interest rates, where the short rate is given by a time-homogenous, one-dimensional affine process in the sense of Duffie, Filipović, and Schachermayer. We show that in such a model yield curves can only be normal, inverse or humped (i.e. endowed with a single local maximum). Each case can be characterized by simple conditions on the present short rate r_t. We give conditions under which the short rate process will converge to a limit distribution and describe the limit distribution in terms of its cumulant generating function. We apply our results to the Vasiček model, the CIR model, a CIR model with added jumps and a model of Ornstein-Uhlenbeck type.

1. Introduction

We consider a model for the term structure of interest rates, where the short rate $(r_t)_{t\geq 0}$ is a one-dimensional conservative affine process in the sense of Duffie, Filipović, and Schachermayer [2003]. An affine short rate process of this type will lead to an exponentially-affine structure of zero-coupon bond prices and thus also to an affine term structure of yields and forward rates.

We emphasize here that the definition of Duffie et al. [2003] is not limited to diffusions, but also includes processes with jumps and even with jumps whose intensity depends in an affine way on the short rate process itself. The class of models we consider naturally includes the Vasiček model, the CIR model and variants of them that are obtained by adding jumps, such as the JCIR-model of Brigo and Mercurio [2006, Section 22.8]. Since they are the best-known, the two 'classical' models of Vasiček and Cox-Ingersoll-Ross will serve as the starting point for our discussion of yield curve shapes:

A common criticism of the (time-homogenous) CIR and the Vasiček model is that they are not flexible enough to accommodate more complex shapes of yield curves, such as curves with a dip (a local minimum), curves with a dip and a hump, or other shapes that are frequently observed in the markets. Often these shortcomings

Date: March 14, 2007.

2000 *Mathematics Subject Classification.* 60J25, 91B28.

Key words and phrases. affine process, term structure of interest rates, Ornstein-Uhlenbeck process, yield curve.

Supported by the Austrian Science Fund (FWF) through project P18022 and the START programm Y328.

Supported by the module M5 "Modelling of Fixed Income Markets" of the PRisMa Lab, financed by Bank Austria and the Republic of Austria through the Christian Doppler Research Association.

Both authors would like to thank Josef Teichmann for most valuable discussions and encouragement. We also thank various proof-readers at FAM for their comments.

are explained by 'too few parameters' in the model (cf. Carmona and Tehranchi [2006, Section 2.3.5] or Brigo and Mercurio [2006, Section 3.2]). However if jumps are added to the mentioned models, additional parameters (potentially infinitely many) are introduced through the jump part, while the model still remains in the scope of affine models. It is not clear per se what consequences the introduction of jumps will have for the range of attainable yield curves, and this is one question we intend to answer in this article.
Moreover, there seems to be some confusion about what shapes of yield curves are actually attainable even in well-studied models like the CIR-model. While most sources (including the original paper of Cox et al. [1985]) mention inverse, normal and humped shapes, Carmona and Tehranchi [2006, Section 2.3.5] write that *'tweaking the parameters [of the CIR model] can produce yield curves with one hump or one dip'*, and Brigo and Mercurio [2006, Section 3.2] state that *'some typical shapes, like that of an inverted yield curve, may not be reproduced by the [CIR or Vasiček] model.'* In our main result, Theorem 3.10, we settle this question and prove that in any time-homogenous, affine one-factor model the attainable yield curves are either inverse, normal or humped. The proof will rely only on tools of elementary analysis and on the characterization of affine processes through the generalized Riccati equations of Duffie et al. [2003].
Another related problem is how the shape of the yield curve is determined by the parameters of the model, and also how – when the parameters are fixed – the yield curve is determined by the level of the current short rate. We show in Section 4.2 that also in this respect the CIR model has not been completely understood and discuss a misconception that originates in [Cox et al., 1985] and is repeated for example in [Rebonato, 1998].

In Section 3.3 we give some results on the limit distributions of the short rate process $(r_t)_{t\geq 0}$. It is well-known that in the Vasiček model the short rate process converges in law to a Gaussian distribution. For general processes of OU-type corresponding results have been given by Jurek and Vervaat [1983] and Sato and Yamazato [1984]. We extend these results in Theorem 3.18 to the class of affine processes.
We conclude our article in Section 4 by applying the theoretical results to several interest rate models, such as the Vasiček model, the CIR model, the JCIR model and a Ornstein-Uhlenbeck-type model.

2. Preliminaries

In this section we collect some key results on affine processes from Duffie et al. [2003]. In their article affine processes are defined on the $(m+n)$-dimensional state space $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, and we will try to simplify notation where this is possible in the one-dimensional case.

Definition 2.1 (One-dimensional affine process)**.** A time-homogenous Markov process $(r_t)_{t\geq 0}$ with state space $D = \mathbb{R}_{\geqslant 0}$ or $\mathbb{R}$ and its semi-group $(P_t)_{t\geq 0}$ are called *affine*, if the characteristic function of its transition kernel $p_t(x, .)$, given by

$$\widehat{p_t}(x,u) = \int_D e^{u\xi}\; p_t(x, d\xi)$$

and defined (at least) on

$$\mathcal{U} = \begin{cases} \{u \in \mathbb{C} : \operatorname{Re} u \leq 0\} & \text{for} \quad D = \mathbb{R}_{\geqslant 0} \,, \\ \{u \in \mathbb{C} : \operatorname{Re} u = 0\} & \text{for} \quad D = \mathbb{R} \,, \end{cases} \tag{2.1}$$

is *exponentially affine* in x. That is, there exist $\mathbb{C}$-valued functions $\phi(t,u)$ and $\psi(t,u)$, defined on $\mathbb{R}_{\geqslant 0} \times \mathcal{U}$, such that

$$\widehat{p}_t(x,u) = \exp\left(\phi(t,u) + x\psi(t,u)\right) \quad \text{for all} \quad x \in D, u \in \mathcal{U} \,. \tag{2.2}$$

We will say that a one-dimensional affine process is of

- **Type A** if it has state space $\mathbb{R}_{\geqslant 0}$,
- **Type B** if it has state space $\mathbb{R}$.

Results on one-dimensional affine processes of type A can also be found in Filipović [2001].

For subsequent results the following regularity condition for $(r_t)_{t \geq 0}$ will be needed:

Definition 2.2. An affine process is called *regular* if it is stochastically continuous and the right hand derivatives

$$\partial_t^+ \phi(t,u)|_{t=0} \qquad \text{and} \qquad \partial_t^+ \psi(t,u)|_{t=0}$$

exist for all $u \in \mathcal{U}$ and are continuous at $u = 0$.

Definition 2.3. The parameters $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ are called *admissible* for a process of type A if

$$a = 0,$$
$$\alpha, b, c, \gamma \in \mathbb{R}_{\geqslant 0} \,,$$
$$\beta \in \mathbb{R} \,,$$

m, μ are Lévy measures on $(0, \infty)$, where m satisfies

$$\int_{(0,\infty)} (x \wedge 1)\, m(dx) < \infty \,,$$

and admissible for a process of type B if

$$a, c \in \mathbb{R}_{\geqslant 0} \,,$$
$$b, \beta \in \mathbb{R} \,,$$

m is a Lévy measure on $\mathbb{R} \setminus \{0\}$,

$$\alpha = 0, \gamma = 0, \mu \equiv 0 \,.$$

Moreover let h_F, h_R be truncation functions defined on $D \setminus \{0\}$ – where h_F can be chosen as 0 in case A – and define the functions $F(u)$, $R(u)$ for $u \in \mathcal{U}$ as

$$F(u) = au^2 + bu - c + \int_{D\setminus\{0\}} \left(e^{u\xi} - 1 - uh_F(\xi)\right) m(d\xi) \,, \tag{2.3}$$

$$R(u) = \alpha u^2 + \beta u - \gamma + \int_{D\setminus\{0\}} \left(e^{u\xi} - 1 - uh_R(\xi)\right) \mu(d\xi) \,. \tag{2.4}$$

The next result is a one-dimensional version of the key result of Duffie et al. [2003]:

Theorem 2.4 (Duffie, Filipović, and Schachermayer, Theorem 2.7)**.** *Suppose* $(r_t)_{t\geq 0}$ *is a one-dimensional regular affine process. Then it is a Feller process. Let* $\mathcal{A}$ *be its infinitesimal generator. Then* $C_c^\infty(D)$ *is a core of* $\mathcal{A}$*,* $C_c^2(D) \subseteq \mathcal{D}(\mathcal{A})$ *and there exist some admissible parameters* $(a, \alpha, b, \beta, c, \gamma, m, \mu)$ *such that, for* $f \in C_c^2(D)$*,*

$$\begin{aligned}\mathcal{A}f(x) = (a+\alpha x)f''(x) + (b+\beta x)f'(x) - (c+\gamma x)f(x) + \\ + \int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - f'(x)h_F(\xi)\right) m(d\xi) + \\ + x\int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - f'(x)h_R(\xi)\right) \mu(d\xi)\ .\end{aligned} \tag{2.5}$$

Moreover (2.2) *is defined for all* $(t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U}$ *where* $\phi(t,u)$ *and* $\psi(t,u)$ *are the unique solutions of the* generalized Riccati equations

$$\begin{aligned}\partial_t\,\phi(t,u) &= F\left(\psi(t,u)\right), \qquad \phi(0,u) = 0\ ,\\ \partial_t\,\psi(t,u) &= R\left(\psi(t,u)\right), \qquad \psi(0,u) = u\ .\end{aligned} \tag{2.6}$$

Conversely let $(a, \alpha, b, \beta, c, \gamma, \mu, m)$ *be some admissible parameters. Then there exists a unique regular affine semigroup* $(P_t)_{t\geq 0}$ *with infinitesimal generator* (2.5) *and* (2.2) *holds for all* $(t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U}$ *where* $\phi(t,u)$ *and* $\psi(t,u)$ *are given by* (2.6)*.*

Closely related to affine processes is the notion of a Ornstein-Uhlenbeck (OU-)type process. These processes are of some importance, since they usually offer good analytic tractability and have been studied for longer than affine processes. Following Sato [1999, Chapter 17] a (D-valued) OU-type process $(X_t)_{t\geq 0}$ can be defined as the solution of the Langevin SDE

$$dX_t = -\lambda X_t\,dt + dL_t, \qquad \lambda \in \mathbb{R}, X_t \in D,$$

where $(L_t)_{t\geq 0}$ is a Lévy process, often called background driving Lévy process (BDLP). In an equivalent definition, a OU-type process is a time-homogenous Markov process, whose transition kernel $p_t(x,.)$ has the characteristic function

$$\widehat{p_t}(x,u) = \exp\left(\int_0^t F(e^{-\lambda s}u)\,ds + xe^{-\lambda t}u\right)\ ,$$

where $F(u)$ is the characteristic exponent of $(L_t)_{t\geq 0}$. From the last equation it is immediately seen that every OU-type process is an affine process in the sense of Definition 2.1. It is also seen that in the generalized Riccati equations (2.6) for a OU-type process necessarily $R(u) = -\lambda u$. Comparing this with (2.4) and Definition 2.3, it is seen that *any* affine process of type B is a process of OU-type. In addition, a closer look at the admissibility conditions of Definition 2.3 reveals that there exist also affine processes of type A that are of OU-type. Such processes have state space $\mathbb{R}_{\geqslant 0}$ and consequently we call them non-negative OU-type processes. We will give an example of such a process in Section 4.4.

Naturally we will not only be interested in the short rate process $(r_t)_{t\geq 0}$ itself, but also in its integral $\int_0^t r_s\,ds$ and in quantities of the type

$$\mathcal{Q}_t\,f(x) := \mathbb{E}\left[\exp\left(-\int_0^t r_s\,ds\right) f(r_t)\middle|\, r_0 = x\right]\ , \tag{2.7}$$

where f is a bounded function on D. Let θ_t be the shift operator, then by the Markov property of $(r_t)_{t\geq 0}$

$$\begin{aligned}\mathcal{Q}_{t+s}f(x) = \mathbb{E}_x\left[\exp\left(-\int_0^{t+s} r_u\,du\right) f(r_{t+s})\right] = \\ = \mathbb{E}_x\left[\exp\left(-\int_0^{t} r_u\,du\right)\cdot \mathbb{E}\left[\exp\left(-\int_0^{s} r_u\,du\right) f(r_s)\circ\theta_t \middle|\, \mathcal{F}_t\right]\right] = \\ = \mathbb{E}_x\left[\exp\left(-\int_0^{t} r_u\,du\right)\mathcal{Q}_s f(r_t)\right] = \mathcal{Q}_t\mathcal{Q}_s f(x) \qquad \text{for all } f\in B(D)\end{aligned}$$

and $(\mathcal{Q}_t)_{t\geq 0}$ forms a semigroup. The next result is an application of the Feynman-Kac formula for Feller semigroups (cf. Rogers and Williams [1994, Section III.19]) and can also be found in Duffie et al. [2003]. It relies on the positivity of $(r_t)_{t\geq 0}$ and is therefore only applicable if $(r_t)_{t\geq 0}$ is of type A.

Proposition 2.5 (Duffie, Filipović, and Schachermayer, Proposition 11.1)**.** *Let $(r_t)_{t\geq 0}$ be a one-dimensional, regular affine process of type A. Then the family $(\mathcal{Q}_t)_{t\geq 0}$ defined by* (2.7) *forms a regular, affine semigroup with infinitesimal generator*

$$\mathcal{B}f(x) = \mathcal{A}f(x) - xf(x) \quad \textit{for all} \quad f\in C_c^2(D)\;.$$

We will make extensive use of the convexity and continuous differentiability of the functions F and R from Definition 2.3. These properties are established in this Lemma:

Lemma 2.6. *If $c=\gamma=0$ then F, R as defined in Definition 2.3 have the following properties:*

(i) $R(0)=0$ and $F(0)=0$.
(ii) If $F(u)<\infty$ on $(c_1,c_2)\subseteq\mathbb{R}$, then F is either strictly convex on (c_1,c_2) or $F(u)=bu$ for all $u\in\mathbb{R}$. The same holds for R with b replaced by β.
(iii) If $F(u)<\infty$ on $(c_1,c_2)\subseteq\mathbb{R}$, then F is continuously differentiable on (c_1,c_2). Also the one-sided derivatives at c_1 and c_2 are defined but may take the values $-\infty$ (at c_1) and $+\infty$ (at c_2). The same holds for R.

Proof. Property (i) is obvious. For Property (ii) note that by the Lévy-Khintchine formula there exists an infinitely divisible random variable X, such that F is its cumulant generating function, i.e.

$$F(u) = \log\mathbb{E}\left[e^{uX}\right] \quad \text{for} \quad u\in(c_1,c_2)\;. \tag{2.8}$$

Choosing two distinct numbers $u,v\in(c_1,c_2)$, we apply the Cauchy-Schwarz inequality to

$$F\left(\frac{u+v}{2}\right) = \log\mathbb{E}\left[e^{\frac{uX}{2}}\cdot e^{\frac{vX}{2}}\right] \leq \log\sqrt{\mathbb{E}[e^{uX}]\cdot\mathbb{E}[e^{vX}]} = \frac{F(u)+F(v)}{2}$$

which shows convexity of F. The inequality is strict unless there exists some $c\neq 0$ such that $e^{uX}=ce^{vX}$ almost surely. This can only be the case if X is constant a.s., in which case F is linear. The same argument applies to R. Property (iii) follows from the convexity and from the fact that F and R are analytic on $\{u\in\mathbb{C}:\operatorname{Re}u\in(c_1,c_2)\}$ (cf. Lukacs [1960, Chapter 7]). ∎

3. Theoretical Results

We will now use the theory from the last section to calculate bond prices, yields and other quantities in an interest rate model where the short rate follows a one-dimensional regular affine process $(r_t)_{t\geq 0}$. Naturally we will also make the assumption that $(r_t)_{t\geq 0}$ is conservative, i.e. that $p_t(x, D) = 1$ for all $(t, x) \subseteq \mathbb{R}_{\geqslant 0} \times D$. This implies by Duffie et al. [2003, Proposition 9.1] that $c = \gamma = 0$ in Definition 2.3. We will need some additional assumptions which are summarized in the following condition:

Condition 3.1. The one-dimensional affine process $(r_t)_{t\geq 0}$ is assumed to be regular and conservative. In addition, if the process is of type B, such that we have $R(u) = \beta u$ by Definition 2.3, we require that

$$F(u) < \infty \quad \text{for all} \quad u \in \begin{cases} (1/\beta, 0] & \text{if} \quad \beta < 0 \ , \\ (-\infty, 0] & \text{else} \ . \end{cases}$$

It will be seen that the condition on F is necessary to guarantee global existence of bond prices in the term structure model. Before we calculate the bond prices we define a quantity that will generalize the coefficient of mean reversion from OU-type processes:

Definition 3.2 (quasi-mean-reversion)**.** Given a one-dimensional conservative affine process $(r_t)_{t\geq 0}$, define the *quasi-mean-reversion* λ as the positive solution of

$$R(-1/\lambda) = 1 \ . \tag{3.1}$$

If there is no positive solution we set $\lambda := 0$.

Since R is by Lemma 2.6 a convex function satisfying $R(0) = 0$, it is easy to see that (3.1) can have at most one solution and thus λ is well-defined. The name *quasi-mean-reversion* is derived from the fact that if $(r_t)_{t\geq 0}$ is a process of OU-type with positive mean-reversion, then $R(u) = \beta u$ and the quasi-mean-reversion $\lambda = -\beta$ is exactly the coefficient of mean reversion of $(r_t)_{t\geq 0}$. When the process $(r_t)_{t\geq 0}$ satisfies Condition 3.1, it is also seen that both F and R are at least defined on $(-1/\lambda, 0]$.

3.1. Bond Prices.

We consider now the price $P(t, t + x)$ of a zero-coupon bond with time to maturity x, at time $t \in \mathbb{R}_{\geqslant 0}$, given by

$$P(t, t + x) = \mathbb{E}\left[\exp\left(-\int_t^{t+x} r_s \, ds\right) \middle| \mathcal{F}_t\right] \ .$$

The affine structure of $(r_t)_{t\geq 0}$ carries over to the bond prices, and we get the following result:

Proposition 3.3. *Let the short rate be given by a one-dimensional affine process $(r_t)_{t\geq 0}$ satisfying Condition 3.1.*
Then the bond price $P(t, t + x)$ exists for all $t, x \geq 0$ and is given by

$$P(t, t + x) = \exp\left(A(x) + r_t B(x)\right) \tag{3.2}$$

where A and B are the unique, globally defined solutions of the generalized Riccati equations

$$\partial_x A(x) = F\left(B(x)\right) \qquad A(0) = 0 \ , \tag{3.3}$$

$$\partial_x B(x) = R\left(B(x)\right) - 1 \qquad B(0) = 0 \ . \tag{3.4}$$

Proof. If $(r_t)_{t\geq 0}$ is of type A, then the proposition follows directly from Proposition 2.5 by noting that $P(t,t+x) = \mathcal{Q}_x\, 1$.
If $(r_t)_{t\geq 0}$ is of type B then, as discussed after Theorem 2.4, it is a process of OU-type and $R(u)$ has the simple structure $R(u) = \beta u$. In this case the statement can be proved by a direct approach: As a OU-type process, $(r_t)_{t\geq 0}$ is the solution of the SDE

$$dr_t = \beta r_t\, dt + dL_t, \quad t \geq 0 \tag{3.5}$$

where $(L_t)_{t\geq 0}$ is a Lévy process with characteristic exponent $F(u)$. The solution of (3.5) is given by

$$r_t = e^{\beta t} r_0 + \int_0^t e^{\beta(t-s)}\, dL_s, \quad t \geq 0$$

(cf. Sato [1999, Chapter 17]), and the integrated process by

$$-\int_t^{t+x} r_s\, ds = B(x) r_t + \int_0^x B(x-s)\, dL_s, \quad t, x \geq 0 \tag{3.6}$$

where $B(x) = (1 - e^{\beta x})/\beta$ if $\beta \neq 0$ and $B(x) = -x$ when $\beta = 0$. It is easily checked that B is continuously differentiable and indeed solves the generalized Riccati equation (3.4). Also note that B is strictly decreasing from $B(0) = 0$ to its limit

$$\lim_{x\to\infty} B(x) = \begin{cases} 1/\beta & \text{if} \quad \beta < 0\,, \\ -\infty & \text{else}\,. \end{cases}$$

Let now $0 = t_0 < t_1 < \cdots < t_{N+1} = x$ be a partition of the interval $[0, x]$. Then it holds by independence and stationarity of the increments of the Lévy process $(L_t)_{t\geq 0}$ that

$$\mathbb{E}\left[\exp\left(\sum_{k=0}^N B(x-t_k)(L_{t_{k+1}} - L_{t_k})\right)\right] = \prod_{k=0}^N \mathbb{E}\left[\exp\left(B(x-t_k)(L_{t_{k+1}} - L_{t_k})\right)\right] =$$
$$= \prod_{k=0}^N \exp\left(F\left(B(x-t_k)\right)(t_{k+1}-t_k)\right) = \exp\left(\sum_{k=0}^N F\left(B(x-t_k)\right)(t_{k+1}-t_k)\right)\,.$$

As the mesh of the partition goes to zero, the exponential on the left hand side converges in probability – and thus also in distribution – to $\exp\left(\int_0^x B(x-s)\, dL_s\right)$. The right hand side converges to

$$\exp\left(\int_0^x F(B(x-s))\, ds\right) = \exp\left(\int_0^x F(B(s))\, ds\right)$$

which is finite by Condition 3.1. By standard results on non-negative random variables (cf. Kallenberg [1997, Lemma 3.11]), convergence of expectations is sufficient for uniform integrability, and the left hand side converges to $\mathbb{E}[\exp\left(\int_0^x B(x-s)\, dL_s\right)]$. Applying this to (3.6) we have

$$\mathbb{E}\left[\exp\left(-\int_t^{t+x} r_s\, ds\right)\right] = \exp\left(\int_0^x F(B(s))\, ds + r_t B(x)\right)\,,$$

and the bond prices are indeed of the form (3.2). □

Corollary 3.4. *Let $(r_t)_{t\geq 0}$ satisfy Condition 3.1 and have quasi-mean-reversion λ. Then the function $B(x)$ from Proposition 3.3 is strictly decreasing and satisfies*

$$\lim_{x\to\infty} B(x) = -1/\lambda \ .$$

Proof. Assume that $\lambda > 0$ and that $B(x) \in (-1/\lambda, 0]$ for some $x \in \mathbb{R}_{\geqslant 0}$. Then by the generalized Riccati equation (3.4) $\partial_x B(x) < 0$ and $B(y) < B(x)$ for some $y > x$. Therefore

$$\lim_{x\to\infty} B(x) \leq -\frac{1}{\lambda} \ .$$

On the other hand, if $B(x) = -1/\lambda$ for some $x \in \mathbb{R}_{\geqslant 0}$, then $\partial_x B(x) = 0$. The assertion follows now by $B(0) = 0$ and continuity of B.
In the case that $\lambda = 0$, find a sequence $(\lambda_k)_{k\in\mathbb{N}} \downarrow 0$. By the same argument as above

$$\lim_{x\to\infty} B(x) \leq -\frac{1}{\lambda_k} \quad \forall\, k \in \mathbb{N} \ .$$

The assertion follows by letting $k \to \infty$. □

3.2. **The Yield Curve and the Forward Rate Curve.** The next results are the central theoretical results of our paper and describe the global shapes of attainable yield curves in any affine one-factor term structure model.

Definition 3.5. The (zero-coupon) yield $Y(r_t, x)$ is given by $Y(r_t, 0) := r_t$ and

$$Y(r_t, x) := -\frac{\log P(t, t+x)}{x} = -\frac{A(x)}{x} - r_t \frac{B(x)}{x} \quad \text{for all} \quad x > 0 \ . \tag{3.7}$$

For r_t fixed, we call the function $Y(r_t, .)$ the **yield curve**.
The (instantaneous) forward rate $f(r_t, x)$ is given by $f(r_t, 0) := r_t$ and

$$f(r_t, x) := -\partial_x \log P(t, t+x) = -A'(x) - r_t B'(x) \quad \text{for all} \quad x > 0 \ . \tag{3.8}$$

For r_t fixed, we call the function $f(r_t, .)$ the **forward rate curve**.

By l'Hospital's rule and the generalized Riccati equations (3.3) and (3.4) it is seen that both the yield and the forward rate curve are continuous at 0.

The first quantity associated to the yield curve that we consider, is the asymptotic level b_{asymp} of the yield curve as $x \to \infty$, also known as long-term yield, consol yield or simply 'long end'.

Theorem 3.6. *Let the short rate process be given by a one-dimensional affine process $(r_t)_{t\geq 0}$ satisfying Condition 3.1 with quasi-mean-reversion λ. If $\lambda > 0$ then*

$$b_{\text{asymp}} = \lim_{x\to\infty} Y(r_t, x) = \lim_{x\to\infty} f(r_t, x) = -F(-1/\lambda) \ .$$

If $\lambda = 0$ then

$$b_{\text{asymp}}(r_t) = -F(-\infty) + r_t\,(1 - R(-\infty))$$

where $F(-\infty)$ and $R(-\infty)$ are understood as limits.

Remark 3.7. We will write $b_{\text{asymp}}(r_t)$ if the long-term yield depends on the current short rate r_t, and use the notation b_{asymp} in the (typical) case where it is independent of r_t.

Proof. By Corollary 3.4 we know that

$$\lim_{x\to\infty} B(x) = -1/\lambda \ .$$

Applying l'Hospital's rule we have

$$\lim_{x\to\infty} \frac{A(x)}{x} = \lim_{x\to\infty} A'(x) = \lim_{x\to\infty} \frac{1}{x}\int_0^x F(B(s))\,ds = \lim_{x\to\infty} F(B(x)) = F(-1/\lambda) \ .$$

If $\lambda > 0$ then by Corollary 3.4 we also have

$$\lim_{x\to\infty} \frac{B(x)}{x} = 0 \quad \text{and} \quad \lim_{x\to\infty} B'(x) = 0 \ .$$

If $\lambda = 0$ then $\lim_{x\to\infty} B(x) = -\infty$ and

$$\lim_{x\to\infty} \frac{B(x)}{x} = \lim_{x\to\infty} B'(x) = R(-\infty) - 1 \ . \qquad \square$$

The next result demonstrates the Dybvig-Ingersoll-Ross-Theorem in the case of affine processes. See Dybvig et al. [1996] for the original result and Hubalek et al. [2002] for a more general proof.

Corollary 3.8 ('Long forward rates never fall')**.** *Under the conditions of Theorem 3.6, the long forward rate $b_{\text{asymp}}(r_t)$ is non-decreasing in t almost surely.*

Proof. Suppose first that $R'(u_0) < 0$ for some $u_0 \in (-\infty, 0)$. Then there exists an affine minorant to R with negative slope, and we have that $\lim_{u\to-\infty} R(u) = \infty$. Since R is continuous and by Definition 3.2 this implies that $\lambda > 0$. In this case, $b_{\text{asymp}}(r_t)$ is constant by Theorem 3.6 and therefore trivially non-decreasing in $t \geq 0$.
Suppose now that $F'(u_0) < 0$ for some $u_0 \in (-\infty, 0)$ and that $R'(u) \geq 0$ for all $u \in (-\infty, 0]$. Then by the second assumption $\lambda = 0$ and by the same argument as before $\lim_{u\to-\infty} F(u) = \infty$. In this case $b_{\text{asymp}}(r_t) = -\infty$ by Theorem 3.6 and therefore also trivially non-decreasing in $t \geq 0$.
Suppose finally that $R(u_0) > 0$ for some $u_0 \in (-\infty, 0)$. Then since $R(0) = 0$ and R is convex we have $R'(u_0) < 0$, and this case is reduced to the first case.
For the remaining cases we can now assume that $R(u) \leq 0$ and that $F'(u)$ and $R'(u)$ are non-negative on $(-\infty, 0)$. This implies of course that

$$\lim_{u\to-\infty} F'(u) \geq 0 \qquad \text{and} \qquad \lim_{u\to-\infty} R'(u) \geq 0 \ . \tag{3.9}$$

By (2.3) and (2.4) it is known that both F and R are cumulant generating functions of some infinitely divisible random variables X and Y. By Lukacs [1960, Theorem 7.2.2] the limits in (3.9) are equal to the 'left extremities' of the random variables X and Y respectively, i.e. to the left boundary of their supports. It follows that the support of X and Y is a subset of $\mathbb{R}_{\geqslant 0}$. By standard results on non-negative infinitely divisible random variables (cf. Sato [1999, Theorem 21.5]) F and R must thus be of the form

$$F(u) = bu + \int_{(0,\infty)} \left(e^{u\xi} - 1\right)\, m(d\xi) \ ,$$

$$R(u) = \beta u + \int_{(0,\infty)} \left(e^{u\xi} - 1\right)\, \mu(d\xi)$$

where $b, \beta \in \mathbb{R}_{\geqslant 0}$. It is seen that in this case $(r_t)_{t\geq 0}$ has no diffusive part, positive jumps only, and is non-decreasing between jumps. Thus $(r_t)_{t\geq 0}$ is an almost surely non-decreasing process, which implies by Theorem 3.6, together with the assumption $R(-\infty) \leq 0$, that also $b_{\text{asymp}}(r_t)$ is non-decreasing almost surely. □

From Theorem 3.6 it is clear that for practical purposes mostly models with $\lambda > 0$ will be useful. So far we know that in this case the short end of the yield curve is given by $Y(r_t, 0) = r_t$ and the long end by $Y(r_t, \infty) = b_{\text{asymp}}$. We will now examine what happens between these two endpoints.

Definition 3.9. The yield curve $Y(r_t, x)$ is called

- **normal** if it is a strictly increasing function of x,
- **inverse** if it is a strictly decreasing function of x,
- **humped** if it has exactly one local maximum and no minimum on $(0, \infty)$.

In addition we call the yield curve **flat** if it is constant over all $x \in \mathbb{R}_{\geqslant 0}$.

This is our main result on the shapes of yield curves in affine one-factor models:

Theorem 3.10 (Yield Curve Shapes – Regular Case)**.** *Let the short rate process be given by a one-dimensional affine process $(r_t)_{t\geq 0}$ satisfying Condition 3.1. In addition suppose that $\lambda > 0$, $F \neq 0$ and that either F or R is non-linear. Then the following holds:*

- *The yield curve $Y(r_t, .)$ can only be normal, inverse or humped.*
- *Define*

$$b_{\text{norm}} := -\frac{F'(-1/\lambda)}{R'(-1/\lambda)} \qquad \text{and} \qquad b_{\text{inv}} := \begin{cases} -\dfrac{F'(0)}{R'(0)} & \text{if } R'(0) < 0 \\ +\infty & \text{if } R'(0) \geq 0\ . \end{cases}$$

  The yield curve is normal if $r_t \leq b_{\text{norm}}$, humped if $b_{\text{norm}} < r_t < b_{\text{inv}}$ and inverse if $r_t \geq b_{\text{inv}}$.

The above theorem is visualized in Figure 1. For its proof we will use a simple Lemma. We state the Lemma without proof since it follows in an elementary way from the usual definition of a convex function on $\mathbb{R}$.

Lemma 3.11. *A strictly convex or a strictly concave function on $\mathbb{R}$ intersects an affine function in at most two points. In the case of two intersection points $p_1 < p_2$, the convex function lies strictly below the affine function on the interval (p_1, p_2); if the function is concave it lies strictly above the affine function on (p_1, p_2).*

Proof of Theorem 3.10. Define the function $H(x)\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ by

$$H(x) := Y(r_t, x)x = -A(x) - r_t B(x)\ . \tag{3.10}$$

We will see that the convexity behavior of H will be crucial for the shape of the yield curve $Y(r_t, .)$. From the generalized Riccati equations (3.3) and (3.4) the first derivative of H is calculated as

$$\partial_x H(x) = -F(B(x)) - r_t \left(R(B(x)) - 1\right) \tag{3.11}$$

and the second as

$$\partial_{xx} H(x) = -B'(x) \left(F'(B(x)) + r_t R'(B(x))\right)\ . \tag{3.12}$$

Note that F and R are continuously differentiable by Lemma 2.6, and also B by (3.4), such that the second derivative of H is well-defined and continuous. Since B

is strictly decreasing by Corollary 3.4, the factor $-B'(x)$ is positive for all $x \in \mathbb{R}_{\geqslant 0}$. The sign of $\partial_{xx}H(x)$ therefore equals the sign of

$$k(x) := F'(B(x)) + r_t R'(B(x)) \ . \tag{3.13}$$

From the fact that B is decreasing and F and R are convex it is obvious that k must be decreasing. We will now show that k has at most a single zero in $[0, \infty)$:

(a) $D = \mathbb{R}_{\geqslant 0}$: We have assumed that either F or R is non-linear. By Lemma 2.6 this implies that either F or R is strictly convex, and thus that either F' or R' is strictly increasing. If $r_t > 0$, then it follows that k is strictly decreasing and thus has at most a single zero. If $r_t = 0$, an additional argument is needed: It could happen that F is of the form $F = bu$ such that $k(x) = b$ and k is no longer strictly decreasing. However, by assumption, $F \neq 0$ such that in this case k has no zero in $[0, \infty)$.

(b) $D = \mathbb{R}$: In this case, by the admissibility conditions in Definition 2.3, we have necessarily $R(u) = \beta u$. Also, since either F or R is non-linear, F must be non-linear and thus by Lemma 2.6 strictly convex. It follows that $k(x) = F'(B(x)) + r_t\beta$ is strictly decreasing and thus has at most a single zero in $[0, \infty)$.

We have shown that k is decreasing and has at most a single zero; to determine whether it has a zero for some value of r_t, we consider the two 'endpoints' $k(0)$ and $\lim_{x\to\infty} k(x)$. First we show that

$$k(0) \geq 0 \quad \text{if and only if} \quad r_t \leq b_{\text{inv}} := \begin{cases} -\dfrac{F'(0)}{R'(0)} & \text{if } R'(0) < 0 \\ +\infty & \text{if } R'(0) \geq 0 \ . \end{cases} \tag{3.14}$$

Since $B(0) = 0$ by Corollary 3.4 we have that

$$k(0) = F'(0) + r_t R'(0) \ .$$

We distinguish two cases:

(a) If $R'(0) < 0$ then the assertion (3.14) follows immediately.

(b) Consider the case that $R'(0) \geq 0$: Assume that $D = \mathbb{R}$. Then we have $R(u) = \beta u$ and $R'(0) = \beta \geq 0$. This, however, stands in contradiction to our assumption $\lambda > 0$, which implies that $\beta = -\lambda < 0$ (cf. Definition 3.2). Thus we must have $D = \mathbb{R}_{\geqslant 0}$ and $r_t \geq 0$; in this case it follows that $k(0) \geq 0$, *for all* $r_t \in D$, and we set $b_{\text{inv}} = +\infty$.

Next we consider the right end of $k(x)$ and show that

$$\lim_{x\to\infty} k(x) \leq 0 \quad \text{if and only if} \quad r_t \geq b_{\text{norm}} := -\frac{F'(-1/\lambda)}{R'(-1/\lambda)} \ . \tag{3.15}$$

Since $\lim_{x\to\infty} B(x) = -1/\lambda$ by Corollary 3.4 we have that

$$\lim_{x\to\infty} k(x) = F'(-1/\lambda) + r_t R'(-1/\lambda) \ . \tag{3.16}$$

By assumption $\lambda > 0$, and by Definition 3.2 it holds that $R(-1/\lambda) = 1$. Also $R(0) = 0$, and by the mean value theorem

$$1 = R(-1/\lambda) - R(0) = -\frac{1}{\lambda} R'(\xi)$$

for some $\xi \in (-1/\lambda, 0)$. Since R' is increasing, it follows that $R'(-1/\lambda) \leq -\lambda < 0$, and we can deduce (3.15) directly from (3.16).

We summarize our results on the function k so far: k stays negative on $(0, \infty)$ if

$r_t \geq b_{\text{inv}}$ and positive if $r_t \leq b_{\text{norm}}$. It has a single zero on $(0, \infty)$ if and only if $b_{\text{norm}} < r_t < b_{\text{inv}}$. If k has a zero on $(0, \infty)$, since k is decreasing, the sign of k will be positive to the left of the zero and negative to the right of the zero.

Since $\partial_{xx} H$ has the same sign as k, the statements above translate in the obvious way to the convexity behavior of H. We will now use the convexity behavior of H to derive our results about the yield curve.
Consider the equation

$$H(x) = cx, \qquad x \in [0, \infty) \tag{3.17}$$

for some fixed $c \in \mathbb{R}$. Since $H(0) = 0$ this equation has at least one solution, $x_0 = 0$. If $r_t \geq b_{\text{inv}}$ then $H(x)$ is strictly concave on $(0, \infty)$, and by Lemma 3.11 the equation (3.17) has at most one additional solution x_1. Also, when the solution exists, $H(x)$ crosses cx from above at x_1. Similarly if $r_t \leq b_{\text{norm}}$ then $H(x)$ is strictly convex, and there exists at most one additional solution x_2 to (3.17) on $[0, \infty)$. If the solution exists, then cx is crossed from below at x_2. In the last case $b_{\text{norm}} < r_t < b_{\text{inv}}$, there exists a x_* – the zero of $k(x)$ – such that $H(x)$ is strictly convex on $(0, x_*)$ and strictly concave on (x_*, ∞). Now there can exist at most two additional solutions x_1, x_2 to (3.17) with $x_1 < x^* < x_2$, such that cx is crossed from below at x_1 and from above at x_2.
Because of definition (3.10), every solution to (3.17), excluding $x_0 = 0$, is also a solution to

$$Y(r_t, x) = c, \qquad x \in (0, \infty) \tag{3.18}$$

with r_t fixed. Also the properties of crossing from above/below are preserved since x is positive. This means that in the case $r_t \geq b_{\text{inv}}$, equation (3.18) has at most a single solution, or in other words, that every horizontal line is crossed by the yield curve at most in a single point. If it is crossed, it is crossed from above. This implies that $Y(x)$ is a strictly decreasing function of x, or following Definition 3.9, that the yield curve is inverse. In the case $r_t \leq b_{\text{norm}}$ we have again that (3.18) has at most a single solution and that every horizontal line is crossed from below by the yield curve, if it is crossed. In other words, the yield curve is normal. In the last case of $b_{\text{norm}} < r_t < b_{\text{inv}}$, the yield curve crosses every horizontal line at most twice, in which case it crosses first from below, then from above. Thus in this case the yield curve is humped. ☐

Corollary 3.12. *Under the conditions of Theorem 3.10 the instantaneous forward rate curve has the same global behavior as the yield curve, i.e.*

$Y(r_t, .)$ *is inverse*	$\iff$	$f(r_t, .)$ *is strictly decreasing*
$Y(r_t, .)$ *is humped*	$\iff$	$f(r_t, .)$ *has exactly one local maximum and no local minimum*
$Y(r_t, .)$ *is normal*	$\iff$	$f(r_t, .)$ *is strictly increasing .*

In the second case the maximum of the forward rate curve is $f(r_t, x_*)$, where x_* solves

$$r_t = -\frac{F'(B(x))}{R'(B(x))}, \quad x \in (0, \infty)\ . \tag{3.19}$$

Proof. This follows from the fact that $\partial_x H(x)$ as given in (3.11) is exactly the forward rate $f(r_t, x)$. The derivative of the forward rate is therefore $\partial_{xx} H(x)$, which is given in (3.12) as

$$\partial_x f(r_t, x) = \partial_{xx} H(x) = -B'(x) \cdot k(x) \ .$$

The factor $-B'(x) \neq 0$ is always positive, and the possible sign changes and zeroes of $k(x)$ are discussed in the proof of Theorem 3.10, leading exactly to the stated equivalences. Equation (3.19) is simply the condition $k(x_*) = 0$. □

Corollary 3.13. *Under the conditions of Theorem 3.10 it holds that*

$$b_{\text{norm}} < b_{\text{asymp}} < b_{\text{inv}} \tag{3.20}$$

whenever the quantities are finite. In addition it holds that

$$D \cap (b_{\text{norm}}, b_{\text{inv}}) \neq \emptyset \ . \tag{3.21}$$

Remark 3.14. Note that equation (3.21) implies that there is always some $r_t \in D$ such that the yield curve $Y(r_t, .)$ is humped.

Proof. By the mean value theorem there exists a $\xi \in (-1/\lambda, 0)$ such that

$$b_{\text{asymp}} = -F(-1/\lambda) = F(0) - F(-1/\lambda) = \frac{1}{\lambda} F'(\xi) \ .$$

Since F is convex and thus F' increasing, it holds that

$$\frac{F'(-1/\lambda)}{\lambda} \leq b_{\text{asymp}} \leq \frac{F'(0)}{\lambda} \ . \tag{3.22}$$

Applying the mean value theorem to R, there exists another $\xi \in (-1/\lambda, 0)$ such that

$$1 = R(-1/\lambda) - R(0) = -\frac{1}{\lambda} R'(\xi) \ .$$

Since R' is increasing we deduce that $R'(-1/\lambda) \leq -\lambda < 0$. Assuming that also $R'(0) < 0$ we get

$$-\frac{1}{R'(-1/\lambda)} \leq \frac{1}{\lambda} \leq -\frac{1}{R'(0)} \ . \tag{3.23}$$

Since either F or R is non-linear, one of the functions is strictly convex by Lemma 2.6. Consequently either both inequalities in (3.22) or in (3.23) are strict. Putting them together we get

$$-\frac{F'(-1/\lambda)}{R'(-1/\lambda)} < b_{\text{asymp}} < -\frac{F'(0)}{R'(0)} \ ,$$

proving (3.20) under the assumption that $R'(0) < 0$.
If $R'(0) \geq 0$ then by definition $b_{\text{inv}} = \infty$. Equation (3.22) still holds, but in (3.23) only the left inequality sign remains valid. Together this still proves that $b_{\text{norm}} < b_{\text{asymp}}$ and we have shown (3.20).

To prove (3.21) we distinguish two cases:

(a) $D = \mathbb{R}$. In this case it is sufficient to prove $-\infty < b_{\text{inv}}$ and $b_{\text{norm}} < \infty$. Consider first b_{inv}. If $R'(0) \geq 0$ then by definition $b_{\text{inv}} = \infty$ and nothing is to prove. If $R'(0) < 0$ then $b_{\text{inv}} = -F'(0)/R'(0)$. By convexity $F'(0) > -\infty$ and the assertion follows. Consider now $b_{\text{norm}} = -F'(-1/\lambda)/R'(-1/\lambda)$. From (3.23) we know that $R'(-1/\lambda) \leq -\lambda < 0$. By convexity $F'(-1/\lambda) < \infty$ and it follows that $b_{\text{norm}} < \infty$.

(b) $D = \mathbb{R}_{\geqslant 0}$. In this case it is sufficient to prove $0 \le b_{\text{norm}}$ and to apply (3.20). As above we have that $b_{\text{norm}} = -F'(-1/\lambda)/R'(-1/\lambda)$ and that $R'(-1/\lambda) \le -\lambda < 0$. By Definition 2.3 it holds that

$$F'(u) = b + \int_{(0,\infty)} \xi e^{u\xi}\, m(d\xi)$$

with $b \ge 0$. It follows that $F'(-1/\lambda) \ge 0$, proving the assertion. □

The last Corollary of this section shows the interesting fact that the occurrence of a humped yield curve is a necessary and sufficient sign of randomness in the short rate model.

Corollary 3.15. *Let the short rate process be given by a one-dimensional affine process $(r_t)_{t\ge 0}$ satisfying Condition 3.1 with $F \ne 0$ and $\lambda > 0$. Then the following statements are equivalent:*

(i) There exists a $r_t \in D$ such that $Y(r_t, .)$ is flat.
(ii) There exists no $r_t \in D$ such that $Y(r_t, .)$ is humped.
(iii) The short rate process $(r_t)_{t\ge 0}$ is deterministic.
(iv) $F(u) = bu$ and $R(u) = \beta u$.

Proof. Theorem 3.10, together with Corollary 3.13, shows already that $\neg(iv)$ implies $\neg(i)$ and $\neg(ii)$. Also, from the form of the generator in (2.5), it is seen that (iii) and (iv) are equivalent. It remains to show that (iv) implies (i) and (ii). Proceeding as in the proof of Theorem 3.10 we obtain instead of (3.13) simply

$$k(x) = b + r_t\beta\ .$$

The yield curve will be humped if and only if k has a single (isolated) zero in $[0, \infty)$. Since k is a constant function, this cannot be the case for any $r_t \in D$ and we have shown (ii). By the same arguments as in the proof of Theorem 3.10 the yield curve is flat if and only if k is constant and equal to 0. This is the case if $r_t = -\frac{b}{\beta}$. It remains to show that $r_t \in D$. Note that $\beta = -\lambda < 0$. In particular $\beta \ne 0$, such that for $D = \mathbb{R}$ we are already done. If $D = \mathbb{R}_{\geqslant 0}$ we have by the admissibility conditions in Definition 2.3 that $b \ge 0$. Thus $r_t = -\frac{b}{\beta} \ge 0$ and we have shown (i). □

3.3. **The Limit Distribution of the Short Rate Process.** It is well-known that the Gaussian Ornstein-Uhlenbeck process, for example, converges in law to a limit distribution and that this distribution is Gaussian. Before we state our corresponding result on limit distributions of affine processes, we want to recall that a real-valued random variable L is called self-decomposable if for every $c \in (0, 1)$ there exists a random variable L_c, independent of L, such that

$$L = cL + L_c \quad \text{for all} \quad c \in (0, 1)\ .$$

Since self-decomposability is a distributional property, we will often identify L and its law, and refer to both as self-decomposable.

For OU-type processes, limit distributions have been studied for some time; the first results can be found in Jurek and Vervaat [1983] and Sato and Yamazato [1984]. The next theorem summarizes these results, and can be found in similar form in Sato [1999, Theorem 17.5]:

Theorem 3.16. *Let $(r_t)_{t\geq 0}$ a OU-type process on $\mathbb{R}$ (i.e. a conservative, regular affine process with $R(u) = \beta u$). If $\beta < 0$ and*

$$\exists\, \epsilon > 0 \text{ such that } \int_0^\epsilon \frac{|F(s)|}{s}\, ds < \infty \tag{3.24}$$

then $(r_t)_{t\geq 0}$ converges to a limit distribution L which is independent of r_0 and has the following properties:

(i) L is self-decomposable.
(ii) The cumulant generating function $\kappa(u) = \log \mathbb{E}\left[e^{uL}\right]$ exists for all $u \in \mathcal{U}$ and is given by

$$\kappa(u) = \frac{1}{\beta} \int_u^0 \frac{F(s)}{s} ds \qquad u \in \mathcal{U}\ . \tag{3.25}$$

Conversely, if L is a self-decomposable distribution on $\mathbb{R}$ and $\beta < 0$, there exists a unique triplet (a, b, m) satisfying the admissibility conditions of Definition 2.3, such that L is the limit distribution of the affine process (of OU-type) given by the parameters (a, b, m, β).

Remark 3.17. Instead of condition (3.24) a 'log-moment condition' like $\mathbb{E} \log(1 + |X|) < \infty$, where X is a random variable with cumulant generating function F, is often found in the literature. This and other similar conditions are in fact equivalent to (3.24), as shown in Steutel and van Harn [2004, Proposition A.2.7].

We now state our corresponding result for affine processes on $\mathbb{R}_+$:

Theorem 3.18. *Let $(r_t)_{t\geq 0}$ be a one-dimensional, regular, conservative affine process with state space $\mathbb{R}_+$. If $R'(0) < 0$ and*

$$\exists\, \epsilon > 0 \text{ such that } \int_{-\epsilon}^0 \frac{F(s)}{R(s)}\, ds < \infty \tag{3.26}$$

then $(r_t)_{t\geq 0}$ converges in law to a limit distribution L which is independent of r_0 and has the following properties:

(i) L is infinitely divisible.
(ii) The cumulant generating function $\kappa(u) = \log \mathbb{E}\left[e^{uL}\right]$ of L exists for all $u \in \mathcal{U}$ and is given by

$$\kappa(u) = \int_u^0 \frac{F(s)}{R(s)} ds \qquad u \in \mathcal{U}\ . \tag{3.27}$$

Proof. By Theorem 2.4 the transition kernel $p_t(x, .)$ of the process $(r_t)_{t\geq 0}$ has the characteristic function

$$\widehat{p_t}(x, u) = \exp\left(\phi(t, u) + x\psi(t, u)\right) \tag{3.28}$$

where ϕ and ψ satisfy the generalized Riccati equations (2.6) for all $u \in \mathcal{U}$, and thus in particular for all $u \in (-\infty, 0]$. From the assumption $R'(0) < 0$ it follows by the convexity of R that $R(u) > 0$ for all $u \in (-\infty, 0)$. The positivity of R implies through the second Riccati equation in (2.6) that $\psi(t, u)$ is a strictly increasing function in t for all $u \in (-\infty, 0)$ and that

$$\lim_{t\to\infty} \psi(t, u) = 0 \qquad \forall\, u \in (-\infty, 0)\ .$$

Consequently,

$$\lim_{t\to\infty} \log \widehat{p}_t(x,u) = \lim_{t\to\infty} \phi(t,u) = \int_0^\infty F(\psi(r,u))\,dr = \int_u^0 \frac{F(s)}{R(s)}\,ds \tag{3.29}$$

where the last two equalities follow from applying the generalized Riccati equations (2.6) and the transformation $s = \psi(r,u)$. All that remains to show for assertion (ii) is that the last term in (3.29) is a cumulant generating function of a random variable L. First note that since $F(u) < \infty$ for all $u \in (-\infty, 0]$ and $R(u) \neq 0$ for all $u \in (-\infty, 0)$, the integral can only diverge at 0. Also, for $u \in (-\infty, 0)$ we have that $R(u) > 0$, since $R'(0) < 0$, and that $F(u) \leq 0$, because of Definition 2.3. Thus $-F(u)/R(u)$ is positive on $(-\infty, 0)$, and condition 3.26 guarantees the convergence of the integral (3.29) at 0 and therefore also for all $u \in (-\infty, 0]$. In addition the integral is a continuous function of u and zero for $u = 0$. By standard results on Laplace transforms of probability measures (cf. Steutel and van Harn [2004, Appendix A]), pointwise convergence of cumulant generating functions to a function that is (left-)continuous at 0 implies convergence in distribution of $(r_t)_{t\geq 0}$ to a limit distribution L with cumulant generating function given by (3.29). Also, if a cumulant generating function exists on the negative reals, it can be analytically extended to all $\mathcal{U} = \{u \in \mathbb{C} : \operatorname{Re} u \leq 0\}$ and we have shown assertion (ii).
It remains to show that L is infinitely divisible: From (2.3) and the Lévy-Khintchine formula it is seen that F is the cumulant generating function of some infinitely divisible random variable X. Thus for any $n \in \mathbb{N}$ there exist i.i.d random variables $X_1^n, \ldots, X_n^n$ with cumulant generating function F_n such that $X \stackrel{d}{=} X_1^n + \cdots + X_n^n$ and $F(u) = nF_n(u)$. Now each (F_n, R) defines an affine process that satisfies the conditions of Theorem 3.18 and therefore converges to a limit distribution L_n with cumulant generating function

$$\kappa_n(u) = \int_u^0 \frac{F_n(s)}{R(s)}\,ds\ .$$

We now have

$$\kappa(u) = \int_u^0 \frac{F(s)}{R(s)}\,ds = n\int_u^0 \frac{F_n(s)}{R(s)}\,ds = n\kappa_n(u)\ ,$$

showing that L is infinitely divisible. □

Comparing Theorem 3.16 and 3.18, there are two obvious questions: First, is it possible to show also in the affine case that the limit distribution is self-decomposable; and second, is it possible to show a converse to Theorem 3.18 for some class of distributions? The second question we must leave open, but to the first question we have an answer: No. We will give in Section 4.3 an example of an affine process of type A, converging to a limit distribution that is infinitely divisible, but not self-decomposable. This result is interesting, since it leaves open the possibility of some unexpected properties of the limit distribution of an affine process. For example a self-decomposable distribution is always unimodal, whereas an infinitely divisible distribution might be not.

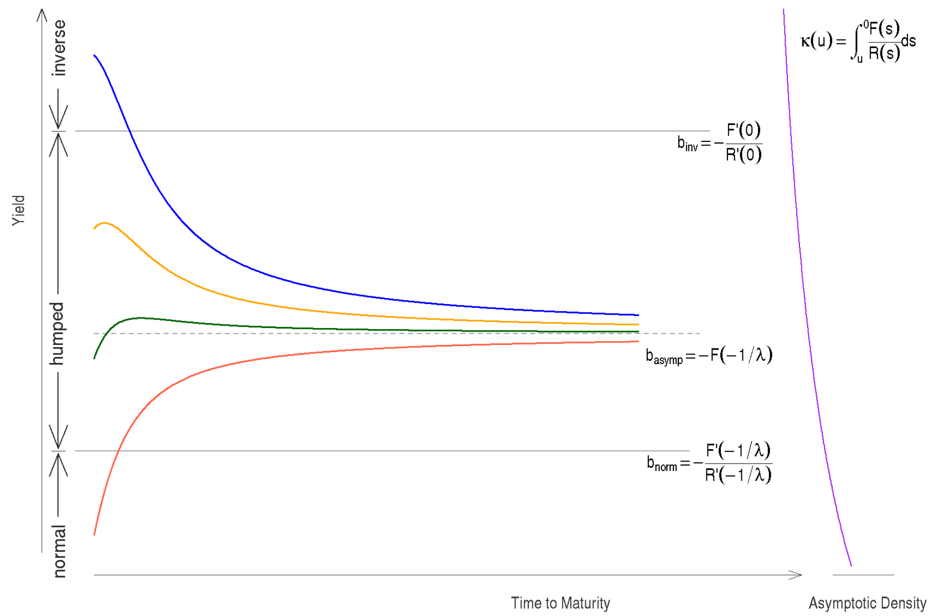

FIGURE 1. This Figure shows a graphical summary of Theorems 3.6, 3.10 and 3.18, as well as the definitions of the key quantities b_{norm}, b_{asymp}, b_{inv} and $\kappa(u)$. In *any* affine model satisfying the conditions of Theorem 3.10, the shapes of yield curves will follow the picture given here. They will be normal if r_0 is below b_{norm}, humped if r_0 is between b_{norm} and b_{inv} and inverse if r_0 is above b_{inv}. Also all yield curves will tend asymptotically to the same level b_{asymp}. If the conditions of Theorem 3.18 are satisfied, the short rate will converge in law to an asymptotic distribution described by the cumulant generating function $\kappa(u)$.

4. Applications

4.1. **The Vasiček model.** We apply the results of the last section to the classical Vasiček model

$$dr_t = -\lambda(r_t - \theta)\, dt + \sigma\, dW_t, \quad r_0 \in \mathbb{R} \tag{4.1}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion and $\lambda, \theta, \sigma > 0$. The Vasiček model is arguably the simplest affine model, and no surprises are to be expected here. In fact all results that we state here can already be found in the original paper of Vasiček [1977]. We advise the reader to view this paragraph as a warm-up for the CIR-model.
Clearly $(r_t)_{t\geq 0}$ is a conservative affine process with

$$F(u) = \lambda\theta u + \frac{\sigma^2}{2}u^2\ , \tag{4.2}$$

$$R(u) = -\lambda u\ . \tag{4.3}$$

From the quadratic term in F and Definition 2.3, it is seen that $(r_t)_{t\geq 0}$ is of type B, i.e. has state space $\mathbb{R}$. This property is often criticized, since it allows the short rate to become negative.
From Theorem 3.10 we calculate

$$b_{\mathrm{inv}} = \theta \qquad \text{and} \qquad b_{\mathrm{norm}} = \theta - \frac{\sigma^2}{\lambda^2} \ ,$$

such that the yield curve in the Vasiček model is normal if $r_t \leq \theta - \sigma^2/\lambda^2$, inverse if $r_t \geq \theta$ and humped in the remaining cases.
The long term yield is calculated from (3.6) as

$$b_{\mathrm{asymp}} = -F(-1/\lambda) = \theta - \frac{\sigma^2}{2\lambda^2} \ ,$$

in this case exactly the *arithmetic* mean of b_{inv} and b_{norm}.
By Theorem 3.18 the cumulant generating function of a possible limit distribution is given by

$$\kappa(u) = \int_u^0 \frac{F(s)}{R(s)}\, ds = \int_0^u \theta + \frac{\sigma^2}{2\lambda} s\, ds = u\theta + \frac{u^2}{2}\frac{\sigma^2}{2\lambda} \ .$$

The integral converges for all $u \in \mathbb{R}$, and we have also $R'(0) = -\lambda < 0$, such that the limit distribution exists by Theorem 3.18. From the form of $\kappa(u)$, it is also immediately seen that the limit distribution is Gaussian with mean θ and variance $\frac{\sigma^2}{2\lambda}$.

4.2. **The Cox-Ingersoll-Ross model.** The Cox-Ingersoll-Ross (CIR)-model was introduced by Cox et al. [1985]. In this model the short rate process $(r_t)_{t\geq 0}$ is given by the SDE

$$dr_t = -a(r_t - \theta)dt + \sigma\sqrt{r_t}\, dW_t, \qquad r_0 \geq 0 \tag{4.4}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian Motion and $a, \theta, \sigma > 0$. The process $(r_t)_{t\geq 0}$ is a conservative affine process with

$$F(u) = a\theta u \ , \tag{4.5}$$

$$R(u) = \frac{\sigma^2}{2}u^2 - au \ . \tag{4.6}$$

From Definition 2.3 it is seen that $(r_t)_{t\geq 0}$ is a process of type A, i.e. has state space $\mathbb{R}_{\geqslant 0}$. The fact that interest rates stay non-negative in the CIR-model is often cited as an advantage of the model over the Vasiček model. Calculating the quasi-mean-reversion (cf. Definition 3.2), we find that

$$\lambda = \frac{1}{2}\left(\sqrt{a^2 + 2\sigma^2} + a\right) \ .$$

From Theorem 3.6 we find that the long-term yield is given by

$$b_{\mathrm{asymp}} = -F(-1/\lambda) = \frac{2a\theta}{\sqrt{a^2 + 2\sigma^2} + a} \ .$$

The boundary between humped and inverse behavior b_{inv} is calculated from Theorem 3.10 as

$$b_{\mathrm{inv}} = -\frac{F'(0)}{R'(0)} = \theta.$$

Both quantities b_{asymp} and b_{inv} can also be found in [Cox et al., 1985, Eq. (26) and following paragraph]. Before we consider b_{norm}, we quote (with notation adapted to (4.4)) from page 394 of [Cox et al., 1985] where the shape of the yield curve is discussed:

> *'When the spot rate is below the long-term yield [= b_{asymp}], the term structure is uniformly rising. With an interest rate in excess of θ [= b_{inv}], the term structure is falling. For intermediate values of the interest rate, the yield curve is humped.'*

In our terminology, they claim that the yield curve is normal for $r_t \le b_{\text{asymp}}$, humped for $b_{\text{asymp}} < r_t < b_{\text{inv}}$ and inverse for $r_t \ge b_{\text{inv}}$. This stands in clear contradiction to Theorem 3.10 and Corollary 3.13 where we have obtained that yield curves are normal if and only if $r_t \le b_{\text{norm}}$ and that $b_{\text{norm}} < b_{\text{asymp}}$, or – in other words – that there are yield curves starting strictly *below* the long-term yield that are still humped.

The claims of Cox et al. [1985] are repeated in [Rebonato, 1998, p. 244f], where even several plots of 'yield surfaces' (the yield as a function of r_t and x) are presented as evidence. However Rebonato fails to indicate the level of b_{asymp} in the plots, such that the conclusion remains ambiguous.

To clarify the scope of humped yield curves in the CIR-model we calculate b_{norm} from Theorem 3.10:

$$b_{\text{norm}} = -\frac{F'(-1/\lambda)}{R'(-1/\lambda)} = \frac{a\theta}{\sqrt{a^2 + 2\sigma^2}} \,. \tag{4.7}$$

The relation $b_{\text{norm}} < b_{\text{asymp}} < b_{\text{inv}}$ is immediately confirmed by noting that b_{asymp} is the *harmonic* mean of b_{norm} and b_{inv}. For a graphical illustration we refer to the second yield curve from below in Figure 1. The plot actually shows CIR yield curves with parameters

$$a = 0.5, \quad \sigma = 0.5, \quad \theta = 6\%$$

plotted over a time scale of 25 years. The second curve from below starts at $r_0 = 4.2\%$, i.e. below the long-term yield, but is visibly humped.

To calculate the limit distribution of $(r_t)_{t\ge 0}$, we apply Theorem 3.18: The cumulant generating function $\kappa(u)$ of the limit distribution is given by

$$\kappa(u) = \int_u^0 \frac{F(s)}{R(s)}\,ds = \int_0^u \frac{\theta}{1 - \dfrac{\sigma^2}{2a}s}\,ds = -\frac{2a\theta}{\sigma^2}\log\left(1 - \frac{\sigma^2}{2a}u\right) \,.$$

This is the cumulant generating function of a gamma distribution with shape parameter $2a\theta/\sigma^2$ and scale parameter $\sigma^2/2a$. Again this result can already be found in Cox et al. [1985, p. 392].

4.3. **An extension of the CIR model.** To illustrate the power of the affine setting, we consider now an extension of the CIR model that is obtained by adding jumps to (4.4). We define

$$dr_t = -a(r_t - \theta)dt + \sigma\sqrt{r_t}\,dW_t + dJ_t, \qquad r_0 \ge 0 \tag{4.8}$$

where $(J_t)_{t\ge 0}$ is a compound Poisson process with intensity $c > 0$ and exponentially distributed jumps of mean $\nu > 0$. This model has been introduced by Duffie and Gârleanu [2001] as a model for default intensity and is used by Filipović [2001] as a short rate model. It can also be found in Brigo and Mercurio [2006] under the name JCIR model. It is easily calculated that

$$F(u) = a\theta u + \frac{cu}{\nu - u}, \quad u \in (-\infty, \nu)\ , \tag{4.9}$$

$$R(u) = \frac{\sigma^2}{2} u^2 - au\ . \tag{4.10}$$

Solving the generalized Riccati equations (3.3) and (3.4) for $A(x)$ and $B(x)$ becomes quite tedious, but the quantities $b_{\text{inv}}, b_{\text{asymp}}, b_{\text{norm}}$ can be calculated from Theorem 3.6 and Theorem 3.10 in a few lines: The quasi-mean reversion λ stays the same as in the CIR model, since R does not change. From

$$F'(u) = a\theta + \frac{c\nu}{(\nu - u)^2}$$

we derive immediately

$$\begin{aligned} b_{\text{inv}} &= \theta + \frac{c}{a\nu} \\ b_{\text{asymp}} &= \frac{2a\theta}{a+\gamma} + \frac{2c}{\nu(a+\nu)+2} \\ b_{\text{norm}} &= \frac{a\theta}{\gamma} + \frac{c\nu\sigma^4}{\gamma(\sigma^2\nu + \gamma - a)^2} \end{aligned}$$

where $\gamma = \sqrt{a^2 + 2\sigma^2}$. Note that by setting the jump intensity c to zero, the expressions of the (original) CIR model are recovered.

Next we calculate the limit distribution of the model. Using the abbreviations $\rho := \sigma^2/2$ and $\Delta := a - \nu\rho$ we obtain

$$\begin{aligned} \kappa(u) &= \int_u^0 \frac{F(s)}{R(s)}\, ds = \int_0^u \frac{\theta}{1 - s\rho/a}\, ds + c\int_0^u \frac{ds}{(s-\nu)(\rho s - a)} = \\ &= -\frac{a\theta}{\rho} \log\left(1 - \frac{u\rho}{a}\right) + \frac{c}{\Delta} \log\left(\frac{1 - u\rho/a}{1 - u/\nu}\right) = \\ &= \left(\frac{c}{\Delta} - \frac{a\theta}{\rho}\right) \log\left(1 - \frac{\rho}{a}u\right) - \frac{c}{\Delta}\log\left(1 - \frac{u}{\nu}\right) \end{aligned}$$

as the cumulant generating function of the limit distribution L.

We now take a closer look at the distribution L, since this will answer the question raised at the end of Section 3.3: For certain parameters, L is an example for a limit distribution of an affine process that is infinitely divisible, but not self-decomposable. First we define

$$\begin{aligned} A &:= \frac{a\theta}{\rho} - \frac{c}{\Delta}\quad , \quad B := \frac{c}{\Delta} \quad \text{and} \\ l(x) &:= A\exp\left(-\frac{a}{\rho}x\right) + B\exp(-\nu x) \qquad (x \in \mathbb{R}_{\geqslant 0})\ . \end{aligned}$$

Consulting your favorite integral table, it can be verified that

$$\kappa(u) = -A\log\left(1 - \frac{\rho}{a}u\right) - B\log\left(1 - \frac{u}{\nu}\right) = \int_0^\infty \left(e^{ux} - 1\right)\frac{l(x)}{x}\, dx \tag{4.11}$$

for all $u \in (-\infty, 0]$. If l is non-negative on $\mathbb{R}_{\geqslant 0}$, then $l(x)/x$ is the density of a Lévy measure and (4.11) is seen to be the Lévy-Khintchine representation for the cumulant generating function of the non-negative infinitely divisible random variable L. However, if in addition l is non-increasing, then L is also self-decomposable (cf. Sato [1999, Corollary 15.11]). We distinguish the following cases:

$A \geq 0$ **and** $B \geq 0$**::** In this case $l(x)$ is positive and non-increasing for all $x \geq 0$; thus L is self-decomposable. It is also easily seen, that in fact L can be written as the convolution of two (possibly degenerate) gamma distributions with parameters $(A, \rho/a)$ and (B, ν^{-1}).

$A < 0$ **and** $B \geq 0$**::** Since $B \geq 0$, also $\Delta > 0$, and from the definition of Δ we get that $a/\rho > \nu$. Thus we have

$$l(x) = Ae^{-\frac{a}{\rho}x} + Be^{-\nu x} > (A+B)e^{-\frac{a}{\rho}x} = \frac{a\theta}{\rho}e^{-\frac{a}{\rho}x} > 0$$

for all $x \geq 0$, such that $l(x)/x$ is a density and L is infinitely divisible (which, of course, it has to be by Theorem 3.18). However, setting for example

$$a = 2, \quad c = 1, \quad \nu = 1, \quad \rho = \frac{1}{2}, \quad \theta = \frac{1}{12},$$

we get $B = 2/3$, $A = -1/3$, and consequently

$$l'(0) = -\frac{a}{\rho}A - \nu B = \frac{2}{3} > 0\ ,$$

such that $l(x)$ is increasing in a neighborhood of 0 and consequently L is *not self-decomposable*.

$A \geq 0$ **and** $B < 0$**::** If $B < 0$ also $\Delta < 0$ and consequently $a/\rho < \nu$. By a similar argument as above, $l(x)$ is positive for all $x \geq 0$ and thus a density. Again we can find parameters

$$a = 1, \quad c = 1, \quad \nu = 2, \quad \rho = 1, \quad \theta = \frac{1}{2}$$

such that $A = 3/2$, $B = -1$ and $k'(0) = 1/2 > 0$. Thus l is increasing in a neighborhood of 0 and L is not self-decomposable for the given parameters.

$A < 0$ **and** $B < 0$**::** We have $A + B = a\theta/\rho > 0$ such that this case can never happen.

4.4. **The gamma model.** Instead of analyzing the properties of a known model, we will now follow a different route and construct a model that satisfies some given properties. We want to construct an affine process that has the same limit distribution as the CIR model (i.e. a gamma distribution), but is a process of OU-type. The second property is equivalent to $R(u) = \beta u$. Considering Theorem 3.16 (or equivalently Theorem 3.18), we know that if we want to obtain a limit distribution, we need $\beta < 0$. To keep with the notation of the Vasiček model, we will write $R(u) = -\lambda u$ where $\lambda > 0$. Now by (3.25) or (3.27) the cumulant generating function of the limit distribution is given by

$$\kappa(u) = \frac{1}{\lambda}\int_0^u \frac{F(s)}{s}\,ds\ . \tag{4.12}$$

Let the limit distribution be a gamma distribution with shape parameter $k > 0$ and scale parameter $\theta > 0$. Then

$$\kappa(u) = -k \log(1 - \theta u)$$

and by (4.12)

$$F(u) = \frac{\lambda \theta k u}{1 - \theta u} \ .$$

Setting $c = \lambda k$ and $\nu = 1/\theta$ it is seen that $F(u)$ is equal to the last term in (4.9). This means that the driving Lévy process of the OU-type process $(r_t)_{t \geq 0}$ is of the same kind as the process $(J_t)_{t \geq 0}$ in (4.8), or in other words, that if the parameters are correspondingly chosen, the jumps of the gamma model have the same distribution as the jumps of the JCIR model.
We interpret the affine process we have constructed as a short rate process. It is clear from Proposition 3.3 that the bond prices are of the form (3.2). From the second generalized Riccati equation (3.4) we calculate

$$B(x) = \frac{e^{-\lambda x} - 1}{\lambda} \ .$$

From the first equation (3.3) we calculate

$$\begin{aligned} A(x) &= \int_0^x F(B(s))\, ds = -\int_0^{B(x)} \frac{F(r)}{(\lambda r + 1)}\, dr = \\ &= \frac{\lambda k}{\theta + \lambda} \left(\log(\theta B(x) - 1) - \theta x \right) \ , \end{aligned}$$

such that the bond prices are given by

$$P(t, t+x) = \exp\left\{ -\lambda x \frac{\theta k}{\theta + \lambda} + r_t B(x) \right\} (\theta B(x) - 1)^{\frac{\lambda k}{\theta + \lambda}} \ .$$

The global shape of the yield curve is described by the quantities

$$b_{\text{inv}} = k\theta, \quad b_{\text{asymp}} = \frac{k}{1/\theta + 1/\lambda}, \quad b_{\text{norm}} = \frac{k/\theta}{(1/\theta + 1/\lambda)^2}$$

and it is seen that for the gamma-OU-process b_{asymp} is the *geometric* average of b_{inv} and b_{norm}.

5. Conclusions

In this article we have given a complete characterization of the yield curve shapes that are attainable in term structure models where the short rate is given by a time-homogenous, one-dimensional affine process. Even though the parameter space for this class of models is infinite-dimensional, the scope of attainable yield curves is very narrow, with only three possible global shapes. In addition we have given conditions under which the short rate process converges to a limit distribution, and we have characterized the limit distribution in terms of its cumulant generating function, extending some known results on OU-type processes.
The most obvious question for future research is the extension of these results to multi-factor models. It is evident from numerical results that in two-factor models yield curves with e.g. a dip, or also with a dip and a hump, can be obtained. It would be interesting to see if more complex shapes can also be produced, or if there are similar limitations as in the single-factor case. Also, in the one-factor case the dependence of the yield curve shape on the current short rate is basically described

by the intervals $D \cap (-\infty, b_{\text{norm}}], (b_{\text{norm}}, b_{\text{inv}})$ and $[b_{\text{inv}}, \infty)$. In the two-factor case the partitioning of the state-space might be much more complex, and we expect to see more interesting transitions between yield curve types. Another aspect is, that since affine processes as a general framework become better understood, extensions of classical models e.g. by adding jumps, like in the JCIR model described in Section 4.2, become more feasible and attractive for applications.

References

Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice.* Springer Finance. Springer, 2nd edition, 2006.

René Carmona and Michael Tehranchi. *Interest Rate Models: An Infinite Dimensional Stochastic Analysis Perspective.* Springer Finance. Springer, 2006.

John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A theory on the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

Darrell Duffie and Nicolae Gârleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41 – 59, 2001.

Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

Philip Dybvig, Jonathan Ingersoll, and Stephen Ross. Long forward and zero-coupon rates can never fall. *The Journal of Business*, 69(1):1–25, 1996.

Damir Filipović. A general characterization of one factor affine term structure models. *Finance and Stochastics*, 5:389–412, 2001.

Friedrich Hubalek, Irene Klein, and Josef Teichmann. A general proof of the Dybvig-Ingersoll-Ross-Theorem: Long forward rates can never fall. *Mathematical Finance*, 12(4):447–451, 2002.

Zbigniew J. Jurek and Wim Vervaat. An integral representation for self-decomposable Banach space valued random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62:247–262, 1983.

Olav Kallenberg. *Foundations of Modern Probability.* Springer, 1997.

Eugene Lukacs. *Characteristic Functions.* Charles Griffin & Co Ltd., 1960.

Riccardo Rebonato. *Interest-Rate Option Models.* Wiley, 2nd edition, 1998.

L.C.G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales, Volume 1.* Cambridge Mathematical Library, 2nd edition, 1994.

Ken-iti Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, 1999.

Ken-iti Sato and M. Yamazato. Operator-selfdecomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Processes and Applications*, 17:73–100, 1984.

Fred Steutel and Klaas van Harn. *Infinite Divisibility of Probability Distributions on the Real Line.* Marcel Dekker Inc., 2004.

Oldrich Vasiček. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

Vienna University of Technology, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria
E-mail address: `mkeller@fam.tuwien.ac.at`

Vienna University of Technology, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria
E-mail address: `thomas@fam.tuwien.ac.at`