Yield curve shapes and the asymptotic short rate distribution in affine one-factor models

Martin Keller-Ressel · Thomas Steiner

Received: 15 March 2007 / Accepted: 25 November 2007 / Published online: 17 January 2008 © Springer-Verlag 2008

Abstract We consider a model for interest rates where the short rate is given under the risk-neutral measure by a time-homogeneous 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 converges to a limit distribution and describe the risk-neutral 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.

Keywords Affine process · Term structure of interest rates · Ornstein–Uhlenbeck process · Yield curve

Mathematics Subject Classification (2000) 60J25 · 91B28

JEL Classification E43 · G12

Supported by the Austrian Science Fund (FWF) through project P18022 and the START program Y328. Supported by the module M5 “Modeling 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 proofreaders at FAM and the anonymous referee for their comments.

M. Keller-Ressel (✉) · T. Steiner
Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
e-mail: mkeller@fam.tuwien.ac.at
T. Steiner
e-mail: thomas@fam.tuwien.ac.at
1 Introduction

We consider a model for the term structure of interest rates where the short rate \( r_t \), \( t \geq 0 \) is given under the martingale measure by a one-dimensional conservative 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. 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. [6] is not limited to diffusions, but also includes processes with jumps and even with jumps whose intensity depends in an affine way on the state of the 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 [1, Sect. 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-homogeneous) CIR and Vasiček models 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. These shortcomings are often explained by ‘too few parameters’ in the model (cf. [2, Sect. 2.3.5] or [1, Sect. 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. [4]) mention inverse, normal, and humped shapes, Carmona and Tehrachi [2, Sect. 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 [1, Sect. 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 3.9, we settle this question and prove that, in any time-homogeneous affine one-factor model, the attainable yield curves are either inverse, normal, or humped. The proof relies only on tools of elementary analysis and on the characterization by Duffie et al. [6] of affine processes through the generalized Riccati equations.

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 Sect. 4.2 that also in this respect the CIR model has not been completely understood and discuss a misconception that originates in [4] and is repeated, for example, in [11].

In Sect. 3.3 we provide conditions under which an affine process converges to a limit distribution. We also characterize the limit distribution in terms of its cumulant generating function, extending results of Jurek and Vervaat [8] and Sato and Yamazato [14] for OU-type processes to the class of affine processes. These results can
again be interpreted in the context of interest rates, where they can be used to derive the risk-neutral asymptotic distribution of the short rate \((r_t)_{t \geq 0}\) as \(t\) goes to infinity.

We conclude our article in Sect. 4 by applying the theoretical results to several interest rate models, such as the Vasiček model, the CIR model, the JCIR model, and an Ornstein–Uhlenbeck-type model.

2 Preliminaries

In this section we collect some key results on affine processes from Duffie et al. [6]. In their article affine processes are defined on the \((m + n)\)-dimensional state space \(\mathbb{R}_m^+ \times \mathbb{R}^n\), and we will try to simplify the notation when possible in the one-dimensional case. Results on affine processes with state space \(\mathbb{R}_+\) can also be found in [7].

**Definition 2.1** (One-dimensional affine process) A time-homogeneous Markov process \((r_t)_{t \geq 0}\) with state space \(D = \mathbb{R}_+\) 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

\[
\hat{p}_t(x, u) = \int_D e^{u\xi} p_t(x, d\xi)
\]

and defined (at least) on

\[
\mathcal{U} = \begin{cases} 
\{ u \in \mathbb{C} : \Re u \leq 0 \} & \text{if } D = \mathbb{R}_+, \\
\{ u \in \mathbb{C} : \Re u = 0 \} & \text{if } D = \mathbb{R},
\end{cases}
\]

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

\[
\hat{p}_t(x, u) = \exp(\phi(t, u) + x\psi(t, u)) \quad \text{for all } x \in D, \ (t, u) \in \mathbb{R}_+ \times \mathcal{U}.
\]  

(2.1)

For subsequent results, we will need the following regularity condition for \((r_t)_{t \geq 0}\).

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

\[
\partial_t^+ \phi(t, u)_{|t=0} \quad \text{and} \quad \partial_t^+ \psi(t, u)_{|t=0}
\]

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 with state space \(\mathbb{R}_+\) if

\[
a = 0,
\]

\[
\alpha, b, c, \gamma \in \mathbb{R}_+,
\]

\[
\beta \in \mathbb{R},
\]
\( m, \mu \) are Lévy measures on \((0, \infty)\), where \( m \) satisfies
\[
\int_{(0, \infty)} (\xi \wedge 1) m(d\xi) < \infty,
\]
and \textit{admissible} for a process with state space \( \mathbb{R} \) if
\[
a, c \in \mathbb{R}_+, \quad \beta \in \mathbb{R},
\]
\( m \) is a Lévy measure on \( \mathbb{R} \setminus \{0\} \),
\[
\alpha = 0, \quad \gamma = 0, \quad \mu \equiv 0.
\]
Moreover, define the truncation functions
\[
h_F(\xi) = \begin{cases} 0 & \text{if } D = \mathbb{R}_+, \\ \frac{\xi}{\frac{\xi}{1+\xi^2}} & \text{if } D = \mathbb{R}, \end{cases}
\]
and finally the functions \( F(u) \) and \( R(u) \) for \( u \in \mathbb{C} \) as
\[
F(u) = au^2 + bu - c + \int_{D \setminus \{0\}} (e^{u\xi} - 1 - uh_F(\xi)) m(d\xi), \quad (2.2)
\]
\[
R(u) = \alpha u^2 + \beta u - \gamma + \int_{D \setminus \{0\}} (e^{u\xi} - 1 - uh_R(\xi)) \mu(d\xi). \quad (2.3)
\]

The next result is a one-dimensional version of the key result of Duffie et al. [6].

**Theorem 2.4** ([6, Theorem 2.7]) Suppose that \((r_t)_{t \geq 0}\) is a one-dimensional regular affine process. Then it is a Feller process. Let \( A \) be its infinitesimal generator. Then \( C_c^\infty(D) \) is a core of \( A \), \( C_c^\infty(D) \subseteq D(A) \) and there exist some admissible parameters \((a, \alpha, b, \beta, c, \gamma, m, \mu)\) such that, for \( f \in C_c^2(D) \),
\[
Af(x) = (a + \alpha x) f''(x) + (b + \beta x) f'(x) - (c + \gamma x) f(x) + \int_{D \setminus \{0\}} (f(x + \xi) - f(x) - f'(x)h_F(\xi)) m(d\xi) + x \int_{D \setminus \{0\}} (f(x + \xi) - f(x) - f'(x)h_R(\xi)) \mu(d\xi). \quad (2.4)
\]
Moreover, \( \phi(t, u) \) and \( \psi(t, u) \) defined by (2.1) solve the generalized Riccati equations
\[
\partial_t \phi(t, u) = F(\psi(t, u)), \quad \phi(0, u) = 0, \quad (2.5a)
\]
\[
\partial_t \psi(t, u) = R(\psi(t, u)), \quad \psi(0, u) = u. \quad (2.5b)
\]
Moreover, 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 (2.4), and (2.1) holds with \( \phi(t, u) \) and \( \psi(t, u) \) given by (2.5).
Closely related to affine processes is the notion of an Ornstein–Uhlenbeck-type (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 [13, Chap. 17] an OU-type process \((X_t)_{t \geq 0}\) can be defined as the solution of the Langevin SDE

\[
dX_t = -\lambda X_t \, dt + dL_t, \quad \lambda \in \mathbb{R}, \ X_0 \in \mathbb{R},
\]

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

\[
\hat{p}_t(x, u) = \exp \left( \int_0^t F(e^{-\lambda s} u) \, ds + x e^{-\lambda t} u \right),
\]

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 (2.5) for an OU-type process necessarily \(R(u) = -\lambda u\). Comparing this with (2.3) and Definition 2.3, it is seen that any regular affine process with state space \(\mathbb{R}\) is a process of OU-type. The converse, however, is not true as there also exist OU-type processes with state space \(\mathbb{R}_+\). We will give an example of such a process in Sect. 4.4.

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

\[
Q_t f(x) := \mathbb{E} \left[ \exp \left( -\int_0^t r_s \, ds \right) f(r_t) \bigg| r_0 = x \right], \quad (2.6)
\]

where \(f\) is a bounded function on \(D\). The next result is an application of the Feynman–Kac formula for Feller semigroups (cf. [12, Sect. III.19]) and can be found in [6]. It relies on the positivity of \((r_t)_{t \geq 0}\) and is therefore only applicable if \(D = \mathbb{R}_+\).

**Proposition 2.5** ([6, Proposition 11.1]) \(\) Let \((r_t)_{t \geq 0}\) be a one-dimensional regular affine process with state space \(\mathbb{R}_+\). Then the family \((Q_t)_{t \geq 0}\) defined by (2.6) forms a regular affine semigroup with infinitesimal generator

\[
Bf(x) = Af(x) - xf(x) \quad \text{for all } f \in \mathcal{C}^2_c(D).
\]

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 the following:

**Lemma 2.6** If \(c = \gamma = 0\), then \(F\) and \(R\) defined in Definition 2.3 have the following properties:

(i) \(R(0) = 0\) and \(F(0) = 0\);

(ii) \(R(u) < \infty\) for all \(u \in (-\infty, 0]\);
(iii) 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 $\beta$;
(iv) If $F(u) < \infty$ on $(c_1, c_2) \subseteq \mathbb{R}$, then $F$ is continuously differentiable on $(c_1, c_2)$; moreover, 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. If $D = \mathbb{R}$, then by Definition 2.3 $R(u) = \beta u$, so that (ii) follows immediately. If $D = \mathbb{R}_+$, we use the estimate
\[
|e^{u\xi} - 1 - uh_R(\xi)| \leq |u| \left( O(\xi^2) \land 1 \right)
\]
for all $u \in (-\infty, 0]$ and $\xi \in \mathbb{R}_+$, and (ii) follows from (2.3). For property (iii), 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., $F(u) = \log \mathbb{E}[e^{uX}]$ for $u \in (c_1, c_2)$. Choosing two distinct numbers $u, v \in (c_1, c_2)$, we apply the Cauchy–Schwarz inequality to obtain
\[
F\left(\frac{u + v}{2}\right) = \log \mathbb{E}\left[ e^{\frac{uX + vX}{2}} \right] \leq \log \sqrt{\mathbb{E}\left[ e^{uX} \right] \mathbb{E}\left[ e^{vX} \right]} = \frac{F(u) + F(v)}{2},
\]
which shows the 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 (iv) follows from the convexity and from the fact that $F$ and $R$ are analytic on $\{u \in \mathbb{C} : \Re u \in (c_1, c_2)\}$ (cf. [9, Chap. 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 $(r_t)_{t \geq 0}$ under the martingale measure. 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 $(t, x) \subseteq \mathbb{R}_+ \times D$. This implies by Duffie et al. [6, 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 has state space $\mathbb{R}$, so that by Definition 2.3 $R(u) = \beta u$, we require that
\[
F(u) < \infty \text{ for all } u \in \begin{cases} (1/\beta, 0] & \text{if } \beta < 0, \\ (-\infty, 0] & \text{otherwise.} \end{cases}
\]

It will be seen that the condition on $F$ is necessary to guarantee the existence of bond prices for all maturities in the term structure model. By Sato [13,
Theorem 25.17] we get an equivalent formulation of Condition 3.1 if we replace $F(u) < \infty$ by $\int_{|\xi|>1} e^{u\xi} m(d\xi) < \infty$. Next we define a quantity that generalizes 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 $\lambda$ as the positive solution of

$$R\left(-\frac{1}{\lambda}\right) = 1. \quad (3.2)$$

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 (3.2) can have at most one solution and thus $\lambda$ 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 seen that $F$ must be defined at least on $(-1/\lambda, 0]$.

We will encounter several times the condition that $\lambda > 0$. The next result gives an equivalent formulation in terms of $(\alpha, \beta, \mu)$.

**Proposition 3.3** The quasi-mean-reversion $\lambda$ is strictly positive if and only if $\alpha > 0$, or $\int_{D \setminus \{0\}} h_R(\xi) \mu(d\xi) = \infty$, or $\beta - \int_{D \setminus \{0\}} h_R(\xi) \mu(d\xi) < 0$.

**Proof** First note that by Lemma 2.6 $R(u) < \infty$ for all $u \in (-\infty, 0]$. Using estimate (2.7) and a dominated convergence argument, it is seen from (2.3) that

$$\lim_{u \to -\infty} \frac{R(u)}{u^2} = \alpha, \quad (3.3)$$

$$\lim_{u \to -\infty} \frac{R(u) - \alpha u^2}{u} = \beta_0 := \beta - \int_{D \setminus \{0\}} h_R(\xi) \mu(d\xi), \quad (3.4)$$

where $\beta_0$ can also take the value $-\infty$. Suppose now that $\alpha > 0$. Then by (3.3) we get $\lim_{u \to -\infty} R(u) = \infty$. Since $R(0) = 0$ and $R$ is continuous, it follows that there exists $\lambda > 0$ such that $R(-1/\lambda) = 1$. Similarly, if $\alpha = 0$ but $\beta_0 < 0$, it follows from (3.4) that $\lim_{u \to -\infty} R(u) = \infty$ and thus again that $\lambda > 0$.

Conversely, suppose that $\alpha = 0$ and $\beta_0 \geq 0$. Then

$$\lim_{u \to -\infty} R'(u) = \lim_{u \to -\infty} \frac{R(u)}{u} = \beta_0 \geq 0.$$

By the convexity of $R$ it follows that $R'(u) \geq 0$ for all $u \in (-\infty, 0)$. Since $R(0) = 0$, this implies that $R(u) \leq 0$ for all $u \in (-\infty, 0)$ and consequently that $\lambda = 0$. \hfill \square
3.1 Bond prices

We now consider the price \( P(t, t + x) \) at time \( t \) of a zero-coupon bond with time to maturity \( x \) given by

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

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

**Proposition 3.4** Let the risk-neutral 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

\[
P(t, t + x) = \exp \left( A(x) + r_t B(x) \right), \tag{3.5}
\]

where \( A \) and \( B \) solve the generalized Riccati equations

\[
\begin{align*}
\partial_x A(x) &= F(B(x)), \quad A(0) = 0, \tag{3.6a} \\
\partial_x B(x) &= R(B(x)) - 1, \quad B(0) = 0. \tag{3.6b}
\end{align*}
\]

**Proof** If \( D = \mathbb{R}_+ \), the assertion directly follows from Proposition 2.5 by noting that \( P(t, t + x) = Q_{x, 1} \).

If \( D = \mathbb{R} \), then, as discussed after Theorem 2.4, \((r_t)_{t \geq 0}\) is a process of OU-type, and \( R(u) \) has the simple structure \( R(u) = \beta u \). By Sato [13, (17.2)–(17.3)] we obtain in this case directly that

\[
\mathbb{E} \left[ \exp \left( - \int_t^{t+x} r_s \, ds \right) \right] = \exp \left( \int_0^x F(B(s)) \, ds + r_t B(x) \right), \tag{3.7}
\]

with \( B(x) = (1 - e^{\beta x})/\beta \) if \( \beta \neq 0 \) and \( B(x) = -x \) if \( \beta = 0 \). As a function of \( x \in \mathbb{R}_+ \), \( B \) is continuously decreasing from 0 to \( 1/\beta \) if \( \beta < 0 \) and from 0 to \( -\infty \) if \( \beta \geq 0 \). It is therefore seen that the integral on the right side of (3.7) is finite for all \( x \in \mathbb{R}_+ \) if and only if \( F \) satisfies (3.1), as required by Condition 3.1. \( \square \)

**Corollary 3.5** Let \((r_t)_{t \geq 0}\) satisfy Condition 3.1 and have quasi-mean-reversion \( \lambda \). Then the function \( B(x) \) from Proposition 3.4 is strictly decreasing and satisfies

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

**Proof** The result follows from a qualitative analysis of the autonomous ODE (3.6b). Let \( \lambda > 0 \). Since \( R(-1/\lambda) - 1 = 0 \), the point \( x_* := -1/\lambda \) is a critical point of (3.6b).

By the convexity of \( R \) and the fact that \( R(0) = 0 \) it follows that \( R'(x_*) < 0 \), so that \( x_* \) is asymptotically stable, i.e., solutions entering a small enough neighborhood of \( x_* \) must converge to \( x_* \). Since \( R(x) - 1 < 0 \) for \( x \in (x_*, 0] \) and there is no other critical
point in \((x_*, 0]\), we conclude that \(B(x)\), the solution of (3.6b) starting at 0, is strictly decreasing and converges to \(x_*\).

If \(\lambda = 0\), then there is no critical point in \((-\infty, 0]\), and \(R(x) - 1 < 0\) for \(x \in (-\infty, 0]\). It follows that \(B(x)\) is strictly decreasing and diverges to \(-\infty\). \(\square\)

3.2 The yield curve and the forward rate curve

The next results are the central theoretical results of this article and describe the global shapes of attainable yield curves in any affine one-factor term structure model.

**Definition 3.6** The (zero-coupon) yield \(Y(r_t, x)\) is given by
\[
Y(r_t, x) := -\frac{\log P(t, t + x)}{x} = -\frac{A(x)}{x} - r_t \frac{B(x)}{x} \quad \text{for all } x > 0.
\]
(3.8)

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
\[
f(r_t, x) := -\partial_x \log P(t, t + x) = -A'(x) - r_t B'(x) \quad \text{for all } x > 0.
\]
(3.9)

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 (3.6) it is seen that both the yield and the forward rate curve are continuous at 0.

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

**Theorem 3.7** Let the risk-neutral short rate process be given by a one-dimensional affine process \((r_t)_{t \geq 0}\) satisfying Condition 3.1 with quasi-mean-reversion \(\lambda\). 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
\[
b_{\text{asymp}} = \lim_{u \to -\infty} -F(u) + r_t (1 - R(u)).
\]

**Proof** From (3.6a) we obtain that
\[
\lim_{x \to \infty} \frac{A(x)}{x} = \lim_{x \to \infty} A'(x) = \lim_{x \to \infty} F(B(x)).
\]
(3.10)

If \(\lambda > 0\), then by Corollary 3.5
\[
\lim_{x \to \infty} B(x) = -1/\lambda, \quad \lim_{x \to \infty} \frac{B(x)}{x} = 0, \quad \text{and} \quad \lim_{x \to \infty} B'(x) = 0,
\]
(3.11)
and the assertion follows by combining (3.8)–(3.11). If \( \lambda = 0 \), then \( \lim_{x \to \infty} B(x) = -\infty \) and

\[
\lim_{x \to \infty} \frac{B(x)}{x} = \lim_{x \to \infty} B'(x) = \lim_{x \to \infty} R(B(x)) - 1.
\]

By setting \( u := B(x) \) we obtain the desired result.

From Theorem 3.7 it is clear that for practical purposes only 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 \( Y(r_t, \infty) = b_{\text{asympt}} \). We will now examine what happens between these two endpoints.

**Definition 3.8** 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 \( x \in \mathbb{R}_+ \).

The next theorem is our main result on the shapes of yield curves in affine one-factor models.

**Theorem 3.9** Let the risk-neutral short rate process be given by a one-dimensional affine process \((r_t)_{t \geq 0}\) satisfying Condition 3.1 and with quasi-mean-reversion \( \lambda > 0 \). In addition, suppose that \( F \neq 0 \) and that either \( F \) or \( R \) is nonlinear. Then the following hold:

– The yield curve \( Y(r_t, \cdot) \) can only be normal, inverse, or humped.
– Define \( b_{\text{norm}} := -\frac{F'(-1/\lambda)}{R'(-1/\lambda)} \) and \( b_{\text{inv}} := \begin{cases} -\frac{F'(0)}{R'(0)} & \text{if } R'(0) < 0, \\ +\infty & \text{if } R'(0) \geq 0. \end{cases} \)

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 Fig. 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 3.10** 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 3.9** Define the function \( H(x) : \mathbb{R}_+ \to \mathbb{R} \) by

\[
H(x) := Y(r_t, x)x = -A(x) - r_t B(x). \tag{3.12}
\]
Fig. 1 This figure shows a graphical summary of Theorems 3.7 and 3.9, as well as the definitions of the key quantities \( b_{\text{norm}} \), \( b_{\text{asymp}} \), and \( b_{\text{inv}} \). In any affine model satisfying the conditions of Theorem 3.9, the shapes of yield curves will follow the picture given here. They will be normal if \( r_0 \) is below \( b_{\text{norm}} \), humped if \( r_0 \) is between \( b_{\text{norm}} \) and \( b_{\text{inv}} \), and inverse if \( r_0 \) is above \( b_{\text{inv}} \). Also all yield curves will tend asymptotically to the same level \( b_{\text{asymp}} \).

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 (3.6) the first derivative of \( H \) is calculated as

\[
\partial_x H(x) = -F(B(x)) - r_t \left( R(B(x)) - 1 \right),
\]

and the second as

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

Note that \( F \) and \( R \) are continuously differentiable by Lemma 2.6 and also \( B \) by (3.6b), so that the second derivative of \( H \) is well defined and continuous. Since \( B \) is strictly decreasing by Corollary 3.5, the factor \(-B'(x)\) is positive for all \( x \in \mathbb{R}_+ \). The sign of \( \partial_{xx} H(x) \) therefore equals the sign of

\[
k(x) := F'(B(x)) + r_t R'(B(x)).
\]

From the facts 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) \( D = \mathbb{R}_+ \): We have assumed that either \( F \) or \( R \) is nonlinear. 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 can happen that \( F \) is of the form \( F(u) = bu \), so that \( k(x) = b \) and \( k \) is no longer strictly decreasing. However, by assumption, \( F \neq 0 \), so 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 necessarily have \( R(u) = \beta u \). Also, since either \( F \) or \( R \) is nonlinear, \( F \) must be nonlinear 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

\[
k(0) \geq 0 \quad \text{if and only if} \quad r_t \leq b_{\text{inv}} := \begin{cases} -\frac{F'(0)}{R'(0)} & \text{if } R'(0) < 0, \\ +\infty & \text{if } R'(0) \geq 0. \end{cases}
\]

(3.16)

Since \( B(0) = 0 \), by Proposition 3.4 it follows that

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

We distinguish two cases:

(a) If \( R'(0) < 0 \), then assertion (3.16) follows immediately.

(b) Consider the case \( R'(0) \geq 0 \). 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 \( D = \mathbb{R}_+ \) 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

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

(3.17)

Since \( \lim_{x \to \infty} B(x) = -1/\lambda \), by Corollary 3.5 we have that

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

(3.18)

By the assumption \( \lambda > 0 \) and by Definition 3.2 we have 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 (3.17) directly from (3.18).

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
If $k$ has a zero in $(0, \infty)$, 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

$$H(x) = cx, \quad x \in [0, \infty),$$

(3.19)

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 3.10 (3.19) 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 (3.19) in $[0, \infty)$. 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 $x_*$, the zero of $k(x)$, such that $H(x)$ is strictly convex on $(0, x_*)$ and strictly concave on $(x_*, \infty)$. Now there can exist at most two additional solutions $x_1, x_2$ to (3.19) with $x_1 < x_* < x_2$, so that $cx$ is crossed from below at $x_1$ and from above at $x_2$.

Because of definition (3.12), every solution to (3.19), excluding $x_0 = 0$, is also a solution to

$$Y(r_t, x) = c, \quad x \in (0, \infty),$$

(3.20)

with $r_t$ fixed. Also the properties of crossing from above/below are preserved, since $x$ is positive. This means that, in the case $r_t \geq b_{\text{inv}}$, (3.20) 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 3.8 that the yield curve is inverse. In the case $r_t \leq b_{\text{norm}}$, we have again that (3.20) 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 $b_{\text{norm}} < r_t < b_{\text{inv}}$, the yield curve crosses every horizontal line at most twice, in which case it crosses first from below and then from above. Thus, in this case, the yield curve is humped.

Corollary 3.11 Under the conditions of Theorem 3.9, the instantaneous forward rate curve has the same global behavior as the yield curve, i.e.,

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

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

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

(3.21)
Proof This follows from the fact that $\partial_x H(x)$ as given in (3.13) 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 (3.14) as

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

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 3.9, leading to the stated equivalences. Equation (3.21) is simply the condition $k(x_\ast) = 0$. \hfill \Box

Corollary 3.12 Under the conditions of Theorem 3.9, it holds that

$$b_{\text{norm}} < b_{\text{asym}} < b_{\text{inv}}$$

whenever the quantities are finite. In addition, it holds that

$$D \cap (b_{\text{norm}}, b_{\text{inv}}) \neq \emptyset.$$  \hfill (3.23)

Remark 3.13 Note that (3.23) implies that there is always some $r_t \in D$ such that the yield curve $Y(r_t, \cdot)$ is humped.

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

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

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

$$\frac{F'(-1/\lambda)}{\lambda} \leq b_{\text{asym}} \leq \frac{F'(0)}{\lambda}.$$  \hfill (3.24)

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

$$\frac{-1}{R'(-1/\lambda)} \leq \frac{1}{\lambda} \leq -\frac{1}{R'(0)}.$$  \hfill (3.25)

Since either $F$ or $R$ is nonlinear, one of the functions is strictly convex by Lemma 2.6. Consequently, both inequalities either in (3.24) or in (3.25) are strict. Putting them together, we get

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

proving (3.22) under the assumption that $R'(0) < 0$.

If $R'(0) \geq 0$, then by definition $b_{\text{inv}} = \infty$. Equation (3.24) still holds, but in (3.25) only the left inequality sign remains valid. Together this still proves that $b_{\text{norm}} < b_{\text{asym}}$, and we have shown (3.22).
To prove (3.23) we distinguish two cases:

(a) \( D = \mathbb{R} \). In this case, it is sufficient to prove that \( -\infty < b_{\text{inv}} \) and \( b_{\text{norm}} < \infty \). Consider first \( b_{\text{inv}} \). If \( R'(0) \geq 0 \), then by definition \( b_{\text{inv}} = \infty \), and nothing needs proving. 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 (3.25) we know that \( R'(-1/\lambda) \leq -\lambda < 0 \). By convexity \( F'(-1/\lambda) < \infty \), and it follows that \( b_{\text{norm}} < \infty \).

(a) \( D = \mathbb{R}_+ \). In this case, it is sufficient to prove that \( 0 \leq b_{\text{norm}} \) and to apply (3.22). As above, we have that \( b_{\text{norm}} = -F'(-1/\lambda)/R'(-1/\lambda) \) and that \( R'(-1/\lambda) \leq -\lambda < 0 \). By Definition 2.3

\[
F'(-1/\lambda) = b + \int_{(0,\infty)} \xi e^{-\xi/\lambda} m(d\xi)
\]

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

The last result 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 3.14** Let the risk-neutral short rate process be given by a one-dimensional affine process \( (r_t)_{t \geq 0} \) satisfying Condition 3.1 with \( F \neq 0 \) and quasi-mean-reversion \( \lambda > 0 \). Then the following statements are equivalent:

(i) There exists a \( r_t \in D \) such that \( Y(r_t, \cdot) \) is flat.

(ii) There exists no \( r_t \in D \) such that \( Y(r_t, \cdot) \) is humped.

(iii) The short rate process \( (r_t)_{t \geq 0} \) is deterministic.

(iv) \( F(u) = bu \) and \( R(u) = \beta u \).

**Proof** Theorem 3.9, together with Corollary 3.12, shows already that \( \neg \text{(iv)} \) implies \( \neg \text{(i)} \) and \( \neg \text{(ii)} \). Also, from the form of the generator in (2.4), 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 3.9, instead of (3.15), we obtain simply

\[
k(x) = b + r_t \beta.
\]

The yield curve is 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 3.9 the yield curve is flat if and only if \( k \) is constant and equal to 0. This is the case if \( r_t = -b/\beta \). It remains to show that \( r_t \in D \). Note that \( \beta = -\lambda < 0 \). In particular, \( \beta \neq 0 \), so that, for \( D = \mathbb{R} \), we are already done. If \( D = \mathbb{R}_+ \), we have by the admissibility conditions in Definition 2.3 that \( b \geq 0 \). Thus, \( r_t = -b/\beta \geq 0 \), and we have shown (i). \( \square \)
3.3 The limit distribution of an affine 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. The goal of this section is to establish the corresponding result for affine processes. While calculating the marginal distributions of an affine process involves solving the generalized Riccati equations \((2.5)\), it will be seen that the limit distribution is much easier to obtain and can be determined directly from the functions \(F\) and \(R\).

In the interest rate model considered in the preceding section, the short rate follows an affine process under the martingale measure, so that the results will allow us to characterize the risk-neutral asymptotic short rate distribution. Often the limit distribution under the objective measure also is of interest, but the affine property is in general not preserved by an equivalent change of measure, so that the results are not directly applicable. Nevertheless, for the sake of tractability, one can impose conditions on the measure change such that the model is affine under both the objective and risk-neutral measures. (See Nicolato and Venardos \([10]\) for an example from option pricing and Cheridito et al. \([3]\) for more general results.) In such a setting, the results can also be applied under the objective measure.

Before we state the result, 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

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

Since self-decomposability is a distributional property, we will 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 \([8, 14]\). The next theorem summarizes these results and can be found in a similar form in \([13, \text{Theorem 17.5}]\).

**Theorem 3.15** Let \((r_t)_{t \geq 0}\) be an OU-type process on \(\mathbb{R}\). If

\[ \beta < 0 \quad \text{and} \quad \int_{|\xi| > 1} \log |\xi| m(d\xi) < \infty, \]

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 self-decomposable;
(ii) The cumulant generating function \(\kappa(u) = \log \int_{\mathbb{R}} e^{ux} dL(x)\) satisfies

\[ \kappa(iu) = -\frac{1}{\beta} \int_u^0 \frac{F(is)}{s} ds \quad \text{for all } u \in \mathbb{R}. \]  

(3.26)

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 \((a, b, m, \beta)\).
As discussed in Sect. 2, every regular affine process with state space \( \mathbb{R} \) is of OU-type, so that the above theorem applies. We now state our corresponding result for affine processes on \( \mathbb{R}_+ \).

**Theorem 3.16** Let \( (r_t)_{t \geq 0} \) be a one-dimensional, regular, and conservative affine process with state space \( \mathbb{R}_+ \). If

\[
R'(0) < 0 \quad \text{and} \quad \int_{\xi > 1} \log \xi \, m(d\xi) < \infty,
\]

then \( (r_t)_{t \geq 0} \) converges in law to a limit distribution \( L \) which is independent of \( r_0 \) and whose cumulant generating function \( \kappa \) is given by

\[
\kappa(u) = \int_0^\infty \frac{F(s)}{R(s)} \, ds \quad \text{for all } u \in (-\infty, 0].
\]

**(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

\[
\hat{p}_t(x,u) = \exp(\phi(t,u) + x\psi(t,u)),
\]

where \( \phi \) and \( \psi \) satisfy the generalized Riccati equations (2.5) for all \( u \in \mathcal{U} \) and thus, in particular, for all \( u \in (-\infty, 0] \). Since \( R(0) = 0 \), 0 is a critical point of the autonomous ODE (2.5b), and by the assumption \( R'(0) < 0 \) it is asymptotically stable. By the convexity of \( R \), \( R'(0) < 0 \) also implies that \( R(u) > 0 \) for all \( u \in (-\infty, 0) \), so that \( \psi(t,u) \) is a strictly increasing function in \( t \) for all \( u \in (-\infty, 0) \). Since 0 is the only critical point of (2.5b) in \( (-\infty, 0] \), it follows that

\[
\lim_{t \to \infty} \psi(t,u) = 0 \quad \text{for all } u \in (-\infty, 0].
\]

Consequently,

\[
\lim_{t \to \infty} \log \hat{p}_t(x,u) = \lim_{t \to \infty} \phi(t,u) = \int_0^\infty F(\psi(r,u)) \, dr = \int_0^\infty \frac{F(s)}{R(s)} \, ds,
\]

where the last two equalities follow from (2.5) and the transformation \( s = \psi(t,u) \). We will now show that the last integral in (3.28) converges absolutely for all \( u \in (-\infty, 0] \). Since \( R(u) \geq 0 \) and \( F(u) \leq 0 \) for all \( u \in (-\infty, 0] \), we obtain

\[
\int_u^0 \frac{F(s)}{R(s)} \, ds = -\int_u^0 \frac{F(s)}{R(s)} \, ds \leq -\frac{1}{R'(0)} \int_u^0 \frac{F(s)}{s} \, ds, \quad u \in (-\infty, 0],
\]

where the inequality follows from the fact that the convex function \( R \) is supported by its tangent at 0. From the definition of \( F(u) \) in (2.2) it is clear that the convergence of the last integral depends only on the jump part of \( F \), i.e., the integral converges if and only if

\[
\int_u^0 \frac{1}{s} \int_{(0,\infty)} (e^{s\xi} - 1) \, m(d\xi) \, ds < \infty \quad \text{for all } u \in (-\infty, 0].
\]

(3.29)
Define $M(u, \xi) = \int_0^u e^{\xi s - 1} sds$. For a fixed $u \in (-\infty, 0]$, it is easily verified that $M(u, \xi) = O(\xi)$ as $\xi \to 0$ and that $M(u, \xi) = O(\log \xi)$ as $\xi \to \infty$. Since the Lévy measure $m(d\xi)$ integrates $(\xi \wedge 1)$ by Definition 2.3 and $1_{\xi > 1} \log \xi$ by assumption, it must also integrate $M(u, \xi)$. Applying Fubini’s theorem, (3.29) follows, so that $\kappa(u) := \int_0^u \frac{F(s)}{R(s)} ds$ converges for all $u \in (-\infty, 0]$. In particular, $\lim_{u \uparrow 0} \kappa(u) = 0$, so that the limit in (3.28) is a function that is left-continuous at 0. By standard results on Laplace transforms of probability measures (cf. [15, Theorem A.3.1]), the pointwise convergence of cumulant generating functions to a function that is left-continuous at 0 implies the convergence in distribution of $(r_t)_{t \geq 0}$ to a limit distribution $L$ with cumulant generating function given by (3.28).

Since the marginal distributions of an affine process are infinitely divisible, the limit distribution $L$ also must be infinitely divisible if it exists. In Theorem 3.15, a stronger result is given for an affine process on $\mathbb{R}$; in this case, $L$ is also self-decomposable. An obvious question is whether this result can be extended to the state space $\mathbb{R}_+$. We will see that the answer is negative. In Sect. 4.3, we give an example of an affine process with state space $\mathbb{R}_+$ which converges to an infinitely divisible limit distribution that is 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 need not be.

### 4 Applications

#### 4.1 The Vasiček model

We apply the results of the last section to the classical Vasiček model

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

(4.1)

where $(W_t)_{t \geq 0}$ is a standard Brownian motion under the risk-neutral measure 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 [16]. We advise the reader to view this paragraph as a warm-up for the following examples.

Clearly, $(r_t)_{t \geq 0}$ is a conservative affine process with

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

(4.2)

$$R(u) = -\lambda u.$$  

(4.3)

From the quadratic term in $F$ and Definition 2.3, it is seen that $(r_t)_{t \geq 0}$ has the state space $\mathbb{R}$. This property is often criticized, since it allows the short rate to become negative.
From Theorem 3.9 we calculate

\[ b_{\text{inv}} = \theta \quad \text{and} \quad b_{\text{norm}} = \theta - \frac{\sigma^2}{\lambda^2}, \]

so that the yield curve in the Vasiček model is normal if \( r_t \leq \theta - \frac{\sigma^2}{\lambda^2} \), inverse if \( r_t \geq \theta \), and humped in the remaining cases.

The long term yield is calculated from (3.7) as

\[ b_{\text{asymp}} = -F(-1/\lambda) = \theta - \frac{\sigma^2}{2\lambda^2}, \]

which is the arithmetic mean of \( b_{\text{inv}} \) and \( b_{\text{norm}} \).

Theorem 3.15 applies, and the cumulant generating function \( \kappa \) of the risk-neutral limit distribution \( L \) satisfies

\[ \kappa(iu) = -\frac{1}{\lambda} \int_0^u \frac{F(is)}{s} \, ds = \int_0^u \left( i\theta - \frac{\sigma^2}{2\lambda}s \right) \, ds = ui\theta - \frac{u^2}{2} \frac{\sigma^2}{2\lambda} \]

for \( u \in \mathbb{R} \). Hence, \( L \) is Gaussian with mean \( \theta \) 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. [4]. In this model, the short rate process \( (r_t)_{t \geq 0} \) is given by the SDE

\[ dr_t = -a(r_t - \theta) \, dt + \sigma \sqrt{r_t} \, dW_t, \quad r_0 \in \mathbb{R}^+, \] (4.4)

where \( (W_t)_{t \geq 0} \) is a standard Brownian motion under the risk-neutral measure and \( a, \theta, \sigma > 0 \). The process \( (r_t)_{t \geq 0} \) is a conservative affine process with

\[ F(u) = a\theta u, \] (4.5)

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

From Definition 2.3 it is seen that \( (r_t)_{t \geq 0} \) has the state space \( \mathbb{R}^+ \). The fact that interest rates stay nonnegative in the CIR-model is often cited as an advantage of the model over the Vasiček model. Calculating the quasi-mean-reversion (see Definition 3.2), we find that

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

From Theorem 3.7 we find that the long-term yield is given by

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

The boundary between humped and inverse behavior \( b_{\text{inv}} \) is calculated from Theorem 3.9 as

\[ b_{\text{inv}} = -\frac{F'(0)}{R'(0)} = \theta. \]
Both quantities $b_{\text{asymp}}$ and $b_{\text{inv}}$ can also be found in [4, (26) and following paragraph]. Before we consider $b_{\text{norm}}$, we quote (with the notation adapted to (4.4)) from page 394 of [4], where the shape of the yield curve is discussed:

‘When the spot rate is below the long-term yield [$= b_{\text{asymp}}$], the term structure is uniformly rising. With an interest rate in excess of $\theta$ [$= b_{\text{inv}}$], the term structure is falling. For intermediate values of the interest rate, the yield curve is humped.’

In our terminology, the authors claim that the yield curve is normal for $r_t \leq b_{\text{asymp}}$, humped for $b_{\text{asymp}} < r_t < b_{\text{inv}}$, and inverse for $r_t \geq b_{\text{inv}}$. This stands in clear contradiction to Theorem 3.9 and Corollary 3.12, where we have obtained that yield curves are normal if and only if $r_t \leq b_{\text{norm}}$ and that $b_{\text{norm}} < b_{\text{asymp}}$, or—in plain words—that there are yield curves starting strictly below the long-term yield that are still humped.

The claims of Cox et al. [4] are repeated in [11, p. 244f], where even several plots of ‘yield surfaces’ (the yield as a function of $r_t$ and $\lambda$) are presented as evidence. However, [11] fails to indicate the level of $b_{\text{asymp}}$ in the plots, so that the conclusion remains ambiguous.

To clarify the scope of humped yield curves in the CIR-model we calculate $b_{\text{norm}}$ from Theorem 3.9 to get

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

The relation $b_{\text{norm}} < b_{\text{asymp}} < b_{\text{inv}}$ is immediately confirmed by noting that $b_{\text{asymp}}$ is the harmonic mean of $b_{\text{norm}}$ and $b_{\text{inv}}$. For a graphical illustration, we refer to the second yield curve from below in Fig. 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 $(r_t)_{t \geq 0}$, we apply Theorem 3.16. The cumulant generating function $\kappa(u)$ of the limit distribution is given by

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

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. [4, 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 the risk-neutral short rate process by

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

where $J_t$ represents the jump process.
where \((J_t)_{t \geq 0}\) 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 \([5]\) as a model for default intensity and is used by Filipović \([7]\) as a short rate model. It can also be found in \([1]\) under the name JCIR model. It is easily calculated that

\[
F(u) = a\theta u + \frac{cu}{\nu - u}, \quad u \in (-\infty, \nu),
\]

\[
R(u) = \frac{\sigma^2}{2}u^2 - au.
\]

Solving the generalized Riccati equations \((3.6)\) 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 Theorems 3.7 and 3.9 in a few lines. The quasi-mean-reversion \(\lambda\) stays the same as in the CIR model, since \(R\) does not change. From

\[
F'(u) = a\theta + \frac{cv}{(v-u)^2}
\]

we immediately derive

\[
b_{\text{inv}} = \theta + \frac{c}{av}, \quad b_{\text{asymp}} = \frac{2a\theta}{a + \gamma} + \frac{2c}{\nu(a + \nu) + 2}, \quad b_{\text{norm}} = \frac{a\theta}{\gamma} + \frac{cv\sigma^4}{\gamma(\sigma^2\nu + \gamma - a)^2},
\]

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

\[
\kappa(u) = \int_u^0 \frac{F(s)}{R(s)} ds = \int_u^0 \frac{\theta}{1 - s\rho/a} ds + c \int_u^0 \frac{ds}{(s - \nu)(\rho s - a)}
\]

\[
= \begin{cases} \left( \frac{\Delta}{\Delta} - \frac{a\theta}{\rho} \right) \log(1 - \frac{\rho}{a} u) - \frac{c}{a} \log(1 - \frac{u}{\nu}) & \text{if } \Delta \neq 0, \\ -\theta \nu \log(1 - \frac{u}{\nu}) + \frac{c}{a} \frac{u}{\nu - u} & \text{if } \Delta = 0, \end{cases}
\]

as the cumulant generating function of the limit distribution \(L\) under the martingale measure.

We now take a closer look at the distribution \(L\), since this will answer the question raised at the end of Sect. 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. We consider the case \(\Delta = 0\) and define

\[
\ell(x) := \left( \theta + \frac{c}{a} x \right) \nu e^{-\nu x}, \quad x \in \mathbb{R}_+.
\]
By Frullani’s integral formula,
\[ \kappa(u) = \int_0^\infty \left( e^{ux} - 1 \right) \frac{\ell(x)}{x} \, dx \] (4.11)
for all \( u \in (-\infty, \nu) \). Since \( \ell \) is nonnegative on \( \mathbb{R}_+ \), \( \ell(x)/x \) 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 infinitely divisible distribution \( L \). In addition, \( L \) is self-decomposable if and only if \( \ell \) is nonnegative and nonincreasing on \( \mathbb{R}_+ \) (cf. [13, Corollary 15.11]).

In the case of \( \ell(x) \) given by (4.10), it is easily calculated that \( \ell(x) \) has a single maximum at \( x^* = \frac{\nu}{\nu - \frac{a\theta}{c}} \). Thus, if \( c \leq a\theta\nu \), then \( x^* \leq 0 \), so that \( \ell \) is nonincreasing on \( \mathbb{R}_+ \), and \( L \) is self-decomposable. If \( c > a\theta\nu \), then \( \ell \) is increasing in the interval \([0, x^*)\), and the limit distribution \( L \) is infinitely divisible but not self-decomposable.

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 on \( \mathbb{R}_+ \) 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, 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 (3.27) the cumulant generating function of the limit distribution is given by
\[ \kappa(u) = \frac{1}{\lambda} \int_0^u \frac{F(s)}{s} \, ds \quad \text{for all } u \in (-\infty, 0]. \] (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)
\[ F(u) = \frac{\lambda k u}{1 - \theta u}. \]

Setting \( c = \lambda k \) and \( \nu = 1/\theta \), it is seen that \( F(u) \) is equal to the last term in (4.8). This means that the driving Lévy process of \((r_t)_{t \geq 0}\) is of the same kind as the process \((J_t)_{t \geq 0}\) in (4.7), i.e., \((r_t)_{t \geq 0}\) is a pure jump OU-type process with exponentially distributed jump heights of mean \( 1/\theta \) and with jump intensity \( \lambda k \).

We interpret the affine process we have constructed as a risk-neutral short rate process. It is clear that the bond prices are of the exponentially affine form (3.5). From the generalized Riccati equation (3.6b) we obtain
\[ B(x) = \frac{e^{-\lambda x} - 1}{\lambda}. \]

From (3.6a) we calculate
\[ A(x) = \int_0^x F(B(s)) \, ds = \frac{\lambda k}{\theta + \lambda} \left( \log(1 - \theta B(x)) - \theta x \right), \]
so that the bond prices are given by

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

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, \( b_{\text{asymp}} \) is the geometric average of \( b_{\text{inv}} \) and \( b_{\text{norm}} \).

5 Conclusions

In this article, we have given, under very general conditions, a characterization of the yield curve shapes that are attainable in term structure models where the risk-neutral short rate is given by a time-homogeneous 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 an affine 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 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 Sect. 4.3, become more feasible and attractive for applications.

References