Chapter 10
Affine Processes

We have seen in Sects. 5.3 and 9.3 above that an affine diffusion induces an affine term-structure. In this chapter, we discuss the class of affine processes in more detail. Their nice analytical properties make them favorite for a broad range of financial applications, including term-structure modeling, option pricing and credit risk modeling.

10.1 Definition and Characterization of Affine Processes

As in Sect. 9.1, we fix a dimension $d \geq 1$ and a closed state space $\mathcal{X} \subset \mathbb{R}^d$ with non-empty interior. We let $b : \mathcal{X} \to \mathbb{R}^d$ be continuous, and $\rho : \mathcal{X} \to \mathbb{R}^{d\times d}$ be measurable and such that the diffusion matrix

$$a(x) = \rho(x)\rho(x)^\top$$

is continuous in $x \in \mathcal{X}$ (see Remark 4.2). Let W denote a d-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Throughout, we assume that for every $x \in \mathcal{X}$ there exists a unique solution $X = X^x$ of the stochastic differential equation

$$\begin{aligned} dX(t) &= b(X(t))\,dt + \rho(X(t))\,dW(t), \\ X(0) &= x. \end{aligned} \tag{10.1}$$

Definition 10.1 We call X *affine* if the $\mathcal{F}_t$-conditional characteristic function of $X(T)$ is exponential affine in $X(t)$, for all $t \leq T$. That is, there exist $\mathbb{C}$- and $\mathbb{C}^d$-valued functions $\phi(t,u)$ and $\psi(t,u)$, respectively, with jointly continuous t-derivatives such that $X = X^x$ satisfies

$$\mathbb{E}\left[\mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] = \mathrm{e}^{\phi(T-t,u)+\psi(T-t,u)^\top X(t)} \tag{10.2}$$

for all $u \in \mathrm{i}\mathbb{R}^d$, $t \leq T$ and $x \in \mathcal{X}$.

Since the conditional characteristic function is bounded by one, the real part of the exponent $\phi(T-t,u)+\psi(T-t,u)^\top X(t)$ in (10.2) has to be negative. Note that $\phi(t,u)$ and $\psi(t,u)$ for $t \geq 0$ and $u \in \mathrm{i}\mathbb{R}^d$ are uniquely[1] determined by (10.2), and satisfy the initial conditions $\phi(0,u) = 0$ and $\psi(0,u) = u$, in particular.

We first derive necessary and sufficient conditions for X to be affine.

[1] In fact, $\phi(t,u)$ may be altered by multiples of $2\pi\mathrm{i}$. We uniquely fix the continuous function $\phi(t,u)$ by $\phi(t,0) = 0$.

D. Filipović, *Term-Structure Models*,
Springer Finance,
DOI 10.1007/978-3-540-68015-4_10,

Theorem 10.1 *Suppose X is affine. Then the diffusion matrix $a(x)$ and drift $b(x)$ are affine in x. That is,*

$$\begin{aligned} a(x) &= a + \sum_{i=1}^{d} x_i \alpha_i, \\ b(x) &= b + \sum_{i=1}^{d} x_i \beta_i = b + \mathcal{B}x \end{aligned} \tag{10.3}$$

for some $d \times d$-matrices a and α_i, and d-vectors b and β_i, where we denote by

$$\mathcal{B} = (\beta_1, \dots, \beta_d)$$

the $d \times d$-matrix with ith column vector β_i, $1 \le i \le d$. Moreover, ϕ and $\psi = (\psi_1, \dots, \psi_d)^\top$ solve the system of Riccati equations

$$\begin{aligned} \partial_t \phi(t,u) &= \frac{1}{2}\psi(t,u)^\top a \psi(t,u) + b^\top \psi(t,u), \\ \phi(0,u) &= 0, \\ \partial_t \psi_i(t,u) &= \frac{1}{2}\psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad 1 \le i \le d, \\ \psi(0,u) &= u. \end{aligned} \tag{10.4}$$

In particular, ϕ is determined by ψ via simple integration:

$$\phi(t,u) = \int_0^t \left(\frac{1}{2}\psi(s,u)^\top a \psi(s,u) + b^\top \psi(s,u) \right) ds.$$

Conversely, suppose the diffusion matrix $a(x)$ and drift $b(x)$ are affine of the form (10.3) *and suppose there exists a solution (ϕ, ψ) of the Riccati equations* (10.4) *such that $\phi(t,u) + \psi(t,u)^\top x$ has a nonpositive real part for all $t \ge 0$, $u \in \mathrm{i}\mathbb{R}^d$ and $x \in \mathcal{X}$. Then X is affine with conditional characteristic function* (10.2).

Proof Suppose X is affine. For $T > 0$ and $u \in \mathrm{i}\mathbb{R}^d$ define the complex-valued Itô process

$$M(t) = \mathrm{e}^{\phi(T-t,u) + \psi(T-t,u)^\top X(t)}.$$

We can apply Itô's formula, separately to the real and imaginary parts of M, and obtain

$$dM(t) = I(t)\,dt + \psi(T-t,u)^\top \rho(X(t))\,dW(t), \quad t \le T,$$

with

$$\begin{aligned} I(t) = &-\partial_T \phi(T-t,u) - \partial_T \psi(T-t,u)^\top X(t) \\ &+ \psi(T-t,u)^\top b(X(t)) + \frac{1}{2}\psi(T-t,u)^\top a(X(t)) \psi(T-t,u). \end{aligned}$$

Since M is a martingale, we have $I(t) = 0$ for all $t \le T$ a.s. Letting $t \to 0$, by continuity of the parameters, we thus obtain

$$\partial_T \phi(T,u) + \partial_T \psi(T,u)^\top x = \psi(T,u)^\top b(x) + \frac{1}{2}\psi(T,u)^\top a(x)\psi(T,u)$$

for all $x \in \mathcal{X}$, $T \ge 0$, $u \in \mathrm{i}\mathbb{R}^d$. Since $\psi(0,u) = u$, this implies that a and b are affine of the form (10.3). Plugging this back into the above equation and separating first-order terms in x yields (10.4).

Conversely, suppose a and b are of the form (10.3). Let (ϕ, ψ) be a solution of the Riccati equations (10.4) such that $\phi(t,u) + \psi(t,u)^\top x$ has a nonpositive real part for all $t \ge 0$, $u \in \mathrm{i}\mathbb{R}^d$ and $x \in \mathcal{X}$. Then M, defined as above, is a uniformly bounded[2] local martingale, and hence a martingale, with $M(T) = \mathrm{e}^{u^\top X(T)}$. Therefore $\mathbb{E}[M(T) \mid \mathcal{F}_t] = M(t)$, for all $t \le T$, which is (10.2), and the theorem is proved. □

In the sequel, we will often deal with systems of Riccati equations of the type (10.4). Therefore, we now recall an important global existence, uniqueness and regularity result for differential equations, which will be used throughout without further mention. We let K be a placeholder for either $\mathbb{R}$ or $\mathbb{C}$.

Lemma 10.1 *Consider the system of ordinary differential equations*

$$\begin{aligned} \partial_t f(t,u) &= R(f(t,u)), \\ f(0,u) &= u, \end{aligned} \tag{10.5}$$

where $R : K^d \to K^d$ is a locally Lipschitz continuous function. Then the following holds:

(a) *For every $u \in K^d$, there exists a life time $t_+(u) \in (0,\infty]$ such that there exists a unique solution $f(\cdot,u) : [0, t_+(u)) \to K \times K^d$ of* (10.5).
(b) *The domain*

$$\mathcal{D}_K = \{(t,u) \in \mathbb{R}_+ \times K^d \mid t < t_+(u)\}$$

is open in $\mathbb{R}_+ \times K^d$ and maximal in the sense that either $t_+(u) = \infty$ or

$$\lim_{t \uparrow t_+(u)} \|f(t,u)\| = \infty,$$

respectively, for all $u \in K^d$.
(c) *For every $t \ge 0$, the t-section*

$$\mathcal{D}_K(t) = \{u \in K^d \mid (t,u) \in \mathcal{D}_K\}$$

[2] We note that the uniform boundedness of the local martingale M is substantial here to infer that M is a true martingale and the transform formula (10.2) holds. See also Exercise 10.4 below.

is open in K^d, *and non-expanding in* t *in the following sense*:

$$K^d = \mathcal{D}_K(0) \supseteq \mathcal{D}_K(t_1) \supseteq \mathcal{D}_K(t_2), \quad 0 \le t_1 \le t_2.$$

In fact, we have $f(s, \mathcal{D}_K(t_2)) \subseteq \mathcal{D}_K(t_1)$ *for all* $s \le t_2 - t_1$.

(d) *If* R *is analytic on* K^d *then* f *is an analytic function on* $\mathcal{D}_K$.

Proof Part (a) follows from the basic theorems for ordinary differential equations, e.g. [3, Theorem 7.4]. It is proved in [3, Theorems 7.6 and 8.3] that $\mathcal{D}_K$ is maximal and open, which is part (b). This also implies that all t-sections $\mathcal{D}_K(t)$ are open in K^d. The inclusion $\mathcal{D}_K(t_1) \supseteq \mathcal{D}_K(t_2)$ is a consequence of the maximality property from part (b), and $f(s, \mathcal{D}_K(t_2)) \subseteq \mathcal{D}_K(t_1)$ follows from the flow property $f(t_1, f(s,u)) = f(t_1 + s, u)$, whence part (c) follows. For a proof of part (d) see [55, Theorem 10.8.2]. □

It is obvious to what extent Lemma 10.1 applies to the system of Riccati equations (10.4). In particular, it is easily checked that $t_+(0) = \infty$ and thus $\mathcal{D}_K(t)$ contains 0 for all $t \ge 0$. We will provide in Sect. 10.7 and Theorem 10.3 below some substantial improvements of the properties for (10.4) stated in Lemma 10.1 for the canonical state space $\mathcal{X}$ introduced in the following section.

10.2 Canonical State Space

There is an implicit trade-off between the parameters a, α_i, b, β_i in (10.3) and the state space $\mathcal{X}$:

- a, α_i, b, β_i must be such that X does not leave the set $\mathcal{X}$;
- a, α_i must be such that $a + \sum_{i=1}^d x_i \alpha_i$ is symmetric and positive semi-definite for all $x \in \mathcal{X}$.

To gain further explicit insight into this interplay, we now and henceforth assume that the state space is of the following canonical form:

$$\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n$$

for some integers $m, n \ge 0$ with $m + n = d$. This canonical state space covers essentially all applications appearing in the finance literature.[3]

[3]Note, however, that other choices for the state space of an affine process are possible. For instance, the trivial example for $d = 1$,

$$dX = -X\,dt, \quad X(0) = x \in \mathcal{X},$$

admits as state space any closed interval $\mathcal{X} \subset \mathbb{R}$ containing 0. This degenerate diffusion process is affine, since $\mathrm{e}^{uX(T)} = \mathrm{e}^{u\mathrm{e}^{-(T-t)}X(t)}$ for all $t \le T$. A non-degenerate example is provided in Exercise 10.1. See also the discussion in [61, Sect. 12]. Moreover, semi-definite matrix-valued affine processes have recently been studied and successfully applied to finance in [30, 32, 48, 49, 84, 86].

For the above canonical state space, we can give necessary and sufficient admissibility conditions on the parameters. The following terminology will be useful in the sequel. We define the index sets

$$I = \{1, \dots, m\} \quad \text{and} \quad J = \{m+1, \dots, m+n\}.$$

For any vector μ and matrix ν, and index sets M, N, we denote by

$$\mu_M = (\mu_i)_{i \in M}, \qquad \nu_{MN} = (\nu_{ij})_{i \in M, j \in N}$$

the respective sub-vector and -matrix.

Theorem 10.2 *The process X on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ is affine if and only if $a(x)$ and $b(x)$ are affine of the form (10.3) for parameters a, α_i, b, β_i which are admissible in the following sense*:

$$\begin{aligned}
& a, \alpha_i \text{ are symmetric positive semi-definite}, \\
& a_{II} = 0 \quad (\text{and thus } a_{IJ} = a_{JI}^\top = 0), \\
& \alpha_j = 0 \quad \text{for all } j \in J, \\
\alpha_{i,kl} = \alpha_{i,lk} &= 0 \quad \text{for } k \in I \setminus \{i\}, \text{for all } 1 \le i, l \le d, \\
& b \in \mathbb{R}_+^m \times \mathbb{R}^n, \\
& \mathcal{B}_{IJ} = 0, \\
& \mathcal{B}_{II} \text{ has nonnegative off-diagonal elements}.
\end{aligned} \tag{10.6}$$

In this case, the corresponding system of Riccati equations (10.4) simplifies to

$$\begin{aligned}
\partial_t \phi(t,u) &= \frac{1}{2} \psi_J(t,u)^\top a_{JJ} \psi_J(t,u) + b^\top \psi(t,u), \\
\phi(0,u) &= 0, \\
\partial_t \psi_i(t,u) &= \frac{1}{2} \psi(t,u)^\top \alpha_i \psi(t,u) + \beta_i^\top \psi(t,u), \quad i \in I, \\
\partial_t \psi_J(t,u) &= \mathcal{B}_{JJ}^\top \psi_J(t,u), \\
\psi(0,u) &= u,
\end{aligned} \tag{10.7}$$

and there exists a unique global solution $(\phi(\cdot,u), \psi(\cdot,u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ for all initial values $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$. In particular, the equation for ψ_J forms an autonomous linear system with unique global solution $\psi_J(t,u) = \mathrm{e}^{\mathcal{B}_{JJ}^\top t} u_J$ for all $u_J \in \mathbb{C}^n$.

Before we prove the theorem, let us illustrate the admissibility conditions (10.6) for the diffusion matrix $\alpha(x)$ for dimension $d = 3$ and the corresponding cases $m =$

0, 1, 2, 3. For the first case $m=0$ we have

$$\alpha(x) \equiv a$$

for an arbitrary positive semi-definite symmetric 3×3-matrix a. For $m=1$, we have

$$a = \begin{pmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{pmatrix}, \qquad \alpha_1 = \begin{pmatrix} + & * & * \\ & + & * \\ & & + \end{pmatrix},$$

for $m=2$,

$$a = \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{pmatrix}, \qquad \alpha_1 = \begin{pmatrix} + & 0 & * \\ & 0 & 0 \\ & & + \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{pmatrix},$$

and for $m=3$,

$$a=0, \qquad \alpha_1 = \begin{pmatrix} + & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 \\ & + & 0 \\ & & 0 \end{pmatrix}, \qquad \alpha_3 = \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{pmatrix},$$

where we leave the lower triangle of symmetric matrices blank; $+$ denotes a non-negative real number and $*$ any real number such that positive semi-definiteness holds.

Proof Suppose X is affine. That $a(x)$ and $b(x)$ are of the form (10.3) follows from Theorem 10.1. Obviously, $a(x)$ is symmetric positive semi-definite for all $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$ if and only if $\alpha_j = 0$ for all $j \in J$, and a and α_i are symmetric positive semi-definite for all $i \in I$.

We extend the diffusion matrix and drift continuously to $\mathbb{R}^d$ by setting

$$a(x) = a + \sum_{i\in I} x_i^+ \alpha_i \quad \text{and} \quad b(x) = b + \sum_{i\in I} x_i^+ \beta_i + \sum_{j\in J} x_j \beta_j.$$

Now let x be a boundary point of $\mathbb{R}^m_+ \times \mathbb{R}^n$. That is, $x_k = 0$ for some $k \in I$. The stochastic invariance Lemma 10.11 below implies that the diffusion must be "parallel to the boundary",

$$e_k^\top \left(a + \sum_{i\in I\setminus\{k\}} x_i \alpha_i \right) e_k = 0,$$

and the drift must be "inward pointing",

$$e_k^\top \left(b + \sum_{i\in I\setminus\{k\}} x_i \beta_i + \sum_{j\in J} x_j \beta_j \right) \geq 0.$$

Since this has to hold for all $x_i \geq 0$, $i \in I \setminus \{k\}$, and $x_j \in \mathbb{R}$, $j \in J$, we obtain the following set of admissibility conditions:

$$\begin{aligned}
a, \alpha_i &\text{ are symmetric positive semi-definite,}\\
a e_k &= 0 \quad \text{for all } k \in I,\\
\alpha_i e_k &= 0 \quad \text{for all } i \in I \setminus \{k\}, \text{ for all } k \in I,\\
\alpha_j &= 0 \quad \text{for all } j \in J,\\
b &\in \mathbb{R}^m_+ \times \mathbb{R}^n,\\
\beta_i^\top e_k &\geq 0 \quad \text{for all } i \in I \setminus \{k\}, \text{ for all } k \in I,\\
\beta_j^\top e_k &= 0 \quad \text{for all } j \in J, \text{ for all } k \in I,
\end{aligned}$$

which is equivalent to (10.6). The form of the system (10.7) follows by inspection.

Now suppose a, α_i, b, β_i satisfy the admissibility conditions (10.6). We show below that there exists a unique global solution $(\phi(\cdot, u), \psi(\cdot, u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$ of (10.7), for all $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$. In particular, $\phi(t, u) + \psi(t, u)^\top x$ has nonpositive real part for all $t \geq 0$, $u \in \mathrm{i}\mathbb{R}^d$ and $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$. Thus the first part of the theorem follows from Theorem 10.1.

It view of the admissibility conditions for a and b, it remains to show that $\psi(t, u)$ is $\mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$-valued and has life time $t_+(u) = \infty$ for all $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$. For $i \in I$, denote the right-hand side of the equation for ψ_i by

$$R_i(u) = \frac{1}{2} u^\top \alpha_i u + \beta_i^\top u,$$

and observe that

$$\Re R_i(u) = \frac{1}{2} \Re u^\top \alpha_i \Re u - \frac{1}{2} \Im u^\top \alpha_i \Im u + \beta_i^\top \Re u.$$

Let us define $x_I^+ = (x_1^+, \ldots, x_m^+)^\top$. Since $\Re \psi_J(t, u) = 0$, it follows from the admissibility conditions (10.6) and Corollary 10.5 below, setting $f(t) = -\Re \psi(t, u)$,

$$b_i(t, x) = -\frac{1}{2} \alpha_{i,ii} \left(x_i^+\right)^2 + \frac{1}{2} \Im \psi(t, u)^\top \alpha_i \Im \psi(t, u) + \beta_{i,I}^\top x_I^+, \quad i \in I,$$

and $b_j(t, x) = 0$ for $j \in J$, that the solution $\psi(t, u)$ of (10.7) has to take values in $\mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$ for all initial points $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$.

Further, for $i \in I$ and $u \in \mathbb{C}^d$, one verifies that

$$\begin{aligned}
\Re(\overline{u_i} R_i(u)) &= \frac{1}{2} \alpha_{i,ii} |u_i|^2 \Re u_i + \Re(\overline{u_i} u_i \alpha_{i,iJ} u_J) + \frac{1}{2} \Re(\overline{u_i} u_J^\top \alpha_{i,JJ} u_J) + \Re(\overline{u_i} \beta_i^\top u)\\
&\leq \frac{K}{2} \left(1 + \|(\Re u_I)^+\| + \|u_J\|^2\right) \left(1 + \|u_I\|^2\right)
\end{aligned}$$

for some finite constant K which does not depend on u. We thus obtain

$$\begin{aligned}\partial_t \|\psi_I(t,u)\|^2 &= 2\Re\left(\overline{\psi_I(t,u)}^\top R_I\left(\psi_I(t,u), \mathrm{e}^{\mathcal{B}_{JJ}^\top t} u_J\right)\right)\\ &\le g(t)\left(1+\|\psi_I(t,u)\|^2\right)\end{aligned}$$

for

$$g(t) = K\left(1 + \|(\Re\psi_I(t,u))^+\| + \|\mathrm{e}^{\mathcal{B}_{JJ}^\top t} u_J\|^2\right).$$

Gronwall's inequality[4] applied to $f(t) = (1+\|\psi_I(t,u)\|^2)$ and $h(t) \equiv f(0)$, yields

$$\|\psi_I(t,u)\|^2 \le \|u_I\|^2 + \left(1+\|u_I\|^2\right)\int_0^t g(s)\mathrm{e}^{\int_s^t g(\xi)\,d\xi}\,ds. \tag{10.8}$$

From above, for all initial points $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$, we know that $(\Re\psi_I(t,u))^+ = 0$ and therefore $t_+(u) = \infty$ by (10.8). Hence the theorem is proved. □

Now suppose X is affine with characteristics (10.3) satisfying the admissibility conditions (10.6). In what follows we show that not only can the functions $\phi(t,u)$ and $\psi(t,u)$, given as solutions of (10.7), be extended beyond $u \in \mathrm{i}\mathbb{R}^d$, but also the validity of the affine transform formula (10.2) carries over. In fact, we will show that (10.2) holds for $u \in \mathbb{R}^d$ if either side is well defined. This asserts exponential moments of $X(t)$ in particular and will prove most useful for deriving pricing formulas in affine factor models.

For any set $U \subset \mathbb{R}^k$ ($k \in \mathbb{N}$), we define the strip

$$\mathcal{S}(U) = \left\{z \in \mathbb{C}^k \mid \Re z \in U\right\}$$

in $\mathbb{C}^k$. The proof of the following theorem is postponed to Sect. 10.7.3. It builds on results that are developed in Sects. 10.6 and 10.7 below.

Theorem 10.3 *Suppose X is affine with admissible parameters as given in* (10.6). *Let $\mathcal{D}_K$ ($K = \mathbb{R}$ or $\mathbb{C}$) denote the maximal domain for the system of Riccati equations* (10.7), *and let $\tau > 0$. Then*:

---

[4]Let f, g and h be nonnegative continuous functions $[0,T] \to \mathbb{R}_+$ with

$$f(t) \le h(t) + \int_0^t g(s) f(s)\,ds, \quad t \in [0,T].$$

Then

$$f(t) \le h(t) + \int_0^t h(s) g(s)\mathrm{e}^{\int_s^t g(\xi)\,d\xi}\,ds, \quad t \in [0,T],$$

see [55, (10.5.1.3)].

(a) $\mathcal{S}(\mathcal{D}_{\mathbb{R}}(\tau)) \subset \mathcal{D}_{\mathbb{C}}(\tau)$.
(b) $\mathcal{D}_{\mathbb{R}}(\tau) = M(\tau)$ *where*

$$M(\tau) = \left\{ u \in \mathbb{R}^d \mid \mathbb{E}\left[\mathrm{e}^{u^\top X^x(\tau)} \right] < \infty \textit{ for all } x \in \mathbb{R}_+^m \times \mathbb{R}^n \right\}.$$

(c) $\mathcal{D}_{\mathbb{R}}(\tau)$ *and* $\mathcal{D}_{\mathbb{R}}$ *are convex sets.*

Moreover, for all $0 \le t \le T$ *and* $x \in \mathbb{R}_+^m \times \mathbb{R}^n$:

(d) (10.2) *holds for all* $u \in \mathcal{S}(\mathcal{D}_{\mathbb{R}}(T-t))$.
(e) (10.2) *holds for all* $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$.
(f) $M(t) \supseteq M(T)$.

As a corollary we may thus formulate the following key message of Theorem 10.3 parts (a), (b) and (d).

Corollary 10.1 *Suppose that either side of* (10.2) *is well defined for some* $t \le T$ *and* $u \in \mathbb{R}^d$. *Then* (10.2) *holds, implying that both sides are well defined in particular, for* u *replaced by* $u + \mathrm{i}v$ *for any* $v \in \mathbb{R}^d$.

Part (f) of Theorem 10.3 states that integrability of $\mathrm{e}^{u^\top X^x(T)}$ for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$, for some given T and $u \in \mathbb{R}^d$, implies integrability of $\mathrm{e}^{u^\top X^x(t)}$ for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$ and $t \le T$. In other words, the set of exponential moment parameters $M(t)$ is non-expanding in t.

10.3 Discounting and Pricing in Affine Models

We let X be affine on the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$ with admissible parameters a, α_i, b, β_i as given in (10.6). Since we are interested in pricing, we interpret

$$\mathbb{P} = \mathbb{Q}$$

as risk-neutral measure and $W = W^*$ as $\mathbb{Q}$-Brownian motion in this section.[5]

A short-rate model of the form

$$r(t) = c + \gamma^\top X(t), \tag{10.9}$$

for some constant parameters $c \in \mathbb{R}$ and $\gamma \in \mathbb{R}^d$, is called an affine short-rate model. Special cases, for dimension $d = 1$, are the Vasiček and CIR short-rate models. We recall from (9.11) that an affine term-structure model always induces an affine short-rate model.

[5]Note, however, that the affine property of X is not preserved under an equivalent change of measure in general. Measure changes which preserve the affine structure are studied in detail in Cheridito, Filipović and Yor [42].

Now let $T > 0$, and consider a T-claim with payoff of the form $f(X(T))$ which meets the required integrability condition

$$\mathbb{E}\left[\mathrm{e}^{-\int_0^T r(s)\,ds} |f(X(T))|\right] < \infty,$$

see (7.3). Its arbitrage price at time $t \le T$ is then given by

$$\pi(t) = \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds} f(X(T)) \mid \mathcal{F}_t\right]. \tag{10.10}$$

Compare this also to (5.6). A particular example is the T-bond with $f \equiv 1$. Our aim is to derive an analytic, or at least numerically tractable, pricing formula for (10.10).

As a first step, we derive a formula for the $\mathcal{F}_t$-conditional characteristic function of $X(T)$ under the T-forward measure, which equals, up to normalization with $P(t,T)$ (see Sect. 7.1),

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds} \mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] \tag{10.11}$$

for $u \in \mathrm{i}\mathbb{R}^d$. Note that the following integrability condition (a) is satisfied in particular if r is uniformly bounded from below, that is, if $\gamma \in \mathbb{R}_+^m \times \{0\}$ (see also Exercise 10.4).

Theorem 10.4 *Let $\tau > 0$. The following statements are equivalent*:

(a) $\mathbb{E}[\mathrm{e}^{-\int_0^\tau r(s)\,ds}] < \infty$ *for all* $x \in \mathbb{R}_+^m \times \mathbb{R}^n$.
(b) *There exists a unique solution* $(\Phi(\cdot,u), \Psi(\cdot,u)) : [0,\tau] \to \mathbb{C} \times \mathbb{C}^d$ *of*

$$\begin{aligned}
\partial_t \Phi(t,u) &= \frac{1}{2}\Psi_J(t,u)^\top a_{JJ}\Psi_J(t,u) + b^\top \Psi(t,u) - c,\\
\Phi(0,u) &= 0,\\
\partial_t \Psi_i(t,u) &= \frac{1}{2}\Psi(t,u)^\top \alpha_i \Psi(t,u) + \beta_i^\top \Psi(t,u) - \gamma_i, \quad i \in I,\\
\partial_t \psi_J(t,u) &= \mathcal{B}_{JJ}^\top \Psi_J(t,u) - \gamma_J,\\
\Psi(0,u) &= u
\end{aligned} \tag{10.12}$$

for $u = 0$.

Moreover, let $\mathcal{D}_K$ ($K = \mathbb{R}$ or $\mathbb{C}$) denote the maximal domain for the system of Riccati equations (10.12)*. If either* (a) *or* (b) *holds then $\mathcal{D}_\mathbb{R}(S)$ is a convex open neighborhood of* 0 *in $\mathbb{R}^d$, and $\mathcal{S}(\mathcal{D}_\mathbb{R}(S)) \subset \mathcal{D}_\mathbb{C}(S)$, for all $S \le \tau$. Further,* (10.11) *allows the following affine representation*:

$$\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds} \mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] = \mathrm{e}^{\Phi(T-t,u) + \Psi(T-t,u)^\top X(t)} \tag{10.13}$$

for all $u \in \mathcal{S}(\mathcal{D}_\mathbb{R}(S))$, $t \le T \le t + S$ and $x \in \mathbb{R}_+^m \times \mathbb{R}^n$.

Proof We first enlarge the state space and consider the real-valued process

$$Y(t) = y + \int_0^t \Big(c + \gamma^\top X(s)\Big)\, ds, \quad y \in \mathbb{R}.$$

A moment's reflection reveals that $X' = \binom{X}{Y}$ is an $\mathbb{R}_+^m \times \mathbb{R}^{n+1}$-valued diffusion process with diffusion matrix $a' + \sum_{i\in I} x_i \alpha_i'$ and drift $b' + \mathcal{B}' x'$ where

$$a' = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}, \qquad \alpha_i' = \begin{pmatrix} \alpha_i & 0 \\ 0 & 0 \end{pmatrix}, \qquad b' = \begin{pmatrix} b \\ c \end{pmatrix}, \qquad \mathcal{B}' = \begin{pmatrix} \mathcal{B} & 0 \\ \gamma^\top & 0 \end{pmatrix}$$

form admissible parameters. We claim that X' is an affine process.

Indeed, the candidate system of Riccati equations reads, for $i \in I$:

$$\begin{aligned}
\partial_t \phi'(t,u,v) &= \frac{1}{2}\psi_J'(t,u,v)^\top a_{JJ} \psi_J'(t,u,v) + b^\top \psi_{\{1,\dots,d\}}'(t,u,v) + \boxed{cv}, \\
\phi'(0,u,v) &= 0, \\
\partial_t \psi_i'(t,u,v) &= \frac{1}{2}\psi'(t,u,v)^\top \alpha_i \psi'(t,u,v) + \beta_i^\top \psi_{\{1,\dots,d\}}'(t,u,v) + \boxed{\gamma_i v}, \\
\partial_t \psi_J'(t,u,v) &= \mathcal{B}_{JJ}^\top \psi_J'(t,u,v) + \boxed{\gamma_J v}, \\
\partial_t \psi_{d+1}'(t,u,v) &= 0, \\
\psi'(0,u,v) &= \begin{pmatrix} u \\ v \end{pmatrix}.
\end{aligned} \tag{10.14}$$

Here we replaced the constant solution $\psi_{d+1}'(\cdot,u,v) \equiv v$ by v in the boxes. Theorem 10.2 carries over and asserts a unique global $\mathbb{C}_- \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^{n+1}$-valued solution $(\phi'(\cdot,u,v), \psi'(\cdot,u,v))$ of (10.12) for all $(u,v) \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n \times \mathrm{i}\mathbb{R}$. The second part of Theorem 10.1 thus asserts that X' is affine with conditional characteristic function

$$\mathbb{E}\left[\mathrm{e}^{u^\top X(T) + vY(T)} \mid \mathcal{F}_t\right] = \mathrm{e}^{\phi'(T-t,u,v) + \psi_{\{1,\dots,d\}}'(T-t,u,v)^\top X(t) + vY(t)}$$

for all $(u,v) \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n \times \mathrm{i}\mathbb{R}$ and $t \le T$.

The theorem now follows from Theorem 10.3 once we set $\Phi(t,u) = \phi'(t,u,-1)$ and $\Psi(t,u) = \psi_{\{1,\dots,d\}}'(t,u,-1)$. Indeed, it is clear by inspection that $\mathcal{D}_K(S) = \{u \in K^d \mid (u,-1) \in \mathcal{D}_K'(S)\}$ where $\mathcal{D}_K'$ denotes the maximal domain for the system of Riccati equations (10.14). □

As immediate consequence of Theorem 10.4, we obtain the following explicit price formulas for T-bonds in terms of Φ and Ψ.

Corollary 10.2 *For any maturity $T \le \tau$, the T-bond price at $t \le T$ is given as*

$$P(t,T) = \mathrm{e}^{-A(T-t) - B(T-t)^\top X(t)}$$

where we define, in accordance with Sect. 5.3,

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

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

$$\mathbb{E}_{\mathbb{Q}^S}\left[\mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] = \frac{\mathrm{e}^{-A(S-T)+\Phi(T-t,u-B(S-T))+\Psi(T-t,u-B(S-T))^\top X(t)}}{P(t,S)} \tag{10.15}$$

for all $u \in \mathcal{S}(\mathcal{D}_{\mathbb{R}}(T) + B(S-T))$, which contains $\mathrm{i}\mathbb{R}^d$.

Proof The bond price formula follows from (10.13) with $u = 0$.

Now let $t \le T \le S \le \tau$. In view of the flow property $\Psi(T, -B(S-T)) = -B(S)$, we know that $-B(S-T) \in \mathcal{D}_{\mathbb{R}}(T)$, and thus $\mathcal{S}(\mathcal{D}_{\mathbb{R}}(T) + B(S-T))$ contains $\mathrm{i}\mathbb{R}^d$. Moreover, for $u \in \mathcal{S}(\mathcal{D}_{\mathbb{R}}(T) + B(S-T))$, we obtain from (10.13) by nested conditional expectation

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{-\int_t^S r(s)\,ds}\mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] &= \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds}\mathbb{E}\left[\mathrm{e}^{-\int_T^S r(s)\,ds} \mid \mathcal{F}_T\right]\mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t\right] \\
&= \mathrm{e}^{-A(S-T)}\mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds}\mathrm{e}^{(u-B(S-T))^\top X(T)} \mid \mathcal{F}_t\right] \\
&= \mathrm{e}^{-A(S-T)+\Phi(T-t,u-B(S-T))+\Psi(T-t,u-B(S-T))^\top X(t)}.
\end{aligned}$$

Normalizing by $P(t,S)$ yields (10.15). □

For more general payoff functions f, we can proceed as follows. Either we recognize the $\mathcal{F}_t$-conditional distribution, say $Q(t,T,dx)$, of $X(T)$ under the T-forward measure $\mathbb{Q}^T$ from its characteristic function in (10.15). Then compute the price (10.10) by (numerical) integration of f

$$\pi(t) = P(t,T)\int_{\mathbb{R}^d} f(x) Q(t,T,dx). \tag{10.16}$$

Or we employ Fourier transform techniques as the following two consecutive affine pricing theorems indicate.

Theorem 10.5 *Suppose either condition* (a) *or* (b) *of Theorem* 10.4 *is met for some $\tau \ge T$, and let $\mathcal{D}_{\mathbb{R}}$ denote the maximal domain for the system of Riccati equations* (10.12). *Assume that f satisfies*

$$f(x) = \int_{\mathbb{R}^q} \mathrm{e}^{(v+\mathrm{i}L\lambda)^\top x}\widetilde{f}(\lambda)\,d\lambda, \quad dx\text{-}a.s. \tag{10.17}$$

for some $v \in \mathcal{D}_{\mathbb{R}}(T)$ *and* $d \times q$*-matrix* L*, and some integrable function* $\widetilde{f} : \mathbb{R}^q \to \mathbb{C}$*, for a positive integer* $q \le d$*. Then the price* (10.10) *is well defined and given by the formula*

$$\pi(t) = \int_{\mathbb{R}^q} \mathrm{e}^{\Phi(T-t, v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)} \widetilde{f}(\lambda)\, d\lambda. \tag{10.18}$$

From the Riemann–Lebesgue theorem ([156, Chap. I, Theorem 1.2]) we know that the right-hand side of (10.17) is continuous in x. Hence the equality (10.17) necessarily holds for all x if f is continuous.

Proof By assumption, we have

$$\mathbb{E}\left[\mathrm{e}^{-\int_0^T r(s)\,ds} |f(X(T))|\right] \le \mathbb{E}\left[\int_{\mathbb{R}^q} \mathrm{e}^{-\int_0^T r(s)\,ds} \mathrm{e}^{v^\top X(T)} |\widetilde{f}(\lambda)|\, d\lambda\right] < \infty.$$

Hence we may apply Fubini's theorem to change the order of integration, which gives

$$\begin{aligned}\pi(t) &= \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds} \int_{\mathbb{R}^q} \mathrm{e}^{(v+\mathrm{i}L\lambda)^\top X(T)} \widetilde{f}(\lambda)\, d\lambda \mid \mathcal{F}_t\right] \\ &= \int_{\mathbb{R}^q} \mathbb{E}\left[\mathrm{e}^{-\int_t^T r(s)\,ds} \mathrm{e}^{(v+\mathrm{i}L\lambda)^\top X(T)} \mid \mathcal{F}_t\right] \widetilde{f}(\lambda)\, d\lambda \\ &= \int_{\mathbb{R}^q} \mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)} \widetilde{f}(\lambda)\, d\lambda,\end{aligned}$$

which is (10.18). □

Next, we give a more constructive and alternative approach, respectively, to the representation (10.17).

Theorem 10.6 *Suppose either condition* (a) *or* (b) *of Theorem* 10.4 *is met for some* $\tau \ge T$*, and let* $\mathcal{D}_{\mathbb{R}}$ *denote the maximal domain for the system of Riccati equations* (10.12)*. Assume that* f *is of the form*

$$f(x) = \mathrm{e}^{v^\top x} h(L^\top x)$$

for some $v \in \mathcal{D}_{\mathbb{R}}(T)$ *and* $d \times q$*-matrix* L*, and some integrable function* $h : \mathbb{R}^q \to \mathbb{R}$*, for a positive integer* $q \le d$*. Define the bounded function*

$$\widetilde{f}(\lambda) = \frac{1}{(2\pi)^q} \int_{\mathbb{R}^q} \mathrm{e}^{-\mathrm{i}\lambda^\top y} h(y)\, dy, \quad \lambda \in \mathbb{R}^q.$$

(a) *If* $\widetilde{f}(\lambda)$ *is an integrable function in* $\lambda \in \mathbb{R}^q$ *then the assumptions of Theorem* 10.5 *are met.*

(b) *If* $v = Lw$, *for some* $w \in \mathbb{R}^q$, *and* $\mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}$ *is an integrable function in* $\lambda \in \mathbb{R}^q$ *then the* $\mathcal{F}_t$*-conditional distribution of the* $\mathbb{R}^q$*-valued random variable* $Y = L^\top X(T)$ *under the* T*-forward measure* $\mathbb{Q}^T$ *admits the continuous density function*

$$q(t,T,y) = \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q} \mathrm{e}^{-(w+\mathrm{i}\lambda)^\top y}\frac{\mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}}{P(t,T)}\,d\lambda. \tag{10.19}$$

In either case, the integral in (10.18) *is well defined and the price formula* (10.18) *holds.*

Proof We recall the fundamental inversion formula from Fourier analysis ([156, Chap. I, Corollary 1.21]): let $g : \mathbb{R}^q \to \mathbb{C}$ be an integrable function with integrable Fourier transform

$$\hat{g}(\lambda) = \int_{\mathbb{R}^q} \mathrm{e}^{-\mathrm{i}\lambda^\top y} g(y)\,dy.$$

Then the inversion formula

$$g(y) = \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q} \mathrm{e}^{\mathrm{i}\lambda^\top y}\hat{g}(\lambda)\,d\lambda \tag{10.20}$$

holds for dy-almost all $y \in \mathbb{R}^q$.

Under the assumption of (a), the Fourier inversion formula (10.20) applied to $h(y)$ yields the representation (10.17). Hence Theorem 10.5 applies.

As for (b), we denote by $q(t,T,dy)$ the $\mathcal{F}_t$-conditional distribution of $Y = L^\top X(T)$ under the T-forward measure $\mathbb{Q}^T$. From (10.15) we infer the characteristic function of the bounded (why?) measure $\mathrm{e}^{w^\top y} q(t,T,dy)$:

$$\begin{aligned}\int_{\mathbb{R}^q} \mathrm{e}^{(w+\mathrm{i}\lambda)^\top y} q(t,T,dy) &= \mathbb{E}\left[\mathrm{e}^{(w+\mathrm{i}\lambda)^\top L^\top X(T)} \mid \mathcal{F}_t\right]\\ &= \frac{\mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}}{P(t,T)}, \quad \lambda \in \mathbb{R}^q.\end{aligned}$$

By assumption, this is an integrable function in λ on $\mathbb{R}^q$. The Fourier inversion formula (10.20) thus applies and the injectivity of the characteristic function (see e.g. [161, Sect. 16.6]) yields that $q(t,T,dy)$ admits the continuous density function (10.19). Moreover, we then obtain

$$\begin{aligned}&P(t,T)\int_{\mathbb{R}^q} |\mathrm{e}^{w^\top y} h(y)| q(t,T,y)\,dy\\ &\quad\le \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q}\int_{\mathbb{R}^q} |h(y)| \left|\mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}\right| d\lambda\,dy < \infty.\end{aligned}$$

Hence again we can apply Fubini's theorem to change the order of integration, which gives

$$\begin{aligned}\pi(t) &= P(t,T)\int_{\mathbb{R}^q} \mathrm{e}^{w^\top y} h(y) q(t,T,y)\,dy \\ &= \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q}\int_{\mathbb{R}^q} \mathrm{e}^{w^\top y} h(y) \mathrm{e}^{-(w+\mathrm{i}\lambda)^\top y} \mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}\,d\lambda\,dy \\ &= \frac{1}{(2\pi)^q}\int_{\mathbb{R}^q}\left(\int_{\mathbb{R}^q} h(y)\mathrm{e}^{-\mathrm{i}\lambda^\top y}\,dy\right) \mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}\,d\lambda,\end{aligned}$$

which is (10.18). □

The integral in the pricing formula (10.18), as well as $\widetilde{f}$, has to be computed numerically in general. In this regard, it is remarkable that this integral is over $\mathbb{R}^q$, where q may be much smaller than the dimension d of the state process X. In fact, we will see that $\widetilde{f}$ is given in closed form and $q=1$ for bond options.

Let us reflect for a moment on the representation (10.17): the payoff $f(X(T))$ is decomposed into a linear combination of (a continuum) of complex-valued basis "payoffs"[6] $\mathrm{e}^{(v+\mathrm{i}L\lambda)^\top X(T)}$ with weights $\widetilde{f}(\lambda)$. By the very nature of the affine process X, these basis claims admit closed-form complex-valued "prices"

$$\pi_{v+\mathrm{i}L\lambda}(t) = \mathrm{e}^{\Phi(T-t,v+\mathrm{i}L\lambda)+\Psi(T-t,v+\mathrm{i}L\lambda)^\top X(t)}.$$

Linearity of pricing thus implies that the price of $f(X(T))$ is given as linear combination of the $\pi_{v+\mathrm{i}L\lambda}(t)$ with the same weights $\widetilde{f}(\lambda)$. But this is just formula (10.18). This reflection unfolds the power of affine diffusion processes. It suggests that we explore other types of diffusion processes that admit closed-form prices for some well specified basis of payoff functions. This approach has been pursued in e.g. [21, 40] and others. It is currently an open area of research.

Our affine pricing theorems 10.5 and 10.6 would not have much practical implications unless we find some interesting payoff functions of the form (10.17). Luckily, such representations do indeed exist for a broad range of the most important payoff functions as we shall see in the following section.

10.3.1 Examples of Fourier Decompositions

We first show that the functions $(\mathrm{e}^y - K)^+$ and $(K-\mathrm{e}^y)^+$ related to the European call and put option payoffs can be explicitly represented in the form (10.17).

[6] Admitting for the sake of reflection that there is such as a complex-valued currency.

Lemma 10.2 *Let* $K>0$. *For any* $y\in\mathbb{R}$ *the following identities hold*:

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

The middle case $(0<w<1)$ *obviously also equals* $(K-\mathrm{e}^y)^+-K$.

We will give some applications of this formula for bond and stock options in Sects. 10.3.2 and 10.3.3 below.

Proof Let $w<0$. Then the function $h(y)=\mathrm{e}^{-wy}(K-\mathrm{e}^y)^+$ is integrable on $\mathbb{R}$. An easy calculation shows that its Fourier transform

$$\hat{h}(\lambda)=\int_{\mathbb{R}}\mathrm{e}^{-(w+\mathrm{i}\lambda)y}(K-\mathrm{e}^y)^+\,dy=\frac{K^{-(w-1+\mathrm{i}\lambda)}}{(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)} \tag{10.21}$$

is also integrable on $\mathbb{R}$. Hence the Fourier inversion formula (10.20) applies, and we conclude that the claimed identity holds for $w<0$. The other cases follow by similar arguments ($\rightarrow$ Exercise 10.5). □

Nothing prevents us from choosing $K=\mathrm{e}^z$ in Lemma 10.2. This way, we obtain the following useful formula related to the payoff of an exchange option.

Corollary 10.3 *For any* $y,z\in\mathbb{R}$ *the following identities hold*:

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

Suppose two asset prices are modeled as $S_i=\mathrm{e}^{X_{m+i}}$, $i=1,2$, where X denotes our affine state diffusion on $\mathbb{R}^m_+\times\mathbb{R}^n$. Then the payoff of the option to exchange c_2 units of asset S_2 against c_1 units of asset S_1 at some date T is[7]

$$f(X(T))=\left(c_1\mathrm{e}^{X_{m+1}(T)}-c_2\mathrm{e}^{X_{m+2}(T)}\right)^+. \tag{10.22}$$

In view of Corollary 10.3, this payoff function can be represented as (10.17) where $q=1$, $v=we_{m+1}+(1-w)e_{m+2}$, $L=e_{m+1}-e_{m+2}$, and

$$\widetilde{f}(\lambda)=\frac{c_1^{w+\mathrm{i}\lambda}c_2^{-(w-1+\mathrm{i}\lambda)}}{2\pi(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)},$$

for some $w>1$.

---

[7] An exchange option is also called Margrabe option. The price formula was derived by Margrabe [121] and Fischer [76] for the case of two jointly lognormal stock price processes.

In a similar way, but now including double integration, we can find an explicit Fourier decomposition of the spread option payoff

$$f(X(T)) = \left(\mathrm{e}^{X_{m+1}(T)} - \mathrm{e}^{X_{m+2}(T)} - K\right)^+, \tag{10.23}$$

for some strike price $K > 0$. Indeed, Lemma 10.3 below implies that this payoff function can be represented as (10.17) where $q = 2$, $v = w_1 e_{m+1} + w_2 e_{m+2}$, $L = (e_{m+1}, e_{m+2})$, and

$$\widetilde{f}(\lambda) = \frac{\Gamma(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))\Gamma(-w_2 - \mathrm{i}\lambda_2)}{(2\pi)^2 K^{w_1+w_2+\mathrm{i}(\lambda_1+\lambda_2)}\Gamma(w_1 + 1 + \mathrm{i}\lambda_1)},$$

for some $w_2 < 0$ and $w_1 > 1 - w_2$.

It remains to be checked from case to case for both payoffs (10.22) and (10.23) whether $v \in \mathcal{D}_{\mathbb{R}}(T)$ holds for Theorem 10.5 to apply ($\rightarrow$ Exercise 10.19).

The following representation including a double Fourier integral is due to Hurd and Zhou [98].

Lemma 10.3 *Let $w = (w_1, w_2)^\top \in \mathbb{R}^2$ be such that $w_2 < 0$ and $w_1 + w_2 > 1$. Then for any $y = (y_1, y_2)^\top \in \mathbb{R}^2$ the following identity holds*:

$$\begin{aligned}
&\left(\mathrm{e}^{y_1} - \mathrm{e}^{y_2} - 1\right)^+ \\
&\quad = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \mathrm{e}^{(w+\mathrm{i}\lambda)^\top y}\frac{\Gamma(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))\Gamma(-w_2 - \mathrm{i}\lambda_2)}{\Gamma(w_1 + 1 + \mathrm{i}\lambda_1)}\, d\lambda_1\, d\lambda_2,
\end{aligned}$$

where the Gamma function $\Gamma(z) = \int_0^\infty t^{-1+z}\mathrm{e}^{-t}\, dt$ is defined for all complex z with $\Re(z) > 0$.

Proof By assumption the function $h(y) = \mathrm{e}^{-w^\top y}(\mathrm{e}^{y_1} - \mathrm{e}^{y_2} - 1)^+$ is integrable on $\mathbb{R}^2$. Its Fourier transform can be calculated, using (10.21) for $K = \mathrm{e}^{y_1} - 1 > 0$ if $y_1 > 0$, as follows:

$$\begin{aligned}
\widehat{h}(\lambda) &= \int_{\mathbb{R}^2} \mathrm{e}^{-(w+\mathrm{i}\lambda)^\top y}\left(\mathrm{e}^{y_1} - \mathrm{e}^{y_2} - 1\right)^+ dy_1\, dy_2 \\
&= \int_0^\infty \mathrm{e}^{-(w_1+\mathrm{i}\lambda_1)y_1}\int_{\mathbb{R}} \mathrm{e}^{-(w_2+\mathrm{i}\lambda_2)y_2}\left(\mathrm{e}^{y_1} - 1 - \mathrm{e}^{y_2}\right)^+ dy_2\, dy_1 \\
&= \int_0^\infty \mathrm{e}^{-(w_1+\mathrm{i}\lambda_1)y_1}\frac{(\mathrm{e}^{y_1} - 1)^{-(w_2-1+\mathrm{i}\lambda_2)}}{(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)}\, dy_1 \\
&= \frac{1}{(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)} \\
&\qquad \times \int_0^\infty (\mathrm{e}^{-y_1})^{w_1+w_2-1+\mathrm{i}(\lambda_1+\lambda_2)}(1 - \mathrm{e}^{-y_1})^{-w_2+1-\mathrm{i}\lambda_2}\, dy_1.
\end{aligned}$$

The change of variables $z = \mathrm{e}^{-y_1}$, with $dz/z = -dy_1$, then yields

$$\widehat{h}(\lambda) = \frac{1}{(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)} \int_0^1 z^{w_1+w_2-2+\mathrm{i}(\lambda_1+\lambda_2)} (1-z)^{-w_2+1-\mathrm{i}\lambda_2}\, dz.$$

Recall that the beta function

$$\mathrm{B}(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

is defined for any complex a, b with $\Re(a), \Re(b) > 0$ by

$$\mathrm{B}(a,b) = \int_0^1 z^{a-1}(1-z)^{b-1}\, dz.$$

Since $w_1 + w_2 > 1$, we obtain from this and the property $\Gamma(z) = (z-1)\Gamma(z-1)$ that

$$\begin{aligned}
\widehat{h}(\lambda) &= \frac{\mathrm{B}(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2), -w_2 + 2 - \mathrm{i}\lambda_2)}{(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)} \\
&= \frac{\Gamma(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))\Gamma(-w_2 + 2 - \mathrm{i}\lambda_2)}{\Gamma(w_1 + 1 + \mathrm{i}\lambda_1)(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)} \\
&= \frac{\Gamma(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))\Gamma(-w_2 - \mathrm{i}\lambda_2)}{\Gamma(w_1 + 1 + \mathrm{i}\lambda_1)}. \qquad (10.24)
\end{aligned}$$

It remains to be checked whether $\widehat{h}(\lambda)$ is integrable in $\lambda \in \mathbb{R}^2$. From the definition of the beta function it follows that $|\mathrm{B}(a,b)| \le \mathrm{B}(\Re(a), \Re(b))$. Hence

$$|\widehat{h}(\lambda)| \le \frac{\mathrm{B}(w_1 + w_2 - 1, -w_2 + 2)}{|(w_2 + \mathrm{i}\lambda_2)(w_2 - 1 + \mathrm{i}\lambda_2)|}. \qquad (10.25)$$

On the other hand, factorizing the first factor in the nominator of the third line in (10.24), we can rewrite $\widehat{h}(\lambda)$ as

$$\widehat{h}(\lambda) = \frac{\mathrm{B}(w_1 + w_2 + 1 + \mathrm{i}(\lambda_1 + \lambda_2), -w_2 - \mathrm{i}\lambda_2)}{(w_1 + w_2 + \mathrm{i}(\lambda_1 + \lambda_2))(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))}.$$

Hence

$$|\widehat{h}(\lambda)| \le \frac{\mathrm{B}(w_1 + w_2 + 1, -w_2)}{|(w_1 + w_2 + \mathrm{i}(\lambda_1 + \lambda_2))(w_1 + w_2 - 1 + \mathrm{i}(\lambda_1 + \lambda_2))|}. \qquad (10.26)$$

The two bounds (10.25) and (10.26) imply that $\widehat{h}(\lambda)$ is integrable in $\lambda \in \mathbb{R}^2$. The Fourier inversion formula (10.20) now yields the claim. □

The two bounds (10.25) and (10.26) assert that the numerical integration be feasible for many models. It can be made efficient by fast Fourier transform, as outlined in [98].

10.3.2 Bond Option Pricing in Affine Models

Let us elaborate further on the pricing of bond options. We assume that either condition (a) or (b) of Theorem 10.4 is met, and fix some maturities $T < S \leq \tau$. A straightforward modification of Lemma 10.2 implies that the payoff function of the European call option on the S-bond with expiry date T and strike price K admits the integral representation

$$\left(\mathrm{e}^{-A(S-T)-B(S-t)^\top x} - K\right)^+ = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-(w+\mathrm{i}\lambda)B(S-t)^\top x}\,\widetilde{f}(w,\lambda)\,d\lambda$$

where we define

$$\widetilde{f}(w,\lambda) = \frac{1}{2\pi}\mathrm{e}^{-(w+\mathrm{i}\lambda)A(S-T)}\frac{K^{-(w-1+\mathrm{i}\lambda)}}{(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)}, \tag{10.27}$$

for any real $w > 1$. A similar formula results for put options. We thus obtain from Theorem 10.5 the following master pricing formula for European call and put bond options.

Corollary 10.4 *There exists some $w_- < 0$ and $w_+ > 1$ such that $-B(S-T)w \in \mathcal{D}_{\mathbb{R}}(T)$ for all $w \in (w_-, w_+)$, where $\mathcal{D}_{\mathbb{R}}$ denotes the maximal domain for the system of Riccati equations (10.12). Define $\widetilde{f}(w,\lambda)$ as in (10.27). Then the line integral*

$$\Pi(w,t) = \int_{\mathbb{R}} \mathrm{e}^{\Phi(T-t,-(w+\mathrm{i}\lambda)B(S-T))+\Psi(T-t,-(w+\mathrm{i}\lambda)B(S-T))^\top X(t)}\,\widetilde{f}(w,\lambda)\,d\lambda$$

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

$$\begin{aligned}
\pi_{call}(t) &= \begin{cases} \Pi(w,t), & \text{if } w \in (1, w_+),\\ \Pi(w,t) + P(t,S), & \text{if } w \in (0,1) \end{cases}\\
&= P(t,S)q(t,S,\mathcal{I}) - KP(t,T)q(t,T,\mathcal{I}),\\
\pi_{put}(t) &= \begin{cases} \Pi(w,t) + KP(t,T), & \text{if } w \in (0,1),\\ \Pi(w,t), & \text{if } w \in (w_-, 0) \end{cases}\\
&= KP(t,T)q(t,T,\mathbb{R}\setminus\mathcal{I}) - P(t,S)q(t,S,\mathbb{R}\setminus\mathcal{I}),
\end{aligned} \tag{10.28}$$

where $\mathcal{I} = (A(S-T) + \log K, \infty)$, and $q(t,S,dy)$ and $q(t,T,dy)$ denote the $\mathcal{F}_t$-conditional distributions of the real-valued random variable $Y = -B(S-T)^\top X(T)$ under the S- and T-forward measure, respectively.

Proof From the flow property $\Psi(T, -B(S-T)) = -B(S)$, we know that $-B(S-T) \in \mathcal{D}_{\mathbb{R}}(T)$. Since $\mathcal{D}_{\mathbb{R}}(T)$ is a convex open neighborhood of 0 in $\mathbb{R}^d$, we obtain that $-B(S-T)w \in \mathcal{D}_{\mathbb{R}}(T)$ for all $w \in (w_-, w_+)$, for some $w_- < 0$ and $w_+ > 1$.

It then follows by inspection that $\mathrm{e}^{\Phi(T-t,-(w+\mathrm{i}\lambda)B(S-T))+\Psi(T-t,-(w+\mathrm{i}\lambda)B(S-T))^\top X(t)}$ is uniformly bounded and $\widetilde{f}(w,\lambda)$ is integrable in $\lambda\in\mathbb{R}$, for any fixed $w\in(w_-,w_+)\setminus\{0,1\}$. Hence the line integral $\Pi(w,t)$ is well defined for all $w\in(w_-,w_+)\setminus\{0,1\}$ and $t\le T$.

Further, we recall from Chap. 7 that we can decompose (10.10), according to (7.7). For the call option we thus obtain

$$\pi(t)=P(t,S)\mathbb{Q}^S[E\mid\mathcal{F}_t]-KP(t,T)\mathbb{Q}^T[E\mid\mathcal{F}_t],$$

for the exercise event $E=\{-B(S-T)^\top X(T)>A(S-T)+\log K\}$, and similarly for the put option.

The price formulas (10.28) now follow from the above discussion, and Theorem 10.5 and Lemma 10.2 ($\to$ Exercise 10.6). This proves the corollary. $\square$

Thus the pricing of European call and put bond options in the present d-dimensional affine factor model boils down to the computation of a line integral $\Pi(w,t)$, which is a simple numerical task. Moreover, in case the distributions $q(t,S,dy)$ and $q(t,T,dy)$ are explicitly known, the pricing is reduced to the computation of the respective probabilities in (10.28) of the exercise events $\mathcal{I}$ and $\mathbb{R}\setminus\mathcal{I}$.

In the following two subsections, we illustrate this approach for the Vasiček and CIR short-rate models.

10.3.2.1 Example: Vasiček Short-Rate Model

The state space is $\mathbb{R}$, and we set $r=X$ for the Vasiček short-rate model

$$dr=(b+\beta r)\,dt+\sigma\,dW.$$

The system (10.12) reads

$$\begin{aligned}
\Phi(t,u)&=\frac{1}{2}\sigma^2\int_0^t\Psi^2(s,u)\,ds+b\int_0^t\Psi(s,u)\,ds,\\
\partial_t\Psi(t,u)&=\beta\Psi(t,u)-1,\\
\Psi(0,u)&=u,
\end{aligned}$$

which admits a unique global solution with

$$\begin{aligned}
\Phi(t,u)&=\frac{1}{2}\sigma^2\left(\frac{u^2}{2\beta}(\mathrm{e}^{2\beta t}-1)+\frac{1}{2\beta^3}(\mathrm{e}^{2\beta t}-4\mathrm{e}^{\beta t}+2\beta t+3)\right.\\
&\qquad\left.-\frac{u}{\beta^2}(\mathrm{e}^{2\beta t}-2\mathrm{e}^{\beta t}+2\beta)\right)+b\left(\frac{\mathrm{e}^{\beta t}-1}{\beta}u-\frac{\mathrm{e}^{\beta t}-1-\beta t}{\beta^2}\right),\\
\Psi(t,u)&=\mathrm{e}^{\beta t}u-\frac{\mathrm{e}^{\beta t}-1}{\beta}
\end{aligned}$$

for all $u\in\mathbb{C}$. Hence (10.13) holds for all $u\in\mathbb{C}$ and $t\le T$.

Moreover, we see that the exponent of the $\mathcal{F}_t$-conditional characteristic function of $r(T)$ under the S-forward measure (10.15) is a quadratic polynomial in u. Hence, under the S-forward measure, $r(T)$ is $\mathcal{F}_t$-conditionally Gaussian distributed with variance $\sigma^2 \frac{\mathrm{e}^{2\beta(T-t)}-1}{2\beta}$. This is in line with (5.11) (why?). A straightforward calculation yields the $\mathcal{F}_t$-conditional $\mathbb{Q}^S$-mean of $r(T)$. The bond option price formula for the Vasiček short-rate model from Proposition 7.2 and Sect. 7.2.1 can now be derived via (10.28) (→ Exercise 10.7).

10.3.2.2 Example: CIR Short-Rate Model

The state space is $\mathbb{R}_+$, and we set $r = X$ for the CIR short-rate model

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

The system (10.12) reads

$$\begin{aligned} \Phi(t,u) &= b \int_0^t \Psi(s,u)\, ds, \\ \partial_t \Psi(t,u) &= \frac{1}{2}\sigma^2 \Psi^2(t,u) + \beta \Psi(t,u) - 1, \\ \Psi(0,u) &= u. \end{aligned} \tag{10.29}$$

By Lemma 10.12 below, there exists a unique solution $(\Phi(\cdot,u), \Psi(\cdot,u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}_-$, and thus identity (10.13) holds, for all $u \in \mathbb{C}_-$ and $t \leq T$. In fact, the solution is given explicitly as

$$\begin{aligned} \Phi(t,u) &= \frac{2b}{\sigma^2} \log \left(\frac{2\theta \mathrm{e}^{\frac{(\theta-\beta)t}{2}}}{L_3(t) - L_4(t)u} \right), \\ \Psi(t,u) &= -\frac{L_1(t) - L_2(t)u}{L_3(t) - L_4(t)u}, \end{aligned}$$

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

$$\begin{aligned} L_1(t) &= 2\left(\mathrm{e}^{\theta t} - 1\right), \\ L_2(t) &= \theta\left(\mathrm{e}^{\theta t} + 1\right) + \beta\left(\mathrm{e}^{\theta t} - 1\right), \\ L_3(t) &= \theta\left(\mathrm{e}^{\theta t} + 1\right) - \beta\left(\mathrm{e}^{\theta t} - 1\right), \\ L_4(t) &= \sigma^2\left(\mathrm{e}^{\theta t} - 1\right). \end{aligned}$$

Some tedious but elementary algebraic manipulations show that the $\mathcal{F}_t$-conditional characteristic function of $r(T)$ under the S-forward measure $\mathbb{Q}^S$ is given by

$$\mathbb{E}_{\mathbb{Q}^S}\left[\mathrm{e}^{ur(T)} \mid \mathcal{F}_t\right] = \frac{\mathrm{e}^{\frac{C_2(t,T,S)r(t)C_1(t,T,S)u}{1-C_1(t,T,S)u}}}{(1-C_1(t,T,S)u)^{\frac{2b}{\sigma^2}}},$$

where

$$C_1(t,T,S) = \frac{L_3(S-T)L_4(T-t)}{2\theta L_3(S-t)}, \qquad C_2(t,T,S) = \frac{L_2(T-t)}{L_4(T-t)} - \frac{L_1(S-t)}{L_3(S-t)}.$$

Comparing this with Lemma 10.4 below, we conclude that the $\mathcal{F}_t$-conditional distribution of the random variable

$$Z(t,T) = \frac{2r(T)}{C_1(t,T,S)}$$

under the S-forward measure $\mathbb{Q}^S$ is noncentral χ^2 with $\frac{4b}{\sigma^2}$ degrees of freedom and parameter of noncentrality $2C_2(t,T,S)r(t)$. The noncentral χ^2-distribution is a generalization of the distribution of the sum of the squares of independent normal distributed random variables ($\to$ Exercise 10.8). It is good to know that the noncentral χ^2-distribution is hard coded in most statistical software packages.[8] Combining this with Corollary 10.4, we obtain explicit European bond option price formulas for the CIR model.

Lemma 10.4 (Noncentral χ^2-Distribution) *The noncentral χ^2-distribution with $\delta > 0$ degrees of freedom and noncentrality parameter $\zeta > 0$ has density function*

$$f_{\chi^2(\delta,\zeta)}(x) = \frac{1}{2}\mathrm{e}^{-\frac{x+\zeta}{2}}\left(\frac{x}{\zeta}\right)^{\frac{\delta}{4}-\frac{1}{2}} I_{\frac{\delta}{2}-1}(\sqrt{\zeta x}), \quad x \geq 0$$

and characteristic function

$$\int_{\mathbb{R}_+} \mathrm{e}^{ux} f_{\chi^2(\delta,\zeta)}(x)\,dx = \frac{\mathrm{e}^{\frac{\zeta u}{1-2u}}}{(1-2u)^{\frac{\delta}{2}}}, \quad u \in \mathbb{C}_-.$$

Here $I_\nu(x) = \sum_{j\geq 0} \frac{1}{j!\Gamma(j+\nu+1)}\left(\frac{x}{2}\right)^{2j+\nu}$ denotes the modified Bessel function of the first kind of order $\nu > -1$.

Proof See e.g. [104, Chap. 29]. □

For illustration, we now fix the following CIR model parameters

$$\sigma^2 = 0.033, \qquad b = 0.08, \qquad \beta = -0.9, \qquad r_0 = 0.08. \tag{10.30}$$

[8] The sampling from a noncentral χ^2-distribution is described in [79, Sect. 3.4.1].

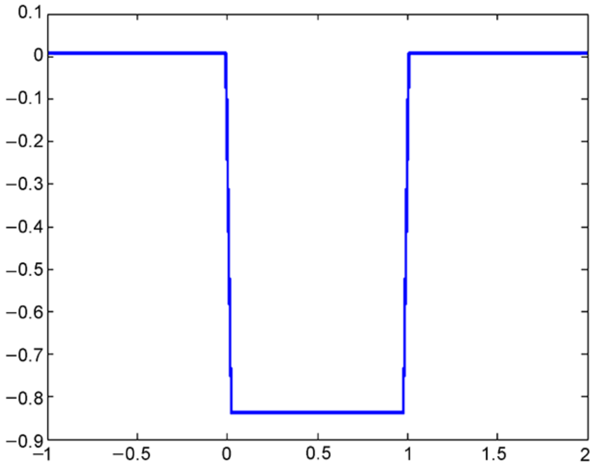


Fig. 10.1 Line integral $\Pi(w,0)$ as a function of w

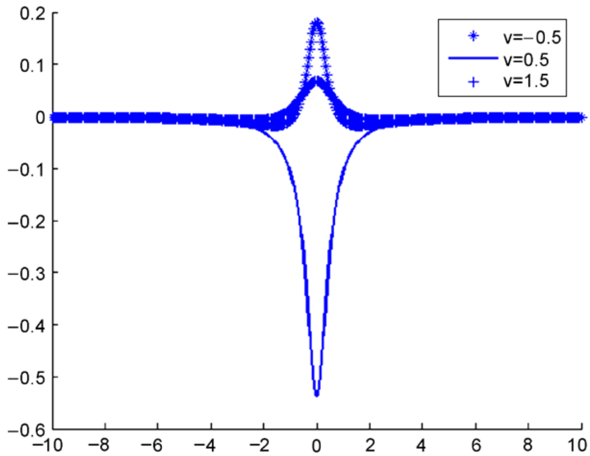


Fig. 10.2 Real part of the integrand of $\Pi(w,0)$, for $w=-0.5, 0.5, 1.5$, as a function of λ

Moreover, we set $t=0$, $T=1$ and $S=2$. Using any software capable of numerical integration, we see that the line integral $\Pi(w,0)$ in Corollary 10.4 behaves numerically stable for w ranging between $(-1,2)\setminus\{0,1\}$ (see Fig. 10.1). On the other hand, we know that $\Pi(w,0)$ diverges for $w\to+\infty$ (why?). The real part of the integrand of $\Pi(w,0)$, for $w=-0.5, 0.5, 1.5$, is plotted as function of λ in Fig. 10.2. The resulting ATM call and put option strike price is $K=0.9180$. The call and put option prices, $\pi_{call}(0)=\pi_{put}(0)=0.0078$, can now be computed by any of the formulas in (10.28) ($\rightarrow$ Exercise 10.13).

As an application, we next compute ATM cap prices and implied Black volatilities ($\rightarrow$ Exercise 10.14). The tenor is as follows: $t=0$ (today), $T_0=1/4$ (first reset date), and $T_i-T_{i-1}\equiv 1/4$, $i=1,\dots,119$ (the maturity of the last cap is $T_{119}=30$). Table 10.1 and Fig. 10.3 show the ATM cap prices and implied Black volatilities for a range of maturities. Like the Vasiček model (see Fig. 7.1), the CIR model seems incapable of producing humped volatility curves.

Table 10.1 CIR ATM cap prices and Black volatilities

Maturity	ATM prices	ATM vols
1	0.0073	0.4506
2	0.0190	0.3720
3	0.0302	0.3226
4	0.0406	0.2890
5	0.0501	0.2647
6	0.0588	0.2462
7	0.0668	0.2316
8	0.0742	0.2198
10	0.0871	0.2017
12	0.0979	0.1886
15	0.1110	0.1744
20	0.1265	0.1594
30	0.1430	0.1442

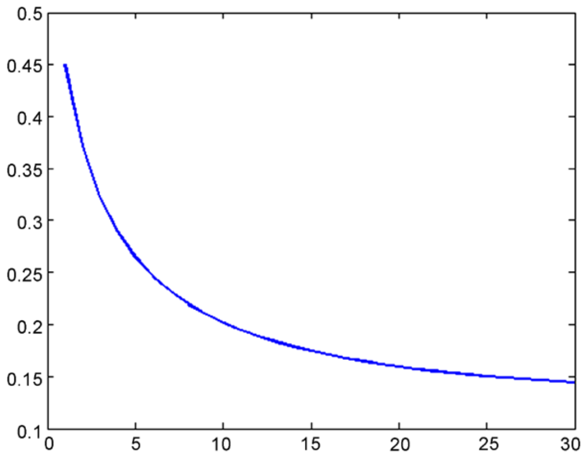


Fig. 10.3 CIR ATM cap Black volatilities

10.3.3 Heston Stochastic Volatility Model

This affine model, proposed by Heston [91], generalizes the Black–Scholes model (see Exercise 4.7 and Sect. 7.3) by assuming a stochastic volatility.

Interest rates are assumed to be constant $r(t) \equiv r \geq 0$, and there is one risky asset (stock) $S = \mathrm{e}^{X_2}$, where $X = (X_1, X_2)$ is the affine process with state space $\mathbb{R}_+ \times \mathbb{R}$ and dynamics

$$
\begin{aligned}
dX_1 &= (k + \kappa X_1)\,dt + \sigma\sqrt{2X_1}\,dW_1,\\
dX_2 &= (r - X_1)\,dt + \sqrt{2X_1}\left(\rho dW_1 + \sqrt{1-\rho^2}dW_2\right)
\end{aligned}
$$

for some constant parameters $k, \sigma \geq 0$, $\kappa \in \mathbb{R}$, and some $\rho \in [-1, 1]$.

The implied risk-neutral stock dynamics read

$$dS = Sr\,dt + S\sqrt{2X_1}\,d\mathcal{W}$$

for the Brownian motion $\mathcal{W} = \rho W_1 + \sqrt{1-\rho^2}W_2$. We see that $\sqrt{2X_1}$ is the stochastic volatility of the price process S. They have possibly non-zero covariation

$$d\langle S, X_1\rangle = 2\rho\sigma S X_1\,dt.$$

The corresponding system of Riccati equations (10.7) is equivalent to ($\rightarrow$ Exercise 10.15)

$$\begin{aligned}
\phi(t,u) &= k\int_0^t \psi_1(s,u)\,ds + ru_2 t,\\
\partial_t\psi_1(t,u) &= \sigma^2\psi_1^2(t,u) + (2\rho\sigma u_2 + \kappa)\psi_1(t,u) + u_2^2 - u_2,\\
\psi_1(0,u) &= u_1,\\
\psi_2(t,u) &= u_2,
\end{aligned} \tag{10.31}$$

which, in view of Lemma 10.12(b) below admits an explicit global solution if $u_1 \in \mathbb{C}_-$ and $0 \leq \Re u_2 \leq 1$. In particular, for $u = (0, 1)$, we obtain

$$\phi(t, 0, 1) = rt, \qquad \psi(t, 0, 1) = (0, 1)^\top.$$

Theorem 10.3 thus implies that $S(T)$ has a finite first moment, for any $T \in \mathbb{R}_+$, and

$$\mathbb{E}[\mathrm{e}^{-rT}S(T) \mid \mathcal{F}_t] = \mathrm{e}^{-rT}\mathbb{E}[\mathrm{e}^{X_2(T)} \mid \mathcal{F}_t] = \mathrm{e}^{-rT}\mathrm{e}^{r(T-t)+X_2(t)} = \mathrm{e}^{-rt}S(t),$$

for $t \leq T$, which is just the martingale property of S.

We now want to compute the price

$$\pi(t) = \mathrm{e}^{-r(T-t)}\mathbb{E}\left[(S(T)-K)^+ \mid \mathcal{F}_t\right]$$

of a European call option on $S(T)$ with maturity T and strike price K. Fix some $w > 1$ small enough with $(0, w) \in \mathcal{D}_\mathbb{R}(T)$, where $\mathcal{D}_\mathbb{R}$ denotes the maximal domain for the system of Riccati equations (10.31). Formula (10.18) combined with Lemma 10.2 then yields ($\rightarrow$ Exercise 10.16)

$$\pi(t) = \mathrm{e}^{-r(T-t)}\int_\mathbb{R} \mathrm{e}^{\phi(T-t,0,w+\mathrm{i}\lambda)+\psi_1(T-t,0,w+\mathrm{i}\lambda)X_1(t)+(w+\mathrm{i}\lambda)X_2(t)}\widetilde{f}(\lambda)\,d\lambda \tag{10.32}$$

with

$$\widetilde{f}(\lambda) = \frac{1}{2\pi}\frac{K^{-(w-1+\mathrm{i}\lambda)}}{(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)}.$$

Table 10.2 Call option prices in the Heston model

T–*K*	0.8	0.9	1.0	1.1	1.2
0.2	0.2016	0.1049	0.0348	0.0074	0.0012
0.4	0.2037	0.1120	0.0478	0.0168	0.0053
0.6	0.2061	0.1183	0.0571	0.0245	0.0100
0.8	0.2088	0.1239	0.0646	0.0310	0.0144
1.0	0.2115	0.1291	0.0711	0.0368	0.0186

Table 10.3 Black–Scholes implied volatilities for the call option prices in the Heston model

T–*K*	0.8	0.9	1.0	1.1	1.2
0.2	0.1715	0.1786	0.1899	0.2017	0.2126
0.4	0.1641	0.1712	0.1818	0.1930	0.2033
0.6	0.1585	0.1656	0.1755	0.1858	0.1954
0.8	0.1544	0.1612	0.1704	0.1799	0.1889
1.0	0.1513	0.1579	0.1664	0.1751	0.1835

Alternatively, we may fix any $0 < w < 1$ and then

$$\pi(t) = S(t) + \mathrm{e}^{-r(T-t)} \int_{\mathbb{R}} \mathrm{e}^{\phi(T-t,0,w+\mathrm{i}\lambda)+\psi_1(T-t,0,w+\mathrm{i}\lambda)X_1(t)+(w+\mathrm{i}\lambda)X_2(t)}\,\widetilde{f}(\lambda)\,d\lambda. \tag{10.33}$$

For illustration, we choose the model parameters

$$X_1(0) = 0.02, \qquad X_2(0) = 0, \qquad \sigma = 0.1, \qquad \kappa = -2.0,$$
$$k = 0.02, \qquad r = 0.01, \qquad \rho = 0.5.$$

Table 10.2 shows European call option prices at $t = 0$ for various strikes K and maturities T. The corresponding implied Black–Scholes volatilities are shown in Table 10.3 and Fig. 10.4 ($\to$ Exercise 10.17).

10.4 Affine Transformations and Canonical Representation

As in the beginning of Sect. 10.3, we let X be affine on the canonical state space $\mathbb{R}^m_+ \times \mathbb{R}^n$ with admissible parameters a, α_i, b, β_i. Hence, in view of (10.1), for any $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$ the process $X = X^x$ satisfies

$$\begin{aligned} dX &= (b + \mathcal{B}X)\,dt + \rho(X)\,dW, \\ X(0) &= x, \end{aligned} \tag{10.34}$$

and $\rho(x)\rho(x)^\top = a + \sum_{i\in I} x_i\alpha_i$.

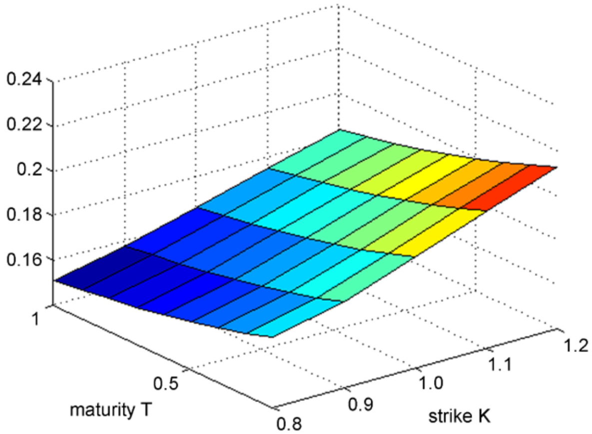


Fig. 10.4 Implied volatility surface for the Heston model

It can easily be checked (→ Exercise 10.20) that for every invertible $d \times d$-matrix Λ, the linear transform

$$Y = \Lambda X$$

satisfies

$$dY = \left(\Lambda b + \Lambda \mathcal{B} \Lambda^{-1} Y\right) dt + \Lambda \rho \left(\Lambda^{-1} Y\right) dW, \quad Y(0) = \Lambda x. \tag{10.35}$$

Hence, Y has again an affine drift and diffusion matrix

$$\Lambda b + \Lambda \mathcal{B} \Lambda^{-1} y \quad \text{and} \quad \Lambda \alpha(\Lambda^{-1} y) \Lambda^{\top}, \tag{10.36}$$

respectively.

On the other hand, the affine short-rate model (10.9) can be expressed in terms of $Y(t)$ as

$$r(t) = c + \gamma^{\top} \Lambda^{-1} Y(t). \tag{10.37}$$

This shows that Y and (10.37) specify an affine short-rate model producing the same short rates, and thus bond prices, as X and (10.9). That is, an invertible linear transformation of the state process changes the particular form of the stochastic differential equation (10.34). But it leaves observable quantities, such as short rates and bond prices invariant.

This motivates the question whether there exists a classification method ensuring that affine short-rate models with the same observable implications have a unique canonical representation. This topic has been addressed in [43, 44, 50, 105], see also the notes section. We now elaborate on this issue and show that the diffusion matrix $\alpha(x)$ can always be brought into block-diagonal form by a regular linear transform Λ with $\Lambda(\mathbb{R}^m_+ \times \mathbb{R}^n) = \mathbb{R}^m_+ \times \mathbb{R}^n$.

We denote by

$$\operatorname{diag}(z_1, \ldots, z_m)$$

the diagonal matrix with diagonal elements $z_1, \dots, z_m$, and we write I_m for the $m \times m$-identity matrix.

Lemma 10.5 *There exists some invertible $d \times d$-matrix Λ with $\Lambda(\mathbb{R}^m_+ \times \mathbb{R}^n) = \mathbb{R}^m_+ \times \mathbb{R}^n$ such that $\Lambda\alpha(\Lambda^{-1}y)\Lambda^\top$ is block-diagonal of the form*

$$\Lambda\alpha(\Lambda^{-1}y)\Lambda^\top = \begin{pmatrix} \operatorname{diag}(y_1, \dots, y_q, 0, \dots, 0) & 0 \\ 0 & p + \sum_{i\in I} y_i \pi_i \end{pmatrix}$$

for some integer $0 \le q \le m$ and symmetric positive semi-definite $n \times n$ matrices $p, \pi_1, \dots, \pi_m$. Moreover, Λb and $\Lambda \mathcal{B} \Lambda^{-1}$ meet the respective admissibility conditions (10.6) *in lieu of b and $\mathcal{B}$.*

Proof From (10.3) we know that $\Lambda\alpha(x)\Lambda^\top$ is block-diagonal for all $x = \Lambda^{-1}y$ if and only if $\Lambda a \Lambda^\top$ and $\Lambda\alpha_i\Lambda^\top$ are block-diagonal for all $i \in I$. By permutation and scaling of the first m coordinate axes (this is a linear bijection from $\mathbb{R}^m_+ \times \mathbb{R}^n$ onto itself, which preserves the admissibility of the transformed b and $\mathcal{B}$), we may assume that there exists some integer $0 \le q \le m$ such that $\alpha_{1,11} = \dots = \alpha_{q,qq} = 1$ and $\alpha_{i,ii} = 0$ for $q < i \le m$. Hence a and α_i for $q < i \le m$ are already block-diagonal of the special form

$$a = \begin{pmatrix} 0 & 0 \\ 0 & a_{JJ} \end{pmatrix}, \qquad \alpha_i = \begin{pmatrix} 0 & 0 \\ 0 & \alpha_{i,JJ} \end{pmatrix}.$$

For $1 \le i \le q$, we may have non-zero off-diagonal elements in the ith row $\alpha_{i,iJ}$. We thus define the $n \times m$-matrix $D = (\delta_1, \dots, \delta_m)$ with ith column $\delta_i = -\alpha_{i,iJ}$ and set

$$\Lambda = \begin{pmatrix} I_m & 0 \\ D & I_n \end{pmatrix}.$$

One checks by inspection that D is invertible and maps $\mathbb{R}^m_+ \times \mathbb{R}^n$ onto $\mathbb{R}^m_+ \times \mathbb{R}^n$. Moreover,

$$D\alpha_{i,II} = -\alpha_{i,JI}, \quad i \in I.$$

From here we easily verify that

$$\Lambda\alpha_i = \begin{pmatrix} \alpha_{i,II} & \alpha_{i,IJ} \\ 0 & D\alpha_{i,IJ} + \alpha_{i,JJ} \end{pmatrix},$$

and thus

$$\Lambda\alpha_i\Lambda^\top = \begin{pmatrix} \alpha_{i,II} & 0 \\ 0 & D\alpha_{i,IJ} + \alpha_{i,JJ} \end{pmatrix}.$$

Since $\Lambda a \Lambda^\top = a$, the first assertion is proved.

The admissibility conditions for Λb and $\Lambda\mathcal{B}\Lambda^{-1}$ can easily be checked as well. □

In view of (10.36), (10.37) and Lemma 10.5 we thus obtain the following result.

Theorem 10.7 (Canonical Representation) *Any affine short-rate model* (10.9), *after some modification of* γ *if necessary, admits an* $\mathbb{R}^m_+ \times \mathbb{R}^n$*-valued affine state process* X *with block-diagonal diffusion matrix of the form*

$$\alpha(x) = \begin{pmatrix} \mathrm{diag}(x_1, \dots, x_q, 0, \dots, 0) & 0 \\ 0 & a + \sum_{i \in I} x_i \alpha_{i,JJ} \end{pmatrix} \tag{10.38}$$

for some integer $0 \le q \le m$.

10.5 Existence and Uniqueness of Affine Processes

All we said about the affine process X so far was under the premise that there exists a unique solution $X = X^x$ of the stochastic differential equation (10.1) on some appropriate state space $\mathcal{X} \subset \mathbb{R}^d$. However, if the diffusion matrix $\rho(x)\rho(x)^\top$ is affine then $\rho(x)$ cannot be Lipschitz continuous in x in general. This raises the question whether (10.1) admits a solution at all.

In this section, we show how X can always be realized as unique solution of the stochastic differential equation (10.1), which is (10.34), in the canonical affine framework $\mathcal{X} = \mathbb{R}^m_+ \times \mathbb{R}^n$ and for particular choices of $\rho(x)$.

We recall from Theorem 10.1 that the affine property of X imposes explicit conditions on $\rho(x)\rho(x)^\top$, but not on $\rho(x)$ as such. Indeed, for any orthogonal $d \times d$-matrix D, the function $\rho(x)D$ yields the same diffusion matrix, $\rho(x)DD^\top\rho(x)^\top = \rho(x)\rho(x)^\top$, as $\rho(x)$ (see also Remark 4.2).

On the other hand, from Theorem 10.2 we know that any admissible parameters a, α_i, b, β_i in (10.3) uniquely determine the functions $(\phi(\cdot, u), \psi(\cdot, u)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$ as solution of the Riccati equations (10.7), for all $u \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$. These in turn uniquely determine the law of the process X. Indeed, for any $0 \le t_1 < t_2$ and $u_1, u_2 \in \mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n$, we infer by iteration of (10.2)

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}^{u_1^\top X(t_1) + u_2^\top X(t_2)}\right] &= \mathbb{E}\left[\mathrm{e}^{u_1^\top X(t_1)} \mathbb{E}\left[\mathrm{e}^{u_2^\top X(t_2)} \mid \mathcal{F}_{t_1}\right]\right] \\ &= \mathbb{E}\left[\mathrm{e}^{u_1^\top X(t_1)} \mathrm{e}^{\phi(t_2 - t_1, u_2) + \psi(t_2 - t_1, u_2)^\top X(t_1)}\right] \\ &= \mathrm{e}^{\phi(t_2 - t_1, u_2) + \phi(t_1, u_1 + \psi(t_2 - t_1, u_2)) + \psi(t_1, u_1 + \psi(t_2 - t_1, u_2))^\top x}. \end{aligned}$$

Hence the joint distribution of $(X(t_1), X(t_2))$ is uniquely determined by the functions ϕ and ψ. By further iteration of this argument, we conclude that every finite-dimensional distribution, and thus the law, of X is uniquely determined by the parameters a, α_i, b, β_i.

We conclude that the law of an affine process X, while uniquely determined by its characteristics (10.3), can be realized by infinitely many variants of the stochastic differential equation (10.34) by replacing $\rho(x)$ by $\rho(x)D$, for any orthogonal $d \times d$-matrix D. We now propose a canonical choice of $\rho(x)$ as follows:

- In view of (10.35) and Lemma 10.5, every affine process X on $\mathbb{R}^m_+ \times \mathbb{R}^n$ can be written as $X = \Lambda^{-1} Y$ for some invertible $d \times d$-matrix Λ and some affine process Y on $\mathbb{R}^m_+ \times \mathbb{R}^n$ with block-diagonal diffusion matrix. It is thus enough to consider such $\rho(x)$ where $\rho(x)\rho(x)^\top$ is of the form (10.38). Obviously, $\rho(x) \equiv \rho(x_I)$ is a function of x_I only.
- Set $\rho_{IJ}(x) \equiv 0$, $\rho_{JI}(x) \equiv 0$, and

$$\rho_{II}(x_I) = \operatorname{diag}(\sqrt{x_1}, \dots, \sqrt{x_q}, 0, \dots, 0).$$

Choose for $\rho_{JJ}(x_I)$ any measurable $n \times n$-matrix-valued function satisfying

$$\rho_{JJ}(x_I)\rho_{JJ}(x_I)^\top = a + \sum_{i \in I} x_i \alpha_{i,JJ}. \tag{10.39}$$

In practice, one would determine $\rho_{JJ}(x_I)$ via Cholesky factorization, see e.g. [129, Theorem 2.2.5]. If $a + \sum_{i \in I} x_i \alpha_{i,JJ}$ is positive definite, then $\rho_{JJ}(x_I)$ turns out to be the unique lower triangular matrix with positive diagonal elements and which satisfies (10.39). If $a + \sum_{i \in I} x_i \alpha_{i,JJ}$ is merely positive semi-definite, then the algorithm becomes more involved. In any case, $\rho_{JJ}(x_I)$ will depend measurably on x_I.
- The stochastic differential equation (10.34) now reads

$$\begin{aligned}
dX_I &= (b_I + \mathcal{B}_{II} X_I)\, dt + \rho_{II}(X_I)\, dW_I,\\
dX_J &= (b_J + \mathcal{B}_{JI} X_I + \mathcal{B}_{JJ} X_J)\, dt + \rho_{JJ}(X_I)\, dW_J,\\
X(0) &= x.
\end{aligned} \tag{10.40}$$

Lemma 10.6 below asserts the existence and uniqueness of an $\mathbb{R}^m_+ \times \mathbb{R}^n$-valued solution $X = X^x$, for any $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$.

We thus have shown:

Theorem 10.8 *Let a, α_i, b, β_i be admissible parameters. Then there exists a measurable function $\rho : \mathbb{R}^m_+ \times \mathbb{R}^n \to \mathbb{R}^{d \times d}$ with $\rho(x)\rho(x)^\top = a + \sum_{i \in I} x_i \alpha_i$, and such that, for any $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$, there exists a unique $\mathbb{R}^m_+ \times \mathbb{R}^n$-valued solution $X = X^x$ of* (10.34).

Moreover, the law of X is uniquely determined by a, α_i, b, β_i, and does not depend on the particular choice of ρ.

The proof of the following lemma uses the concept of a weak solution, which is beyond the scope of this book and therefore mentioned without further explanation. The interested reader will find detailed background in e.g. [106, Sect. 5.3]. At first reading, the following result may simply be taken for granted.

Lemma 10.6 *For any $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$, there exists a unique $\mathbb{R}^m_+ \times \mathbb{R}^n$-valued solution $X = X^x$ of* (10.40).

Proof First, we extend ρ continuously to $\mathbb{R}^d$ by setting $\rho(x) = \rho(x_1^+, \dots, x_m^+)$, where we define $x_i^+ = \max(0, x_i)$.

Now observe that X_I solves the autonomous equation

$$dX_I = (b_I + \mathcal{B}_{II} X_I)\, dt + \rho_{II}(X_I)\, dW_I, \quad X_I(0) = x_I. \tag{10.41}$$

Obviously, there exists a finite constant K such that the linear growth condition

$$\|b_I + \mathcal{B}_{II} x_I\|^2 + \|\rho(x_I)\|^2 \le K(1 + \|x_I\|^2)$$

is satisfied for all $x \in \mathbb{R}^m$. By [99, Theorems 2.3 and 2.4] there exists a weak solution[9] of (10.41). On the other hand, (10.41) is exactly of the form as assumed in [162, Theorem 1], which implies that pathwise uniqueness[10] holds for (10.41). The Yamada–Watanabe theorem, see [162, Corollary 3] or [106, Corollary 5.3.23], thus implies that there exists a unique solution $X_I = X_I^{x_I}$ of (10.41), for all $x_I \in \mathbb{R}^m$.

Given $X_I^{x_I}$, it is then easily seen that

$$X_J(t) = \mathrm{e}^{\mathcal{B}_{JJ} t} \left(x_J + \int_0^t \mathrm{e}^{-\mathcal{B}_{JJ} s} (b_J + \mathcal{B}_{JI} X_I(s))\, ds \right.$$
$$\left. + \int_0^t \mathrm{e}^{-\mathcal{B}_{JJ} s} \rho_{JJ}(X_I(s))\, dW_J(s) \right)$$

is the unique solution to the second equation in (10.40).

Admissibility of the parameters b and β_i and the stochastic invariance Lemma 10.11 eventually imply that $X_I = X_I^{x_I}$ is $\mathbb{R}^m_+$-valued for all $x_I \in \mathbb{R}^m_+$. Whence the lemma is proved. □

10.6 On the Regularity of Characteristic Functions

This auxiliary section provides some analytic regularity results for characteristic functions, which are of independent interest. These results enter the main text only via the proof of Theorem 10.3 in Sect. 10.7.3 below. This section may thus be skipped at the first reading.

---

[9] A weak solution consists of a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ carrying a continuous adapted process X_I and a Brownian motion W_I such that (10.41) is satisfied. The crux of a weak solution is that X_I is not necessarily adapted to the filtration generated by the Brownian motion W_I. See [162, Definition 1] or [106, Definition 5.3.1].

[10] Pathwise uniqueness holds if, for any two weak solutions (X_I, W_I) and (X_I', W_I) of (10.41) defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with common Brownian motion W_I and with common initial value $X_I(0) = X_I'(0)$, the two processes are indistinguishable: $\mathbb{P}[X_I(t) = X_I'(t)$ for all $t \ge 0] = 1$. See [162, Definition 2] or [106, Sect. 5.3].

Let ν be a bounded measure on $\mathbb{R}^d$, and denote by

$$G(z) = \int_{\mathbb{R}^d} \mathrm{e}^{z^\top x}\, \nu(dx)$$

its characteristic function[11] for $z \in \mathrm{i}\mathbb{R}^d$. Note that $G(z)$ is actually well defined for $z \in \mathcal{S}(V)$ where

$$V = \left\{ y \in \mathbb{R}^d \mid \int_{\mathbb{R}^d} \mathrm{e}^{y^\top x}\, \nu(dx) < \infty \right\}.$$

We first investigate the interplay between the (marginal) moments of ν and the corresponding (partial) regularity of G.

Lemma 10.7 *Define $g(y) = G(iy)$ for $y \in \mathbb{R}^d$, and let $k \in \mathbb{N}$ and $1 \le i \le d$. If $\partial_{y_i}^{2k} g(0)$ exists then*

$$\int_{\mathbb{R}^d} |x_i|^{2k}\, \nu(dx) < \infty.$$

On the other hand, if $\int_{\mathbb{R}^d} \|x\|^k\, \nu(dx) < \infty$ then $g \in C^k$ and

$$\partial_{y_{i_1}} \cdots \partial_{y_{i_l}} g(y) = \mathrm{i}^l \int_{\mathbb{R}^d} x_{i_1} \cdots x_{i_l} \mathrm{e}^{\mathrm{i}y^\top x}\, \nu(dx)$$

for all $y \in \mathbb{R}^d$, $1 \le i_1, \dots, i_l \le d$ and $1 \le l \le k$.

Proof As usual, let e_i denote the ith standard basis vector in $\mathbb{R}^d$. Observe that $s \mapsto g(se_i)$ is the characteristic function of the image measure of ν on $\mathbb{R}$ by the mapping $x \mapsto x_i$. Since $\partial_s^{2k} g(se_i)|_{s=0} = \partial_{y_i}^{2k} g(0)$, the assertion follows from the one-dimensional case, see [120, Theorem 2.3.1].

The second part of the lemma follows by differentiating under the integral sign, which is allowed by dominated convergence. □

Lemma 10.8 *The set V is convex. Moreover, if $U \subset V$ is an open set in $\mathbb{R}^d$, then G is analytic on the open strip $\mathcal{S}(U)$ in $\mathbb{C}^d$.*

Proof Since $G : \mathbb{R}^d \to [0, \infty]$ is a convex function, its domain $V = \{y \in \mathbb{R}^d \mid G(y) < \infty\}$ is convex, and so is every level set $V_l = \{y \in \mathbb{R}^d \mid G(y) \le l\}$ for $l \ge 0$.

Now let $U \subset V$ be an open set in $\mathbb{R}^d$. Since any convex function on $\mathbb{R}^d$ is continuous on the open interior of its domain, see [136, Theorem 10.1], we infer that G is continuous on U. We may thus assume that $U_l = \{y \in \mathbb{R}^d \mid G(y) < l\} \cap U \subset V_l$ is open in $\mathbb{R}^d$ and non-empty for $l > 0$ large enough.

---

[11] This is a slight abuse of terminology, since the characteristic function $g(y) = G(\mathrm{i}y)$ of ν is usually defined on real arguments $y \in \mathbb{R}^d$. However, it facilitates the subsequent notation.

Let $z \in \mathcal{S}(U_l)$ and (z_n) be a sequence in $\mathcal{S}(U_l)$ with $z_n \to z$. For n large enough, there exists some $p > 1$ such that $pz_n \in \mathcal{S}(U_l)$. This implies $p\Re z_n \in V_l$ and hence

$$\int_{\mathbb{R}^d} \left| \mathrm{e}^{z_n^\top x} \right|^p \nu(dx) \le l.$$

Hence the class of functions $\{\mathrm{e}^{z_n^\top x} \mid n \in \mathbb{N}\}$ is uniformly integrable with respect to ν, see [161, 13.3]. Since $\mathrm{e}^{z_n^\top x} \to \mathrm{e}^{z^\top x}$ for all x, we conclude by Lebesgue's convergence theorem that

$$|G(z_n) - G(z)| \le \int_{\mathbb{R}^d} \left| \mathrm{e}^{z_n^\top x} - \mathrm{e}^{z^\top x} \right| \nu(dx) \to 0.$$

Hence G is continuous on $\mathcal{S}(U_l)$.

It thus follows from the Cauchy formula, see [55, Sect. IX.9], that G is analytic on $\mathcal{S}(U_l)$ if and only if, for every $z \in \mathcal{S}(U_l)$ and $1 \le i \le d$, the function $\zeta \mapsto G(z + \zeta e_i)$ is analytic on $\{\zeta \in \mathbb{C} \mid z + \zeta e_i \in \mathcal{S}(U_l)\}$. Here, as usual, we denote by e_i the ith standard basis vector in $\mathbb{R}^d$.

We thus let $z \in \mathcal{S}(U_l)$ and $1 \le i \le d$. Then there exists some $\varepsilon_- < 0 < \varepsilon_+$ such that $z + \zeta e_i \in \mathcal{S}(U_l)$ for all $\zeta \in \mathcal{S}([\varepsilon_-, \varepsilon_+])$. In particular, $|\mathrm{e}^{(z+\varepsilon_- e_i)^\top x}|\,\nu(dx)$ and $|\mathrm{e}^{(z+\varepsilon_+ e_i)^\top x}|\,\nu(dx)$ are bounded measures on $\mathbb{R}^d$. By dominated convergence, it follows that the two summands

$$\begin{aligned} G(z + \zeta e_i) &= \int_{\{x_i < 0\}} \mathrm{e}^{(\zeta - \varepsilon_-)x_i} \mathrm{e}^{(z+\varepsilon_- e_i)^\top x} \nu(dx) \\ &\quad + \int_{\{x_i \ge 0\}} \mathrm{e}^{(\zeta - \varepsilon_+)x_i} \mathrm{e}^{(z+\varepsilon_+ e_i)^\top x} \nu(dx), \end{aligned}$$

are complex differentiable, and thus G is analytic, in $\zeta \in \mathcal{S}((\varepsilon_-, \varepsilon_+))$. Whence G is analytic on $\mathcal{S}(U_l)$. Since $\mathcal{S}(U) = \bigcup_{l>0} \mathcal{S}(U_l)$, the lemma follows. □

In general, V does not have an open interior in $\mathbb{R}^d$. The next lemma provides sufficient conditions for the existence of an open set $U \subset V$ in $\mathbb{R}^d$.

Lemma 10.9 *Let U' be an open neighborhood of* 0 *in $\mathbb{C}^d$ and h an analytic function on U'. Suppose that $U = U' \cap \mathbb{R}^d$ is star-shaped around* 0 *and $G(z) = h(z)$ for all $z \in U' \cap \mathrm{i}\mathbb{R}^d$. Then $U \subset V$ and $G = h$ on $U' \cap \mathcal{S}(U)$.*

Proof We first suppose that $U' = P_\rho$ for the open polydisc

$$P_\rho = \left\{ z \in \mathbb{C}^d \mid |z_i| < \rho_i,\ 1 \le i \le d \right\},$$

for some $\rho = (\rho_1, \dots, \rho_d) \in \mathbb{R}^d_{++}$. Note the symmetry $\mathrm{i}P_\rho = P_\rho$.

As in Lemma 10.7, we define $g(y) = G(iy)$ for $y \in \mathbb{R}^d$. By assumption, $g(y) = h(\mathrm{i}y)$ for all $y \in P_\rho \cap \mathbb{R}^d$. Hence g is analytic on $P_\rho \cap \mathbb{R}^d$, and the Cauchy formula, [55, Sect. IX.9], yields

$$g(y) = \sum_{i_1,\dots,i_d \in \mathbb{N}_0} c_{i_1,\dots,i_d} y_1^{i_1} \cdots y_d^{i_d} \quad \text{for } y \in P_\rho \cap \mathbb{R}^d$$

where $\sum_{i_1,\dots,i_d \in \mathbb{N}_0} c_{i_1,\dots,i_d} z_1^{i_1} \cdots z_d^{i_d} = h(\mathrm{i}z)$ for all $z \in P_\rho$. This power series is absolutely convergent on P_ρ, that is,

$$\sum_{i_1,\dots,i_d \in \mathbb{N}_0} |c_{i_1,\dots,i_d}||z_1^{i_1} \cdots z_d^{i_d}| < \infty \quad \text{for all } z \in P_\rho.$$

From the first part of Lemma 10.7, we infer that ν possesses all moments, that is, $\int_{\mathbb{R}^d} \|x\|^k \, \nu(dx) < \infty$ for all $k \in \mathbb{N}$. From the second part of Lemma 10.7 thus

$$c_{i_1,\dots,i_d} = \frac{\mathrm{i}^{i_1+\cdots+i_d}}{i_1! \cdots i_d!} \int_{\mathbb{R}^d} x_1^{i_1} \cdots x_d^{i_d} \, \nu(dx).$$

From the inequality $|x_i|^{2k-1} \le (x_i^{2k} + x_i^{2k-2})/2$, for $k \in \mathbb{N}$, and the above properties, we infer that for all $z \in P_\rho$,

$$\int_{\mathbb{R}^d} \mathrm{e}^{\sum_{i=1}^d |z_i||x_i|} \, \nu(dx) = \sum_{i_1,\dots,i_d \in \mathbb{N}_0} \frac{|z_1^{i_1} \cdots z_d^{i_d}|}{i_1! \cdots i_d!} \int_{\mathbb{R}^d} |x_1^{i_1} \cdots x_d^{i_d}| \, \nu(dx) < \infty.$$

Hence $P_\rho \cap \mathbb{R}^d \subset V$, and Lemma 10.8 implies that G is analytic on $\mathcal{S}(P_\rho \cap \mathbb{R}^d)$. Since the power series for G and h coincide on $P_\rho \cap \mathrm{i}\mathbb{R}^d$, we conclude that $G = h$ on P_ρ, and the lemma is proved for $U' = P_\rho$.

Now let U' be an open neighborhood of 0 in $\mathbb{C}^d$. Then there exists some open polydisc $P_\rho \subset U'$ with $\rho \in \mathbb{R}^d_{++}$. By the preceding case, we have $P_\rho \cap \mathbb{R}^d \subset V$ and $G = h$ on P_ρ. In view of Lemma 10.8 it thus remains to show that $U = U' \cap \mathbb{R}^d \subset V$.

To this end, let $a \in U$. Since U is star-shaped around 0 in $\mathbb{R}^d$, there exists some $s_1 > 1$ such that $sa \in U$ for all $s \in [0, s_1]$ and $h(sa)$ is analytic in $s \in (0, s_1)$. On the other hand, there exists some $0 < s_0 < s_1$ such that $sa \in P_\rho \cap \mathbb{R}^d$ for all $s \in [0, s_0]$, and $G(sa) = h(sa)$ for $s \in (0, s_0)$. This implies

$$\int_{\{a^\top x \ge 0\}} \mathrm{e}^{sa^\top x} \, \nu(dx) = h(sa) - \int_{\{a^\top x < 0\}} \mathrm{e}^{sa^\top x} \, \nu(dx)$$

for $s \in (0, s_0)$. By Lemma 10.8, the right-hand side is an analytic function in $s \in (0, s_1)$. We conclude by Lemma 10.10 below, for μ defined as the image measure of ν on $\mathbb{R}_+$ by the mapping $x \mapsto a^\top x$, that $a \in V$. Hence the lemma is proved. □

Lemma 10.10 *Let μ be a bounded measure on $\mathbb{R}_+$, and h an analytic function on $(0, s_1)$, such that*

$$\int_{\mathbb{R}_+} \mathrm{e}^{sx} \, \mu(dx) = h(s) \tag{10.42}$$

for all $s \in (0, s_0)$, *for some numbers* $0 < s_0 < s_1$. *Then* (10.42) *also holds for* $s \in (0, s_1)$.

Proof Define $f(s) = \int_{\mathbb{R}_+} \mathrm{e}^{sx}\,\mu(dx)$ and $s_\infty = \sup\{s > 0 \mid f(s) < \infty\} \geq s_0$, such that

$$f(s) = +\infty \quad \text{for } s > s_\infty. \tag{10.43}$$

We assume, by contradiction, that $s_\infty < s_1$. Then there exists some $s_* \in (0, s_\infty)$ and $\varepsilon > 0$ such that $s_* < s_\infty < s_* + \varepsilon$ and such that h can be developed in an absolutely convergent power series

$$h(s) = \sum_{k \geq 0} \frac{c_k}{k!}(s - s_*)^k \quad \text{for } s \in (s_* - \varepsilon, s_* + \varepsilon).$$

In view of Lemma 10.8, f is analytic, and thus $f = h$, on $(0, s_\infty)$. Hence we obtain, by dominated convergence,

$$c_k = \frac{d^k}{ds^k} h(s)\Big|_{s=s_*} = \frac{d^k}{ds^k} f(s)\Big|_{s=s_*} = \int_{\mathbb{R}_+} x^k \mathrm{e}^{s_* x}\,\mu(dx) \geq 0.$$

By monotone convergence, we conclude

$$\begin{aligned} h(s) &= \sum_{k \geq 0} \int_{\mathbb{R}_+} \frac{x^k}{k!}(s - s_*)^k \mathrm{e}^{s_* x}\,\mu(dx) = \int_{\mathbb{R}_+} \sum_{k \geq 0} \frac{x^k}{k!}(s - s_*)^k \mathrm{e}^{s_* x}\,\mu(dx) \\ &= \int_{\mathbb{R}_+} \mathrm{e}^{sx}\,\mu(dx) \end{aligned}$$

for all $s \in (s_*, s_* + \varepsilon)$. But this contradicts (10.43). Whence $s_\infty \geq s_1$, and the lemma is proved. □

10.7 Auxiliary Results for Differential Equations

In this section we deliver invariance and comparison results for stochastic and ordinary differential equations, which are used in the proofs of the main Theorems 10.2, 10.3 and 10.4 and Lemma 10.6. This section can be skipped at the first reading.

10.7.1 Some Invariance Results

We start with an invariance result for the stochastic differential equation (10.1).

Lemma 10.11 *Suppose b and ρ in (10.1) admit a continuous and measurable extension to $\mathbb{R}^d$, respectively, and such that a is continuous on $\mathbb{R}^d$. Let $u \in \mathbb{R}^d \setminus \{0\}$ and define the half space*

$$H = \{x \in \mathbb{R}^d \mid u^\top x \geq 0\},$$

its interior $H^0 = \{x \in \mathbb{R}^d \mid u^\top x > 0\}$, and its boundary $\partial H = \{x \in H \mid u^\top x = 0\}$.

(a) *Fix $x \in \partial H$ and let $X = X^x$ be a solution of (10.1). If $X(t) \in H$ for all $t \geq 0$, then necessarily*

$$u^\top a(x) u = 0, \tag{10.44}$$

$$u^\top b(x) \geq 0. \tag{10.45}$$

(b) *Conversely, if (10.44) and (10.45) hold for all $x \in \mathbb{R}^d \setminus H^0$, then any solution X of (10.1) with $X(0) \in H$ satisfies $X(t) \in H$ for all $t \geq 0$.*

Intuitively speaking, (10.44) means that the diffusion must be "parallel to the boundary", and (10.45) says that the drift must be "inward pointing" at the boundary of H.

Proof Fix $x \in \partial H$ and let $X = X^x$ be a solution of (10.1). Hence

$$u^\top X(t) = \int_0^t u^\top b(X(s))\, ds + \int_0^t u^\top \rho(X(s))\, dW(s).$$

Since a and b are continuous, there exists a stopping time $\tau_1 > 0$ and a finite constant K such that

$$|u^\top b(X(t \wedge \tau_1))| \leq K$$

and

$$\|u^\top \rho(X(t \wedge \tau_1))\|^2 = u^\top a(X(t \wedge \tau_1)) u \leq K$$

for all $t \geq 0$. In particular, the stochastic integral part of $u^\top X(t \wedge \tau_1)$ is a martingale. Hence

$$\mathbb{E}\left[u^\top X(t \wedge \tau_1)\right] = \mathbb{E}\left[\int_0^{t \wedge \tau_1} u^\top b(X(s))\, ds\right], \quad t \geq 0.$$

We now argue by contradiction, and assume first that $u^\top b(x) < 0$. By continuity of b and $X(t)$, there exists some $\varepsilon > 0$ and a stopping time $\tau_2 > 0$ such that $u^\top b(X(t)) \leq -\varepsilon$ for all $t \leq \tau_2$. In view of the above this implies

$$\mathbb{E}\left[u^\top X(\tau_2 \wedge \tau_1)\right] < 0.$$

This contradicts $X(t) \in H$ for all $t \geq 0$, whence (10.45) holds.

As for (10.44), let $C > 0$ be a finite constant and define the stochastic exponential $Z_t = \mathcal{E}(-C\int_0^t u^\top \rho(X)\,dW)$. Then Z is a positive local martingale. Integration by parts yields

$$u^\top X(t)Z(t) = \int_0^t Z(s)\left(u^\top b(X(s)) - Cu^\top a(X(s))u\right)ds + M(t)$$

where M is a local martingale. Hence there exists a stopping time $\tau_3 > 0$ such that for all $t \geq 0$,

$$\mathbb{E}\left[u^\top X(t\wedge\tau_3)Z(t\wedge\tau_3)\right] = \mathbb{E}\left[\int_0^{t\wedge\tau_3} Z(s)\left(u^\top b(X(s)) - Cu^\top a(X(s))u\right)ds\right].$$

Now assume that $u^\top a(x)u > 0$. By continuity of a and $X(t)$, there exists some $\varepsilon > 0$ and a stopping time $\tau_4 > 0$ such that $u^\top a(X(t))u \geq \varepsilon$ for all $t \leq \tau_4$. For $C > K/\varepsilon$, this implies

$$\mathbb{E}\left[u^\top X(\tau_4\wedge\tau_3\wedge\tau_1)Z(\tau_4\wedge\tau_3\wedge\tau_1)\right] < 0.$$

This contradicts $X(t) \in H$ for all $t \geq 0$. Hence (10.44) holds, and part (a) is proved.

As for part (b), suppose (10.44) and (10.45) hold for all $x \in \mathbb{R}^d \setminus H^0$, and let X be a solution of (10.1) with $X(0) \in H$. For $\delta, \varepsilon > 0$ define the stopping time

$$\tau_{\delta,\varepsilon} = \inf\left\{t \mid u^\top X(t) \leq -\varepsilon \text{ and } u^\top X(s) < 0 \text{ for all } s \in [t-\delta, t]\right\}.$$

Then on $\{\tau_{\delta,\varepsilon} < \infty\}$ we have $u^\top \rho(X(s)) = 0$ for $\tau_{\delta,\varepsilon} - \delta \leq s \leq \tau_{\delta,\varepsilon}$ and thus

$$0 > u^\top X(\tau_{\delta,\varepsilon}) - u^\top X(\tau_{\delta,\varepsilon} - \delta) = \int_{\tau_{\delta,\varepsilon}-\delta}^{\tau_{\delta,\varepsilon}} u^\top b(X(s))\,ds \geq 0,$$

a contradiction. Hence $\tau_{\delta,\varepsilon} = \infty$. Since $\delta, \varepsilon > 0$ were arbitrary, we conclude that $u^\top X(t) \geq 0$ for all $t \geq 0$, as desired. Whence the lemma is proved. □

It is straightforward to extend Lemma 10.11 towards a polyhedral convex set $\bigcap_{i=1}^k H_i$ with half-spaces $H_i = \{x \in \mathbb{R}^d \mid u_i^\top x \geq 0\}$, for some elements $u_1, \dots, u_k \in \mathbb{R}^d \setminus \{0\}$ and some $k \in \mathbb{N}$. This holds in particular for the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$. Moreover, Lemma 10.11 includes time-inhomogeneous[12] ordinary differential equations as special case. The proofs of the following two corollaries are left to the reader (→ Exercise 10.22).

Corollary 10.5 *Let $H_i = \{x \in \mathbb{R}^d \mid x_i \geq 0\}$ denote the ith canonical half space in $\mathbb{R}^d$, for $i = 1, \dots, m$. Let $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a continuous map satisfying, for*

[12]Time-inhomogeneous differential equations can be made homogeneous by enlarging the state space.

all $t \geq 0$,

$$b(t,x) = b(t, x_1^+, \dots, x_m^+, x_{m+1}, \dots, x_d) \quad \text{for all } x \in \mathbb{R}^d, \text{ and}$$

$$b_i(t,x) \geq 0 \quad \text{for all } x \in \partial H_i,\ i = 1, \dots, m.$$

Then any solution f *of*

$$\partial_t f(t) = b(t, f(t))$$

with $f(0) \in \mathbb{R}_+^m \times \mathbb{R}^n$ *satisfies* $f(t) \in \mathbb{R}_+^m \times \mathbb{R}^n$ *for all* $t \geq 0$.

Corollary 10.6 *Let* $B(t)$ *and* $C(t)$ *be continuous* $\mathbb{R}^{m\times m}$*- and* $\mathbb{R}_+^m$*-valued parameters, respectively, such that* $B_{ij}(t) \geq 0$ *whenever* $i \neq j$. *Then the solution* f *of the linear differential equation in* $\mathbb{R}^m$

$$\partial_t f(t) = B(t) f(t) + C(t)$$

with $f(0) \in \mathbb{R}_+^m$ *satisfies* $f(t) \in \mathbb{R}_+^m$ *for all* $t \geq 0$.

Here and subsequently, we let $\succeq$ denote the partial order on $\mathbb{R}^m$ induced by the cone $\mathbb{R}_+^m$. That is, $x \succeq y$ if $x - y \in \mathbb{R}_+^m$. Then Corollary 10.6 may be rephrased, for $C(t) \equiv 0$, by saying that the operator $\mathrm{e}^{\int_0^t B(s)\,ds}$ is $\succeq$-order preserving, i.e. $\mathrm{e}^{\int_0^t B(s)\,ds}\mathbb{R}_+^m \subseteq \mathbb{R}_+^m$.

10.7.2 Some Results on Riccati Equations

We first provide the explicit solution for the one-dimensional Riccati equation.

Lemma 10.12 *Consider the Riccati differential equation*

$$\partial_t G = AG^2 + BG - C, \qquad G(0,u) = u, \tag{10.46}$$

where $A, B, C \in \mathbb{C}$ *and* $u \in \mathbb{C}$, *with* $A \neq 0$ *and* $B^2 + 4AC \in \mathbb{C} \setminus \mathbb{R}_-$. *Let* $\sqrt{\cdot}$ *denote the analytic extension of the real square root to* $\mathbb{C} \setminus \mathbb{R}_-$, *and define* $\theta = \sqrt{B^2 + 4AC}$.

(a) *The function*

$$G(t,u) = -\frac{2C(\mathrm{e}^{\theta t} - 1) - (\theta(\mathrm{e}^{\theta t} + 1) + B(\mathrm{e}^{\theta t} - 1))u}{\theta(\mathrm{e}^{\theta t} + 1) - B(\mathrm{e}^{\theta t} - 1) - 2A(\mathrm{e}^{\theta t} - 1)u} \tag{10.47}$$

is the unique solution of equation (10.46) *on its maximal interval of existence* $[0, t_+(u))$. *Moreover,*

$$\int_0^t G(s,u)ds = \frac{1}{A}\log\left(\frac{2\theta \mathrm{e}^{\frac{\theta - B}{2}t}}{\theta(\mathrm{e}^{\theta t} + 1) - B(\mathrm{e}^{\theta t} - 1) - 2A(\mathrm{e}^{\theta t} - 1)u}\right). \tag{10.48}$$

(b) *If, moreover,* $A > 0$, $B \in \mathbb{R}$, $\Re(C) \geq 0$ *and* $u \in \mathbb{C}_-$ *then* $t_+(u) = \infty$ *and* $G(t,u)$ *is* $\mathbb{C}_-$*-valued.*

Proof Recall that the square root $\sqrt{z} = \mathrm{e}^{1/2 \log(z)}$ is the well-defined analytic extension of the real square root to $\mathbb{C} \setminus \mathbb{R}_-$, through the main branch of the logarithm which can be written in the form $\log(z) = \int_{[0,z]} \frac{dz}{z}$. Hence we may write (10.46) as

$$\partial_t G = A(G - \theta_+)(G - \theta_-), \qquad G(0,u) = u,$$

where $\theta_\pm = \frac{-B \pm \sqrt{B^2 + 4AC}}{2A}$, and it follows that

$$G(t,u) = \frac{\theta_+(u - \theta_-) - \theta_-(u - \theta_+)\mathrm{e}^{\theta t}}{(u - \theta_-) - (u - \theta_+)\mathrm{e}^{\theta t}},$$

which can be seen to be equivalent to (10.47). As $\theta_+ \neq \theta_-$, numerator and denominator cannot vanish at the same time t, and certainly not for t near zero. Hence, by the maximality of $t_+(u)$, (10.47) is the solution of (10.46) for $t \in [0, t_+(u))$. Finally, the integral (10.48) is checked by differentiation. This proves (a).

As for (b), we show along the lines of the proof of Theorem 10.2, that for this choice of coefficients global solutions exist for initial data $u \in \mathbb{C}_-$ and stay in $\mathbb{C}_-$. To this end, write $R(G) = AG^2 + BG - C$, then

$$\Re(R(G)) = A(\Re(G))^2 - A(\Im(G))^2 + B\Re(G) - \Re(C) \leq A(\Re(G))^2 + B\Re(G)$$

and since $A, B \in \mathbb{R}$ we have that $\Re(G(t,u)) \leq 0$ for all times $t \in [0, t_+(u))$, see Corollary 10.5 below. Furthermore, we see that $\Re(\overline{G} R(G)) \leq (1 + |G|^2)(|B| + |C|)$, hence $\partial_t |G(t,u)|^2 \leq 2(1 + |G(t,u)|^2)(|B| + |C|)$. This implies, by Gronwall's inequality ([55, (10.5.1.3)]), that $t_+(u) = \infty$. Hence the lemma is proved. □

Next, we consider time-inhomogeneous Riccati equations in $\mathbb{R}^m$ of the special form

$$\partial_t f_i(t) = A_i f_i(t)^2 + B_i^\top f(t) + C_i(t), \quad i = 1, \ldots, m, \tag{10.49}$$

for some parameters $A, B, C(t)$ satisfying the following admissibility conditions:

$$\begin{aligned} A &= (A_1, \ldots, A_m) \in \mathbb{R}^m, \\ B_{i,j} &\geq 0 \quad \text{for } 1 \leq i \neq j \leq m, \\ C(t) &= (C_1(t), \ldots, C_m(t)) \quad \text{continuous } \mathbb{R}^m\text{-valued.} \end{aligned} \tag{10.50}$$

The following lemma provides a comparison result for (10.49). It shows, in particular, that the solution of (10.49) is uniformly bounded from below on compacts with respect to $\succeq$ if $A \succeq 0$.

Lemma 10.13 *Let* $A^{(k)}, B, C^{(k)}$, $k = 1, 2$, *be parameters satisfying the admissibility conditions* (10.50), *and*

$$A^{(1)} \preceq A^{(2)}, \qquad C^{(1)}(t) \preceq C^{(2)}(t). \tag{10.51}$$

Let $\tau > 0$ *and* $f^{(k)} : [0, \tau) \to \mathbb{R}^m$ *be solutions of* (10.50) *with* A *and* C *replaced by* $A^{(k)}$ *and* $C^{(k)}$, *respectively,* $k = 1, 2$. *If* $f^{(1)}(0) \preceq f^{(2)}(0)$ *then* $f^{(1)}(t) \preceq f^{(2)}(t)$ *for all* $t \in [0, \tau)$. *If, moreover,* $A^{(1)} = 0$ *then*

$$e^{Bt}\left(f^{(1)}(0) + \int_0^t e^{-Bs} C^{(1)}(s)\, ds\right) \preceq f^{(2)}(t)$$

for all $t \in [0, \tau)$.

Proof The function $f = f^{(2)} - f^{(1)}$ solves

$$\begin{aligned}
\partial_t f_i &= A_i^{(2)} \left(f_i^{(2)}\right)^2 - A_i^{(1)} \left(f_i^{(1)}\right)^2 + B_i^\top f + C_i^{(2)} - C_i^{(1)} \\
&= \left(A_i^{(2)} - A_i^{(1)}\right) \left(f_i^{(2)}\right)^2 + A_i^{(1)} \left(f_i^{(2)} + f_i^{(1)}\right) f_i + B_i^\top f + C_i^{(2)} - C_i^{(1)} \\
&= \widetilde{B}_i^{\top} f + \widetilde{C}_i,
\end{aligned}$$

where we write

$$\begin{aligned}
\widetilde{B}_i &= \widetilde{B}_i(t) = B_i + A_i^{(1)} \left(f_i^{(2)}(t) + f_i^{(1)}(t)\right) e_i, \\
\widetilde{C}_i &= \widetilde{C}_i(t) = \left(A_i^{(2)} - A_i^{(1)}\right) \left(f_i^{(2)}(t)\right)^2 + C_i^{(2)}(t) - C_i^{(1)}(t).
\end{aligned}$$

Note that $\widetilde{B} = (\widetilde{B}_{i,j})$ and $\widetilde{C}$ satisfy the assumptions of Corollary 10.6 in lieu of B and C, and $f(0) \in \mathbb{R}_+^m$. Hence Corollary 10.6 implies $f(t) \in \mathbb{R}_+^m$ for all $t \in [0, \tau)$, as desired. The last statement of the lemma follows by the variation of constants formula for $f^{(1)}(t)$. □

After these preliminary comparison results for the Riccati equation (10.49), we now can state and prove an important result for the system of Riccati equations (10.7).

Lemma 10.14 *Let* $\mathcal{D}_\mathbb{R}$ *denote the maximal domain for the system of Riccati equations* (10.7). *Let* $(\tau, u) \in \mathcal{D}_\mathbb{R}$. *Then*:

(a) $\mathcal{D}_\mathbb{R}(\tau)$ *is star-shaped around zero.*
(b) $\theta^* = \sup\{\theta \geq 0 \mid \theta u \in \mathcal{D}_\mathbb{R}(\tau)\}$ *satisfies either* $\theta^* = \infty$ *or*

$$\lim_{\theta \uparrow \theta^*} \|\psi_I(t, \theta u)\| = \infty.$$

In the latter case, there exists some $x^* \in \mathbb{R}_+^m \times \mathbb{R}^n$ *such that*

$$\lim_{\theta \uparrow \theta^*} \phi(\tau, \theta u) + \psi(\tau, \theta u)^\top x^* = \infty.$$

Proof We first assume that the matrices α_i are block-diagonal, such that $\alpha_{i,iJ} = 0$, for all $i = 1, \dots, m$.

Fix $\theta \in (0, 1]$. We claim that $\theta u \in \mathcal{D}_{\mathbb{R}}(\tau)$. It follows by inspection that $f^{(\theta)}(t) = \frac{\psi_I(t,\theta u)}{\theta}$ solves (10.49) with

$$A_i^{(\theta)} = \frac{1}{2}\theta\alpha_{i,ii}, \qquad B = \mathcal{B}_{II}^\top,$$

$$C_i^{(\theta)}(t) = \beta_{i,J}^\top \psi_J(t,u) + \frac{1}{2}\psi_J(t,u)^\top \theta\alpha_{i,JJ}\psi_J(t,u),$$

and $f(0) = u$. Lemma 10.13 thus implies that $f^{(\theta)}(t)$ is nicely behaved, as

$$e^{\mathcal{B}_{II}^\top t}\left(u + \int_0^t e^{-\mathcal{B}_{II}^\top s} C^{(0)}(s)\,ds\right) \preceq f^{(\theta)}(t) \preceq \psi_I(t,u), \tag{10.52}$$

for all $t \in [0, t_+(\theta u)) \cap [0, \tau]$. By the maximality of $\mathcal{D}_{\mathbb{R}}$ we conclude that $\tau < t_+(\theta u)$, which implies $\theta u \in \mathcal{D}_{\mathbb{R}}(\tau)$, as desired. Hence $\mathcal{D}_{\mathbb{R}}(\tau)$ is star-shaped around zero, which is part (a).

Next suppose that $\theta^* < \infty$. Since $\mathcal{D}_{\mathbb{R}}(\tau)$ is open, this implies $\theta^* u \notin \mathcal{D}_{\mathbb{R}}(\tau)$ and thus $t_+(\theta^* u) \leq \tau$. From part (a) we know that $(t, \theta u) \in \mathcal{D}_{\mathbb{R}}$ for all $t < t_+(\theta^* u)$ and $0 \leq \theta \leq \theta^*$. On the other hand, there exists a sequence $t_n \uparrow t_+(\theta^* u)$ such that $\|\psi_I(t_n, \theta^* u)\| > n$ for all $n \in \mathbb{N}$. By continuity of ψ on $\mathcal{D}_{\mathbb{R}}$, we conclude that there exists some sequence $\theta_n \uparrow \theta^*$ with $\|\psi_I(t_n, \theta_n u) - \psi_I(t_n, \theta^* u)\| \leq 1/n$ and hence

$$\lim_n \|\psi_I(t_n, \theta_n u)\| = \infty. \tag{10.53}$$

Applying Lemma 10.13 as above, where initial time $t = 0$ is shifted to t_n, yields

$$g_n = e^{\mathcal{B}_{II}^\top(\tau - t_n)}\left(f^{(\theta_n)}(t_n) + \int_{t_n}^\tau e^{\mathcal{B}_{II}^\top(t_n - s)} C^{(0)}(s)\,ds\right) \preceq f^{(\theta_n)}(\tau).$$

Corollary 10.6 implies that $e^{\mathcal{B}_{II}^\top(\tau - t_n)}$ is $\succeq$-order preserving. That is, $e^{\mathcal{B}_{II}^\top(\tau - t_n)}\mathbb{R}_+^m \subseteq \mathbb{R}_+^m$. Hence, in view of (10.52) for $f^{(\theta_n)}(t_n)$,

$$\begin{aligned} g_n &\succeq e^{\mathcal{B}_{II}^\top(\tau - t_n)}\left(e^{\mathcal{B}_{II}^\top t_n}\left(u + \int_0^{t_n} e^{-\mathcal{B}_{II}^\top s} C^{(0)}(s)\,ds\right) + \int_{t_n}^\tau e^{\mathcal{B}_{II}^\top(t_n - s)} C^{(0)}(s)\,ds\right) \\ &= e^{\mathcal{B}_{II}^\top \tau}\left(u + \int_0^\tau e^{-\mathcal{B}_{II}^\top s} C^{(0)}(s)\,ds\right). \end{aligned}$$

On the other hand, elementary operator norm inequalities yield

$$\|g_n\| \geq e^{-\|\mathcal{B}_{II}\|\tau}\|f^{(\theta_n)}(t_n)\| - e^{\|\mathcal{B}_{II}\|\tau}\tau \sup_{s\in[0,\tau]}\|C^{(0)}(s)\|.$$

Together with (10.53), this implies $\|g_n\| \to \infty$. From Lemma 10.15 below we conclude that $\lim_n f^{(\theta_n)}(\tau)^\top y^* = \infty$ for some $y^* \in \mathbb{R}_+^m$. Moreover, in view of Lemma 10.13, we know that $f^{(\theta)}(\tau)^\top y^*$ is nondecreasing in θ. Therefore $\lim_{\theta\uparrow\theta^*} f^{(\theta)}(\tau)^\top y^* = \infty$. Applying (10.52) and Lemma 10.15 below again, this also implies that

$$\lim_{\theta\uparrow\theta^*} \|f^{(\theta)}(\tau)\| = \infty.$$

It remains to set $x^* = (y^*, 0)$ and observe that $b_I \in \mathbb{R}^m_+$ and thus

$$\phi(\tau,\theta u) = \int_0^\tau \left(\frac{1}{2}\psi_J(t,\theta u)^\top a_{JJ}\,\psi_J(t,\theta u) + b_I^\top \psi_I(t,\theta u) + b_J^\top \psi_J(t,\theta u)\right) dt$$

is uniformly bounded from below for all $\theta \in [0,\theta^*)$. Thus the lemma is proved under the premise that the matrices α_i are block-diagonal for all $i = 1,\dots,m$.

The general case of admissible parameters a, α_i, b, β_i is reduced to the preceding block-diagonal case by a linear transformation along the lines of Lemma 10.5. Indeed, define the invertible $d \times d$-matrix Λ

$$\Lambda = \begin{pmatrix} I_m & 0 \\ D & I_n \end{pmatrix}, \tag{10.54}$$

where the $n \times m$-matrix $D = (\delta_1,\dots,\delta_m)$ has ith column vector

$$\delta_i = \begin{cases} -\frac{\alpha_{i,iJ}}{\alpha_{i,ii}}, & \text{if } \alpha_{i,ii} > 0, \\ 0, & \text{else.} \end{cases}$$

It is then not hard to see ($\rightarrow$ Exercise 10.23) that $\Lambda(\mathbb{R}^m_+ \times \mathbb{R}^n) = \mathbb{R}^m_+ \times \mathbb{R}^n$, and

$$\widetilde{\phi}(t,u) = \phi(t,\Lambda^\top u), \qquad \widetilde{\psi}(t,u) = \left(\Lambda^\top\right)^{-1} \psi(t,\Lambda^\top u) \tag{10.55}$$

satisfy the system of Riccati equations (10.7) with a, α_i, b, and $\mathcal{B} = (\beta_1,\dots,\beta_d)$ replaced by the admissible parameters

$$\widetilde{a} = \Lambda a \Lambda^\top, \qquad \widetilde{\alpha_i} = \Lambda \alpha_i \Lambda^\top, \qquad \widetilde{b} = \Lambda b, \qquad \widetilde{\mathcal{B}} = \Lambda \mathcal{B} \Lambda^{-1}. \tag{10.56}$$

Moreover, $\widetilde{\alpha_i}$ are block-diagonal, for all $i = 1,\dots,m$.

By the first part of the proof, the corresponding maximal domain $\widetilde{\mathcal{D}_\mathbb{R}}(\tau)$, and hence also $\mathcal{D}_\mathbb{R}(\tau) = \Lambda^\top \widetilde{\mathcal{D}_\mathbb{R}}(\tau)$, is star-shaped around zero. Moreover, if $\theta^* < \infty$, then

$$\lim_{\theta\uparrow\theta^*} \|\psi_I(\tau,\theta u)\| = \lim_{\theta\uparrow\theta^*} \left\| \widetilde{\psi}_I\left(\tau, \theta\left(\Lambda^\top\right)^{-1} u\right)\right\| = \infty,$$

and there exists some $x^* \in \mathbb{R}^m_+ \times \mathbb{R}^n$ such that

$$\begin{aligned} &\lim_{\theta\uparrow\theta^*} \phi(\tau,\theta u) + \psi(\tau,\theta u)^\top x^* \\ &\quad = \lim_{\theta\uparrow\theta^*} \widetilde{\phi}\left(\tau,\theta\left(\Lambda^\top\right)^{-1} u\right) + \widetilde{\psi}\left(\tau,\theta\left(\Lambda^\top\right)^{-1} u\right)^\top \Lambda x^* = \infty. \end{aligned}$$

Hence the lemma is proved. □

Lemma 10.15 *Let $c \in \mathbb{R}^m$, and (c_n) and (d_n) be sequences in $\mathbb{R}^m$ such that*

$$c \preceq c_n \preceq d_n$$

for all $n \in \mathbb{N}$. Then the following are equivalent:

(a) $\|c_n\| \to \infty$.
(b) $c_n^\top y^* \to \infty$ *for some* $y^* \in \mathbb{R}_+^m \setminus \{0\}$.

In either case, $\|d_n\| \to \infty$ *and* $d_n^\top y^* \to \infty$.

Proof (a) ⇒ (b): since $\|c_n\|^2 = \sum_{i=1}^m (c_n^\top e_i)^2$ and $c_n^\top e_i \geq c^\top e_i$, we conclude that $c_n^\top e_i \to \infty$ for some $i = 1, \dots, m$.

(b) ⇒ (a): this follows from $\|c_n^\top y^*\| \leq \|c_n\| \|y^*\|$.

The last statement now follows since $d_n^\top y^* \geq c_n^\top y^*$. □

We now have all the ingredients needed for the proof of Theorem 10.3.

10.7.3 Proof of Theorem 10.3

We first claim that, for every $u \in \mathbb{C}^d$ with $t_+(u) < \infty$, there exists some $i \in I$ and some sequence $t_n \uparrow t_+(u)$ such that

$$\lim_n (\Re \psi_i(t_n, u))^+ = \infty. \tag{10.57}$$

Indeed, otherwise we would have $\sup_{t \in [0, t_+(u))} \|(\Re \psi_I(t, u))^+\| < \infty$. But then (10.8) would imply $\sup_{t \in [0, t_+(u))} \|\psi_I(t, u)\| < \infty$, which is absurd. Whence (10.57) is proved.

In the following, we write

$$G(u, t, x) = \mathbb{E}\left[\mathrm{e}^{u^\top X^x(t)}\right], \qquad V(t, x) = \left\{u \in \mathbb{R}^d \mid G(u, t, x) < \infty\right\}.$$

Since X is affine, by definition we have $\mathbb{R}_+ \times \mathrm{i}\mathbb{R}^d \subset \mathcal{D}_\mathbb{C}$ and (10.2) implies

$$G(u, t, x) = \mathrm{e}^{\phi(t,u) + \psi(t,u)^\top x} \tag{10.58}$$

for all $u \in \mathrm{i}\mathbb{R}^d$, $t \in \mathbb{R}_+$ and $x \in \mathbb{R}_+^m \times \mathbb{R}^n$. Moreover, by Lemma 10.14, $\mathcal{D}_\mathbb{R}(t) = \mathcal{D}_\mathbb{C}(t) \cap \mathbb{R}^d$ is open and star-shaped around 0 in $\mathbb{R}^d$. Hence Lemma 10.9 implies that $\mathcal{D}_\mathbb{R}(t) \subset V(t, x)$ and (10.58) holds for all $u \in \mathcal{D}_\mathbb{C}(t) \cap \mathcal{S}(\mathcal{D}_\mathbb{R}(t))$, for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$ and $t \in [0, \tau]$.

Now let $u \in \mathcal{D}_\mathbb{R}(\tau)$ and $v \in \mathbb{R}^d$, and define

$$\theta^* = \inf\{\theta \in \mathbb{R}_+ \mid u + \mathrm{i}\theta v \notin \mathcal{D}_\mathbb{C}(\tau)\}.$$

We claim that $\theta^* = \infty$. Arguing by contradiction, assume that $\theta^* < \infty$. Since $\mathcal{D}_{\mathbb{C}}(\tau)$ is open, this implies $u + \mathrm{i}\theta^* v \notin \mathcal{D}_{\mathbb{C}}(\tau)$, and thus

$$t_+(u + \mathrm{i}\theta^* v) \leq \tau. \tag{10.59}$$

On the other hand, since $\mathcal{D}_{\mathbb{R}}(\tau)$ is open, $(1+\varepsilon)u \in \mathcal{D}_{\mathbb{R}}(\tau)$ for some $\varepsilon > 0$. Hence (10.58) holds and $G(t, (1+\varepsilon)u, x)$ is uniformly bounded in $t \in [0, \tau]$, by continuity of $\phi(t, (1+\varepsilon)u)$ and $\psi(t, (1+\varepsilon)u)$ in t. We infer that the class of random variables $\{\mathrm{e}^{(u+\mathrm{i}\theta^* v)^\top X(t)} \mid t \in [0, \tau]\}$ is uniformly integrable, see [161, 13.3]. Since $X(t)$ is continuous in t, we conclude by Lebesgue's convergence theorem that $G(t, u + \mathrm{i}\theta^* v, x)$ is continuous in $t \in [0, \tau]$, for all $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$. But for all $t < t_+(u + \mathrm{i}\theta^* v)$ we have $(t, u + \mathrm{i}\theta^* v) \in \mathcal{D}_{\mathbb{C}}(t) \cap \mathcal{S}(\mathcal{D}_{\mathbb{R}}(t))$, and thus (10.58) holds for all $x \in \mathbb{R}^m_+ \times \mathbb{R}^n$. In view of (10.57), this contradicts (10.59). Whence $\theta^* = \infty$ and thus $u + \mathrm{i}v \in \mathcal{D}_{\mathbb{C}}(\tau)$. This proves (a).

Applying the above arguments to[13] $\mathbb{E}[\mathrm{e}^{u^\top X(T)} \mid \mathcal{F}_t] = G(T-t, u, X(t))$ with $T = t + \tau$ yields (d). Part (e) follows, since, by Theorem 10.2, $\mathbb{C}^m_- \times \mathrm{i}\mathbb{R}^n \subset \mathcal{S}(\mathcal{D}_{\mathbb{R}}(t))$ for all $t \in \mathbb{R}_+$.

As for (b), we first let $u \in \mathcal{D}_{\mathbb{R}}(\tau)$. From part (d) it follows that $u \in M(\tau)$. Conversely, let $u \in M(\tau)$, and define $\theta^* = \sup\{\theta \geq 0 \mid \theta u \in \mathcal{D}_{\mathbb{R}}(\tau)\}$. We have to show that $\theta^* > 1$. Assume, by contradiction, that $\theta^* \leq 1$. From Lemma 10.14, we know that there exists some $x^* \in \mathbb{R}^m_+ \times \mathbb{R}^n$ such that

$$\lim_{\theta \uparrow \theta^*} \phi(\tau, \theta u) + \psi(\tau, \theta u)^\top x^* = \infty. \tag{10.60}$$

On the other hand, from part (d) and Jensen's inequality, we obtain

$$\mathrm{e}^{\phi(\tau,\theta u)+\psi(\tau,\theta u)^\top x^*} = G(\tau, \theta u, x^*) \leq G(\tau, u, x^*)^\theta \leq G(\tau, u, x^*) < \infty$$

for all $\theta < \theta^*$. But this contradicts (10.60), hence $u \in \mathcal{D}_{\mathbb{R}}(\tau)$, and part (b) is proved. Since $M(\tau)$ is convex, this also implies (c). Finally, part (f) follows from part (b) and the respective inclusion property $\mathcal{D}_{\mathbb{R}}(t) \supseteq \mathcal{D}_{\mathbb{R}}(T)$. Whence Theorem 10.3 is proved.

10.8 Exercises

Exercise 10.1 This exercise provides an example for a multivariate affine process defined on a state space which is not of the form $\mathbb{R}^m_+ \times \mathbb{R}^n$.

Consider the epigraph $\mathcal{X} = \{x \in \mathbb{R}^2 \mid x_1 \geq x_2^2\}$ of the parabola $x_1 = x_2^2$ in $\mathbb{R}^2$. Let $W = (W_1, W_2)^\top$ be a two-dimensional standard Brownian motion. For every $y \geq 0$, there exists a unique nonnegative affine diffusion process $Y = Y^y$ satisfying

$$dY = 2\sqrt{Y}\, dW_1, \quad Y(0) = y$$

[13] Here we use the Markov property of X, see Theorem 4.5.

(you do not have to prove this fact, it follows from Lemma 10.6). For every $x \in \mathcal{X}$ we define the $\mathcal{X}$-valued diffusion process $X = X^x$ by

$$X_1(t) = (W_2(t) + x_2)^2 + Y^y(t),$$
$$X_2(t) = W_2(t) + x_2,$$

where $y = y(x)$ is the unique nonnegative number with $x_1 = x_2^2 + y$.

(a) Show that X satisfies

$$dX_1 = dt + 2\sqrt{X_1 - X_2^2}\, dW_1 + 2X_2\, dW_2,$$
$$dX_2 = dW_2.$$

Conclude that the drift and diffusion matrix of X are affine functions of x. Verify that the diffusion matrix is positive semi-definite on $\mathcal{X}$.

(b) Verify by solving the corresponding Riccati equations that

$$\mathbb{E}\left[e^{u^\top X(T)} \mid \mathcal{F}_t\right] = \frac{1}{\sqrt{1 - 2u_1(T-t)}} e^{\frac{(T-t)u_2^2 + 2u^\top X(t)}{2(1-2u_2(T-t))}} \quad \text{for } u = (u_1, u_2)^\top \in \mathrm{i}\mathbb{R}^2.$$

Conclude that X is an affine process.

Exercise 10.2 Let B be a Brownian motion and define the $\mathbb{R}_+^2$-valued process X by $X_i(t) = (\sqrt{x_i} + B(t))^2$, for $i = 1, 2$, for some $x \in \mathbb{R}_+^2$.

(a) Show that X satisfies

$$dX_1 = dt + 2\sqrt{X_1}\, dW,$$
$$dX_2 = dt + 2\sqrt{X_2}\, dW,$$
$$X(0) = x,$$

for some Brownian motion W. Is X an affine process? Why (not)?

(b) Compute the characteristic function of $X(t)$ and verify your finding concerning the (supposed) affine property of X.

Exercise 10.3 Let W be a Brownian motion. The aim of this exercise is to find some $\gamma \in \mathcal{L}$ such that the stochastic exponential $\mathcal{E}(\gamma \bullet W)$ is a martingale, while γ does not satisfy Novikov's condition (Theorem 4.7).

(a) Let $c > 0$ be some real constant. Show that

$$\mathbb{E}\left[e^{c\int_0^t W(s)^2\, ds}\right] = \begin{cases} \frac{1}{\sqrt{\cos(t\sqrt{2c})}}, & t < \frac{\pi}{2\sqrt{2c}}, \\ \infty, & t \geq \frac{\pi}{2\sqrt{2c}}. \end{cases}$$

Hint: show that $X_1(t) = (\sqrt{x_1} + W(t))^2$ and $X_2(t) = x_2 + \int_0^t X_1(s)\, ds$ define an affine process in $\mathbb{R}_+^2$.

(b) Define the positive local martingale $M = \mathcal{E}(-W \bullet W)$. Prove that M is a martingale, that is, $\mathbb{E}[M(t)] = 1$ for all finite $t \geq 0$. Hint: show that $\int_0^t W(s)\,dW(s) = W(t)^2/2 - t/2$ and use the affine process (X_1, X_2) from part (a).
(c) Conclude that you have just found a $\gamma \in \mathcal{L}$ such that the stochastic exponential $\mathcal{E}(\gamma \bullet W)$ is a true martingale while γ does not satisfy Novikov's condition.
(d) Finally show that, for any finite time horizon T, $d\mathbb{Q}/d\mathbb{P} = M(T)$ defines an equivalent measure such that W has a mean reverting drift under $\mathbb{Q}$:

$$dW(t) = -W(t)\,dt + dW^*(t),$$

where $W^*(t)$ denotes the Girsanov transformed Brownian motion, for $t \leq T$.

Exercise 10.4 The aim of this exercise is to give a direct proof for the validity of (10.13) in the case where $\gamma \in \mathbb{R}_+^m \times \{0\}$. Let $u \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ and $T \in \mathbb{R}_+$.

(a) Along the lines of the proof of Theorem 10.2 show that there exists a unique solution $(\Phi(\cdot, u), \Psi(\cdot, u)) : \mathbb{R}_+ \to \mathbb{C} \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ of (10.12) with

$$\Re(\Phi(t, u)) = -ct.$$

(b) Now argue as in the proof of Theorem 10.1, and show that

$$M(t) = \mathrm{e}^{-\int_0^t r(s)\,ds} \mathrm{e}^{\Phi(T-t,u)+\Psi(T-t,u)^\top X(t)}, \quad t \leq T,$$

is a martingale with $M(T) = \mathrm{e}^{-\int_0^T r(s)\,ds} \mathrm{e}^{u^\top X(T)}$.
(c) Conclude that (10.13) holds.

Exercise 10.5 Complete the proof of Lemma 10.2.

Exercise 10.6 Complete the proof of Corollary 10.4 by deriving the price formulas (10.28).

Exercise 10.7 Derive explicit call and put bond option price formulas for the Vasiček short-rate model and the results from Sect. 7.2.1 using the approach outlined in Sect. 10.3.2.1.

Exercise 10.8 The aim of this exercise is to derive an intuition for the noncentral χ^2-distribution, and its interplay with affine processes. Fix $\delta \in \mathbb{N}$ and some real numbers $\nu_1, \ldots, \nu_\delta$, and define $\zeta = \sum_{i=1}^{\delta} \nu_i^2$.

(a) Let $N_1, \ldots, N_\delta$ be independent standard normal distributed random variables. Define $Z = \sum_{i=1}^{\delta} (N_i + \nu_i)^2$. Show by direct integration that the characteristic function of Z equals

$$\mathbb{E}[\mathrm{e}^{uZ}] = \frac{\mathrm{e}^{\frac{\zeta u}{1-2u}}}{(1-2u)^{\frac{\delta}{2}}}, \quad u \in \mathbb{C}_-.$$

Conclude by Lemma 10.4 that Z is noncentral χ^2-distributed with δ degrees of freedom and noncentrality parameter ζ.

(b) Now let $W_1, \dots, W_\delta$ be independent standard Brownian motions with respect to some filtration $(\mathcal{F}_t)$, and define the process $X(t) = \sum_{i=1}^{\delta}(W_i(t) + \nu_i)^2$, $t \geq 0$. Using (a), show that the $\mathcal{F}_t$-conditional characteristic function of $X(T)$ equals

$$\mathbb{E}[\mathrm{e}^{uX(T)} \mid \mathcal{F}_t] = \frac{\mathrm{e}^{\frac{u}{1-2(T-t)u}X(t)}}{(1-2(T-t)u)^{\frac{\delta}{2}}}, \quad u \in \mathbb{C}_-, \; t < T. \tag{10.61}$$

Conclude that the $\mathcal{F}_t$-conditional distribution of $X(T)$ is noncentral χ^2 with δ degrees of freedom and noncentrality parameter $\frac{X(t)}{T-t}$.

(c) Along the lines of Exercise 5.4, show that X satisfies the stochastic differential equation

$$dX(t) = \delta\, dt + 2\sqrt{X(t)}\, dB(t), \quad X(0) = \zeta,$$

for the Brownian motion $dB = \sum_{i=1}^{\delta} \frac{W_i + \nu_i}{\sqrt{X}}\, dW_i$.

(d) Conclude by either (b) or (c) that X is an affine process with state space $\mathbb{R}_+$.

(e) Find the explicit solutions of the corresponding Riccati equations

$$\partial_t \phi(t,u) = \cdots, \qquad \partial_t \psi(t,u) = \cdots,$$

for X and verify (10.61) by applying the affine transform formula

$$\mathbb{E}[\mathrm{e}^{uX(T)} \mid \mathcal{F}_t] = \mathrm{e}^{\phi(T-t,u)+\psi(T-t,u)X(t)}.$$

Exercise 10.9 Let $b, \sigma > 0$ and $\beta \in \mathbb{R}$, and consider the affine process

$$dX = (b + \beta X)\, dt + \sigma\sqrt{X}\, dW, \quad X(0) = x \in \mathbb{R}_+,$$

with state space $\mathbb{R}_+$.

(a) Use Lemma 10.12 for finding the explicit solutions ϕ and ψ of the corresponding Riccati equations.

(b) Define $C(\tau) = \frac{\sigma^2(\mathrm{e}^{\beta\tau}-1)}{4\beta}$ ($= \frac{\sigma^2\tau}{4}$ if $\beta = 0$) for $\tau > 0$. Check that $C(\tau)$ is positive and increasing in τ. Let $t < T$, and show that the $\mathcal{F}_t$-conditional distribution of $\frac{X(T)}{C(T-t)}$ is noncentral χ^2 with $\frac{4b}{\sigma^2}$ degrees of freedom and noncentrality parameter $\frac{\mathrm{e}^{\beta(T-t)}X(t)}{C(T-t)}$.

(c) Verify the findings of Exercise 10.8(b).

Exercise 10.10 Let $\sigma > 0$ and $\beta \in \mathbb{R}$, and consider the affine process

$$dX = \beta X\, dt + \sigma\sqrt{X}\, dW, \quad X(0) = x \in \mathbb{R}_+,$$

with state space $\mathbb{R}_+$. Define $C(t)$ as in Exercise 10.9(b).

(a) Show that the characteristic function of $Z(t) = \frac{X(t)}{C(t)}$ equals

$$\mathbb{E}[\mathrm{e}^{uZ(t)}] = \exp\left[\frac{\frac{\mathrm{e}^{\beta t}x}{C(t)}u}{1-2u}\right], \quad u \in \mathbb{C}_-. \tag{10.62}$$

This is the characteristic function of the so-called noncentral χ^2 distribution with zero degrees of freedom[14] and noncentrality parameter $\frac{\mathrm{e}^{\beta t}x}{C(t)}$.

(b) Show that

$$\mathbb{P}[X(t)=0] = \exp\left[-\frac{\mathrm{e}^{\beta t}x}{2C(t)}\right] > 0, \quad t > 0.$$

Conclude that the noncentral χ^2 distribution with zero degrees of freedom admits no density on $\mathbb{R}_+$ (hint: take limit $u \to -\infty$ in (10.62)).

(c) Derive the asymptotic result

$$\lim_{t\to\infty} \mathbb{P}[X(t)=0] = \begin{cases} 1, & \beta \le 0, \\ \mathrm{e}^{-\frac{2\beta}{\sigma^2}}, & \beta > 0. \end{cases}$$

(d) Now suppose $\beta = 0$. Show that, for every $T > 0$, there exists some reals M, $p > 0$ such that

$$\mathbb{E}[\mathrm{e}^{pX(t)}] \le M, \quad t \le T.$$

Conclude that X is a true martingale.

Exercise 10.11 Fix a constant interest rate $r > 0$ and $\sigma > 0$, and consider the affine stock model with risk-neutral dynamics ($\mathbb{P} = \mathbb{Q}$)

$$dS = rS\,dt + \sigma\sqrt{S}\,dW, \quad S(0) = s_0 \ge 0. \tag{10.63}$$

(a) Using Exercise 10.10(d), show that the discounted stock price process $\mathrm{e}^{-rt}S(t)$ is a martingale ≥ 0.

(b) Check, by Exercise 10.10, that $\mathbb{P}[S(t)=0] > 0$ and argue that S may be interpreted as defaultable stock price: once the price hits zero, it remains zero (hint: the solution of (10.63) is unique).

(c) Define $C(t) = \frac{\sigma^2(\mathrm{e}^{rt}-1)}{4r}$, and derive from Exercise 10.10 that $\frac{S(t)}{C(t)}$ has a noncentral χ^2 distribution with zero degrees of freedom and parameter of noncentrality $\frac{\mathrm{e}^{rt}s_0}{C(t)}$.

(d) For the parameters $s_0 = 100$, $r = 0.01$ and $\sigma = 4$ derive the European call option prices and implied volatilities, by inverting the Black–Scholes option price formula in Proposition 7.3, as shown in Tables 10.4–10.5 and Fig. 10.5. Show that the risk-neutral default probability is $\mathbb{P}[S(1)=0] = 0.3501 \times 10^{-5}$.

[14] The noncentral χ^2 distribution with zero degrees of freedom has been defined by Siegel [152].

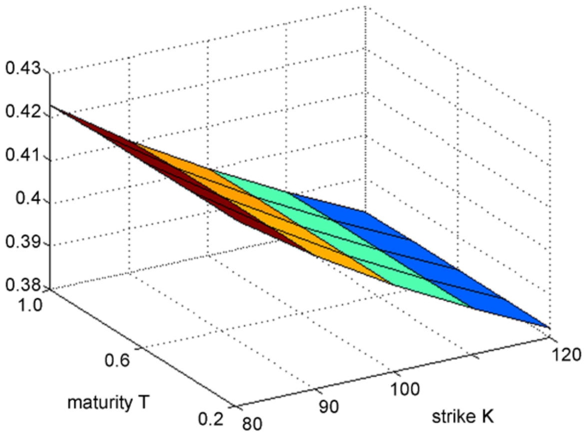
Table 10.4 Call option prices in the affine stock model (10.63)

T–K	80	90	100	110	120
0.2	21.1236	13.1847	7.2226	3.4253	1.3983
0.4	22.9365	15.8324	10.2530	6.2162	3.5274
0.6	24.6300	17.9697	12.5878	8.4644	5.4662
0.8	26.1726	19.8052	14.5570	10.3953	7.2166
1.0	27.5739	21.4231	16.2776	12.1002	8.8050

Table 10.5 Black–Scholes implied volatilities for the affine stock model (10.63)

T–K	80	90	100	110	120
0.2	0.4229	0.4108	0.4001	0.3907	0.3822
0.4	0.4230	0.4109	0.4003	0.3908	0.3823
0.6	0.4232	0.4110	0.4004	0.3909	0.3824
0.8	0.4232	0.4111	0.4004	0.3909	0.3824
1.0	0.4226	0.4106	0.4001	0.3906	0.3821

Fig. 10.5 Implied volatility surface for the affine stock model (10.63)



Exercise 10.12 Let $b \geq 0$, $\beta \in \mathbb{R}$ and $\sigma > 0$, and consider the affine process

$$dX = (b + \beta X)\, dt + \sigma \sqrt{X}\, dW, \qquad X(0) = x_0 > 0,$$

with state space $\mathbb{R}_+$. For any $c \geq 0$ we define the stopping time

$$\tau_c = \inf\{t \geq 0 \mid X(t) = c\} \leq \infty.$$

Thus, $\{\tau_0 = \infty\}$ is the event that X never hits zero. The aim of this exercise is to prove the following claims:

(a) If $b \geq \frac{\sigma^2}{2}$, then $\mathbb{P}[\tau_0 = \infty] = 1$.
(b) If $b < \frac{\sigma^2}{2}$ and $\beta \leq 0$, then $\mathbb{P}[\tau_0 < \infty] = 1$.
(c) If $b < \frac{\sigma^2}{2}$ and $\beta > 0$, then $\mathbb{P}[\tau_0 < \infty] \in (0, 1)$.

Note that Exercise 10.10(b) and (c) is a special case of the claims (b) and (c) (why?).

- Define the function

$$f(x) = \int_1^x \mathrm{e}^{-\frac{2\beta}{\sigma^2} y} y^{-\frac{2b}{\sigma^2}} \, dy, \quad x \geq 0,$$

and show that $f(X)$ is a local martingale (hint: Itô's formula).
- Let $0 < r < x_0 < R$, and define the stopping time $\tau_{r,R} = \tau_r \wedge \tau_R$. Show that

$$f(X(t \wedge \tau_{r,R})) - f(x_0) = \int_0^t f'(X(s))\sigma\sqrt{X(s)} 1_{\{s \leq \tau_{r,R}\}} \, dW(s), \quad t \geq 0.$$

- Taking the second moment on both sides (hint: Itô isometry), derive

$$M_1 \geq \mathbb{E}\left[\left(f(X(t \wedge \tau_{r,R})) - f(x_0)\right)^2\right] \geq M_2 \mathbb{E}[t \wedge \tau_{r,R}],$$

for some real constant $M_1, M_2 > 0$ which do not depend on t (hint: show that $\sigma^2 x f'(x)^2 \geq M_2$ for all $x \geq r$). Conclude that $\mathbb{E}[\tau_{r,R}] < \infty$, and hence $\tau_{r,R} < \infty$ a.s.
- Show that $f(x_0) = \mathbb{E}[f(X(t \wedge \tau_{r,R}))] = \mathbb{E}[f(X(\tau_{r,R}))]$ (hint: show that $f(X(t \wedge \tau_{r,R}))$ is a martingale and use dominated convergence). Derive from this the identity

$$f(x_0) = f(r)\mathbb{P}[\tau_r < \tau_R] + f(R)\mathbb{P}[\tau_r > \tau_R].$$

- Using monotone convergence and the continuity of X, show that $\lim_{r\to 0} \mathbb{P}[\tau_r < \tau_R] = \mathbb{P}[\tau_0 < \tau_R]$, $\lim_{r\to 0} \mathbb{P}[\tau_r > \tau_R] = \mathbb{P}[\tau_0 > \tau_R]$, and $\tau_R \uparrow \infty$ for $R \uparrow \infty$.
- Suppose $b \geq \frac{\sigma^2}{2}$. Show that $\lim_{r\to 0} f(r) = -\infty$, and infer that $\mathbb{P}[\tau_0 = \infty] = 1$. Thus claim (a) is proved.
- Suppose $b < \frac{\sigma^2}{2}$. Show that $f(0) = \lim_{r\to 0} f(r)$ exists in $\mathbb{R}$, and

$$\lim_{R\to\infty} f(R) \begin{cases} = \infty, & \beta \leq 0, \\ \in \mathbb{R}, & \beta > 0. \end{cases}$$

Deduce from this that claims (b) and (c) hold.

Exercise 10.13 For the CIR model with parameters (10.30), compute the at-the-money European call and put option prices $\pi_{call}(t = 0)$ and $\pi_{put}(t = 0)$ on the $(S = 2)$-bond with expiry date $T = 1$, using each of the formulas in (10.28). Compare the results.

Exercise 10.14 Derive the ATM cap prices and Black volatilities in Table 10.1.

Exercise 10.15 Derive (10.31), and check how (10.12) would look like in the Heston stochastic volatility model.

Exercise 10.16 Derive the call option price formulas (10.32) and (10.33) in the Heston model.

Exercise 10.17

(a) Compute the call option prices in Table 10.2, using either formula (10.32) or (10.33) and Lemma 10.12, in the Heston model.
(b) Then derive the corresponding implied volatilities from Table 10.3 by inverting the Black–Scholes option price formula in Proposition 7.3.
(c) Show that the implied volatilities are decreasing with increasing strike price K if the stock price S and the volatility process X_1 have negative covariation $d\langle S, X_1\rangle$, that is if $\rho < 0$.

Exercise 10.18 Find the representation of the affine process X underlying the Heston stochastic volatility model in Sect. 10.3.3 in the form (10.34) in terms of parameters $a, \alpha_1, \alpha_2, b, \mathcal{B} = (\beta_1, \beta_2)$.

(a) Show that $\mathcal{B}$ is singular, and hence cannot have negative eigenvalues (hence X is not strictly mean reverting).
(b) Verify that the diffusion matrix is not of block-diagonal form. Find an invertible 2×2-matrix Λ with $\Lambda(\mathbb{R}_+ \times \mathbb{R}) = \mathbb{R}_+ \times \mathbb{R}$ and an affine process Y on $\mathbb{R}_+ \times \mathbb{R}$ with block-diagonal diffusion matrix such that $X = \Lambda^{-1} Y$. How do you have to adjust the stock price process S to be expressed as a function of Y?

Exercise 10.19 Consider the following multivariate extension of Heston's stochastic volatility model. Interest rates are assumed to be constant $r(t) \equiv r \geq 0$, and there are two risky assets $S_i = \mathrm{e}^{X_{i+1}}$, $i = 1, 2$, where $X = (X_1, X_2, X_3)^\top$ is the affine process with state space $\mathbb{R}_+ \times \mathbb{R}^2$ and dynamics

$$
\begin{aligned}
dX_1 &= (k + \kappa X_1)\,dt + \sigma\sqrt{2X_1}\,dW_1,\\
dX_2 &= (r - \sigma_2 X_1)\,dt + \sigma_2\sqrt{2X_1}\left(\rho_1\,dW_1 + \sqrt{1-\rho_1^2}\,dW_2\right),\\
dX_3 &= (r - \sigma_3 X_1)\,dt + \sigma_2\sqrt{2X_1}\left(\rho_2\,dW_1 + \rho_3\,dW_2 + \sqrt{1-\rho_2^2-\rho_3^2}\,dW_3\right)
\end{aligned}
$$

for some constant parameters $k, \sigma \geq 0$, $\kappa \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, and some $\rho_i \in [-1, 1]$ satisfying $\rho_2^2 + \rho_3^2 \leq 1$.

(a) Verify that this is an arbitrage-free model.
(b) Find and solve the corresponding Riccati equations.
(c) Compute the price of the exchange option with payoff $(c_1 S_1(T) - c_2 S_2(T))^+$ for various parameter specifications.

(d) Compute the price of the spread option with payoff $(S_1(T) - S_2(T) - K)^+$ for various parameter specifications.

Exercise 10.20 Let X be the affine process given in (10.34), and let Λ be a regular $d \times d$-matrix and $\lambda \in \mathbb{R}^d$. Show that the affine transform $Y = \Lambda X + \lambda$ satisfies

$$dY = \left(\Lambda b - \Lambda\beta^\top \Lambda^{-1}\lambda + \Lambda\beta^\top \Lambda^{-1} Y\right) dt + \Lambda\rho\left(\Lambda^{-1}(Y - \lambda)\right) dW.$$

Verify that the drift and diffusion matrix of Y are affine in $Y(t)$.

Exercise 10.21 Derive the corresponding system of Riccati equations (10.7) for the block-diagonal diffusion matrix (10.38).

Exercise 10.22 Derive Corollaries 10.5 and 10.6 as special cases from the invariance Lemma 10.11.

Exercise 10.23 Finish the proof of Lemma 10.14 by showing that:

(a) Λ in (10.54) satisfies $\Lambda(\mathbb{R}^m_+ \times \mathbb{R}^n) = \mathbb{R}^m_+ \times \mathbb{R}^n$.
(b) $\widetilde{a}$, $\widetilde{\alpha_i}$, $\widetilde{b}$, $\widetilde{\mathcal{B}}$, given by (10.56), are admissible and $\widetilde{\alpha_i}$ are block-diagonal, for all $i = 1, \ldots, m$.
(c) $\widetilde{\phi}$ and $\widetilde{\psi}$ in (10.55) satisfy the system of Riccati equations (10.7) with a, α_i, b, and $\mathcal{B} = (\beta_1, \ldots, \beta_d)$ replaced by $\widetilde{a}$, $\widetilde{\alpha_i}$, $\widetilde{b}$, $\widetilde{\mathcal{B}}$.

10.9 Notes

Affine Markov models have been employed in finance for decades, and they have found growing interest due to their computational tractability as well as their capability to capture empirical evidence from financial time series. Their main applications lie in the theory of term-structure of interest rates, stochastic volatility option pricing and the modeling of credit risk (see [61] and the references therein). There is a vast literature on affine models. We mention here explicitly just the few articles [4, 29, 43, 50, 58, 60, 80, 91, 114] and [61] for a broader overview. The generalizations to time-inhomogeneous affine processes have been studied in detail in [71].

A preliminary version of this chapter has been published as a review article [72]. Theorem 10.3(b) and (d) was first proved by Glasserman and Kim [80] for strictly mean reverting affine diffusion processes, which, however, excludes the Heston stochastic volatility model (see Exercise 10.18). The strict mean reversion assumption was subsequently relaxed in [72]. The blow-up property of ψ_I stated in Lemma 10.14(b) is crucial for the proof of Theorem 10.3(b). Its proof is inspired by the line of arguments in [80]. It is yet unclear whether it holds for the class of affine jump-diffusion processes in general. The convexity property of the maximal domain stated in Theorem 10.3(c) represents a non-trivial result for ordinary differential equations. Only in the mid 1990s were corresponding convexity results derived in the analysis literature, see Lakshmikantham et al. [109].

The Fourier transform of an option payoff as shown in Lemma 10.2 was first proposed and utilized for option pricing via fast Fourier transform methods by Carr and Madan [36]. The formula for the exchange option in Corollary 10.3 is obviously related, but seems to be new in the financial literature. The Fourier decomposition of the spread option payoff in Lemma 10.3 has been found and explored by Hurd and Zhou [98]. More examples of payoff functions with explicit Fourier decomposition, including the one in Lemma 10.2, can be found in Hubalek et al. [94].

The classification problem for affine term-structure models raised in Sect. 10.4 has been addressed by the specification analysis in Dai and Singleton [50], which was subsequently extended by Collin-Dufresne et al. [44]. However, it is shown in Cheridito et al. [43] that the Dai–Singleton classification is not exhaustive for state space dimension $d \geq 4$.

The existence and uniqueness proof of affine diffusion processes given in Sect. 10.5 builds on the seminal result by Yamada and Watanabe [162]. We note that this approach is different from the one used in [61], which uses infinite divisibility on the canonical state space and the Markov semigroup theory, thereby asserting existence of weak solutions only.

Section 10.6 is a more elaborated version from the Appendix in [61]. Exercise 10.12 is adapted from [110, Exercise 34].