Affine processes with compact state space[*]

Paul Krühner[†]   Martin Larsson[‡]

Abstract

The behavior of affine processes, which are ubiquitous in a wide range of applications, depends crucially on the choice of state space. We study the case where the state space is compact, and prove in particular that (i) no diffusion is possible; (ii) jumps are possible and enforce a grid-like structure of the state space; (iii) jump components can feed into drift components, but not vice versa. Using our main structural theorem, we classify all bivariate affine processes with compact state space. Unlike the classical case, the characteristic function of an affine process with compact state space may vanish, even in very simple cases.

1 Introduction

Affine processes are ubiquitous in a wide range of applications, in particular in finance, which has motivated a rich literature developing the mathematical theory of affine processes. We refrain from giving a comprehensive overview here; good starting points include [7, 6, 9, 14] and references therein.

Every affine process comes with a state space where it evolves, and the corresponding existence and uniqueness theory depends crucially on the properties of the state space. A complete theory is available for the product space $\mathbb{R}^m \times \mathbb{R}_+^n$ [5] and the convex cone of symmetric positive semidefinite matrices [3], and in the diffusion case for polyhedral and quadratic state spaces [16]. Various other state spaces have also been studied, such as symmetric cones [4]. Our focus in the present paper is on compact state spaces, which so far have not received a systematic treatment.

While slightly different definitions of affine processes exist in the literature, they all have the common feature that affine processes are semimartingales (at least before a

[†]University of Liverpool, Institute of Financial & Actuarial Mathematics, Mathematical Sciences Building, Liverpool, L69 7ZL. E-mail: peisenbe@liverpool.ac.uk
[‡]ETH Zurich, Department of Mathematics, Rämistrasse 101, CH-8092, Zurich, Switzerland.
  E-mail: martin.larsson@math.ethz.ch

killing time in the non-conservative case) whose differential characteristics are affine functions of the current state. Since diffusion coefficients must remain nonnegative on the state space and degenerate on the boundary, state spaces which are cones fit well with the affine structure. A similar remark applies to jump intensities. For compact state spaces, as we will see, the affine structure forces the diffusion coefficient to vanish. Jump intensities may however be nonzero, but only if the state space has a grid-like structure along directions where jumps can occur. In particular, the state space may be a finite discrete set, which after an affine transformation only contains points with integer coordinates. Mathematically, this leads to arguments with a combinatorial rather than analytical flavor.

Classically, affine processes have the property that the conditional Fourier transform is an exponential-affine function of the state (indeed, this is sometimes taken as part of the definition of an affine process). In cases where the state space is a finite set, it turns out that the characteristic function may attain the value zero, which precludes the exponential-affine structure. Instead, the characteristic function is a polynomial, where the state appears *in the exponents*; this leads to a well-defined expression since, up to an affine transformation, the process has integer coordinates.

The paper is organized as follows. In Section 2 we introduce the setup and summarize some of our key results. Then, in Section 3, we prove that affine processes with compact state space cannot have a diffusive component. In Section 4 we discuss the jump (and drift) behavior of affine processes with compact state space. In the final Section 5 we provide several examples and use the results developed so far to identify all possible affine processes with compact state space in $\mathbb{R}$ and $\mathbb{R}^2$. Throughout the paper, $\mathbb{N} = \{0, 1, 2, \ldots\}$ denotes the natural numbers with zero included. For $d \in \mathbb{N}$, $\mathbb{R}^d$ denotes d-dimensional Euclidean space equipped with the usual inner product $\langle \cdot, \cdot \rangle$ and $\mathbb{S}^d$ denotes the set of symmetric $d \times d$ matrices. In particular, we have $\mathbb{R}^0 = \{0\}$. The Dirac measure at a point $x \in \mathbb{R}^d$ is denoted by δ_x. Further unexplained notation follows [11].

2 Setup and main results

Fix a measurable space $(\Omega, \mathscr{F})$ equipped with a right-continuous filtration $\mathbb{F} = (\mathscr{F}_t)_{t \geq 0}$. Let E be a non-empty measurable subset of $\mathbb{R}^d$, $d \geq 1$, and let $X = (X_t)_{t \geq 0}$ be a càdlàg adapted process taking values in E. We set $X_{0-} = X_0$ by convention. Finally, let $(\mathbb{P}_x)_{x \in E}$ be a family of probability measures on $\mathscr{F}$ such that for each $x \in E$,

$$X \text{ is a } \mathbb{P}_x\text{-semimartingale with } \mathbb{P}_x(X_0 = x) = 1.$$

We work with the following notion of an affine process, where $\mathfrak{M}^d_*$ denotes the vector space of all signed Radon measures on $\mathbb{R}^d \setminus \{0\}$.[1]

Definition 2.1. *We say that X is* affine *if its differential characteristics are affine in the following sense: There exist affine maps*

$$b \colon \mathbb{R}^d \to \mathbb{R}^d, \qquad c \colon \mathbb{R}^d \to \mathbb{S}^d, \qquad F \colon \mathbb{R}^d \to \mathfrak{M}^d_*$$

such that for each $x \in E$, $c(x)$ is positive semidefinite, $F(x, d\xi)$ is a positive measure on $\mathbb{R}^d \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^d \setminus \{0\}} (\|\xi\|^2 \wedge 1) F(x, d\xi) < \infty, \tag{2.1}$$

---

[1]That is, $\mathfrak{M}^d_*$ is the vector space of all set functions of the form $\mu = \mu_+ - \mu_-$, where μ_+ and μ_- are positive measures on $\mathbb{R}^d \setminus \{0\}$ that assign finite mass to compact subsets.

and the characteristics[2] (B, C, ν) of X under $\mathbb{P}_x$ are given by

$$B_t = \int_0^t b(X_{s-})ds,$$

$$C_t = \int_0^t c(X_{s-})ds,$$

$$\nu(dt, d\xi) = F(X_{t-}, d\xi)dt.$$

We refer to (b, c, F) as the triplet of the affine process X.

We are interested in the following condition on the state space:

The state space E is compact.

For notational simplicity, we assume throughout this paper that the affine span of E is all of $\mathbb{R}^d$. This does not restrict generality; see Remark 2.3 in [13] for a discussion of this point.

Remark 2.2. The choice of truncation function does not affect Definition 2.1. Changing to a different truncation function yields a different function $b(x)$, but does not affect the affine property. Note also the abuse of terminology: We refer to the process X as affine, although this is rather a property of the family $(\mathbb{P}_x)_{x \in E}$.

Our convention $X_{0-} = X_0$ implies that the differential characteristics $b(X_{t-})$, $c(X_{t-})$ and $F(X_{t-}, d\xi)$ are $\mathbb{P}_x$-almost surely equal to $b(x)$, $c(x)$, and $F(x, d\xi)$, respectively, when $t = 0$. By continuity of b, c, and F, the differential characteristics are therefore right-continuous at $t = 0$, which will be important when we apply the nonnegativity result Lemma B.1.

The following is our main result regarding the structure of affine processes on compact state spaces.

Theorem 2.3. Let X be affine with triplet (b, c, F), and suppose the state space E is compact. Then there is no diffusion, $c = 0$, and there exist an invertible affine map $T \colon \mathbb{R}^d \to \mathbb{R}^d$ as well as $k \in \mathbb{N}$ such that the following conditions hold:

(i) $T(E) \subseteq \mathbb{N}^k \times \mathbb{R}^{d-k}$,

(ii) $Y = (T(X)_1, \ldots, T(X)_k)$ is affine and Markov,

(iii) $Z = (T(X)_{k+1}, \ldots, T(X)_d)$ can only jump when Y jumps, that is,

$$\{t \geq 0 \colon \Delta Z(t) \neq 0\} \subseteq \{t \geq 0 \colon \Delta Y(t) \neq 0\}, \tag{2.2}$$

and its jump characteristic is of the form $\nu^Z(dt, d\zeta) = F^Z(Y_{t-}, d\zeta)dt$ for some affine map $F^Z \colon \mathbb{R}^k \to \mathfrak{M}_*^{d-k}$.

(iv) the canonical coordinate projections $\pi_j \colon (x_1, \ldots, x_d) \mapsto x_j$ are normalized jump counters of the transformed process (Y, Z) for $j = 1, \ldots, k$.[3]

Remark 2.4. Observe that k may be zero in Theorem 2.3. Then Y is trivial since it takes values in $\mathbb{R}^0 = \{0\}$, and (2.2) implies that Z does not jump. Thus X is in fact the (deterministic) solution of a linear ordinary differential equation.

Theorem 2.3 combines the conclusions of Theorems 3.1 and 4.7, whose statements and proofs are given in Sections 3 and 4, respectively. The proofs yield in some respects more detailed information than the theorem itself; see for instance Lemma 4.2 and Proposition 4.5. The usefulness of this added detail is illustrated in Section 5, where we

---

[2]Characteristics of semimartingales are defined in [11, Definition II.2.6]

[3]Normalized jump counters are introduced in Definition 4.1 below.

discuss examples and classify all affine processes on compact state spaces in dimensions $d = 1, 2$. In particular, Theorem 5.5, which treats the case where $E \subseteq \mathbb{N}^2$ is a finite set, has the following corollary:

Theorem 2.5. *Let $X = (X_1, X_2)$ be a 2-dimensional affine process with finite and irreducible state space.*[4] *Assume X has no autonomous components in the sense that no nonzero linear combination $a_1 X_1 + a_2 X_2$ of the components of X is itself an affine process. Then, up to an invertible affine transformation, one has $E = \{x \in \mathbb{N}^2 : x_1 + x_2 \leq N\}$ for some $N \in \mathbb{N}$, and*

$$
\begin{aligned}
F(x, \cdot) &= x_1 \left(\lambda_1 \delta_{(-1,0)} + \lambda_2 \delta_{(-1,1)} \right) + x_2 \left(\lambda_3 \delta_{(0,-1)} + \lambda_4 \delta_{(1,-1)} \right) \\
&\quad + (N - x_1 - x_2) \left(\lambda_5 \delta_{(1,0)} + \lambda_6 \delta_{(0,1)} \right)
\end{aligned}
$$

for some $\lambda_1, \ldots, \lambda_6 \in \mathbb{R}_+$.

The characteristic function of X_t in Theorem 2.5 is given by

$$
\mathbb{E}_x[e^{\mathbf{i} \langle u, X_t \rangle}] = \Phi(u, t) \Psi_1(u, t)^{x_1} \Psi_2(u, t)^{x_2},
$$

where $\Phi(u, t)$ and $\Psi_i(u, t)$, $i = 1, 2$, are the solutions of the Riccati equations

$$
\begin{aligned}
\partial_t \Phi &= N \lambda_5 (\Psi_1 - 1) \Phi + N \lambda_6 (\Psi_2 - 1) \Phi, & \Phi(u, 0) &= 1, \\
\partial_t \Psi_1 &= \lambda_1 - (\lambda_1 + \lambda_2 - \lambda_5 - \lambda_6) \Psi_1 + \lambda_2 \Psi_2 - \lambda_6 \Psi_1 \Psi_2 - \lambda_5 \Psi_1^2, & \Psi_1(u, 0) &= e^{\mathbf{i} u_1}, \\
\partial_t \Psi_2 &= \lambda_3 - (\lambda_3 + \lambda_4 - \lambda_5 - \lambda_6) \Psi_1 + \lambda_4 \Psi_1 - \lambda_5 \Psi_1 \Psi_2 - \lambda_6 \Psi_2^2, & \Psi_2(u, 0) &= e^{\mathbf{i} u_2}.
\end{aligned}
$$

This follows from the fact that $M_u(t) = \Phi(u, T - t) \Psi_1(u, T - t)^{X_1(t)} \Psi_2(u, T - t)^{X_2(t)}$ defines a martingale for any $T \geq 0$, as can be seen by applying Itô's formula along with the definition of the characteristics of X. Observe that since $E \subset \mathbb{N}^2$, only integer powers appear in the above expressions.

Compare this to the classical case, where the $\Psi_i(u, t)$ are of exponential form and it is their exponents that satisfy (generalized) Riccati equations; see for instance [5, Theorem 2.7]. In particular, the $\Psi_i(u, t)$ are then necessarily nonzero. [2] also works with $\Psi_i(u, t)$ of exponential form, but allows $\Phi(u, t)$, and hence the characteristic function of X_t, to become zero; see equation (1.4) and the subsequent remark in [2]. Our situation with finite state space is different, and it turns out that the $\Psi_i(u, t)$ can in fact reach zero for certain arguments; see Proposition 5.1 and the subsequent discussion.

We end this section by noting that any finite state Markov chain can be viewed as an affine process as follows. Let d be the number of states of the Markov chain, let q_{ij} be the intensity of transitioning from state i to state j, and let $E = \{e_1, \ldots, e_d\} \subset \mathbb{R}^d$ consist of the canonical unit vectors in $\mathbb{R}^d$. The affine process with state space E and triplet $(0, 0, F)$ with $F(x, d\xi) = \sum_{i,j=1}^d e_i^\top x \, q_{ij} \delta_{e_j - e_i}(d\xi)$ then has the same law as the original Markov chain, under the identification of its state space with E. This construction leads to an affine process of potentially very large dimension d. If one instead considers a fixed d, only some finite-state Markov chains can be viewed as d-dimensional affine processes. If $d = 1$, the only such Markov chain is the birth–death process; see Proposition 5.1. We thank an anonymous referee for bringing our attention to these observations.

3 Diffusion

In this section we prove that an affine process with compact state space necessarily has no diffusion.

---

[4]That X has irreducible state space means that for every $x \in E$, X has positive $\mathbb{P}_x$-probability of eventually reaching any other state $y \in E$.

Theorem 3.1. *Let X be affine with triplet (b, c, F), and suppose the state space E is compact. Then there is no diffusion, i.e., $c = 0$.*

The proof occupies the remainder of this section, and we start with two auxiliary results. The first is an expression of the intuitive notion that there can be no diffusion perpendicularly to the boundary of E.

Lemma 3.2. *Let the assumptions of Theorem 3.1 be in force. Let $u \in \mathbb{R}^d$, $\overline{x} \in E$, and $\langle u, \overline{x} \rangle = \max_{x \in E} \langle u, x \rangle$. Then $c(\overline{x})u = 0$. Note that the set of possible maximizers $\overline{x}$ depends on the choice of u.*

Proof. Define $f(x) = \langle u, \overline{x} - x \rangle$. Then under $\mathbb{P}_{\overline{x}}$ the process $Z = f(X)$ is a nonnegative semimartingale with $Z_0 = 0$. Its second differential characteristic is [12, Proposition B.1]

$$\widetilde{c}_t = \langle \nabla f(X_{t-}), c(X_{t-}) \nabla f(X_{t-}) \rangle = \langle u, c(X_{t-})u \rangle, \qquad \widetilde{c}_0 = \langle u, c(\overline{x})u \rangle.$$

Lemma B.1 implies $\widetilde{c}_0 = 0$, and hence $c(\overline{x})u = 0$. $\qquad \square$

Let $\mathrm{conv}(E)$ denote the convex hull of E, which is again compact. Since the affine span of E, and hence of $\mathrm{conv}(E)$, is all of $\mathbb{R}^d$, we have

$$\mathrm{int}\,\mathrm{conv}(E) \neq \emptyset.$$

The next result is an application of the [1] theorem, allowing us to replace E by $\mathrm{conv}(E)$ in Lemma 3.2. By definition, this means that c is parallel to $\mathrm{conv}(E)$; see (A.1).

Lemma 3.3. *Let the assumptions of Theorem 3.1 be in force. Then c is parallel to $\mathrm{conv}(E)$.*

Proof. Pick any nonzero $\overline{x} \in \mathrm{conv}(E)$ and any $u \in N_{\mathrm{conv}(E)}(\overline{x})$. This means that $\langle u, \overline{x} \rangle = \max_{x \in \mathrm{conv}(E)} \langle u, x \rangle$. We must show that $c(\overline{x})u = 0$. By Carathéodory's theorem [15, Theorem 17.1], $\overline{x}$ can be expressed as a convex combination of $d + 1$ points in E. Thus there exist $k \leq d + 1$, $x_1, \dots, x_k \in E$, and $\lambda_1, \dots, \lambda_k \in (0, 1)$ such that $\overline{x} = \sum_{i=1}^{k} \lambda_i x_i$. Since

$$\max_i \langle u, x_i \rangle \geq \sum_{i=1}^{k} \lambda_i \langle u, x_i \rangle = \langle u, \overline{x} \rangle = \max_{x \in K} \langle u, x \rangle \geq \max_{x \in E} \langle u, x \rangle \geq \max_i \langle u, x_i \rangle,$$

and since the λ_i are strictly positive, one has $\langle u, x_i \rangle = \max_{x \in E} \langle u, x \rangle$ for all i. Thus Lemma 3.2 yields $c(x_i)u = 0$ for all i, whence, on taking convex combinations, $c(\overline{x})u = 0$. $\qquad \square$

The following result is the key ingredient in the proof of Theorem 3.1. Its proof relies on notions and results from convex analysis that are developed in Section A.

Proposition 3.4. *Let K be a compact convex subset of $\mathbb{R}^d$ with $\mathrm{int}\,K \neq \emptyset$. Let $c : \mathbb{R}^d \to \mathbb{S}^d$ be an affine map parallel to K with $c(x)$ positive semidefinite for all $x \in K$. Then $c = 0$.*

For the proof of this proposition it is convenient to adopt the following coordinate-free notation: For vector spaces V and W, we let $S(V)$ denote the space of symmetric linear maps $V \to V$, and $L(V, W)$ the space of linear maps $V \to W$. Thus $S(V) \simeq \mathbb{S}^n$ and $L(V, W) \simeq \mathbb{R}^{m \times n}$, where $n = \dim V$ and $m = \dim W$.

Proof. After applying a translation, we may suppose that $0 \in \mathrm{int}\,K$. Then $c(x)$ is of the form $c(x) = c_0 + \ell(x)$ for some $c_0 \in \mathbb{S}^d_+$ and some linear map $\ell : \mathbb{R}^d \to \mathbb{S}^d$. Consider the orthogonal direct sum decomposition

$$\mathbb{R}^d = V \oplus W,$$

where $W = \ker c_0$ and $V = W^\perp$ is the range of c_0. With respect to this decomposition, $c(x)$ takes the form

$$c(x) = c_0 + L(x) = \begin{pmatrix} a_0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \ell_{11}(x) & \ell_{12}(x) \\ \ell_{21}(x) & \ell_{22}(x) \end{pmatrix},$$

where $a_0 \in S(V)$ is positive definite and the maps $\ell_{11} : \mathbb{R}^d \to S(V)$, $\ell_{12} : \mathbb{R}^d \to L(W, V)$, and $\ell_{22} : \mathbb{R}^d \to S(W)$ are all linear.

Consider any $w \in W$. The map $x \mapsto \langle w, c(x)w \rangle = \langle w, \ell_{22}(x)w \rangle$ is linear and nonnegative in some open ball $B \subset K$ around the origin, and is thus identically zero. Since $c(x)$, and hence $\ell_{22}(x)$, is positive semidefinite for all $x \in B$, we deduce $\ell_{22}(x) = 0$ for all $x \in \mathbb{R}^d$. Again by positive semidefiniteness of $c(x)$ for $x \in B$, this implies $\ell_{12}(x) = 0$ for all $x \in B$, and hence for all $x \in \mathbb{R}^d$.

We now show that $\dim V = 0$; this will imply that $c = \ell_{22} = 0$ and thus complete the proof. To this end, define

$$\widehat{K} = V \cap K,$$

a compact convex subset of V with $0 \in \text{int}\, \widehat{K}$, where the interior is to be understood relative to V. Consider also the linear map

$$\widehat{c} : V \to S(V), \qquad y \mapsto a_0 + \ell_{11}(y).$$

We claim that $\widehat{c}$ is parallel to $\widehat{K}$. To see this, pick any $y \in \widehat{K}$, $v \in N_{\widehat{K}}(y)$. Lemma A.1 then yields $v + w \in N_K(y)$ for some $w \in W$. Consequently, using that $c(x)w = 0$ and that c is parallel to K by assumption,

$$\widehat{c}(y)v = c(y)(v + w) = 0.$$

Thus $\widehat{c}$ is parallel to $\widehat{K}$ as claimed. Furthermore, $\widehat{c}$ is affine, positive semidefinite on $\widehat{K}$, and $\widehat{c}(0) = a_0$ is invertible. Since $\widehat{K}$ is compact, Lemma A.3 implies $\dim V = 0$, as required. $\qquad\square$

The proof of Theorem 3.1 is now straightforward.

Proof of Theorem 3.1. Since c is affine, the positive semidefiniteness of $c(x)$ for $x \in E$ in fact holds for all $x \in \text{conv}(E)$. Moreover, c is parallel to $\text{conv}(E)$ by Lemma 3.3. The result now follows from Proposition 3.4 with $K = \text{conv}(E)$. $\qquad\square$

4 Jumps and drift

Assume that X is affine with triplet (b, c, F) and compact state space $E \subset \mathbb{R}^d$ whose affine span is all of $\mathbb{R}^d$. Since $c = 0$ by Theorem 3.1 we are left with $(b, 0, F)$, and our goal is to describe its structure. This will be done through a sequence of intermediate results culminating with Theorem 4.7 below.

It is convenient to introduce the set

$$S = \bigcup_{x \in E} \text{supp}(F(x, \cdot)) \setminus \{0\}, \tag{4.1}$$

which can be thought of as the collection of all possible jump sizes of X. If X does not jump at all, then $S = \emptyset$.

A key property of F, beyond its affine structure, is that $F(x, A) = 0$ whenever $x \in E$ and $A \cap (E - x) = \emptyset$, which holds because $F(X_{t-}, d\xi)dt$ is the jump characteristic of X. A more useful way to phrase this property is

$$x \in E, A \subseteq \mathbb{R}^d \text{ open}, F(x, A) > 0 \implies x + \xi \in E \text{ for some } \xi \in A,$$

which we will use repeatedly without explicit mentioning. Together with compactness of E this leads to the existence of *jump counters*, which we now define.

Definition 4.1. *A* jump counter *corresponding to $u \in S$ is an affine function $\psi_u : \mathbb{R}^d \to \mathbb{R}$, not identically zero, such that $\psi_u \geq 0$ on E and*

$$x \in E \text{ and } \psi_u(x) > 0 \quad \Longrightarrow \quad x + u \in E.$$

We call ψ_u normalized if $\psi_u(u) - \psi_u(0) = -1$.

Jump counters are useful because they have rather strong implications for the structure of E. To see why, let ψ_u be a jump counter corresponding to some vector $u \in S$. Each point $x \in E$ satisfies either $\psi_u(x) = 0$ or $\psi_u(x) > 0$. In the latter case, $x + u$ again lies in E. Then by the same token, either $\psi_u(x + u) = 0$, or $x + 2u \in E$. Iterating this argument and observing that the procedure must terminate in finitely many steps since E is compact, we obtain, for every $x \in E$,

$$\{x, x + u, \ldots, x + nu\} \subseteq E \quad \text{and} \quad \psi_u(x + nu) = 0 \quad \text{for some } n \in \mathbb{N}. \tag{4.2}$$

Since ψ_u is not identically zero and E affinely spans $\mathbb{R}^d$, we have $\psi_u(x) > 0$ for some $x \in E$. Thus (4.2) implies that ψ_u is strictly decreasing in direction u, so that $\psi_u(0) - \psi_u(u) > 0$. Therefore, we can always replace ψ_u by $\psi_u/(\psi_u(0) - \psi_u(u))$ to obtain a normalized jump counter.

Lemma 4.2. *Any normalized jump counter ψ_u corresponding to some $u \in S$ satisfies*

  (i) $\psi_u(E) \subseteq \mathbb{N}$,
 (ii) $\psi_u(x) = 1$ *for some $x \in E$,*
(iii) $\ker \psi_u = \mathrm{aff}\{x_1, \ldots, x_d\}$ *for some $x_1, \ldots, x_d \in E$.*

Moreover, if ϕ_u is any other normalized jump counter corresponding to u, then $\phi_u = \psi_u$.

Proof. For any $x \in E$ and with $n \in \mathbb{N}$ as in (4.2) we have

$$\psi_u(x) = \psi_u(x) - \psi_u(x + nu) = \psi_u(0) - \psi_u(nu) = n(\psi_u(0) - \psi_u(u)) = n,$$

where the affine property of ψ_u is used in the second and third equalities. This yields (i). Next, with $x \in E$ such that $\psi_u(x) > 0$, a similar calculation gives $\psi_u(x + (n-1)u) = 1$ which proves (ii). We now argue (iii). Since ψ_u is strictly decreasing along u, the intersection $(x + \mathbb{R}u) \cap \ker \psi_u$ contains exactly one element, where we write $x + \mathbb{R}u = \{x + tu : t \in \mathbb{R}\}$. The projection along u onto $\ker \psi_u$ is therefore given by

$$\pi : \mathbb{R}^d \to \ker \psi_u, \quad x \mapsto (x + \mathbb{R}u) \cap \ker \psi_u,$$

where we identify the intersection on the right-hand side with the single element it contains. In particular, for any $x \in E$ and with $n \in \mathbb{N}$ as in (4.2) we have $\pi(x) = x + nu \in E$ for some $n \in \mathbb{N}$, so that $\pi(E) \subseteq E$. Since the affine span of E is all of $\mathbb{R}^d$, the image $\pi(E)$ affinely spans $\ker \psi_u$. Thus E contains d points $x_1, \ldots, x_d$ whose affine span is $\ker \psi_u$, proving (iii).

It remains to prove the uniqueness statement. For any $x \in E \cap \ker \psi_u$ we have $\psi_u(x + u) = \psi_u(x + u) - \psi_u(x) = -1$, and hence $x + u \notin E$. Therefore $\phi_u(x) = 0$ by the definition of a jump counter. Letting $x_1, \ldots, x_d \in E$ affinely span $\ker \psi_u$, we obtain

$$\ker \psi_u = \mathrm{aff}\{x_1, \ldots, x_d\} \subseteq \ker \phi_u.$$

Thus there exists a constant λ such that $\phi_u = \lambda \psi_u$, and since both ψ_u and ϕ_u are normalized we get $1 = \phi_u(0) - \phi_u(u) = \lambda(\psi_u(0) - \psi_u(u)) = \lambda$. $\qquad \square$

In Proposition 4.5 below we obtain the existence of normalized jump counters and show that they have additional properties. The proof uses the following two lemmas.

Lemma 4.3. *Let $\mathscr{W}^{d-1}$ denote the set of all affine subspaces $W \subseteq \mathbb{R}^d$ with $\dim W \leq d-1$. Then*

$$\inf_{W \in \mathscr{W}^{d-1}} \sup_{x \in E} d(x, W) > 0,$$

where $d(x, W) = \inf\{\|x - y\|\colon y \in W\}$ is the distance from x to W.

Proof. For any $W \in \mathscr{W}^{d-1}$ and $x \in \mathbb{R}^d$ there exist $y \in W$ and $u \in (W - y)^\perp$ with $\|u\| = 1$ such that $d(x, W) = |\langle u, x - y \rangle|$. Thus

$$\inf_{W \in \mathscr{W}^{d-1}} \sup_{x \in E} d(x, W) = \inf\left\{ \sup_{x \in E} |\langle u, x - y \rangle|\colon (u, y) \in \mathscr{S}^{d-1} \times \mathbb{R}^d \right\},$$

where $\mathscr{S}^{d-1}$ is the unit sphere in $\mathbb{R}^d$. Since E is compact, the map $(u, y) \mapsto \sup_{x \in E} |\langle u, x - y \rangle|$ is continuous, and it suffices to let y range over a compact subset $K \subseteq \mathbb{R}^d$. Thus the infimum is attained for some $(\overline{u}, \overline{y}) \in \mathscr{S}^{d-1} \times K$, so that

$$\inf_{W \in \mathscr{W}^{d-1}} \sup_{x \in E} d(x, W) = \sup_{x \in E} |\langle \overline{u}, x - \overline{y} \rangle|.$$

The right-hand side is strictly positive since the affine span of E is all of $\mathbb{R}^d$. $\qquad \square$

Lemma 4.4. *For every $u \in S$ there exist a constant $\varepsilon > 0$ and a probability measure λ on $\mathbb{R}^d$ such that*

$$F(\cdot, A \cap B_\varepsilon(u)) = \lambda(A) F(\cdot, B_\varepsilon(u)) \quad \text{for all measurable } A \subseteq \mathbb{R}^d, \tag{4.3}$$

where $B_\varepsilon(u)$ denotes the open ball with radius ε centered at u. Furthermore, if $x \in E$ and $F(x, B_\varepsilon(u)) > 0$, then $x + u \in E$. In particular, $\phi_u = F(\cdot, B_\varepsilon(u))$ is a jump counter.

Proof. Define the diameter of E by $\operatorname{diam}(E) = \sup\{\|x - y\|\colon x, y \in E\} > 0$. Let $N \in \mathbb{N}$ such that $N > (\operatorname{diam}(E) + 1)/\|u\|$ and pick $\delta > 0$ such that

$$\delta < \inf_{W \in \mathscr{W}^{d-1}} \sup_{x \in E} d(x, W),$$

where $\mathscr{W}^{d-1}$ denotes the set of all affine subspaces $W \subseteq \mathbb{R}^d$ with $\dim W \leq d - 1$; this is possible by Lemma 4.3. Set

$$\varepsilon = \frac{\delta \wedge 1}{N} \wedge \frac{\|u\|}{2} > 0,$$

and define the affine functions $\phi_u = F(\cdot, B_\varepsilon(u))$ and $\phi_A = F(\cdot, A \cap B_\varepsilon(u))$ for any measurable subset $A \subseteq \mathbb{R}^d$. These functions are finite-valued since $B_\varepsilon(u)$ is bounded away from the origin. Consider the affine subspace

$$V = \ker \phi_u \cap \ker \phi_A.$$

We claim that

$$V \neq \emptyset \text{ and } V \neq \mathbb{R}^d. \tag{4.4}$$

To see this, note that $u \in \operatorname{supp} F(x_0, \cdot)$ for some $x_0 \in E$, so that $\phi_u(x_0) = F(x_0, B_\varepsilon(u)) > 0$. Thus there exists $x_1 \in E$ with $x_1 - x_0 \in B_\varepsilon(u)$. Then, recursively, if $x_j \in E$ satisfies $\phi_u(x_j) > 0$, we find $x_{j+1} \in E$ with $x_{j+1} - x_j \in B_\varepsilon(u)$. Note that $\|x_j - x_0\| \geq j(\|u\| - \varepsilon) \geq j\|u\|/2$. Therefore by compactness of E there is a maximal j such that $x_j \in E$ and $\phi_u(x_j) > 0$, and for this j we have $\phi_u(x_{j+1}) \leq 0$. But $F(x, \cdot)$ is a positive measure for all $x \in E$, so we deduce $0 \leq \phi_A(x_{j+1}) \leq \phi_u(x_{j+1}) \leq 0$. This completes the proof of (4.4).

Next, we claim that

$$\ker \phi_u \subseteq \ker \phi_A. \tag{4.5}$$

If $\phi_A \equiv 0$, this certainly holds. Otherwise (4.4) implies $\dim V \le d - 1$ and, in case of equality, $\ker \phi_u = \ker \phi_A$. It remains to exclude the possibility that $\dim V \le d - 2$, so we assume for contradiction that this holds. Then the affine subspace

$$W = V + \mathbb{R}u$$

satisfies $\dim W \le d-1$. By definition of δ, we can then find $x_0 \in E$ such that $d(x_0, W) \ge \delta$. In particular $x_0 \notin W$, whence either $\phi_u(x_0) > 0$ or $\phi_A(x_0) > 0$ (or both). Thus there exists $x_1 \in E$ with $x_1 - x_0 \in B_\varepsilon(u)$, and hence

$$d(x_1, W) = d(x_1 - u, W) \ge d(x_0, W) - d(x_0, x_1 - u) \ge \delta - \varepsilon \ge \delta\left(1 - \frac{1}{N}\right).$$

Again we proceed recursively: If $x_j \in E$ satisfies $d(x_j, W) \ge \delta(1 - j/N) > 0$, we find $x_{j+1} \in E$ with $x_{j+1} - x_j \in B_\varepsilon(u)$ and $d(x_{j+1}, W) \ge \delta(1 - (j+1)/N)$. Consequently,

$$\|x_N - x_0\| = \left\|Nu + \sum_{j=0}^{N-1} (x_{j+1} - x_j - u)\right\| \ge N\|u\| - N\varepsilon > \operatorname{diam}(E),$$

a contradiction. Thus (4.5) is proved.

Next, (4.5) implies that $\phi_A = \lambda(A)\phi_u$ for some constant $\lambda(A)$ that depends on A, which proves (4.3). The fact that λ is a probability measure follows by inspecting (4.3) at a point $x \in E$ for which $F(x, B_\varepsilon(u)) = \phi_u(x) > 0$.

Finally, to prove the last statement, consider $x \in E$ such that $F(x, B_\varepsilon(u)) > 0$. Define $A_n = B_{n^{-1} \wedge \varepsilon}(u)$, so that

$$F(\cdot, A_n) = \lambda(A_n) F(\cdot, B_\varepsilon(u)) \quad \text{for all large } n.$$

Since $u \in S$, the left-hand side is not identically zero, and so $\lambda(A_n) > 0$ for all large n. Evaluating at x then yields $F(x, A_n) > 0$, and thus there exist $u_n \in A_n$ such that $x + u_n \in E$. Since $u_n \to u$ and E is closed, it follows that $x + u \in E$ as claimed. $\qquad\square$

The following proposition shows that each jump size $u \in S$ admits a normalized jump counter with additional properties. Furthermore, the proposition gives information on how jump counters ψ_u and ψ_v corresponding to different $u, v \in S$ interact. Informally, $\psi_u(x)$ counts the number of jumps of size u the process X can perform, starting from $x \in E$, until it reaches $\ker \psi_u$. Now, if $\psi_u(x + v) < \psi_u(x)$, then a jump of size v will bring X closer to $\ker \psi_u$, and the proposition shows that then, in fact, $\psi_u = \psi_v$. If instead $\psi_u(x + v) = \psi_u(x)$, then the jump v is parallel to $\ker \psi_u$. This implies that $\ker \psi_u$ and $\ker \psi_v$ have a nonempty intersection, where jumps of size u or v do not occur. Finally, if $\psi_u(x + v) > \psi_u(x)$ then a jump of size v moves X farther away from $\ker \psi_u$, and in this case it turns out that $v = -u$.

Proposition 4.5. *Every $u \in S$ admits a unique normalized jump counter ψ_u. The jump counter satisfies*

$$F(x, A \cap B_{\varepsilon_u}(u)) = \lambda_u(A)\psi_u(x) \quad \text{for all } x \in E \text{ and all measurable } A \subseteq \mathbb{R}^d, \tag{4.6}$$

where λ_u is a finite measure on $\mathbb{R}^d$ and $\varepsilon_u > 0$ is a constant. Moreover, for any $u, v \in S$ and setting $\alpha = \psi_u(v) - \psi_u(0)$ and $\beta = \psi_v(u) - \psi_v(0)$, one of the following conditions holds:

(i) $\alpha = \beta = -1$ and $\psi_u = \psi_v$,

(ii) $\alpha = \beta = 1$ *and* $u = -v$,

(iii) *one of* α *and* β *equals zero, and* $\alpha, \beta \in \mathbb{N}$.

Proof. Existence of a jump counter ψ_u satisfying (4.6) follows from Lemma 4.4. Since (4.6) is preserved after positive scaling of ψ_u, we may assume that ψ_u is normalized as discussed before Lemma 4.2. This also yields uniqueness.

Consider now $u, v \in S$ with jump counters ψ_u, ψ_v and α, β as stated. There is some $x \in E$ with $\psi_v(x) > 0$ and hence $x + v \in E$. Thus $\alpha = \psi_u(v) - \psi_u(0) = \psi_u(x+v) - \psi_u(x) \in \mathbb{Z}$ due to Lemma 4.2(i), and similarly $\beta \in \mathbb{Z}$ as well. We proceed by examining the possible values of α and β.

Case 1: $\alpha < 0$ or $\beta < 0$. Suppose $\alpha < 0$. Lemma 4.2(iii) yields points $x_1, \ldots, x_d \in E$ that affinely span $\ker \psi_u$. Moreover,

$$\psi_u(x_j + v) = \psi_u(x_j + v) - \psi_u(x_j) = \psi_u(v) - \psi_u(0) = \alpha < 0,$$

and hence $x_j + v \notin E$. We conclude that $\psi_v(x_j) = 0$ for $j = 1, \ldots, d$, whence $x_1, \ldots, x_d$ affinely span $\ker \psi_v$. Thus the kernels of ψ_u and ψ_v coincide, so that $\psi_v = \lambda \psi_u$ for some constant λ. Observing that

$$-1 = \psi_v(v) - \psi_v(0) = \lambda(\psi_u(v) - \psi_u(0)) = \lambda \alpha$$

we have $\lambda = -1/\alpha$, and therefore

$$-1 = \psi_u(u) - \psi_u(0) = -\alpha(\psi_v(u) - \psi_v(0)) = -\alpha\beta.$$

Consequently $\alpha = \beta = -1$ and $\psi_u = \psi_v$, so that (i) holds. The same argument yields the conclusion when $\beta < 0$.

Case 2: $\alpha > 0$ and $\beta > 0$. From Lemma 4.2(ii) there is $x \in E$ such that $\psi_v(x) = 1$. Define

$$J = \{j \in \mathbb{N} \colon x + j(u+v) \in E \text{ and } \psi_v(x + j(u+v)) > 0\}.$$

We claim that $J = \mathbb{N}$. To see this, first note that $0 \in J$ since $\psi_v(x) > 0$. Moreover, for any $j \in J$ we have $x + ju + (j+1)v \in E$ and thus $\psi_u(x + ju + (j+1)v) = \psi_u(x + j(u+v)) + \alpha > 0$. Consequently, $x + (j+1)(u+v) \in E$ and $\psi_v(x + (j+1)(u+v)) = \psi_v(x + ju + (j+1)v) + \beta > 0$. That is, $j + 1 \in J$ and hence $J = \mathbb{N}$ by induction. Compactness of E then forces $u + v = 0$, thus $\alpha = \psi_u(-u) - \psi_u(0) = \psi_u(0) - \psi_u(u) = 1$, and similarly $\beta = 1$. Thus (ii) holds.

Case 3: $\alpha = 0$ and $\beta \geq 0$, or $\alpha \geq 0$ and $\beta = 0$. This directly gives (iii) since α and β are integers. □

Recall that $\mathfrak{M}_*^d$ denotes the vector space of all signed Radon measures on $\mathbb{R}^d \setminus \{0\}$.

Lemma 4.6. *Assuming that $S \neq \emptyset$, there exist vectors $u_1, \ldots, u_k \in S$ and an affine map $F_1 \colon \mathbb{R}^k \to \mathfrak{M}_*^d$ such that*

(i) *the affine map $\Psi \colon \mathbb{R}^d \to \mathbb{R}^k$ given by $\Psi(x) = (\psi_{u_1}(x), \ldots, \psi_{u_k}(x))$ is surjective,*

(ii) *$F_1(\Psi(x), \cdot) = F(x, \cdot)$ for all $x \in \mathbb{R}^d$,*

(iii) *every normalized jump counter ψ_u, $u \in S$, is of the form*

$$\psi_u = c_0 + c_1 \psi_{u_1} + \cdots + c_k \psi_{u_k}$$

for some $c_0, \ldots, c_k \in \mathbb{R}$.

Here the ψ_{u_j} are the normalized jump counters corresponding to the u_j.

Proof. We first claim that

$$\operatorname{span}\{\psi_u \colon u \in S\} = \operatorname{span}\{F(\cdot, A) \colon A \subseteq \mathbb{R}^d \text{ measurable}\}. \qquad (4.7)$$

The inclusion $\supseteq$ follows immediately from (4.6). We now prove the reverse inclusion $\subseteq$. For each $u \in S$, let λ_u and ε_u be as in Proposition 4.5. Then $\{B_{\varepsilon_u}(u) : u \in S\}$ is an open covering of S. Since any open covering of any subset of $\mathbb{R}^d$ admits a countable subcovering, we may choose countably many vectors $u_j \in S$, $j \in \mathbb{N}$, such that $S \subseteq \bigcup_{j=1}^{\infty} B_{\varepsilon_{u_j}}(u_j)$. Defining $C_1 = B_{\varepsilon_{u_1}}(u_1)$ and then recursively $C_j = B_{\varepsilon_{u_j}}(u_j) \setminus \bigcup_{i<j} C_i$ we obtain a pairwise disjoint countable measurable covering of S with $C_j \subseteq B_{\varepsilon_{u_j}}(u_j)$ for all $j \in \mathbb{N}$. For any measurable $A \subseteq \mathbb{R}^d$ we then have

$$F(x, A) = \sum_{j=1}^{\infty} F(x, A \cap C_j) = \sum_{j=1}^{\infty} \lambda_{u_j}(A \cap C_j)\psi_{u_j}(x), \quad x \in \mathbb{R}^d.$$

Since $\operatorname{span}\{\psi_u \colon u \in S\}$ is a subspace of the $(d+1)$-dimensional space of all affine functions from $\mathbb{R}^d$ to $\mathbb{R}$, it is closed under pointwise convergence and therefore contains $F(\cdot, A)$. This proves (4.7).

Write $V = \operatorname{span}\{\psi_u \colon u \in S\}$ for brevity. If V contains the constant function 1, we can find $u_1, \ldots, u_k$ such that $\{1, \psi_{u_1}, \ldots, \psi_{u_k}\}$ is a basis for V. If V does not contain 1, we can find $u_1, \ldots, u_k$ such that $\{\psi_{u_1}, \ldots, \psi_{u_k}\}$ is a basis for V. In either case, the ψ_{u_j} along with 1 are linearly independent, and we have (iii).

We also obtain (i). Indeed, if $c_0 + c_1\psi_{u_1} + \cdots + c_k\psi_{u_k} = 0$ for some constants $c_0, \ldots, c_k$, then $c_0 = c_1 = \cdots = c_k = 0$ by linear independence. Thus the affine space $\Psi(\mathbb{R}^d)$ is not contained in any proper affine subspace of $\mathbb{R}^k$, and so must be all of $\mathbb{R}^k$, which proves (i).

Since Ψ is surjective, there exists an affine map $\Phi : \mathbb{R}^k \to \mathbb{R}^d$ such that $\Psi \circ \Phi = \operatorname{id}$. We define $F_1 : \mathbb{R}^k \to \mathfrak{M}_*^d$ by

$$F_1(y, A) = F(\Phi(y), A).$$

This is affine in y, being the composition of two affine maps. We now argue (ii). Pick $x \in \mathbb{R}^d$ and set $y = \Psi(x)$ and $x' = \Phi(y)$. It follows from (4.7) and the choice of $u_1, \ldots, u_k$ that for any fixed measurable subset $A \subseteq \mathbb{R}^d$, there exist $c_0 \in \mathbb{R}$ and $c \in \mathbb{R}^k$ such that $F(\cdot, A) = c_0 + \langle c, \Psi(\cdot) \rangle$. Since also $\Psi(x') = \Psi \circ \Phi(y) = y = \Psi(x)$, we get

$$F_1(y, A) = F(x', A) = c_0 + \langle c, \Psi(x') \rangle = c_0 + \langle c, \Psi(x) \rangle = F(x, A).$$

This proves (ii). $\qquad \square$

The affine map $\Psi \colon \mathbb{R}^d \to \mathbb{R}^k$ in Lemma 4.6, being surjective, can be extended to an invertible affine map $T \colon \mathbb{R}^d \to \mathbb{R}^d$ whose first k component functions $T_1, \ldots, T_k$ are precisely $\Psi_1, \ldots, \Psi_k$. In particular, $T_j = \psi_{u_j}$ is a normalized jump counter for each $j = 1, \ldots, k$. Note that $k = 0$ is possible, and occurs precisely when X does not jump at all and thus $S = \emptyset$. In this case T is simply an arbitrary invertible affine map, for example the identity map.

Theorem 4.7. *The invertible affine map T satisfies the following properties:*

(i) $T(E) \subseteq \mathbb{N}^k \times \mathbb{R}^{d-k}$,

(ii) $Y = (T(X)_1, \ldots, T(X)_k)$ *is affine and Markov,*

(iii) $Z = (T(X)_{k+1}, \ldots, T(X)_d)$ *can only jump when Y jumps, that is,*

$$\{t \ge 0 \colon \Delta Z(t) \ne 0\} \subseteq \{t \ge 0 \colon \Delta Y(t) \ne 0\} \text{ a.s.,} \qquad (4.8)$$

and its jump characteristic is of the form $\nu^Z(dt, d\zeta) = F^Z(Y_{t-}, d\zeta)dt$ for some affine map $F^Z \colon \mathbb{R}^k \to \mathfrak{M}_*^{d-k}$,

(iv) *the canonical coordinate projections* $\pi_j \colon (x_1, \ldots, x_d) \mapsto x_j$ *are normalized jump counters of the transformed process* (Y, Z) *for* $j = 1, \ldots, k$.

Proof. Part (i) follows directly from the construction of T and Lemma 4.2(i). To prove (ii), let F_1 be as in Lemma 4.6 and define

$$F^Y(y, B) = F_1(y, \Psi^{-1}(\Psi(0) + B))$$

for each $y \in \mathbb{R}^k$ and measurable $B \subseteq \mathbb{R}^k \setminus \{0\}$. For any $y = \Psi(x)$ with $x \in E$, the definition of F^Y along with Lemma 4.6(ii) yield

$$
\begin{aligned}
\int_{\mathbb{R}^k \setminus \{0\}} (\|\eta\|^2 \wedge 1) F^Y(y, d\eta) &= \int_{\{\xi \in \mathbb{R}^d \colon \Psi(\xi) \neq \Psi(0)\}} (\|\Psi(\xi) - \Psi(0)\|^2 \wedge 1) F_1(y, d\xi) \\
&\leq \kappa^2 \int_{\mathbb{R}^d \setminus \{0\}} (\|\xi\|^2 \wedge 1) F(x, d\xi) \\
&< \infty,
\end{aligned}
$$

where κ is the maximum of one and the operator norm of the linear map $\Psi - \Psi(0) \colon \mathbb{R}^d \to \mathbb{R}^k$. Thus, for all $y \in \Psi(E)$, $F^Y(y, \cdot)$ is a positive measure on $\mathbb{R}^k \setminus \{0\}$ satisfying (2.1), and in particular lies in $\mathfrak{M}_*^k$. Since E affinely spans $\mathbb{R}^d$ and Ψ is surjective, it follows that $\Psi(E)$ affinely spans $\mathbb{R}^k$. From the affine dependence on y we then infer $F(y, \cdot) \in \mathfrak{M}_*^k$ for all $y \in \mathbb{R}^k$. Thus F^Y qualifies as the jump kernel of an affine process with state space $\Psi(E)$.

Next, the process $Y = \Psi(X)$ is an $\mathbb{N}^k$-valued semimartingale with characteristics (B^Y, C^Y, ν^Y), say. Since it takes values in a discrete set, we may take $B^Y = 0$ and $C^Y = 0$ (the latter also follows from the fact that X has no diffusion part). In view of [12, Proposition B.1] along with Lemma 4.6(ii), the jump characteristic ν^Y is given by

$$
\begin{aligned}
\nu^Y([0, t] \times B) &= \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \mathbf{1}_{\{\Psi(X_{s-} + \xi) - \Psi(X_{s-}) \in B\}} F(X_{s-}, d\xi) ds \\
&= \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \mathbf{1}_{\{\Psi(\xi) - \Psi(0) \in B\}} F_1(\Psi(X_{s-}), d\xi) ds \\
&= \int_0^t F^Y(Y_{s-}, B) ds
\end{aligned}
$$

for any $t \geq 0$ and measurable $B \subseteq \mathbb{R}^k \setminus \{0\}$. We conclude that Y is an affine process with state space $\Psi(E)$ and triplet $(0, 0, F^Y)$. Since the state space is finite the Markov property follows easily, for instance by an argument based on [8, Theorem 4.4.1]. This completes the proof of (ii).

We now prove (iii). We claim that there exists a nullset $N \subseteq \Omega$, such that for every $\omega \notin N$ one has $\Delta X_t(\omega) \in S \cup \{0\}$ for all $t \geq 0$. Indeed, since the jump characteristic of X is $F(X_{t-}, d\xi) dt$, [11, Theorem II.1.8] yields

$$\mathbb{E}\left[\sum_{t > 0} \mathbf{1}_{\mathbb{R}^d \setminus (S \cup \{0\})}(\Delta X_t) \right] = \mathbb{E}\left[\int_0^\infty F(X_{t-}, \mathbb{R}^d \setminus (S \cup \{0\})) dt \right],$$

which is equal to zero by definition of S. Thus $\sum_{t > 0} \mathbf{1}_{\mathbb{R}^d \setminus (S \cup \{0\})}(\Delta X_t(\omega)) = 0$ for all ω outside some nullset N. In view of the convention $X_0 = X_{0-}$, this is precisely what we claimed. Now, pick any $\omega \notin N$, along with $t \geq 0$ such that $\Delta Z_t(\omega) \neq 0$. Then the vector $u = \Delta X_t(\omega)$ is nonzero and hence lies in S. Let ψ_u be the corresponding normalized jump counter, which satisfies

$$\psi_u(X_t(\omega)) - \psi_u(X_{t-}(\omega)) = \psi_u(u) - \psi_u(0) = -1$$

by definition. On the other hand, Lemma 4.6(iii) yields $c_0, \ldots, c_k \in \mathbb{R}$ such that $\psi_u = c_0 + c_1 \psi_{u_1} + \cdots + c_k \psi_{u_k}$, whence

$$\psi_u(X_t(\omega)) - \psi_u(X_{t-}(\omega)) = c_1 (\Delta Y_t)_1(\omega) + \cdots + c_k (\Delta Y_t)_k(\omega).$$

It follows that $\Delta Y_t(\omega) \neq 0$, which proves (4.8). To obtain the form of the characteristic $\nu^Z(dt, d\zeta)$, define $\Phi \colon \mathbb{R}^d \to \mathbb{R}^{d-k}$ by $\Phi(x)_j = T(x)_{k+j}$ for $j = 1, \ldots, d-k$, so that $Z = \Phi(X)$. Then, as above, we have

$$\nu^Z([0,t] \times C) = \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} \mathbf{1}_{\{\Phi(X_{s-}+\xi) - \Phi(X_{s-}) \in C\}} F(X_{s-}, d\xi) ds = \int_0^t F^Z(Y_{s-}, C) ds$$

for any $t \geq 0$ and measurable $C \subseteq \mathbb{R}^{d-k} \setminus \{0\}$, where $F^Z(y, C) = F_1(y, \Phi^{-1}(\Phi(0) + C))$. That F^Z maps $\mathbb{R}^k$ to $\mathfrak{M}_*^{d-k}$ follows in the same way as the corresponding statement for F^Y above. This completes the proof of (iii).

It remains to prove (iv). For each $j = 1, \ldots, k$, the vector $v_j := T(u_j) - T(0)$ is a possible jump size of (Y, Z). We check that π_j is the corresponding normalized jump counter. Certainly π_j is affine, not identically zero, and nonnegative on $T(E) \subseteq \mathbb{N}^k \times \mathbb{R}^{d-k}$. Since ψ_{u_j} is normalized, so is π_j, because

$$\pi_j(v_j) - \pi_j(0) = \pi_j(v_j) = \psi_{u_j}(u_j) - \psi_{u_j}(0) = -1.$$

Finally, if $\pi_j(T(x)) > 0$ for some $x \in E$, then $\psi_{u_j}(x) > 0$, hence $x + u_j \in E$, and therefore

$$T(x) + v_j = T(x + u_j) \in T(E).$$

This completes the proof of (iv) and of the theorem. $\qquad \square$

5 Examples and further classification

In this section we classify all affine processes with compact state space in one and two dimensions. For an affine process X with triplet (b, c, F), we continue to let S denote the set (4.1) of possible jump sizes.

Dimension $d = 1$. In this case Theorem 2.3 yields, up to an affine transformation, two possible affine processes X with compact state space: either X is deterministic, or X is a finite-state continuous time Markov chain. We only inspect the second possibility, which turns out to result in a simple birth–death process.

Proposition 5.1. *Assume X is a non-deterministic[5] affine process with compact state space $E \subseteq \mathbb{R}$. Then, up to an affine transformation, we have $E = \{0, \ldots, N\}$ for some integer $N \geq 1$, and there are real numbers $\alpha > 0$, $\beta \geq 0$ such that*

$$F(x, \cdot) = x \alpha \delta_{-1} + (N - x) \beta \delta_1, \qquad x \in \mathbb{R}.$$

Moreover, the moment generating function of X is given by

$$\mathbb{E}_x[e^{u X_t}] = \Phi(u, t) \Psi(u, t)^x, \qquad x \in E,$$

for any $u \in \mathbb{C}$ and $t \geq 0$, where we set $0^0 = 1$, and the functions

$$\Phi(u, t) = \left(\frac{\alpha + \beta(e^u + (1 - e^u)e^{-t(\alpha+\beta)})}{\alpha + \beta} \right)^N,$$

$$\Psi(u, t) = 1 + \frac{(\beta + \alpha)(e^u - 1)}{(\beta e^u + \alpha)e^{t(\alpha+\beta)} - \beta(e^u - 1)},$$

---

[5] By non-deterministic we mean that for some $x \in E$, the law of X under $\mathbb{P}_x$ is not a point mass concentrated on one single càdlàg function.

solve the Riccati equations

$$\partial_t \Phi(u,t) = N\beta\Phi(u,t)(\Psi(u,t)-1), \qquad\qquad \Phi(u,0) = 1,$$
$$\partial_t \Psi(u,t) = \alpha + (\beta - \alpha)\Psi(u,t) - \beta\Psi(u,t)^2, \qquad \Psi(u,0) = e^u.$$

Proof. Theorem 2.3 yields, after an affine transformation, that $E \subseteq \mathbb{N}$ and that the coordinate projection $\pi_1 \colon x \mapsto x$, which coincides with the identity since $d = 1$, is a normalized jump counter. In particular $E \subseteq \{0, \dots, N\}$ for some $N \in E$, and since X is non-deterministic, E must contain at least two states, so $N \geq 1$. Let $u \in S$ be a vector whose normalized jump counter is π_1. Then $-1 = \pi_1(u) - \pi_1(0) = u$, so (4.2) implies $E = \{0, \dots, N\}$. If $S = \{-1\}$ then F has the claimed jump structure with $\beta = 0$. Otherwise, consider $v \in S$, $v \neq -1$, and its corresponding normalized jump counter ψ_v. Since $\pi_1(v) - \pi_1(0) = v \neq -1$, Proposition 4.5 yields $v = -u = 1$. Thus $S = \{-1, 1\}$ and F again has the claimed structure with $\beta > 0$.

We now turn to the moment generating function. Fix $u \in \mathbb{R}$ and $t \geq 0$, and define $M_s = \Phi(u, t-s)\Psi(u, t-s)^{X_s}$, $s \in [0, t]$. A calculation using that Φ and Ψ solve the Riccati equations yields that M is a martingale with $M_t = e^{uX_t}$. Hence

$$\mathbb{E}_x[e^{uX_t}] = \mathbb{E}_x[M_t] = M_0 = \Phi(u,t)\Psi(u,t)^x,$$

as claimed. $\qquad\qquad\square$

Observe that, while the affine processes in Proposition 5.1 do admit closed-form moment generating and characteristic functions, there is no exponential-affine transform formula. Indeed, classically one would write

$$\Phi(u,t) = e^{\phi(u,t)} \qquad \text{and} \qquad \Psi(u,t) = e^{\psi(u,t)},$$

where ϕ and ψ solve (generalized) Riccati equations. Here this is not possible in general, as can be seen by taking $N = 1$, $\alpha = 1$, and $\beta = 0$. In this case $\Psi(u,t) = 1 + (e^u - 1)e^{-t}$, which vanishes for $u = \mathrm{i}\pi + \log(e^t - 1)$.

Dimension $d = 2$, case $k = 1$. Up to an affine transformation Theorem 2.3 yields $E \subseteq \mathbb{N}^k \times \mathbb{R}^{2-k}$ for some $k \in \{0, 1, 2\}$. Again $k = 0$ means that X is deterministic, so we ignore this case. If $k = 1$, then $X = (Y, Z)$ where the first component Y is itself a one-dimensional non-deterministic affine process, and therefore of the form described in Proposition 5.1 with state space $\{0, \dots, N\}$ for some nonzero $N \in \mathbb{N}$ and jump kernel $F^Y(y, \cdot) = y\alpha\delta_{-1} + (N-y)\beta\delta_1$ for some parameters $\alpha > 0$ and $\beta \geq 0$. The second component of Z may or may not jump, depending on whether β is zero or not. The following proposition gives the precise statement.

Proposition 5.2. *Consider the affine process $X = (Y, Z)$ just described, with $F^Y(y, \cdot) = y\alpha\delta_{-1} + (N-y)\beta\delta_1$ for some $\alpha > 0$ and $\beta \geq 0$. If $\beta > 0$, then up to an affine transformation of X, Z is continuous and we have*

$$E = \{0, \dots, N\} \times E^Z$$

for some compact subset $E^Z \subseteq \mathbb{R}$. If instead $\beta = 0$, then Z need not be continuous, but can jump at most y times $\mathbb{P}_{(y,z)}$-a.s. for each $(y, z) \in E$.

Proof. Suppose $\beta > 0$. There are compact subsets $F_0, \dots, F_N \subseteq \mathbb{R}$ such that $E = \bigcup_{j=0}^{N} \{j\} \times F_j$. Let $f_j = \min F_j$ for $j = 0, \dots, N$, and define $\phi(y, z) = z - f_0 - y(f_1 - f_0)$ for any $(y, z) \in \mathbb{R}^2$. Then the map $(y, z) \mapsto (y, \phi(y, z))$ is affine and invertible, and the process $(Y, \phi(Y, Z))$ is affine with state space $\bigcup_{j=0}^{N} \{j\} \times G_j$, where $G_j = F_j - f_0 + j(f_1 - f_0)$. In
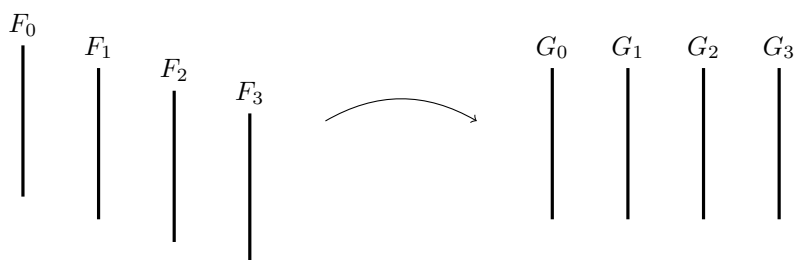
Figure 1: The map $(y, z) \mapsto (y, \phi(y, z))$ in the proof of Proposition 5.2.

particular, $\min G_0 = \min G_1 = 0$. Moreover, this process still satisfies the properties of Theorem 2.3 (with Z replaced by $\phi(Y, Z)$). Therefore, we suppose already from the outset that

$$\min F_0 = \min F_1 = 0 \text{ and the map } T \text{ in Theorem 2.3 is the identity.}$$

This initial transformation is illustrated in Figure 1.

Now, pick $u = (u_1, u_2) \in S$. Since Y jumps by ± 1, we have $u_1 \in \{-1, 0, 1\}$. But due to (4.8), we cannot have $u_1 = 0$ and $u_2 \neq 0$, so in fact $u_1 \in \{-1, 1\}$. Suppose $u_1 = 1$ and let ψ_u be the associated normalized jump counter. Since Y does not jump in direction $u_1 = 1$ when $Y_t = N$, (4.8) implies that Z cannot jump when $Y_t = N$. Thus $\{N\} \times \mathbb{R} \subseteq \ker \psi_u$, which together with the normalization condition uniquely determines ψ_u. Similarly, for $v \in S$ with $v_1 = -1$ we have $\{0\} \times \mathbb{R} \subseteq \ker \psi_v$, and this uniquely determines ψ_v. It follows that ψ_u and ψ_v are the only normalized jump counters.

Next, since $\psi_u(0, z) > 0$ for all $z \in F_0$, we have $F_0 + u \subseteq F_1$, and similarly $F_1 + v \subseteq F_0$. Since $\min F_0 = \min F_1 = 0$ this yields $u_2 \geq 0$ and $v_2 \geq 0$, and hence $u + v = (0, \lambda)$ for some $\lambda \geq 0$. But since $F_0 + (u + v) \subseteq F_0$, and hence $F_0 + (0, n\lambda) \subseteq F_0$ for all $n \in \mathbb{N}$, compactness of F_0 forces $\lambda = 0$. Thus $u = (1, 0)$, $v = (-1, 0)$, $S = \{u, v\}$, and it follows that Z is continuous. It also follows that $F_j = F_0$ for all $j = 1, \ldots, N$, since otherwise a jump in direction u or v would for some point lead out of the state space. This completes the proof of the case $\beta > 0$.

Suppose $\beta = 0$. Then Y only has downward jumps of unit size and stops when it reaches zero. Thus Y jumps exactly y times, $\mathbb{P}_{(y,z)}$-a.s., for each $(y, z) \in E$. Due to (4.8), Z thus jumps at most y times, $\mathbb{P}_{(y,z)}$-a.s. $\qquad \square$

We now give two examples corresponding to the two cases $\beta > 0$ and $\beta = 0$ in Proposition 5.2. The first example shows that while Z is continuous, its drift may still depend on the jump component Y. The second example shows that for $\beta = 0$, the layers $F_0, \ldots, F_N$ need not be equal, and the jumps of Z cannot be eliminated by applying invertible affine transformations.

Example 5.3. Let $N \in \mathbb{N}$ be nonzero and set $E = \{0, \ldots, N\} \times [0, 1]$. Define

$$F((y, z), \cdot) = y\delta_{(-1, 0)} + (N - y)\delta_{(1, 0)} \qquad \text{and} \qquad b(y, z) = (0, y/N - z).$$

For each $x = (y, z) \in E$, the law $\mathbb{P}_x$ of the canonical process $X = (Y, Z)$ is specified as follows. First, Y is the continuous time Markov chain with state space $\{0, \ldots, N\}$, jump intensity measure

$$F^Y(y, \cdot) = y\delta_{-1} + (N - y)\delta_1,$$

and initial condition $Y_0 = y$. Next, Z is the solution of the equation
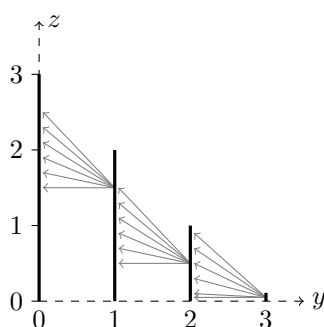
$$dZ_t = b(Y_t, Z_t)dt$$

Figure 2: The process in Example 5.4 jumps from any layer $\{j\} \times [0, N - j]$ to the next layer to the left. The vertical component of the jump size is standard uniform. Within a layer, the process performs a downward linear drift motion.

with initial condition $Z_0 = z$. Then $X = (Y, Z)$ is affine as in Proposition 5.2 with $\beta > 0$.

Example 5.4. Let $N \in \mathbb{N}$ be nonzero and set $E = \bigcup_{j=0}^{N} \{j\} \times [0, N - j]$. Define

$$F((y, z), \cdot) = y(\delta_{-1} \otimes \lambda) \qquad \text{and} \qquad b(y, z) = (0, -z),$$

where λ denotes Lebesgue measure on $[0, 1]$. For each $x = (y, z) \in E$, the law $\mathbb{P}_x$ of the canonical process $X = (Y, Z)$ is specified as follows. Before the first jump and starting from $z \in [0, N - y]$, Z satisfies the equation $dZ_t = -Z_t dt$. When the first jump occurs, which happens with intensity y, Y jumps from y to $y - 1$ and Z makes a positive jump of uniformly distributed size. The process lands in the layer $\{y - 1\} \times [0, N - y + 1]$, and Z continues to perform its downward motion until the next jump, which happens with intensity $y - 1$, and so on. The resulting process is affine as in Proposition 5.2 with $\beta = 0$. Moreover, the set S of possible jump sizes affinely spans $\mathbb{R}^2$, so it is clear that no invertible affine transformation can eliminate the jumps in one of the components. See Figure 2 for an illustration.

Dimension $d = 2$**, case** $k = 2$**.** It remains to consider the case where, up to an affine transformation, we have $E \subseteq \mathbb{N}^2$. That is, we assume $X = T(X) = Y$ in Theorem 2.3, so that $k = 2$ and the coordinate projections $\pi_j \colon x \mapsto x_j$, $j = 1, 2$, are normalized jump counters. The following result classifies this situation, and shows that there are three possibilities.

Theorem 5.5. *Assume $X = (X_1, X_2)$ is affine with state space $E \subseteq \mathbb{N}^2$, and that π_1, π_2 are normalized jump counters. Then X is of one of the following three types:*

(i) *Up to a further affine transformation, E has a layer structure in the sense that $E_0 = \{x \in E : x_2 = 0\}$ is stochastically invariant and $\pi_2(u) \leq 0$ for every $u \in S$ so that there are no upward jumps. Thus there are at most $N = \max\{x_2 : x \in E\}$ downward jumps, after which the process arrives in E_0 and stays there. Moreover, we have*

$$\{(-1, 0)\} \subseteq S \subseteq \{(-1, 0), (1, 0)\} \cup \{(K, -1) : K \in \mathbb{N}\}.$$

(ii) *X_1 and X_2 are independent affine processes, and*

$$\{(-1, 0), (0, -1)\} \subseteq S \subseteq \{(-1, 0), (0, -1), (1, 0), (0, 1)\}.$$

All normalized jump counters apart from π_1 and π_2 are of the form $N - \pi_1$ or $K - \pi_2$ for some $N, K \in \mathbb{N}$.
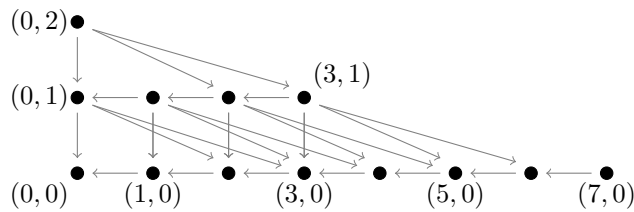
Figure 3: The process in Example 5.6.

(iii) *The set of normalized jump counters is either $\{\pi_1, \pi_2\}$ or $\{\pi_0, \pi_1, \pi_2\}$, where $\pi_0 = N - \pi_1 - \pi_2$ for some $N \in \mathbb{N}$. The set of possible jump sizes S satisfies*

$$S \subseteq \{(-1,0), (-1,1), (0,-1), (0,1), (1,-1), (1,0)\}.$$

Before giving the proof of Theorem 5.5 we give three examples to illustrate the three possibilities. First, Case (ii) is easily constructed by taking two independent one-dimensional affine jump-processes as in Proposition 5.1. Next, we consider an example of Case (i).

Example 5.6. We construct an affine process on a state space with three layers $\{x_2 = 2\}$, $\{x_2 = 1\}$, and $\{x_2 = 0\}$. The situation is illustrated in Figure 3. Let

$$E = \{(0,2), (0,1), (1,1), (2,1), (3,1), (0,0), (1,0), (2,0), (3,0), (4,0), (5,0), (6,0), (7,0)\},$$
$$S = \{(-1,0), (0,-1), (2,-1), (3,-1)\},$$
$$\psi_u = \begin{cases} \pi_1, & u = (-1,0), \\ \pi_2, & \text{any other } u \in S, \end{cases}$$
$$F(x, \cdot) = \sum_{u \in S} \psi_u(x)\delta_u, \quad x \in \mathbb{R}^d.$$

Then there is an affine process with triplet $(0, 0, F)$ and state space E. Note that one could add points $(n, 0)$, $n = 8, 9, \ldots$, to the state space and still obtain an affine process. Note also that the process eventually gets trapped in $E_0 = \{x_2 = 0\}$, and in this particular example even in $(0,0)$.

Finally, we give an example of Case (iii). In fact, we provide a more general example of a d-dimensional affine process with compact and irreducible state space and no autonomous components. We say that a d-dimensional affine process X has *no autonomous components* if there is no invertible affine map $T \colon \mathbb{R}^d \to \mathbb{R}^d$ such that $(T(X)_1, \ldots, T(X)_k)$ is itself an affine process for some $k \le d - 1$.

Example 5.7. Let $d, N \in \mathbb{N}$ with $d, N \ge 1$ and define

$$E = \Big\{ x \in \mathbb{N}^d \colon \sum_{j=1}^d x_j \le N \Big\},$$
$$\pi_0 = N - \sum_{j=1}^d \pi_j,$$
$$S = \{e_j - e_k \colon j, k = 0, \ldots, d, j \ne k\},$$

where as usual $\pi_j \colon x \mapsto x_j$ are the coordinate projections, $e_0 = 0$, and e_j is the jth canonical unit vector for $j = 1, \ldots, d$. The state space E consists of the points of the
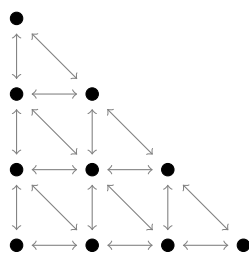
Figure 4: The process in Example 5.7 for $d = 2$ and $N = 3$.

solid simplex with integer coordinates. For each element $u = e_j - e_k \in S$, define also $\psi_u = \pi_k$. Let $\lambda \colon S \to (0, \infty)$ be arbitrary, and define

$$F(x, \cdot) = \sum_{u \in S} \lambda(u) \psi_u(x) \delta_u, \quad x \in \mathbb{R}^d.$$

Then $F(x, \cdot)$ is a finite measure concentrated on $E - x$ for any $x \in E$, and $x \mapsto F(x, \cdot)$ is affine. It is then straightforward to construct an affine process X with triplet $(0, 0, F)$ and state space E. The set S is the set of its possibly jump sizes, and ψ_u is the normalized jump-counter corresponding to $u \in S$. Moreover, X has compact and irreducible state space and no autonomous components. With the special choices $d = 2$ and $N = 3$ we obtain the situation illustrated in Figure 4, and the set of possible jump sizes is

$$S = \{(-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0)\}. \tag{5.1}$$

If we allow $\lambda(u) = 0$ additionally for some u, then the corresponding jump does not occur, and the set of possible jump sizes becomes a subset of the one given in (5.1).

In fact, we believe that Example 5.7 is the only possibility of this type:

Conjecture 5.8. *Let $d \geq 2$ and let X be a d-dimensional affine process with finite and irreducible state space and no autonomous components. Then there is an invertible affine transformation $T \colon \mathbb{R}^d \to \mathbb{R}^d$ such that $T(X)$ coincides with the process constructed in Example 5.7, with the only difference that the function λ may be $[0, \infty)$-valued instead of $(0, \infty)$-valued.*

After this series of examples we now turn to the proof of Theorem 5.5. We first show that "large jumps" imply the layer structure in Theorem 5.5(i); see Lemma 5.9 below. Thereafter, we assume that there are no "large jumps", which has significant further implications.

Lemma 5.9. *Assume that there is $u \in S$ such that $\max\{|u_1|, |u_2|\} \geq 2$. Then any $v \in S$ satisfies $\psi_u(v) - \psi_u(0) \in \{-1, 0\}$. Moreover, for any $v, v' \in S$ such that $\psi_u(v) - \psi_u(0) = \psi_u(v') - \psi_u(0) = 0$ we have $v' \in \{v, -v\}$. In particular, after applying the affine transformation $T = (\pi_1, \psi_u)$ if $|u_1| \geq 2$, or $T = (\pi_2, \psi_u)$ otherwise, we obtain case (i) in Theorem 5.5.*

Proof. We assume without loss of generality that $|u_1| \geq 2$. Let $w \in S$ be a jump size whose normalized jump counter is $\psi_w = \pi_1$. Then Proposition 4.5 with u and w yields

$$\psi_u(w) - \psi_u(0) = 0 \quad \text{and} \quad \pi_1(u) = u_1 \geq 2. \tag{5.2}$$

Define $T = (\pi_1, \psi_u)$ and consider the transformed process $T(X)$. Then $a = Tw - T0 \in T(S) - T(0)$ is in the set of possible jump sizes of $T(X)$ and its normalized jump counter

is π_1, and $b = Tu - T0 \in T(S) - T(0)$ is in the set of possible jump sizes of $T(X)$ and its normalized jump counter is π_2. Moreover, we have

$$\pi_1(a) = \pi_1(w) = -1, \qquad \pi_2(a) = \psi_u(w) - \psi_u(0) = 0,$$
$$\pi_1(b) = \pi_1(u) \geq 2, \qquad \pi_2(b) = \psi_u(u) - \psi_u(0) = -1.$$

That is $a = (-1, 0)$, its normalized jump counter is π_1 and $b = (b_1, -1)$ with $b_1 \geq 2$ and normalized jump counter π_2.

Pick any $c \in T(S) - T(0)$. We claim that $\pi_2(c) \in \{0, -1\}$, and assume for contradiction that $\pi_2(c) \notin \{0, -1\}$. Proposition 4.5 yields that $\pi_2(c) \geq 1$. If $\pi_1(c) \leq -1$, then π_1 would be the normalized jump counter of c and, hence, $c_1 = -1$. Proposition 4.5 applied to b and c would yield $\pi_2(c) = 0$. Thus $c_1 = \pi_1(c) \geq 0$. Thus, $c \neq -b$ and Proposition 4.5 applied to b and c yields $\psi_c(b) - \psi_c(0) = 0$ because $\pi_2(c) \geq 1$. Since a, b are linearly independent and $\psi_c(b) - \psi_c(0) = 0$ we get $\psi_c(a) - \psi_c(0) \neq 0$. Since $a \neq -c$ and $\pi_1(c) \geq 0$ we get from Proposition 4.5 that $\psi_c(a) - \psi_c(0) \geq 1$ and $\pi_1(c) = 0$. Thus $c = (0, c_2)$.

We have $\psi_c = K - \alpha\pi_1 - \beta\pi_2$ for some $K, \alpha, \beta \in \mathbb{R}$. We have $1 \leq \psi_c(a) - \psi_c(0) = \alpha \in \mathbb{N}$. We also have $0 = \psi_c(b) - \psi_c(0) = -\alpha b_1 + \beta$ and, hence, $\beta = \alpha b_1 \geq 2$ and $\beta \in \mathbb{N}$. However, $-1 = \psi_c(c) - \psi_c(0) = -\beta c_2 \leq -2$. A contradiction. $\qquad\square$

We now inspect the possibility that there are exactly two normalized jump counters.

Lemma 5.10. *Assume that $\|u\|_\infty = 1$ for every $u \in S$, and assume that $\{\psi_u : u \in S\}$ has two elements, namely π_1 and π_2. Then, there are $a, b, c, d \in \mathbb{R}_+$ such that*

$$F(x, \cdot) = a\,\pi_1(x)\delta_{(-1,0)} + b\,\pi_1(x)\delta_{(-1,1)} + c\,\pi_2(x)\delta_{(0,-1)} + d\,\pi_2(x)\delta_{(1,-1)}, \quad x \in E.$$

In particular, we have case (iii) *in Theorem 5.5.*

Proof. Let $u \in S$. Then $\pi_1(u) = -1$ or $\pi_2(u) = -1$. Suppose $\pi_1(u) = -1$, i.e. $u = (-1, u_2)$. Then Proposition 4.5 yields $\pi_2(u) \in \{0, 1\}$. The case $\pi_2(u) = -1$ is similar, and we deduce that S can only contain $(-1, 0)$ and $(-1, 1)$ (corresponding to π_1) and $(0, -1)$ and $(1, -1)$ (corresponding to π_2). $\qquad\square$

Next, we consider no "large jumps" and at least three normalized jump counters.

Lemma 5.11. *Assume that $\max\{|u_1|, |u_2|\} = 1$ for every $u \in S$, and assume there exists $w \in S$ such that $\psi_w \notin \{\pi_1, \pi_2\}$. Then, after some affine transformation, either we have case* (i) *in Theorem 5.5, or we have*

$$S \subseteq \{(-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0)\}.$$

Proof. The same argument as in the proof of Lemma 5.10 yields $u \in \{(-1, 0), (-1, 1)\}$ for any $u \in S$ with $\psi_u = \pi_1$, and $u \in \{(0, -1), (1, -1)\}$ for any $u \in S$ with $\psi_u = \pi_2$. Next, choose any $w \in S$ with $\psi_w \notin \{\pi_1, \pi_2\}$. Proposition 4.5 yields $\pi_1(w) \neq -1$. Since also $\pi_1(w) \leq 1$ by the assumption on the jump sizes, we get $\pi_1(w) \in \{0, 1\}$. The same argument yields $\pi_2(w) \in \{0, 1\}$. Therefore, we have $w \in \{(0, 1), (1, 0), (1, 1)\}$.

If $(1, 1) \notin S$, then $S \subseteq \{(-1, 0), (-1, 1), (0, -1), (0, 1), (1, -1), (1, 0)\}$, as claimed. Suppose instead that $(1, 1) \in S$. Let $w = (1, 1)$ and observe that $\psi_w \notin \{\pi_1, \pi_2\}$. Let $u, v \in S$ have normalized jump counters π_1, π_2, respectively. We have seen above that $u \neq -w$, so Proposition 4.5 yields $\psi_w(u) - \psi_w(0) = 0$. Similarly, $\psi_w(v) - \psi_w(0) = 0$. In particular, u and v both lie in $\ker(\psi_w - \psi_w(0))$, which implies that they are proportional and therefore given by $u = -v = (-1, 1)$. Consequently, $S \subseteq \{u, v, w, (1, 0), (0, 1)\}$. Now, since $(-1, 1)$ spans $\ker(\psi_w - \psi_w(0))$, we deduce that $\psi_w = N - (\pi_1 + \pi_2)/2$ for some $N \in \mathbb{N}$. Assume for contradiction that $(1, 0) \in S$. Then $\psi_{(1,0)} = \psi_w$. Moreover, there is $x \in E$ such that

$\psi_w(x) = 1$. We thus have $0 = \psi_w(x + (1,0)) = 1/2$, which is a contradiction. Hence $(1,0) \notin S$, and we similarly deduce $(0,1) \notin S$. Thus $S = \{u, v, w\}$. Now define

$$T = (\pi_1, \psi_w - \psi_w(0)).$$

Then $Tu = (-1, 0)$, $Tw = (1, -1)$ and $Tv = (1, 0)$. Thus, we have (i) in Theorem 5.5 after applying the affine transformation T. □

We are now ready to give the proof of Theorem 5.5.

Proof of Theorem 5.5. We assume that neither case (i) nor case (iii) in Theorem 5.5 holds, and prove that case (ii) must then hold. Lemma 5.9 yields $\|y\|_\infty = 1$ for every $y \in S$, and Lemma 5.10 yields $\psi_z \notin \{\pi_1, \pi_2\}$ for some $z \in S$. Lemma 5.11 then yields that

$$S \subseteq \{(-1,0), (-1,1), (0,-1), (0,1), (1,-1), (1,0)\}.$$

In particular, $z \in \{(0,1), (1,0)\}$ and we consider the case $z = (1,0)$ (the case $z = (0,1)$ is analogous). Let $x, y \in S$ such that $\pi_1(x) = -1$ and $\pi_2(y) = -1$. Then $x \in \{(-1,0), (-1,1)\}$ and $y \in \{(0,-1), (1,-1)\}$. Since Theorem 5.5(i) does not hold, there is $w \in S$ such that $\pi_2(w) = 1$. Hence, $w \in \{(0,1), (-1,1)\}$.

Being an affine function, ψ_z can be written $\psi_z = N - b\pi_1 - a\pi_2$ for some $N, b, a \in \mathbb{R}$. We have $1 = \psi_z(0) - \psi_z(z) = b$ and hence $\psi_z = N - \pi_1 - a\pi_2$. Moreover, since $\pi_2(z) = 0$ we obtain $-y_1 + a = \psi_z(y) - \psi_z(0) \in \mathbb{N}$ which yields $a \in \mathbb{N}$. Since $N - u_1 - au_2 = \psi_z(u) \in \mathbb{N}$ for any $u \in E$, we also have $N \in \mathbb{N}$.

Observe that $-1 \le \psi_z(w) - \psi_z(0) \in \{-a, 1-a\}$, so that $a \in \{0, 1, 2\}$. Assume for contradiction $a = 2$. Then $w = (-1, 1)$ and $\psi_z(w) - \psi_z(0) = -1$, which by Proposition 4.5 yields $\psi_w = \psi_z$. Since $\pi_1(w) = -1$, Proposition 4.5 yields $\psi_w = \pi_1$, a contradiction. Thus $a \in \{0, 1\}$.

Assume for contradiction $a = 1$. In this case, $\psi_z = N - \pi_1 - \pi_2 = \pi_0$. For any $v \in S$ we then have $-1 \in \{\pi_1(v), \pi_2(v), \pi_0(v) - \pi_0(0)\}$ and, hence, $\psi_v \in \{\pi_0, \pi_1, \pi_2\}$. We thus have case (iii). A contradiction.

Thus $a = 0$. Then $\psi_z = N - \pi_1$ and Proposition 4.5 yields $x = -z = (-1, 0)$. Moreover, $\pi_1(w) \ne -1$ because otherwise $w = -z = (-1, 0)$. This gives $w = (0, 1)$. By the same argument as above, we obtain the representation $\psi_w = K - c\pi_1 - \pi_2$ for some $K \in \mathbb{N}$ and $c \in \{0, 1\}$. However, $\psi_w(z) - \psi_w(0) = -c$ implies $c = 0$ because $w \ne -z$. Thus we have $\psi_w = K - \pi_2$. Now, for any $v \in S$ we have $1 \in \{\psi_x(v), \psi_y(v), \psi_z(v) - \psi_z(0), \psi_w(v) - \psi_w(0)\}$ which yields $v \in \{x, y, z, w\}$. Thus $S = \{(-1, 0), (0, -1), (1, 0), (0, 1)\}$ and we have case (ii). □

A Some notions from convex analysis

Fix $n \ge 0$ and let $K \subseteq \mathbb{R}^n$ be a closed convex set. We briefly review some notions and results from convex analysis. For further information and details, see [15].

- The *normal cone* of K at $\overline{x} \in K$ is the closed convex cone

$$N_K(\overline{x}) = \{u \in \mathbb{R}^n \colon \langle u, x - \overline{x}\rangle \le 0 \text{ for all } x \in K\}.$$

  Note in particular that if $\overline{x} \in \operatorname{int} K$, then $N_K(\overline{x}) = \{0\}$.

- A *supporting half-space* of K at $\overline{x} \in K$ is a closed half-space which contains K and whose boundary contains $\overline{x}$. A *supporting hyperplane* of K at $\overline{x}$ is a hyperplane which is the boundary of a supporting half-space of K at $\overline{x}$. Any supporting hyperplane T of K at $\overline{x}$ is of the form

$$T = \{\overline{x} + t \colon t \in \mathbb{R}^n \text{ and } \langle u, t\rangle = 0\}$$

  for some $u \in N_K(\overline{x})$.

- A map $c : \mathbb{R}^n \to \mathbb{S}^n$ is said to be *parallel to K* if

$$c(x)u = 0 \text{ for all } x \in K \text{ and } u \in N_K(x). \tag{A.1}$$

Lemma A.1. *Let $V \subset \mathbb{R}^n$ be a linear subspace, and let $\pi : \mathbb{R}^n \to V$ be the orthogonal projection onto V. Then, for any $\overline{y} \in V \cap K$,*

$$N_{V \cap K}^V(\overline{y}) = \pi(N_K(\overline{y}))$$

where $N_{V \cap K}^V(\overline{y})$ is the normal cone of $V \cap K$ in V.

Proof. First, let $u \in N_K(\overline{y})$. Then $\langle \pi(u), y - \overline{y} \rangle = \langle u, \pi(y - \overline{y}) \rangle = \langle u, y - \overline{y} \rangle \le 0$ holds for any $y \in V \cap K$. Thus $\pi(u) \in N_{V \cap K}(\overline{y})$.

Conversely, let $v \in N_{V \cap K}(\overline{y})$. It suffices to consider the case $v \ne 0$. In this case the set

$$T_v = \{\overline{y} + t : t \in V \text{ and } \langle v, t \rangle = 0\}$$

is a supporting hyperplane in V of $V \cap K$ at $\overline{y}$. Since T_v is disjoint from the interior of K, there exists a supporting hyperplane T in $\mathbb{R}^n$ of K at $\overline{y}$ such that $T_v \subseteq T$. Being a supporting hyperplane, T is given by

$$T = \{\overline{y} + s : s \in \mathbb{R}^n \text{ and } \langle u, s \rangle = 0\}$$

for some $u \in N_K(\overline{y})$. Then, for any $t_1, t_2 \in T_v \subseteq T$,

$$\langle \pi(u), t_1 - t_2 \rangle = \langle u, \pi(t_1 - t_2) \rangle = \langle u, t_1 - t_2 \rangle = 0.$$

Thus $\pi(u) = \lambda v$ for some constant λ. Since both u and v are outward-pointing from K, one sees that $\lambda > 0$. Thus $v \in \pi(N_K(\overline{y}))$ as claimed. $\square$

Lemma A.2. *Suppose that $n \ge 1$ and that K is compact with $0 \in \mathrm{int}\, K$. Then there exists some $x \in K \setminus \{0\}$ with $x \in N_K(x)$.*

Proof. By compactness, there exists $x \in K$ such that $\|x\| = \max_{x' \in K} \|x'\|$. Then for any $x' \in K$, the Cauchy-Schwartz inequality yields

$$\langle x, x' - x \rangle = \langle x, x' \rangle - \|x\|^2 \le \|x\|\|x'\| - \|x\|^2 \le 0.$$

Thus $x \in N_K(x)$. Finally, $x \ne 0$ follows since $n \ge 1$ and $0 \in \mathrm{int}\, K$. $\square$

Lemma A.3. *Suppose that $n \ge 0$ and that K is compact with $0 \in \mathrm{int}\, K$. Let $c : \mathbb{R}^n \to \mathbb{S}^n$ be an affine map parallel to K with $c(0)$ invertible. Then $n = 0$.*

Proof. We may suppose that $c(0) = \mathrm{id}$. To see this, let $A = c(0)^{1/2} \in \mathbb{S}^n$ and define the set $\widehat{K} = A^{-1} K$ as well as the map $\widehat{c} : \mathbb{R}^n \to \mathbb{S}^n$ via $\widehat{c}(y) = A^{-1} c(Ay) A^{-1}$. Then $\widehat{K}$ is again compact and convex with $0 \in \mathrm{int}\, \widehat{K}$. Moreover, by chasing the definitions one verifies that

$$N_{\widehat{K}}(y) = A N_K(Ay).$$

Thus if $v \in N_{\widehat{K}}(y)$, then $v = Au$ for some $u \in N_K(Ay)$, whence $\widehat{c}(y)v = A^{-1} c(Ay) u = 0$. It follows that $\widehat{c}$ is parallel to $\widehat{K}$. In summary, $\widehat{K}$ and $\widehat{c}$ satisfy the same properties as K and c, and in addition $\widehat{c}(0) = \mathrm{id}$. We thus assume without loss of generality that

$$c(x) = \mathrm{id} + \ell(x)$$

for some linear map $\ell : \mathbb{R}^n \to \mathbb{S}^n$.

Assume for contradiction that $n \geq 1$. By Lemma A.2 there exists some $x \in K \setminus \{0\}$ with $x \in N_K(x)$. Since c is parallel to K, one has $0 = c(x)x = x + \ell(x)x$. Hence for any $\lambda \in \mathbb{R}$,

$$c(-\lambda x)x = x - \lambda \ell(x)x = (1 + \lambda)x.$$

By compactness of K, there exists $\lambda > 0$ such that $-\lambda x \in \partial K$. Thus $N_K(-\lambda x)$ contains some nonzero element u, and the set

$$T = \{-\lambda x + t \colon \langle u, t \rangle = 0\}$$

is a supporting hyperplane of K at $-\lambda x$. On the other hand, since c takes values in $\mathbb{S}^n$ and is parallel to K,

$$\langle u, x \rangle = \frac{1}{1 + \lambda} \langle u, c(-\lambda x)x \rangle = \frac{1}{1 + \lambda} \langle c(-\lambda x)u, x \rangle = 0.$$

Thus with $t = \lambda x$ we find that $0 = -\lambda x + t \in T$, which contradicts the hypothesis that $0 \in \operatorname{int} K$. $\qquad\square$

B Nonnegative semimartingales

Variations of the following result are well-known in the literature; see e.g. [16, Proposition 3.1] or [10, Lemma A.1]. We use the convention $Y_{0-} = Y_0$ for the semimartingale Y appearing below.

Lemma B.1. *Let Y be a semimartingale with differential characteristics (b, c, F) with respect to a truncation function χ. Assume $Y \geq 0$, $Y_0 = 0$, and that b_t, c_t, $\int_{\mathbb{R}}(1 - e^{-y} - \chi(y))F_t(dy)$ are right-continuous at $t = 0$. Then $c_0 = 0$.*

Proof. The process $Y' = 1 - e^{-Y}$ is again a semimartingale, whose differential characteristics (b', c', F') with respect to the truncation function $\chi'(y) = y\mathbf{1}_{\{|y| \leq 1\}}$ can be computed using [12, Proposition B.1]. One finds that

$$b_t' = b_t e^{-Y_{t-}} - \frac{1}{2}c_t e^{-Y_{t-}} + e^{-Y_{t-}} \int_{\mathbb{R}} (1 - e^{-y} - \chi(y))F_t(dy),$$
$$c_t' = c_t e^{-2Y_{t-}}.$$

Our assumptions directly imply that

$$b_t' \text{ and } c_t' \text{ are right-continuous at } t = 0. \tag{B.1}$$

Next, note that $\Delta Y'$ takes values in $[0, 1]$, and that $\chi'(y) = y$ for all y in this set. [11, Lemma I.4.24 and Proposition II.2.29] then imply that Y' is a special semimartingale with canonical decomposition $Y' = N + B'$, where N is a local martingale and B' is given by $B_t' = \int_0^t b_s' ds$. Let $N = N^c + N^d$ be the decomposition of N into its continuous and purely discontinuous local martingale parts. Enlarging the probability space if necessary, we can then find a Brownian motion W such that $N_t^c = \int_0^t \sqrt{c_s'}dW_s$. Define

$$Z = \mathscr{E}(-\phi W)$$

for some $\mathscr{F}_0$-measurable random variable $\phi \geq 0$ to be determined later. Integration by parts yields

$$Z_t Y_t' = \int_0^t Y_{s-}' dZ_s + \int_0^t Z_s dY_s' + [Z, Y']_t = M_t + \int_0^t Z_s \left(b_s' - \phi\sqrt{c_s'}\right) ds, \tag{B.2}$$

where $M_t = \int_0^t Y_{s-}' dZ_s + \int_0^t Z_s dN_s$ is a local martingale null at zero. Let σ be a strictly positive reducing stopping time for M, for instance $\sigma = \inf\{t \geq 0 \colon M_t \geq 1 \text{ or } Z_t \geq 2\}$.

Indeed, σ is strictly positive, and since $|\Delta N| = |\Delta Y'| \leq 1$ we have $|M^\sigma| \leq 3$. Now, define the stopping time

$$\tau = \inf\left\{t \geq 0\colon b_t' \geq 1 + b_0' \text{ or } c_t' \leq \frac{c_0'}{4} \text{ or } Z_t \leq \frac{1}{2}\right\} \wedge \sigma \wedge 1.$$

Due to (B.1) we have $\tau > 0$ on $\{c_0' > 0\}$. Set $\phi = 2(2 + 0 \vee b_0')/\sqrt{c_0'}$ on $\{c_0' > 0\}$, and $\phi = 0$ on $\{c_0' = 0\}$. Then, on $\{c_0' > 0\}$ and for all $s \in [0, \tau)$ we have $b_s' - \phi\sqrt{c_s'} \leq -1$ and $Z_s \geq 1/2$. On $\{c_0' = 0\}$ we have $\tau = 0$. Therefore, in view of (B.2), we get

$$0 \leq \mathbb{E}[Z_\tau Y_\tau'] = \mathbb{E}\left[\int_0^1 \mathbf{1}_{[0,\tau)}(s) Z_s \left(b_s' - \phi\sqrt{c_s'}\right) ds\right] \leq -\frac{1}{2}\int_0^1 \mathbb{P}(\tau > s)ds.$$

Thus $\mathbb{P}(c_0' > 0) = \mathbb{P}(\tau > 0) = 0$, which proves the lemma since $c_0 = c_0'$. $\qquad\square$

References

[1] C. Carathéodory, *über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen*, Math. Ann. **64** (1907), no. 1, 95–115. MR-1511425

[2] C. Cuchiero, *Affine and polynomial processes*, Ph.D. thesis, ETH ZURICH, 2011.

[3] Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann, *Affine processes on positive semidefinite matrices*, Ann. Appl. Probab. **21** (2011), no. 2, 397–463. MR-2807963

[4] Christa Cuchiero, Martin Keller-Ressel, Eberhard Mayerhofer, and Josef Teichmann, *Affine processes on symmetric cones*, J. Theoret. Probab. **29** (2016), no. 2, 359–422. MR-3500404

[5] D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab. **13** (2003), no. 3, 984–1053. MR-1994043

[6] Darrell Duffie, Jun Pan, and Kenneth Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica **68** (2000), no. 6, 1343–1376. MR-1793362

[7] Darrell Duffie and Kenneth J Singleton, *Modeling term structures of defaultable bonds*, Review of Financial studies **12** (1999), no. 4, 687–720.

[8] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, Characterization and convergence. MR-0838085

[9] Damir Filipović, *Term-structure models*, Springer Finance, Springer-Verlag, Berlin, 2009, A graduate course. MR-2553163

[10] Damir Filipović and Martin Larsson, *Polynomial diffusions and applications in finance*, Finance Stoch. **20** (2016), no. 4, 931–972. MR-3551857

[11] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR-1943877

[12] Jan Kallsen and Paul Krühner, *On a Heath-Jarrow-Morton approach for stock options*, Finance Stoch. **19** (2015), no. 3, 583–615. MR-3369440

[13] Martin Keller-Ressel and Eberhard Mayerhofer, *Exponential moments of affine processes*, Ann. Appl. Probab. **25** (2015), no. 2, 714–752. MR-3313754

[14] Monika Piazzesi, *Affine term structure models*, Handbook of financial econometrics **1** (2010), 691–766.

[15] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR-0274683

[16] Peter Spreij and Enno Veerman, *Affine diffusions with non-canonical state space*, Stoch. Anal. Appl. **30** (2012), no. 4, 605–641. MR-2946041

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm