Large deviations of affine processes

Matthew Varble

September 17, 2022

Abstract

This is an abstract of the entire dissertation; summarize a history of large deviations and affine processes, then abstractly summarize our large deviations result.

Acknowledgment

This is where I acknowledge how I am useless without others.

Contents

Introduction

This is where I give the reader a little more history and detail regarding affine processes and large deviations, should they read this paper without already being well-versed in the subject.

Notation and conventions

Throughout, unless specifically referenced elsewhere, all notions of this text are formally defined and explored in [Kal02] or [JS03]. Most of our notation will coincide with these texts (as well as most other literature), except in regards to some particular conventions. Let us establish some of these here. A stochastic process X with a marginal-index-set I and state space $(\mathbb{X}, \mathscr{X})$ will be indifferently recognized as:

- a collection $X = (X_t)_{t \in I}$ of marginals $X_t : \Omega \to \mathbb{X}$,
- a map $X : \Omega \times I \to \mathbb{X}$,
- or its curried version $X : \Omega \to \mathbb{X}^I$.

With this convention, we find it appropriate to denote filtrations $\mathscr{F} = (\mathscr{F}_t)_{t \geq 0}$ of increasing σ-algebras $\mathscr{F}_t$. Seeing as $\mathscr{F}$ denotes the actual family of σ-algebras, we denote the joined algebra with an infinity subscript, $\mathscr{F}_\infty \coloneqq \bigvee_{t \geq 0} \mathscr{F}_t$. The blackboard notation will generally correspond to a topological space, including those objects we typically introduce in analysis.

- The real $\mathbb{R}$, the complex $\mathbb{C} = \mathbb{R} \oplus \mathrm{i}\mathbb{R}$, the and non-negative $\mathbb{R}_+ = [0, \infty)$ numbers with the usual Euclidean topologies.
- For real normed vector spaces $\mathbb{V}$, $\mathbb{W}$, the space $\mathbb{L}(\mathbb{V}, \mathbb{W})$ of real linear maps $\mathbb{V} \to \mathbb{W}$, equipped with operator norm.
  $$|T| \coloneqq \sup_{|v|=1} |Tv|$$
  We also concisely denote $\mathbb{L}(\mathbb{V}) \coloneqq \mathbb{L}(\mathbb{V}, \mathbb{V})$.
- For the a separable metric space $\mathbb{X}$ and an interval $I \subseteq \mathbb{R}_+$, the space $\mathbb{D}(I, \mathbb{X})$ of càdlàg functions, equipped with the Skorokhod J1 topology.
- For topological spaces $\mathbb{X}, \mathbb{Y}$, the space $\mathbb{C}(\mathbb{X}, \mathbb{Y})$ of continuous functions, equipped with the supremum norm.

- For finite-dimensional normed vector spaces $\mathbb{V}, \mathbb{W}$ and open $\mathbb{U} \subseteq \mathbb{V}$, the subspace $\mathbb{C}^1(\mathbb{U}, \mathbb{W})$ of functions $f \in \mathbb{C}(\mathbb{U}, \mathbb{W})$ in which there is a derivative map $\mathrm{D}f \in \mathbb{C}(\mathbb{U}, \mathbb{L}(\mathbb{V}, \mathbb{W}))$.

$$\lim_{|v| \to 0} \frac{\big|f(u+v) - f(u) - \mathrm{D}f(u) \cdot v\big|}{|v|} = 0$$

  For $f \in \mathbb{C}^1(\mathbb{U}, \mathbb{R})$, we denote $\nabla f \in \mathbb{C}(\mathbb{U}, \mathbb{V})$ the gradient,

$$\big\langle v, \nabla f(u) \big\rangle \coloneqq \mathrm{D}f(u) \cdot v,$$

  If there is some canonical ordered basis $(e_1, \ldots, e_{\dim \mathbb{V}})$ of $\mathbb{V}$, denote $\mathrm{D}_i f \in \mathbb{C}(\mathbb{U}, \mathbb{R})$ the i-th partial derivative.

$$\mathrm{D}_i f(u) \coloneqq \mathrm{D}f(u) \cdot e_i, \quad i = 1, \ldots, d$$

- For finite-dimensional normed vector space $\mathbb{V}$ and open $\mathbb{U} \subseteq \mathbb{V}$, the subspace $\mathbb{C}^2(\mathbb{U}, \mathbb{R})$ of $f \in \mathbb{C}^1(\mathbb{U}, \mathbb{R})$ in which we also have $\nabla f \in \mathbb{C}^1(\mathbb{U}, \mathbb{V})$. In such a case, we denote $\mathrm{D}^2 f \in \mathbb{C}(\mathbb{U}, \mathbb{L}(\mathbb{V}))$ the Hessian.

$$\mathrm{D}^2 f(u) \coloneqq \mathrm{D}\big(\nabla f(u)\big)$$

  If there is some canonical ordered basis $(e_1, \ldots, e_{\dim \mathbb{V}})$ of $\mathbb{V}$, denote $\mathrm{D}_{ij} f \in \mathbb{C}(\mathbb{U}, \mathbb{R})$ the second-order ij-th partial derivative.

$$\mathrm{D}_{ij} f(u) \coloneqq \big\langle e_i, \mathrm{D}^2 f(u) \cdot e_j \big\rangle, \quad i, j = 1, \ldots, d$$

For any topological space $\mathbb{X}$ and subset $A \subseteq \mathbb{X}$, we denote A° and $\overline{A}$ its interior and closure, respectively. In the case that this topology on $\mathbb{X}$ is induced by some metric d, we denote $B(x, \delta), \overline{B}(x, \delta) \subseteq \mathbb{X}$ the respectively open and closed balls centered at $x \in \mathbb{X}$ with radius $\delta > 0$.

$$B(x, \delta) \coloneqq \Big\{x' \in \mathbb{X} : d(x', x) < \delta\Big\}, \quad \overline{B}(x, \delta) \coloneqq \Big\{x' \in \mathbb{X} : d(x', x) \leq \delta\Big\}$$

When the topology on $\mathbb{X}$ is canonical, we will denote the associated Borel algebra $\mathscr{B}(\mathbb{X})$. Particular examples of this convention are:

- the Borel algebra $\mathscr{B}(\mathbb{V})$ associated to the topology induced from a canonical inner-product $\langle \cdot, \cdot \rangle$ on a vector space $\mathbb{V}$.
- the Borel algebra $\mathscr{B}(\mathbb{X})$ associated to the relative topology of some subset $\mathbb{X}$ of a space $\mathbb{V}$ with itself some canonical topology.

In the case that we are dealing with a finite-dimensional real vector space $\mathbb{V}$ with inner-product $\langle \cdot, \cdot \rangle$, we assume some canonical orthonormal basis $e_1, \ldots, e_{\dim \mathbb{V}} \in \mathbb{V}$ and establish the associated isometric isomorphism $\mathbb{V} \equiv \mathbb{R}^d$.

$$v \in \mathbb{V} \quad \longleftrightarrow \quad \big(v^1, \ldots, v^{\dim \mathbb{V}}\big); \qquad v^i \coloneqq \langle v, e_i \rangle, \quad i = 1, \ldots, \dim \mathbb{V}$$

Similarly identify any map $f : \mathbb{A} \to \mathbb{V}$ with component maps $f_1, \ldots, f_d : \mathbb{A} \to \mathbb{R}$.

$$f : \mathbb{A} \to \mathbb{V} \quad \longleftrightarrow \quad \big(f_1, \ldots, f_d\big) : \mathbb{A} \to \mathbb{R}^d; \qquad f_i(a) \coloneqq \langle f(a), e_i \rangle$$

Extend the inner-product symmetrically to a bilinear form on $\mathbb{V} \oplus \mathrm{i}\mathbb{V}$,

$$\big\langle v_1 + \mathrm{i}w_1, v_2 + \mathrm{i}w_2\big\rangle = \big(\langle v_1, v_2\rangle - \langle w_1, w_2\rangle\big) + \mathrm{i}\big(\langle v_1, w_2\rangle + \langle w_1, v_2\rangle\big),$$

and define the trace of an operator $T \in \mathbb{L}(\mathbb{V})$, as follows.

$$\mathrm{tr}(T) = \sum_{i=1}^{d} \langle e_i, Te_i\rangle$$

We adopt that $(\Omega, \Sigma, \mathrm{P})$ is an abstract probability space that—through the process of enlargement via Kolmogorov's extension theorem—we without loss of generality assume it is equipped with identifications of various quantities $X : \Omega \to \mathbb{X}$ into measurable spaces $(\mathbb{X}, \mathscr{X})$ associated with distributions μ on $(\mathbb{X}, \mathscr{X})$. We typically presume such maps X to be measurable without mention and will otherwise specify this fact explicitly by using the notation $X \in \Sigma/\mathscr{X}$. For each probability measure P on (Ω, Σ), we denote the P-distribution of such X by P_X or pushfoward notation, $X_\#\mathrm{P}$.

$$\mathrm{P}_X\Gamma \coloneqq (X_\#\mathrm{P})(\Gamma) \coloneqq \mathrm{P}(X \in \Gamma) \coloneqq \mathrm{P}\big(X^{-1}\Gamma\big), \quad \Gamma \in \mathscr{X}$$

For intuition, we will also denote integration against this distribution as follows.

$$\int_{\mathbb{X}} \mathrm{P}(X \in \mathrm{d}x) f(x) \coloneqq \int_{\mathbb{X}} \mathrm{P}_X(\mathrm{d}x) f(x) = \int_{\Omega} \mathrm{P}(\mathrm{d}\omega) f\big(X(\omega)\big) \eqqcolon \mathrm{E}_{\mathrm{P}} f(X)$$

Just as E_{P} denotes the expectation operator of the measure P, we will denote $\mathrm{E}_{\mathrm{P}}(\cdot|\mathscr{G})$ the conditional expectation operator of P associated with a filtration $\mathscr{G}$. Should we choose a target space $(\mathbb{Y}, \mathscr{Y})$ and a natural σ-algebra $Y^{-1}\mathscr{Y}$ from some quantity $Y \in \Sigma/\mathscr{Y}$, we denote $\mathrm{E}_{\mathrm{P}}(\cdot|Y = \cdot)$ the factoring of $\mathrm{E}_{\mathrm{P}}(\cdot|Y^{-1}\mathscr{Y})$ through Y.

$$\mathrm{E}_{\mathrm{P}}(X|Y = y) = \mathrm{E}_{\mathrm{P}}\big(X|Y^{-1}\mathscr{Y}\big)\Big|_{Y=y}$$

Also, any quantity $X : \Omega \to \mathbb{X}$ will be identified with the identity map on its codomain, so that we may abusingly use the convenient expectation notation.

$$\mathrm{E}_{\mathrm{P}_X} f(X) \coloneqq \mathrm{E}_{\mathrm{P}_X} f = \int_{\mathbb{X}} f(x)\mathrm{P}_X(\mathrm{d}x) = \int_{\Omega} f\big(X(\omega)\big)\mathrm{P}(\mathrm{d}\omega) = \mathrm{E}_{\mathrm{P}} f(X)$$

This will particularly be useful for when we discuss Markov processes and their associated identities.

Chapter I

Affine processes

Here I put a summary of chapter, along with a short history. It will include the following important notes.

- Chapter addresses important fundamental results of affine processes.
- Chapter addresses consequences of mgf results that are important for us, though not specified exactly much in the literature
- Chapter presents results in full generality, even though we have light-tails assumption; this helps for future extensions.

I.1 Formulation

We start by specifying our affine processes as in [KRM15]. That is to say, we fix a finite-dimensional real vector space $\mathbb{V}$ with inner-product $\langle \cdot, \cdot \rangle$ and select a convex, closed $\mathbb{X} \subseteq \mathbb{V}$ satisfying $0 \in \mathbb{X}$ and $\operatorname{span} \mathbb{X} = \mathbb{V}$. Associate this space with the finite exponentials.

$$\mathcal{U}_{\mathbb{X}} := \left\{ u \in \mathbb{V} \oplus \mathrm{i}\mathbb{V} : \sup_{x \in \mathbb{X}} \exp \langle \Re(u), x \rangle < \infty \right\}$$

We may now define the notion of an affine process on $\mathbb{X}$.

Definition I.1.1. *For a probability space $(\Omega, \Sigma, \mathrm{P})$ with filtration $\mathscr{F} = (\mathscr{F}_t)_{t \geq 0}$, an affine process X on $\mathbb{X}$ is a stochastically continuous, time-homogeneous $(\mathrm{P}, \mathscr{F})$-Markov process on $\mathbb{X}$ in which the bounded moments have the following log-affine dependence on the initial state.*

$$\begin{aligned} \mathrm{E}_{\mathrm{P}_x} \exp \langle u, X_t \rangle &= \exp \Psi(t, u, x) \\ \Psi(t, u, x) &= \psi_0(t, u) + \big\langle \psi(t, u), x \big\rangle, \end{aligned} \qquad t \geq 0,\ u \in \mathcal{U}_{\mathbb{X}} \tag{I.1.2}$$

Above, we are denoting $(\mathrm{P}_x)_{x \in \mathbb{X}}$ the conditional P-distributions of X factored through the initial state $x \in \mathbb{X}$.

Remark I.1.3. *(a) See [KRM15, Remark 2.3] for an argument on how our assumptions on $\mathbb{X}$ are at no loss of generality; $\mathbb{X}$ may as well be any nonempty convex set.*

(b) *Note how (I.1.2) specifies the characteristic function of each transition kernel of the Markov process X; thus, should an affine process exist for choice of Ψ,* only *one will exist, up to distribution.*

(c) *See how our notation (ψ_0, ψ) differs from that of [KRM15] and other papers, which typically use (ϕ, ψ). We choose to do this because affine functions prevail throughout our investigation of affine processes, and we saw this an opportunity to have more cohesive notation of all such affine functions.*

$$\alpha(x) = a_0 + \sum_{i=1}^{d} x^i a_i$$

(d) *If we have a vector space $\mathbb{A}$ and affine map $\alpha : \mathbb{X} \to \mathbb{A}$ determined by $a_0, \ldots, a_d \in \mathbb{A}$ via $\alpha(x) = a_0 + \sum_{i=1}^{d} x^i a_i$, then our linear assumptions $0 \in \mathbb{X}$ and $\operatorname{span} \mathbb{X} = \mathbb{V}$ uniquely determine $a_0, \ldots, a_d \in \mathbb{A}$. In particular, the map Ψ uniquely identifies its parts $\psi_i : \mathbb{R}_+ \times \mathcal{U}_{\mathbb{X}} \to \mathbb{C}$ for $i = 0, \ldots, d$.*

In [Cuc11, Theorem 1.2.7], it is shown that, without loss of generality on conditional distributions $(\mathrm{P}_x)_{x\in\mathbb{X}}$, an affine process X can be chosen to have càdlàg paths. Thus, each distribution P_x may be recognized as a measure on the Borel algebra associated with the space $\mathbb{D}([0,\infty), \mathbb{X})$ of càdlàg functions equipped with the Skorokhod topology. By imposing this regularity, the following theorem tells us that an affine process X as in Definition I.1.1 is a $(\mathrm{P}_x, \mathscr{F})$ jump-diffusion for each $x \in \mathbb{X}$. For relevant definitions and results pertaining to jump-diffusions, we refer the reader to Appendix A.

Theorem I.1.4. *An affine process X on $\mathbb{X}$ is a $(\mathrm{P}_x, \mathscr{F})$ jump-diffusion in which the differential χ-characteristics $(\beta^\chi, \alpha, \mu)$ are affine maps of the following form.*

$$\beta^\chi(x) \coloneqq b_0^\chi + \sum_{i=1}^{d} x^i b_i^\chi, \quad \alpha(x) \coloneqq a_0 + \sum_{i=1}^{d} x^i a_i, \quad \mu(x, \mathrm{d}v) \coloneqq m_0(\mathrm{d}v) + \sum_{i=1}^{d} x^i m_i(\mathrm{d}v)$$

The associated Lévy-Khintchine map Λ *then also affine,*

$$\begin{aligned} \Lambda(u,x) &= \big\langle u, \beta^\chi(x)\big\rangle + \frac{1}{2}\big\langle u, \alpha(x)u\big\rangle + \int_{\mathbb{V}} \Big(e^{\langle u,v\rangle} - 1 - \big\langle u, \chi(v)\big\rangle\Big)\mu(x,\mathrm{d}v) \\ &= L_0(u) + \sum_{i=1}^{d} x^i L_i(u) \\ L_i(u) &\coloneqq \big\langle u, b_i^\chi(x)\big\rangle + \frac{1}{2}\big\langle u, a_i(x)u\big\rangle + \int_{\mathbb{V}} \Big(e^{\langle u,v\rangle} - 1 - \big\langle u, \chi(v)\big\rangle\Big)m_i(x,\mathrm{d}v), \end{aligned}$$

and each $u \in \mathrm{i}\mathbb{V}$ induces the following differential equation.

$$\begin{cases} \psi_0(t,u) = L_0\big(\psi(t,u)\big) & t \geq 0 \\ \psi(t,u) = L\big(\psi(t,u)\big) & t \geq 0 \\ \psi_0(0,u) = 0 \\ \psi(0,u) = u \end{cases} \tag{I.1.5}$$

Proof. This is simply a restatement of [Cuc11, Theorem 1.5.4].

Remark I.1.6. *By Remark I.1.3(d) and linearity of differentiation, the equation in (I.1.5) is equivalent to the following system of equations.*

$$\forall\ x \in \mathbb{X}, \qquad \begin{cases} \dot{\Psi}(t,u,x) = \Lambda\big(\psi(t,u),x\big) & t \geq 0 \\ \Psi(0,u,x) = \langle u, x\rangle \end{cases} \tag{I.1.7}$$

Henceforth, we fix X a càdlàg affine process with conditional distributions $(\mathrm{P}_x)_{x\in\mathbb{X}}$ on $\mathbb{D}([0,\infty),\mathbb{X})$, induced filtration $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$, and moment function Ψ as in Definition I.1.1. We will use the truncation function $\chi(v) = v1_{|v|\leq 1}$ and fix the differential χ-characteristics (β^X, α, μ) and Lévy-Khintchine map Λ as in Theorem I.1.4.

I.2 Existence of real moments

This section elaborates upon the extension of the transform formula in (I.1.2) and equations (I.1.5) and (I.1.7) to real moments $u \in \mathbb{V}$. Clearly, should any extension exist for some $u \in \mathbb{V}$, the value $\Lambda(u,x) = \dot{\Psi}(0,u,x)$ should be well-defined. Throughout this section, we recall our exploration in Section A.4 of the Lévy-Khintchine map Λ and its essential domain of real moments.

$$\mathcal{D}_\Lambda := \Big\{u \in \mathbb{V} : \Lambda(u,x) \text{ is well-defined for all } x \in \mathbb{X}\Big\}$$

These will allow us to establish existence results of $\Psi(\cdot,u,\cdot)$ for real moments $u \in \mathbb{V}$. The following definition will get us started.

Definition I.2.1. *For each $\tau \geq 0$ and $u \in \mathcal{D}_\Lambda$, we say a function $Q^u : [0,\tau]\times\mathbb{X} \to \mathbb{R}$ satisfies* $\mathrm{system}(\Lambda,\tau,u)$ *if the following hold.*

$$\forall\, t \in [0,\tau],\ x \in \mathbb{X}, \quad Q^u(t,x) = q_0^u(t) + \langle q^u(t), x\rangle,$$
$$\forall\, x \in \mathbb{X}, \quad \begin{cases} \dot{Q}^u(t,x) = \Lambda\big(q^u(t),x\big), & t \in [0,\tau] \\ Q^u(0,x) = \langle u, x\rangle \end{cases}$$

Now define the following sets.

$$\mathcal{D}_\Psi(\tau) := \Big\{u \in \mathcal{D}_\Lambda : \textit{there exists a solution to } \mathrm{system}(\Lambda,\tau,u)\Big\}$$
$$\mathcal{D}_\Psi := \bigcup_{\tau\geq 0}\Big(\{\tau\}\times\mathcal{D}_\Psi(\tau)\Big)$$

Theorem I.2.2. *(a) There exists a map $\Psi : \mathcal{D}_\Psi \times \mathbb{X} \to \mathbb{R}$ of the form*

$$\Psi(t,u,x) = \psi_0(t,u) + \big\langle \psi(t,u), x\big\rangle$$

such that, for each $(\tau,u) \in \mathcal{D}_\Psi$, $\Psi(\cdot,u,\cdot)$ is a solution to $\mathrm{system}(\Lambda,\tau,u)$ *dominated by all other such solutions. Moreover, this map satisfies the following for each $(\tau,u) \in \mathcal{D}_\Psi$ and $x \in \mathbb{X}$.*

$$\mathrm{E}_{\mathrm{P}_x} \exp\langle u, X_t\rangle = \exp\Psi(t,u,x), \quad t \in [0,\tau] \tag{I.2.3}$$

(b) If $\tau \geq 0$, $u \in \mathbb{V}$, and $x \in \mathbb{X}^\circ$ are such that $\mathrm{E}_{\mathrm{P}_x} \exp\langle u, X_\tau\rangle < \infty$, then $(\tau,u) \in \mathcal{D}_\Psi$.

Proof. With Remark I.1.6, this is the same as [KRM15, Theorem 2.14].

Now that we have two characterizations for the space $\mathcal{D}_\Psi$, we seek to understand properties of it and the associated moment map $\Psi : \mathcal{D}_\Psi \times \mathbb{X} \to \mathbb{R}$.

Proposition I.2.4. *(a) For each $\tau > 0$, $\mathcal{D}_\Psi(\tau)$ is open in $\mathcal{D}_\Lambda^\circ$,*

(b) For each $\tau > 0$ and $u \in \mathcal{D}_\Psi(\tau) \cap \mathcal{D}_\Lambda^\circ$, $\Psi(\cdot, u, \cdot)$ from Theorem I.2.2 is the unique solution to system(Λ, τ, u).

(c) Ψ is continuously differentiable on $\mathcal{D}_\Psi^\circ \times \mathbb{X}$.

Proof. Fix $\tau > 0$ and $u \in \mathcal{D}_\Psi(\tau) \cap \mathcal{D}_\Lambda^\circ$. Because $u \in \mathcal{D}_\Psi(\tau)$, $\Psi(\cdot, u, \cdot)$ exists on $[0, \tau] \times \mathbb{X}$ as a solution to system(Λ, τ, u). As mentioned in Remark I.1.6, the function $\psi(\cdot, u)$ associated with $\Psi(\cdot, u, \cdot)$ is a solution to the following equation,

$$\begin{cases} \dot{\psi}(t, u) = f\big(t, \psi(t, u)\big) & t \in [0, \tau] \\ \psi(0, u) = u \end{cases} \tag{I.2.5}$$

where $f : \mathbb{R} \times \mathcal{D}_\Lambda^\circ \to \mathbb{V}$ is defined by $f(t, u) \coloneqq L(u)$. Seeing as f is continuously differentiable on $\mathbb{R} \times \mathcal{D}_\Lambda^\circ$ by Lemma A.4.4, we may use [Wal98, III.13 Theorem X] to ensure some $\epsilon > 0$ such that the band

$$S_\epsilon \coloneqq \Big\{(t, v) \in [0, \tau] \times \mathbb{V} : |v - \psi(t, u)| \le \epsilon\Big\}$$

is contained in $\mathbb{R} \times \mathcal{D}_\Lambda^\circ$ and provides us to each $(t_0, v) \in S_\epsilon$ a unique solution $q(\cdot, t_0, v)$ to the following initial value problem,

$$\begin{cases} \dot{q}(t, t_0, v) = f\big(t, q(t, t_0, v)\big) & t \in [t_0, \tau] \\ q(t_0, t_0, v) = v \end{cases}$$

which is continuously differentiable with derivatives $\partial_{t_0} q(t, t_0, v) \in \mathbb{V}$ and $Dq(t, t_0, v) \in \mathbb{L}(\mathbb{V})$ satisfying the following equations.

$$\partial_{t_0} q(t, t_0, v) = -f(t_0, u) + \int_{t_0}^t Df\big(s, q(s, t_0, v)\big)\partial_{t_0} q(s, t_0, v)\mathrm{d}s$$

$$Dq(t, t_0, v) = \mathrm{id}_{\mathbb{V}} + \int_{t_0}^t Df\big(s, q(s, t_0, v)\big)Dq(s, t_0, v)\mathrm{d}s$$

In particular, for each $v \in B(u, \epsilon)$, we have $|v - \psi(0, u)| = |v - u| < \epsilon$, and so $(0, v) \in S_\epsilon$; this allows us to disregard the middle coordinate and have $q : [0, \tau] \times B(u, \epsilon) \to \mathbb{V}$ such that $q(\cdot, v)$ is the unique solution to

$$\begin{cases} \dot{q}(t, v) = L\big(q(t, v)\big), & t \in [0, \tau] \\ q(0, v) = v \end{cases}$$

and the derivative in the second coordinate $Dq(t, v) \in \mathbb{L}(\mathbb{V})$ satisfies the following equation.

$$Dq(t, v) = \mathrm{id}_{\mathbb{V}} + \int_0^t DL\big(q(s, v)\big)Dq(s, v)\mathrm{d}s$$

From here, we may define $Q : [0, \tau] \times B(u, \epsilon) \times \mathbb{X} \to \mathbb{R}$ as follows.

$$Q(t, v, x) \coloneqq q_0(t, v) + \big\langle q(t, v), x \big\rangle$$
$$q_0(t, v) \coloneqq \int_0^t L_0\big(q(s, v)\big)\mathrm{d}s$$
$$L_0(v) \coloneqq \Lambda(v, 0)$$

Because the image of $q(\cdot, v)$ on $[0, \tau]$ remains in $\mathcal{D}_\Lambda^\circ$, on which L is continuously differentiable, q_0 is continuously differentiable with derivatives $\dot{q}_0$ and Dq_0 satisfying the following.

$$\dot{q}_0(t, v) = L_0\big(q(s, v)\big)$$
$$Dq_0(t, v) = \int_0^t DL_0\big(q(s, v)\big)Dq(s, v)\mathrm{d}s$$

By linearity, $Q(\cdot, v, \cdot)$ is a solution to $\mathrm{system}(\Lambda, \tau, v)$ and so $v \in \mathcal{D}_\Psi(\tau)$. We now have $B(u, \epsilon) \subseteq \mathcal{D}_\Psi(\tau)$, concluding part (a). Meanwhile, any solution Q^u to $\mathrm{system}(\Lambda, \tau, u)$ is such that the associated q^u solves (I.2.5) and so $q^u = q(\cdot, u)$. From here, it is thus the case that $Q^u = Q(\cdot, u, \cdot)$. This means Ψ from Theorem I.2.2 is such that $\Psi(\cdot, u, \cdot)$ is the unique solution to $\mathrm{system}(\Lambda, u, \tau)$, concluding part (b). Lastly, for each $x \in \mathbb{X}$, linearity also shows us that $\Psi(\cdot, \cdot, x)$ is continuously differentiable in a neighborhood of (t, u), with derivative in the second coordinate $D\Psi(\cdot, \cdot, x)$ satisfying the following.

$$\begin{aligned}
D\Psi(t, u, x) &= D\psi_0(t, u) + D\psi(t, u) \cdot x \\
&= Dq_0(t, u) + Dq(t, u) \cdot x \\
&= \int_0^t DL_0\big(q(s, u)\big)Dq(s, u)\mathrm{d}s + \left(\mathrm{id}_{\mathbb{V}} + \int_0^t DL\big(q(s, u)\big)Dq(s, u)\mathrm{d}s\right) \cdot x \\
&= x + \int_0^t \left(DL_0\big(q(s, u)\big)Dq(s, u) + \sum_{i=1}^d x_i DL_i\big(q(s, u)\big)Dq(s, u)\right)\mathrm{d}s \\
&= x + \int_0^t D\left(L_0 + \sum_{i=1}^d x_i L_i\right)\big(q(s, u)\big)Dq(s, u)\mathrm{d}s \\
&= x + \int_0^t D\Lambda\big(q(s, u), x\big)Dq(s, u)\mathrm{d}s \\
&= x + \int_0^t D\Lambda\big(\psi(s, u), x\big)D\psi(s, u)\mathrm{d}s
\end{aligned}$$

This concludes part (c).

Corollary I.2.6. *Suppose $0 \in \mathcal{D}_\Lambda^\circ$. For each $\tau > 0$, there exists $\gamma > 0$ with $B(0, \gamma) \subseteq \mathcal{D}_\Psi(\tau)$.*

Proof. Suppose $0 \in \mathcal{D}_\Lambda^\circ$ and fix $\tau > 0$. Theorem I.2.2(b) tells us $0 \in \mathcal{D}_\Psi(\tau)$. Thus, $0 \in \mathcal{D}_\Psi(\tau) \cap \mathcal{D}_\Lambda^\circ$ and so we may use Proposition I.2.4(a) to ensure $0 \in \mathcal{D}_\Psi(\tau)^\circ$. This means there is some $\gamma > 0$ such that $B(0, \gamma) \subseteq \mathcal{D}_\Psi(\tau)$.

Proposition I.2.7. *For each compact set $K \subseteq \mathcal{D}_\Lambda^\circ$, there exists $\delta > 0$ such that $K \subseteq \mathcal{D}_\Psi(\delta)$. Moreover, $\Psi(\cdot, u, \cdot)$ from Theorem I.2.2 is the unique solution to $\mathrm{system}(\Lambda, \delta, u)$ for each $u \in K$.*

Proof. Firstly, we recognize that by virtue of $K \subseteq \mathcal{D}_\Lambda^\circ$ being compact, we have some $\epsilon > 0$ such that the associated open set

$$K^\epsilon := \Big\{ u \in \mathbb{V} : \inf_{u' \in K} |u - u'| < \epsilon \Big\}$$

has closure $\overline{K^\epsilon}$ contained in $\mathcal{D}_\Lambda^\circ$. Note in particular that this provides us with a buffer of radius ϵ around each point in $\mathcal{D}_\Lambda^\circ$.

$$\overline{B}(u, \epsilon) := \Big\{ u' \in \mathbb{V} : |u' - u| \leq \epsilon \Big\}$$
$$\bigcup_{u \in \mathcal{D}_\Lambda} \overline{B}(u, \epsilon) \subseteq \overline{K^\epsilon} \subseteq \mathcal{D}_\Lambda^\circ$$

With these sets established, we mitigate the task of finding a solution Q^u to system(Λ, δ, u) to that of finding a solution q^u to the related equation.

$$\begin{cases} \dot{q}^u(t) = L\big(q^u(t)\big) & t \in [0, \delta] \\ q^u(0) = u \end{cases} \tag{I.2.8}$$

For a fixed $u \in \mathcal{D}_\Lambda^\circ$, the existence of some $\delta_u > 0$ and solution q^u to (I.2.8) may be obtained from the usual fixed-point method (see [Wal98, II.6 Theorem III]). Indeed, Remark I.1.3(d) and Lemma A.4.4 provide us a Lipschitz property for L on $\overline{B}(u, \epsilon)$,

$$\big|L(v) - L(w)| \leq |v - w| C_{u,\epsilon}, \qquad v, w \in \overline{B}(u, \epsilon)$$
$$C_{u,\epsilon} := \sup_{u' \in \overline{B}(u,\epsilon)} \Big| DL\big(u', x\big) \Big|$$

and so a Banach space $(\mathbb{B}_u, \|\cdot\|_{\mathbb{B}_u})$ defined by

$$\delta_u := 1 \wedge \frac{\epsilon}{\sup_{u' \in \overline{B}(u,\epsilon)} |L(u')|}$$
$$\mathbb{B}_u := \mathbb{C}([0, \delta_u], \mathbb{V})$$
$$\|f\|_{\mathbb{B}_u} := \sup_{t \in [0, \delta_u]} |f(t)| e^{-2C_{u,\epsilon} t}$$

is partially equipped with a map $T : \mathbb{C}([0, \delta_u], K) \to \mathbb{C}([0, \delta_u], \overline{K^\epsilon})$ defined by

$$Tf(t) := u + \int_0^t L\big(f(s)\big) \mathrm{d}s,$$

satisfying a contraction property,

$$\|Tf - Tg\|_{\mathbb{B}_u} \leq \frac{1}{2} \|f - g\|_{\mathbb{B}_u},$$

which induces a unique solution $q^u \in \mathbb{C}([0, \delta_u], \overline{K^\epsilon})$ to the associated fixed-point equation, $Tq^u = q^u$. This solution q^u is thus a unique solution to (I.2.8).

From here, we define the following positive δ,

$$\delta := \inf_{u \in K} \delta_u \geq 1 \wedge \inf_{u \in K} \frac{\epsilon}{\sup_{u' \in \overline{B}(u,\epsilon)} |L(u')|} \geq 1 \wedge \frac{\epsilon}{\sup_{u' \in \overline{K^\epsilon}} |L(u')|} > 0$$

so that each $u \in K$ has a unique solution q^u to (I.2.8). This induces the following map $Q^u : [0, \delta] \times \mathbb{X} \to \mathbb{R}$ for each $u \in K$.

$$Q^u(t, x) \coloneqq q_0^u(t) + \big\langle q^u(t), x \big\rangle$$
$$q_0^u(t) \coloneqq \int_0^t L_0\big(q^u(s)\big)\mathrm{d}s$$

By linearity, Q^u is a solution to system(Λ, δ, u) for each $u \in K$, and so $K \subseteq \mathcal{D}_\Psi(\delta)$. For each $u \in K \subseteq \mathcal{D}_\Psi(\delta)$, a solution $\tilde{Q}^u$ to system(Λ, δ, u) is such that the associated $\tilde{q}^u$ solves (I.2.8) and so $\tilde{q}^u = q^u$. From here, it is thus the case that $\tilde{Q}^u = Q^u$. This means Ψ from Theorem I.2.2 is such that $\Psi(\cdot, u, \cdot)$ is the unique solution to system(Λ, u, δ) for all $u \in K$.

Proposition I.2.9. *For any compact subset $K \subseteq \mathcal{D}_\Psi^\circ$, there exists $C_K > 0$ such that the following holds for all $(t, u) \in K$.*

$$\big|\Psi(t, u, x) - \Psi(0, u, x)\big| \leq C_K \cdot t \cdot \big(1 + |x|\big)$$

Proof. Let $K \subseteq \mathcal{D}_\Psi^\circ$ be compact. By Remark I.1.6 and Proposition I.2.4(c), we have that the functions ψ_i for $i = 0, \ldots, d$ are continuously differentiable on $\mathcal{D}_\Psi^\circ$. Thus, we may define the following positive numbers.

$$C_{K,i} \coloneqq \sup_{(t,u)\in K} |\dot{\psi}_i(t, u)|, \quad i = 0, \ldots, d$$
$$C_K \coloneqq \max\Big\{C_{K,0}, C_{K,1}\sqrt{d}, \ldots, C_{K,d}\sqrt{d}\Big\} < \infty$$

Using the fundamental theorem of calculus and that $\Psi(\cdot, u, \cdot)$ solves system(Λ, τ, u), we produce the following bound for all $(t, u) \in K$.

$$\begin{aligned}
\big|\Psi(t, u, x) - \Psi(0, u, x)\big| &= \Big|\psi_0(t, u) + \big\langle \psi(t, u) - u, x\big\rangle\Big| \\
&\leq \big|\psi_0(t, u)\big| + \big|\psi(t, u) - u\big| \cdot |x| \\
&= \left|\int_0^t \dot{\psi}_0(s, u)\mathrm{d}s\right| + \left|\int_0^t \dot{\psi}_i(s, u)\mathrm{d}s\right| \cdot |x| \\
&\leq C_{K,0} \cdot t + \Big(\sum_{i=1}^d C_{K,i}^2\Big)^{1/2} \cdot t \cdot |x| \\
&\leq C_K \cdot t \cdot \big(1 + |x|\big)
\end{aligned}$$

I.3 Finite-dimensional distributions

With a good grasp of the finite real moments associated with our affine process X and their correspondence with Ψ, we now leverage these results to the finite-dimensional distributions. In other words, this section serves to lift Theorem I.2.3 on marginals X_t to one on finite-dimensional distributions $(X_{t_1}, \ldots, X_{t_n})$. Let us establish some notation.

For any space $\mathbb{A}$, positive integer $n \in \mathbb{N}$, and $\underline{a} \in \mathbb{A}^n$, adopt the convention of denoting $\underline{a} = (a_1, \ldots, a_n)$ and

$$\underline{a}_{\ell:m} = (a_\ell, \ldots, a_m) \in \mathbb{A}^{m-\ell+1}, \quad 1 \leq \ell \leq m \leq n.$$

For each $n \in \mathbb{N}$ and $\underline{t} \in [0,\infty)^n$, define the projection map $\pi_{\underline{t}} : \mathbb{X}^{[0,\infty)} \to \mathbb{X}^n$ by

$$\pi_{\underline{t}}(\xi) \coloneqq \xi(\underline{t}) \coloneqq \big(\xi(t_1), \dots, \xi(t_n)\big).$$

Denote $\underline{t} \vdash [0,\infty)$ to mean that $\underline{t}$ is additionally a partition of the following form.

$$0 < t_1 < \cdots < t_n$$

For each such partition $\underline{t} \vdash [0,\infty)$, associate the following notation.

$$\begin{aligned} t_0 &\coloneqq 0 \\ \Delta t_k &\coloneqq t_k - t_{k-1}, \qquad\qquad 1 \le k \le n \\ |\underline{t}| &\coloneqq n \end{aligned}$$

Lastly, for any $A \subseteq [0,\infty)$, denote $\underline{t} \vdash A$ to mean $\underline{t} \vdash [0,\infty)$ and $t_1, \dots, t_{|\underline{t}|} \in A$. For each $n \in \mathbb{N}$, extend the linear operations of $\mathbb{V}$ to $\mathbb{V}^n$, componentwise. Similarly, extend the definition of our inner-product on $\mathbb{V} \oplus \mathrm{i}\mathbb{V}$ to one on $(\mathbb{V} \oplus \mathrm{i}\mathbb{V})^n$, like so.

$$\langle \underline{u}, \underline{v} \rangle \coloneqq \sum_{k=1}^{n} \langle u_k, v_k \rangle$$

We now clearly specify the extension of Ψ to finite-dimensional projections from the perspective of Theorem I.2.2 and equation (I.2.3). Note that this specifically *permits* infinite values.

Definition I.3.1. *To each $\underline{t} \vdash [0,\infty)$, define $\Psi(\underline{t}, \cdot, \cdot) : (\mathbb{V} \oplus \mathrm{i}\mathbb{V})^{|\underline{t}|} \times \mathbb{X} \to (-\infty, \infty]$ as the cumulant generating function of $X_{\underline{t}}$.*

$$\mathrm{E}_{\mathrm{P}_x} \exp \langle \underline{u}, X_{\underline{t}} \rangle \eqqcolon \exp \Psi(\underline{t}, \underline{u}, x)$$

Note that this extends the definition of Ψ in that we may always consider some time $t > 0$ as a partition $t \vdash [0,\infty)$.

Before we investigate real moments, let us establish the easier result on purely complex moments. This will give us intuition for the objects we create in the sequel.

Proposition I.3.2. *For any $\underline{t} \vdash [0,\infty)$, $\underline{u} \in \mathrm{i}\mathbb{V}^{|\underline{t}|}$, and $x \in \mathbb{X}$, we have the following identity, where we denote $n \coloneqq |\underline{t}|$ for brevity.*

$$\begin{aligned} \theta_n &\coloneqq u_n \\ \theta_k &\coloneqq u_k + \psi(\Delta t_{k+1}, \theta_{k+1}), \quad k = n-1, \dots, 1 \\ \Psi(\underline{t}, \underline{u}, x) &= \sum_{k=1}^{|\underline{t}|} \psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_k, \theta_k), x \big\rangle \end{aligned}$$

Proof. We start by recognizing that $\underline{u} \in \mathrm{i}\mathbb{V}$ means the following identity.

$$\left| e^{\langle u_k, x \rangle} \right| = \exp \langle \Re(u_k), x \rangle = 1, \quad k = 1, \dots, n$$

In particular, we have $\theta_n = u_n \in \mathcal{U}_{\mathbb{X}}$; we show $\theta_k \in \mathcal{U}_{\mathbb{X}}$ for $k = n-1, \dots, 1$ by induction.

$$\exp \langle \Re(\theta_k), x \rangle = \left| e^{\langle \theta_k, x \rangle} \right|$$

$$
\begin{aligned}
&= \left|e^{\langle u_k+\psi(\Delta t_{k+1},\theta_{k+1}),x\rangle}\right| \\
&= \left|e^{\langle u_k,x\rangle}\right| \cdot \left|e^{\langle\psi(\Delta t_{k+1},\theta_{k+1}),x\rangle}\right| \\
&= \left|e^{-\psi_0(\Delta t_k,\theta_{k+1})+\Psi(\Delta t_{k+1},\theta_{k+1},x)}\right| \\
&= \left|e^{-\psi_0(\Delta t_{k+1},\theta_{k+1})}\right| \cdot \left|\mathrm{E}_{\mathrm{P}_x} \exp\left\langle\theta_{k+1}, X_{\Delta t_{k+1}}\right\rangle\right| \\
&\leq \left|e^{-\psi_0(\Delta t_{k+1},\theta_{k+1})}\right| \cdot \mathrm{E}_{\mathrm{P}_x}\left|\exp\left\langle\theta_{k+1}, X_{\Delta t_{k+1}}\right\rangle\right| \\
&\leq \left|e^{-\psi_0(\Delta t_{k+1},\theta_{k+1})}\right| \cdot \sup_{x'\in\mathbb{X}} \exp\Re\langle\theta_{k+1}, x'\rangle
\end{aligned}
$$

Now observe the following identity.

$$
\begin{aligned}
&\Psi(\underline{t},\underline{u},x) \\
&= \log \mathrm{E}_{\mathrm{P}_x} \exp\left\langle\underline{u}, X_{\underline{t}}\right\rangle \\
&= \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-1}\left\langle u_k, X_{t_k}\right\rangle \cdot \exp\left\langle\theta_n, X_{t_n}\right\rangle\right) \qquad \text{(I.3.3)}\\
&= \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-1}\left\langle u_k, X_{t_k}\right\rangle \cdot \mathrm{E}_{\mathrm{P}_x}\left(\exp\left\langle\theta_n, X_{t_n}\right\rangle \middle| \mathscr{F}_{t_{n-1}}\right)\right) \\
&= \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-1}\left\langle u_k, X_{t_k}\right\rangle \cdot \exp\Psi(\Delta t_n,\theta_n,X_{t_{n-1}})\right) \\
&= \psi_0(\Delta t_n,\theta_n) \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-1}\left\langle u_k, X_{t_k}\right\rangle \cdot \exp\left(\left\langle\psi(\Delta t_n,\theta_n), X_{t_{n-1}}\right\rangle\right)\right) \\
&= \psi_0(\Delta t_n,\theta_n) \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-2}\left\langle u_k, X_{t_k}\right\rangle \cdot \exp\left(\left\langle u_{n-1}+\psi(\Delta t_n,\theta_n), X_{t_{n-1}}\right\rangle\right)\right) \\
&= \psi_0(\Delta t_n,\theta_n) \qquad \text{(I.3.4)}\\
&\quad + \log \mathrm{E}_{\mathrm{P}_x}\left(\exp\sum_{k=1}^{n-2}\left\langle \theta_k-\psi(\Delta t_{k+1},\theta_{k+1}), X_{t_k}\right\rangle \cdot \exp\left(\left\langle\theta_{n-1}, X_{t_{n-1}}\right\rangle\right)\right)
\end{aligned}
$$

The final term of (I.3.3) is identical to that of (I.3.4) where we have reduced $k=n$ to $k=n-1$. Repeating these equalities inductively $k=n-1,\ldots,1$ will result in the desired identity.

$$
\Psi(\underline{t},\underline{u},x) = \sum_{k=2}^{n}\psi_0(\Delta t_k,\theta_k) + \log \mathrm{E}_{\mathrm{P}_x}\exp\left\langle\theta_1, X_{t_1}\right\rangle = \sum_{k=1}^{n}\psi_0(\Delta t_k,\theta_k) + \left\langle\psi(\Delta t_1,\theta_1), x\right\rangle
$$

As the preceding result shows, the $\mathbb{X}$-affine structure of Ψ allows us to iteratively factor the exponentials in our expectation. The problem with extending this to real moments like in Theorem I.2.2 is that each of the produced quantities θ_k need not produce an integrable exponential on which we apply the transform formula. The next result is our way of coercing such a property to occur; the map $U_{\underline{t}}$ serves to parameterize those moments $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ which the resulting $\underline{\theta}$ is in $\prod_{k=1}^{|\underline{t}|} \mathcal{D}_\Psi(\Delta t_k)$, since this is precisely the set on which we may perform the calculations between (I.3.3) and (I.3.4). This set turns out to be important in our discussion, so we will reserve it special notation.

$$\mathcal{D}_\Psi(\underline{t}) \coloneqq \prod_{k=1}^{|\underline{t}|} \mathcal{D}_\Psi(\Delta t_k), \quad \underline{t} \vdash [0,\infty)$$

Proposition I.3.5. *For each $\underline{t} \vdash [0,\infty)$, the following map $U_{\underline{t}}$ is a continuous injection, where we denote $n \coloneqq |\underline{t}|$ for brevity.*

$$U_{\underline{t}} : \mathcal{D}_\Psi(\underline{t}) \to \mathbb{V}^{|\underline{t}|}, \quad U_{\underline{t}}(\underline{\theta}) \coloneqq \Big(\theta_1 - \psi(\Delta t_2, \theta_2), \ldots, \theta_{n-1} - \psi(\Delta t_n, \theta_n), \theta_n\Big)$$

Moreover, for each $x \in \mathbb{X}$ and $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$, we have the following (finite) identity.

$$\Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x\big) = \sum_{k=1}^{|\underline{t}|} \psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_1, \theta_1), x \big\rangle \tag{I.3.6}$$

Proof. Fix $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$. By definition, this means that to each $k = 1, \ldots, |\underline{t}|$, we have $\theta_k \in \mathcal{D}_\Psi(\Delta t_k)$, and so $\psi(\Delta t_k, \theta_k)$ is well-defined. This ensures that $U_{\underline{t}}$ is well-defined. Now select another point $\underline{\theta}' \in \mathcal{D}_\Psi(\underline{t})$ such that $U_{\underline{t}}(\underline{\theta}) = U_{\underline{t}}(\underline{\theta}')$. The final component of $U_{\underline{t}}$ ensures that $\theta_n = \theta_n'$; by means of induction, we then get $\theta_{k-1} = \theta_{k-1}'$ for $k = n, \ldots, 2$, via the equality on the respective component map.

$$\theta_{k-1} - \psi(\Delta t_k, \theta_k) = U_{\underline{t},k-1}(\underline{\theta}) = U_{\underline{t},k-1}(\underline{\theta}') = \theta_{k-1}' - \psi(\Delta t_k, \theta_k')$$

This indicates to us that $U_{\underline{t}}$ is an injection, and continuity comes simply from continuity of each $\psi(\Delta t_k, \cdot)$ via Proposition I.2.4(c).

It now remains to show the identity in (I.3.6). This reduces down to repeatedly applying iterated expectations; fix $x \in \mathbb{X}$ and observe the following.

$$\begin{aligned}
&\Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x\big) \\
&= \log \mathrm{E}_{\mathrm{P}_x} \exp \big\langle U_{\underline{t}}(\underline{\theta}), X_{\underline{t}} \big\rangle \\
&= \log \mathrm{E}_{\mathrm{P}_x} \left(\exp \sum_{k=1}^{n-1} \big\langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), X_{t_k} \big\rangle \cdot \exp \big\langle \theta_n, X_{t_n} \big\rangle \right) &\text{(I.3.7)} \\
&= \log \mathrm{E}_{\mathrm{P}_x} \left(\exp \sum_{k=1}^{n-1} \big\langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), X_{t_k} \big\rangle \cdot \mathrm{E}_{\mathrm{P}_x} \Big(\exp \big\langle \theta_n, X_{t_n} \big\rangle | \mathscr{F}_{t_{n-1}} \Big) \right) \\
&= \log \mathrm{E}_{\mathrm{P}_x} \left(\exp \sum_{k=1}^{n-1} \big\langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), X_{t_k} \big\rangle \cdot \exp \Psi(\Delta t_n, \theta_n, X_{t_{n-1}}) \right) \\
&= \psi_0(\Delta t_n, \theta_n) \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x} \left(\exp \sum_{k=1}^{n-1} \big\langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), X_{t_k} \big\rangle \cdot \exp \Big(\big\langle \psi(\Delta t_n, \theta_n), X_{t_{n-1}} \big\rangle \Big) \right) \\
&= \psi_0(\Delta t_n, \theta_n) &\text{(I.3.8)} \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x} \left(\exp \sum_{k=1}^{n-2} \langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), X_{t_k} \rangle \cdot \exp \Big(\big\langle \theta_{n-1}, X_{t_{n-1}} \big\rangle \Big) \right)
\end{aligned}$$

The final term of (I.3.8) is identical to that of (I.3.7), where we have reduced $k = n$ to $k = n-1$. Repeating these equalities inductively $k = n-1, \dots, 1$ will result in the desired identity.

$$\Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x\big) = \sum_{k=2}^{n} \psi_0(\Delta t_k, \theta_k) + \log \mathrm{E}_{\mathrm{P}_x} \exp\langle \theta_1, X_{t_1}\rangle = \sum_{k=1}^{n} \psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_1, \theta_1), x\big\rangle$$

We now turn to the analogue of Theorem I.2.2(b), in which P_x-finite moments $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ for initial points $x \in \mathbb{X}^\circ$ are precisely those $\underline{u} \in \mathcal{D}_\Psi(\underline{t})$.

Proposition I.3.9. *Fix $\underline{t} \vdash [0,\infty)$ and denote $n \coloneqq |\underline{t}|$ for brevity. If $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ is such that $\Psi(\underline{t}, \underline{u}, x) < \infty$ for some $x \in \mathbb{X}^\circ$, then the following recursion holds.*

$$\begin{aligned} \theta_n &\coloneqq u_n \in \mathcal{D}_\Psi(\Delta t_n) \\ \theta_k &\coloneqq u_k + \psi\big(\Delta t_{k+1}, \theta_{k+1}\big) \in \mathcal{D}_\Psi(\Delta t_k), \quad k = n-1, \dots, 1 \end{aligned} \tag{I.3.10}$$

Proof. Consider $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ from which we may not construct the recursion in (I.3.10). In other words, there exists maximal $j \in \{1, \dots, n\}$ in the recursion which fails; i.e. $\theta_k \in \mathcal{D}_\Psi(\Delta t_k)$ for all $k = n, \dots, j+1$ and $\theta_j \notin \mathcal{D}_\Psi(\Delta t_j)$. We now repeat the work as in (I.3.7)-(I.3.8) for a fixed $x \in \mathbb{X}^\circ$ to get the following identity.

$$\begin{aligned}
&\log \mathrm{E}_{\mathrm{P}_x} \exp\langle \underline{u}, X_{\underline{t}}\rangle \\
&= \log \mathrm{E}_{\mathrm{P}_x}\Bigg(\exp\Big(\sum_{k=1}^{n-1} \langle u_k, X_{t_k}\rangle \Big) \cdot \mathrm{E}_{\mathrm{P}_x}\Big(\exp\big\langle u_n, X_{t_n}\big\rangle | \mathscr{F}_{t_{n-1}} \Big)\Bigg) \\
&= \psi_0(\Delta t_n, u_n) \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x}\Bigg(\exp\Big(\sum_{k=1}^{n-2} \langle u_k, X_{t_k}\rangle \Big) \cdot \mathrm{E}_{\mathrm{P}_x}\Big(\exp\big\langle u_{n-1} + \psi(\Delta t_n, u_n), X_{t_{n-1}}\big\rangle | \mathscr{F}_{t_{n-2}} \Big)\Bigg) \\
&= \psi_0\big(\Delta t_n, \theta_n\big) \\
&\quad + \log \mathrm{E}_{\mathrm{P}_x}\Bigg(\exp\Big(\sum_{k=1}^{n-2} \langle u_k, X_{t_k}\rangle \Big) \cdot \mathrm{E}_{\mathrm{P}_x}\Big(\exp\big\langle \theta_{n-1}, X_{t_{n-1}}\big\rangle | \mathscr{F}_{t_{n-2}} \Big)\Bigg) \\
&\ \ \vdots \\
&= \sum_{k=j+1}^{n} \psi_0(\Delta t_k, \theta_k) + \log \mathrm{E}_{\mathrm{P}_x}\Bigg(\exp\Big(\sum_{k=1}^{j-1} \langle u_k, X_{t_k}\rangle \Big) \cdot \mathrm{E}_{\mathrm{P}_x}\Big(\exp\big\langle \theta_j, X_{t_j}\big\rangle | \mathscr{F}_{t_{j-1}} \Big)\Bigg) \\
&= \sum_{k=j+1}^{n} \psi_0(\Delta t_k, \theta_k) + \log \mathrm{E}_{\mathrm{P}_x}\Bigg(\exp\Big(\sum_{k=1}^{j-1} \langle u_k, X_{t_k}\rangle \Big) \cdot \mathrm{E}_{\mathrm{P}_{X_{t_{j-1}}}} \exp\big\langle \theta_j, X_{\Delta t_j}\big\rangle \Bigg)
\end{aligned}$$

By Theorem I.2.2, since $\theta_j \notin \mathcal{D}_\Psi(\Delta t_j)$, we have $\mathrm{E}_{\mathrm{P}_{x'}} \exp\langle \theta_j, X_{\Delta t_j}\rangle = \infty$ for all $x' \in \mathbb{X}^\circ$, so the above integrand is infinite on the set $X_{t_{j-1}} \in \mathbb{X}^\circ$. Seeing as $x \in \mathbb{X}^\circ$, this set is P_x non-negligible, and so the quantity is infinite. We conclude that $\underline{u} \notin \mathcal{D}_\Psi(\underline{t})$, which finishes the proof by contrapositive.

Our final result of this section explores more on how finite moments $\underline{u}$ of $X_{\underline{t}}$ relate to those $\underline{\theta}$ of the increments $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$. To see this, we define the following increment cumulant generating function,

$$\varphi(t, \theta, x) \coloneqq \log \mathrm{E}_{\mathrm{P}_x} \exp\langle \theta, X_t - x\rangle = \Psi(t, \theta, x) - \langle \theta, x\rangle$$

Theorem I.3.11. *Fix $\underline{t} \vdash [0,\infty)$ and $x_0 \in \mathbb{X}^\circ$. The map $U_{\underline{t}}$ is a homeomorphism from $\mathcal{D}_\Psi(\underline{t})$ to the collection of $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ for which $\Psi(\underline{t}, \underline{u}, x_0) < \infty$. In particular, this means $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ satisfies $\Psi(\underline{t}, \underline{u}, x_0) < \infty$ if and only if $\underline{u} = U_{\underline{t}}(\underline{\theta})$. Moreover, we have the following identity for all $\underline{x} \in \mathbb{X}^{|\underline{t}|}$.*

$$\langle \underline{u}, \underline{x} \rangle - \Psi(\underline{t}, \underline{u}, x_0) = \sum_{k=1}^{|\underline{t}|} \Big(\langle \theta_k, x_k - x_{k-1} \rangle - \varphi(\Delta t_k, \theta_k, x_{k-1}) \Big), \quad \underline{u} = U_{\underline{t}}(\underline{\theta})$$

Proof. By Proposition I.3.5, we have that U is indeed a continuous map from $\mathcal{D}_\Psi(\underline{t})$ into the finite domain of $\Psi(\underline{t}, \cdot, x_0)$. Conversely, Proposition I.3.9 indicates to us that, on the finite domain of $\Psi(\underline{t}, \cdot, x_0)$, a recursively-defined map $T_{\underline{t}}$ from (I.3.10) exists. Denoting $n \coloneqq |\underline{t}|$, we see that this map is continuous by induction and continuity of compositions.

$$\begin{aligned} T_{\underline{t}}(\underline{u}) = \Big(T_{\underline{t},1}(\underline{u}_{1:n}), \ldots, T_{\underline{t},n}(\underline{u}_{n:n}) \Big), \quad & T_{\underline{t},n}(\underline{u}_{n:n}) = u_n \\ & T_{\underline{t},k}(\underline{u}_{k:n}) = u_k + \psi\big(\Delta t_{k+1}, T_{\underline{t}}(\underline{u}_{k+1:n})\big) \end{aligned}$$

Observe that $T_{\underline{t}}$ is the inverse of $U_{\underline{t}}$. To see this, fix $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$ and $\underline{u} \coloneqq U_{\underline{t}}(\underline{\theta})$. The final coordinate is obvious,

$$T_{\underline{t},n}(\underline{u}_{n:n}) = u_n = U_{\underline{t},n}(\underline{\theta}) = \theta_n,$$

while an inductive hypothesis $T_{\underline{t},k}(\underline{u}_{k:n}) = \theta_k$ gives us the next step.

$$T_{\underline{t},k-1}(\underline{u}_{k-1:n}) = U_{\underline{t},k-1}(\underline{\theta}) + \psi\big(\Delta t_k, T_{\underline{t},k}(\underline{u}_{k:n})\big) = \theta_{k-1} - \psi(\Delta t_k, \theta_k) + \psi(\Delta t_k, \theta_k) = \theta_{k-1}$$

Dual to this, fix $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ for which $\Psi(\underline{t}, \underline{u}, x_0) < \infty$ and define $\underline{\theta} \coloneqq T_{\underline{t}}(\underline{u})$. Again, we immediately have

$$U_{\underline{t},n}(\underline{\theta}) = \theta_n = T_{\underline{t},n}(\underline{u}_{n:n}) = u_n,$$

and an inductive hypothesis of $U_{\underline{t},k}(\underline{\theta}) = u_k$ results in the next step.

$$U_{\underline{t},k-1}(\underline{\theta}) = \theta_{k-1} - \psi(\Delta t_k, \theta_k) = T_{\underline{t},k-1}(\underline{u}_{k-1:n}) - \psi\big(\Delta t_k, T_{\underline{t},k}(\underline{u}_{k:n})\big) = u_{k-1}$$

We have now showed that $U_{\underline{t}}$ is a homeomorphism with inverse $T_{\underline{t}}$. It remains to show our identity for a pairing $\underline{u} = U_{\underline{t}}(\underline{\theta})$.

$$\begin{aligned} &\langle \underline{u}, \underline{x} \rangle - \Psi(\underline{t}, \underline{u}, x_0) \\ &= \big\langle U_{\underline{t}}(\underline{\theta}), \underline{x} \big\rangle - \Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x_0\big) \\ &= \sum_{i=1}^{n-1} \big\langle \theta_k - \psi(\Delta t_{k+1}, \theta_{k+1}), x_k \big\rangle + \langle \theta_n, x_n \rangle - \sum_{i=1}^{n} \psi_0(\Delta t_k, \theta_k) - \big\langle \psi(\Delta t_1, \theta_1), x_0 \big\rangle \\ &= \sum_{i=1}^{n} \Big(\langle \theta_k, x_k \rangle - \psi_0(\Delta t_k, \theta_k) - \big\langle \psi(\Delta t_k, \theta_k), x_k \big\rangle \Big) \\ &= \sum_{i=1}^{n} \Big(\langle \theta_k, x_k \rangle - \Psi(\Delta t_k, \theta_k, x_k) \Big) \\ &= \sum_{i=1}^{n} \Big(\langle \theta_k, x_k - x_{k-1} \rangle - \varphi(\Delta t_k, \theta_k, x_k) \Big) \end{aligned}$$

I.4 Affine jump-diffusions

This section shows how the notions of jump-diffusions explained in Appendix A apply in the affine case. Firstly, we prove the uniform-boundedness property for the affine jump kernel μ associated to our affine process.

Lemma I.4.1. *The jump kernel μ satisfies the following uniform-boundedness property. Any function $f \in \mathscr{B}(\mathbb{V})/\mathscr{B}(\mathbb{R})$ that satisfies*

$$\int_{\mathbb{V}} |f(v)|\mu(x, \mathrm{d}v) < \infty$$

for all $x \in \mathbb{X}$ then satisfies the following.

$$x \mapsto \int_{\mathbb{V}} |f(v)|\mu(x, \mathrm{d}v) \text{ bounded on compact sets}$$

Proof. Seeing as $0 \in \mathbb{X}$ and $\operatorname{span} \mathbb{X} = \mathbb{V}$, we can take appropriate linear combinations to get finite integrals for each of the parts $m_0, \dots, m_d$ of μ.

$$F_i \coloneqq \int_{\mathbb{V}} |f(v)|m_i(\mathrm{d}v) < \infty, \quad i = 0, \dots, d$$

From here, the result is a simple effect of our affine property and boundedness of compact sets.

$$\sup_{|x| \leq M} \left| \int_{\mathbb{V}} |f(v)|\mu(x, \mathrm{d}v) \right| = \sup_{|x| \leq M} \left| F_0 + \sum_{i=1}^{d} x^i F_i \right| \leq F_0 + M \sum_{i=1}^{d} F_i < \infty$$

With this result, we can state succinct versions of the results which exist for general jump-diffusions.

Proposition I.4.2. *If $0 \in \mathcal{D}_\Lambda^\circ$, then X is a $(\mathrm{P}_x, \mathscr{F})$ special jump-diffusion for each $x \in \mathbb{X}$. The resulting drift map $\beta : \mathbb{X} \to \mathbb{V}$ in the special semimartingale decomposition,*

$$X_t = x + \beta(X) \bullet \ell_t + X^{\mathrm{c}} + \mathrm{id}_{\mathbb{V}} * \tilde{q}_t^X$$

is also affine, making all the special differential characteristics (β, α, μ) affine.

Proof. By combining Lemma I.4.1 and Proposition A.4.10, we get that X is special. Now, we perform the algebra to see the affine structure of β.

$$\begin{aligned}
\beta(x) &= \beta^\chi(x) + \int_{\mathbb{V}} \big(v - \chi(v)\big)\mu(x, \mathrm{d}v) \\
&= \left(b_0^\chi + \sum_{i=1}^{d} x^i b_i^\chi\right) + \int_{\mathbb{V}} \big(v - \chi(v)\big)\left(m_0(\mathrm{d}v) + \sum_{i=1}^{d} x^i m_i(\mathrm{d}v)\right) \\
&= \left(b_0^\chi + \int_{\mathbb{V}} \big(v - \chi(v)\big)m_0(\mathrm{d}v)\right) + \sum_{i=1}^{d} x^i \left(b_i^\chi + \int_{\mathbb{V}} \big(v - \chi(v)\big)m_i(\mathrm{d}v)\right)
\end{aligned}$$

Proposition I.4.3. *If the jump kernel satisfies $\mu(x, \mathbb{V}) < \infty$ for all $x \in \mathbb{X}$, then X is $(\mathrm{P}_x, \mathscr{F})$ locally countable for all $x \in \mathbb{X}$. In the resulting factorization,*

$$\mu(x, \mathrm{d}v) = \lambda(x)\kappa(x, \mathrm{d}v),$$

the intensity λ is an affine map and the jump distribution κ is a convex mixture of probability distributions $k_0, \dots, k_d$ whenever $\lambda(x) \neq 0$.

$$\lambda(x) = l_0 + \sum_{i=1}^{d} x^i l_i, \qquad \kappa(x, \mathrm{d}v) = \frac{l_0}{\lambda(x)} k_0(\mathrm{d}v) + \sum_{i=1}^{d} \frac{x^i l_i}{\lambda(x)} k_i(\mathrm{d}v),$$

Proof. By combining Lemmas I.4.1 and A.3.2, we get the desired local countability. Because $0 \in \mathbb{X}$ and $\operatorname{span} \mathbb{X} = \mathbb{V}$, we are able to take appropriate linear combinations to ensure finiteness of the quantities $l_i \coloneqq m_i(\mathbb{V})$ for each $i = 0, \dots, d$. This allows us to define our intensity map.

$$\lambda(x) \coloneqq l_0 + \sum_{i=1}^{d} x^i l_i = m_0(\mathbb{V}) + \sum_{i=1}^{d} x^i m_i(\mathbb{V}) = \mu(x, \mathbb{V})$$

Now, just as in Remark A.3.3, each non-zero l_i will produce a probability distribution $k_i(\mathrm{d}v) \coloneqq m_i(\mathrm{d}v)/l_i$; otherwise, simply define $k_i(\mathrm{d}v) \coloneqq \delta_{e_1}$. This way, we have the factoring $m_i(\mathrm{d}v) = l_i k_i(\mathrm{d}v)$ for each $i = 0, \dots, d$. If $\lambda(x) \neq 0$, we see our other desired identity.

$$\begin{aligned}
\kappa(x, \mathrm{d}v) &\coloneqq \frac{1}{\lambda(x)} \mu(x, \mathrm{d}v) \\
&= \frac{1}{\lambda(x)} \Big(m_0(\mathrm{d}v) + \sum_{i=1}^{d} x^i m_i(\mathrm{d}v) \Big) \\
&= \frac{1}{\lambda(x)} \Big(l_0 k_0(\mathrm{d}v) + \sum_{i=1}^{d} x^i l_i k_i(\mathrm{d}v) \Big) \\
&= \frac{l_0}{\lambda(x)} k_0(\mathrm{d}v) + \sum_{i=1}^{d} \frac{x^i l_i}{\lambda(x)} k_i(\mathrm{d}v)
\end{aligned}$$

Theorem I.4.4. *If $0 \in \mathcal{D}_\Lambda^\circ$, then any $h \in \mathbb{D}([0, \infty), \mathbb{V})$ of finite variation, compact support, and image contained in $\mathcal{D}_\Lambda^\circ$ is such that*

$$\exp\Big(h \bullet X - \Lambda\big(h, X\big) \bullet \ell\Big)$$

is a $(\mathrm{P}_x, \mathscr{F})$ martingale for every $x \in \mathbb{X}$.

Proof. The quantity $M = \exp\big(h \bullet X - \Lambda(h, X) \bullet \ell\big)$ is a $(\mathrm{P}_x, \mathscr{F})$ local martingale by our hypotheses and Theorem A.4.13. To get the remaining martingale property, we first note that the compact support of h means that there exists $\tau > 0$ such that $h(t) = 0$ for all $t > \tau$. This makes $M = M^\tau$, and so we only need to consider the martingale property on the interval $[0, \tau]$. For this, we use [SV10, Theorem 2.6], which requires the maps

$$(s, x) \mapsto \big\langle h(s), \alpha(x) h(s) \big\rangle, \quad (s, x) \mapsto \int_{\mathbb{V}} \big(e^{\langle h(s), v \rangle} - 1 - \langle h(s), v \rangle \big) \mu(x, \mathrm{d}v)$$

are bounded on compact sets of points (s, x). This comes from the fact that the image of h is contained in some compact subset of $\mathcal{D}_\Lambda^\circ$ and that Λ is uniformly bounded on compact subsets of $\mathcal{D}_\Lambda^\circ \times \mathbb{X}$ by Lemma A.4.4.

Chapter II

Large deviations of affine processes

1. Summarize how multiple *frameworks* are utilized: Dembo, Feng, Puhalskii
2. Summarize history of works treating DG+EM differently

II.1 Asymptotic family

We will prove a large deviation principle for a family $(\mathrm{P}^\epsilon_{x_0})_{\epsilon>0}$ of distributions $\mathrm{P}^\epsilon_{x_0}$ of affine processes ϵX^ϵ with initial point $x_0 \in \mathbb{X}^\circ$ in which the special differential characteristics $(\beta^\epsilon, \alpha^\epsilon, \mu^\epsilon)$ of each respective X^ϵ have the following parameterization.

$$\beta^\epsilon(x) = \frac{1}{\epsilon}\beta^1(\epsilon x), \quad \alpha^\epsilon(x) = \frac{1}{\epsilon}\alpha^1(\epsilon x), \quad \mu^\epsilon(x, \mathrm{d}v) = \frac{1}{\epsilon}\mu^1(\epsilon x, \mathrm{d}v), \quad x \in \mathbb{X} \tag{II.1.1}$$

In effect, the family $(\mathrm{P}^\epsilon_x)_{\epsilon>0}$ is induced by a *base distribution* $\mathrm{P}_x \coloneqq \mathrm{P}^1_x$ associated with *base affine process* $X \coloneqq X^1$ and *base special differential characteristics* $(\beta, \alpha, \mu) \coloneqq (\beta^1, \alpha^1, \mu^1)$. This also implies a similar parameterization for the Lévy-Khintchine maps Λ^ϵ associated with $(\beta^\epsilon, \alpha^\epsilon, \mu^\epsilon)$ in terms of the base map Λ from (β, α, μ).

$$\Lambda^\epsilon(u, x) = \frac{1}{\epsilon}\Lambda(u, \epsilon x), \quad u \in \mathbb{V},\ x \in \mathbb{X} \tag{II.1.2}$$

Using the notation of Appendix A.4, we see that the set $\mathcal{D}_\Lambda(x)$ of finite points of $\Lambda(\cdot, x)$ is identical to that $\mathcal{D}_{\Lambda^\epsilon}(\epsilon x)$ of $\Lambda^\epsilon(\cdot, \epsilon x)$. So long that $\mathbb{X}$ is a cone—which is to say that $\mathbb{X}$ an additive set, closed under non-negative-scalar multiplication—we have $\mathbb{X} = \epsilon\mathbb{X}$, and so the following sets agree.

$$\mathcal{D}_\Lambda = \bigcap_{x\in\mathbb{X}} \mathcal{D}_\Lambda(x) = \bigcap_{x\in\mathbb{X}} \mathcal{D}_{\Lambda^\epsilon}(\epsilon x) = \bigcap_{x\in\mathbb{X}} \mathcal{D}_{\Lambda^\epsilon}(x) = \mathcal{D}_{\Lambda^\epsilon}$$

Note that a parameterization like (II.1.1) or (II.1.2) may exist irrespective of the affine property on (β, α, μ). However, affine processes are distinct in the existence (from Theorem I.2.2) of an affine map $\Psi^\epsilon : \mathcal{D}_{\Psi^\epsilon} \to \mathbb{R}$ respective to X^ϵ,

$$\Psi^\epsilon(t, u, x) = \psi^\epsilon_0(t, u) + \big\langle \psi^\epsilon(t, u), x \big\rangle,$$

in which $\Psi^\epsilon(\cdot,u,\cdot)$ is the minimal solution of system$(\Lambda^\epsilon,\tau,u)$ for each $(\tau,u)\in\mathcal{D}_{\Psi^\epsilon}$,

$$\forall\ x\in\mathbb{X},\quad \begin{cases}\dot\Psi^\epsilon(t,u,x)=\Lambda^\epsilon\big(\psi^\epsilon(t,u),x\big), & t\in[0,\tau]\\ \Psi^\epsilon(0,u,x)=\langle u,x\rangle\end{cases}$$

and is the cumulant generating function of each marginal.

$$\mathrm{E}_{\mathrm{P}^\epsilon_x}\exp\langle u,X^\epsilon_\tau\rangle=\exp\Psi^\epsilon(\tau,u,x/\epsilon),\quad (\tau,u)\in\mathcal{D}_\Psi,\ x\in\mathbb{X}$$

Above, note that we are naturally denoting $(\mathrm{P}^\epsilon_x)_{x\in\mathbb{X}}$ the other distributions P^ϵ_x of ϵX^ϵ in which it starts at various $x\in\mathbb{X}$ (hence why we have a x/ϵ in the last coordinate). The following result shows us that our parameterization in (II.1.1) and (II.1.2) applies these cumulant generating functions, where $\Psi:=\Psi^1$ and $\mathcal{D}_\Psi:=\mathcal{D}^1_\Psi$.

Proposition II.1.3. *Assume $\mathbb{X}$ is a cone satisfying $\operatorname{span}\mathbb{X}=\mathbb{V}$. For each $\epsilon>0$, we have $\mathcal{D}_\Psi=\mathcal{D}_{\Psi^\epsilon}$ and the following identities.*

$$\Psi^\epsilon(t,u,x)=\frac{1}{\epsilon}\Psi(t,u,\epsilon x),\quad \psi^\epsilon_0(t,u)=\frac{1}{\epsilon}\psi_0(t,u),\quad \psi^\epsilon(t,u)=\psi(t,u),$$

Proof. Start by selecting $(\tau,u)\in\mathcal{D}_\Psi$. This means that $u\in\mathcal{D}_\Psi(\tau)$ and $\Psi(\cdot,u,\cdot)$ is a solution to system(Λ,u,τ). Observe that this implies the following identity for all $x\in\mathbb{X}$.

$$\frac{\partial}{\partial t}\frac{1}{\epsilon}\Psi(t,u,\epsilon x)=\frac{1}{\epsilon}\dot\Psi(t,u,\epsilon x)=\frac{1}{\epsilon}\Lambda\big(\psi(t,u),\epsilon x\big)=\Lambda^\epsilon(\psi(t,u),x),\quad t\in[0,\tau]$$
$$\frac{1}{\epsilon}\Psi(0,u,\epsilon x)=\frac{1}{\epsilon}\langle u,\epsilon x\rangle=\langle u,x\rangle$$

This means that $\frac{1}{\epsilon}\Psi(\cdot,u,\epsilon\cdot)$ is a solution to system$(\Lambda^\epsilon,\tau,u)$. By definition, existence of a solution means that $u\in\mathcal{D}_{\Psi^\epsilon}(\tau)$, and so $(\tau,u)\in\mathcal{D}_{\Psi^\epsilon}$. Theorem I.2.2 then tells us $\Psi^\epsilon(\cdot,u,\cdot)$ exists and is dominated by the other solution.

$$\Psi^\epsilon(t,u,x)\le\frac{1}{\epsilon}\Psi(t,u,\epsilon x),\quad t\in[0,\tau],\ x\in\mathbb{X}$$

On the other hand, if we have $(\tau,u)\in\mathcal{D}_{\Psi^\epsilon}$, then $u\in\mathcal{D}_{\Psi^\epsilon}(\tau)$, and so $\Psi^\epsilon(\cdot,u,\cdot)$ is a solution to system$(\Lambda^\epsilon,\tau,u)$. Now, we have the following identity for all $x\in\mathbb{X}$,

$$\frac{\partial}{\partial t}\epsilon\Psi^\epsilon(t,u,x/\epsilon)=\epsilon\dot\Psi^\epsilon(t,u,x/\epsilon)=\epsilon\Lambda^\epsilon\big(\psi^\epsilon(t,u),\epsilon x\big)=\Lambda(\psi(t,u),x),\quad t\in[0,\tau]$$
$$\epsilon\Psi^\epsilon(0,u,x/\epsilon)=\epsilon\langle u,x/\epsilon\rangle=\langle u,x\rangle,$$

and so $\epsilon\Psi^\epsilon(\cdot,u,\cdot)$ is a solution to system(Λ,τ,u). Again, we may conclude from this that $(\tau,u)\in\mathcal{D}_\Psi$ and that $\Psi^\epsilon(\cdot,u,\cdot)$ exists and is dominated by the other solution.

$$\Psi(t,u,x)\le\epsilon\Psi^\epsilon(t,u,x/\epsilon),\quad t\in[0,\tau],\ x\in\mathbb{X}$$

In total, we have now shown that $\mathcal{D}_\Psi=\mathcal{D}_{\Psi^\epsilon}$, and inequalities (15) and (15) indicate to us that the following functions agree.

$$\Psi^\epsilon(t,u,x)=\frac{1}{\epsilon}\Psi(t,u,x),\quad (t,u)\in\mathcal{D}_\Psi,\ x\in\mathbb{X}$$

This means equality of the following affine expressions.

$$\begin{aligned}\psi_0^\epsilon(t,u) + \big\langle \psi^\epsilon(t,u), x\big\rangle &= \Psi^\epsilon(t,u,x)\\ &= \frac{1}{\epsilon}\Psi(t,u,\epsilon x)\\ &= \frac{1}{\epsilon}\psi_0(t,u) + \frac{1}{\epsilon}\big\langle \psi(t,u), \epsilon x\big\rangle = \frac{1}{\epsilon}\psi_0(t,u) + \big\langle \psi(t,u), x\big\rangle\end{aligned}$$

Seeing as $\operatorname{span}\mathbb{X} = \mathbb{V}$, we may take appropriate linear combinations to show the remaining identities.

$$\psi_0^\epsilon(t,u) = \frac{1}{\epsilon}\psi_0(t,u), \quad \psi_i^\epsilon(t,u) = \psi_i(t,u), \quad i = 1,\dots,d$$

Remark II.1.4. *Note that the above proof can be applied to complex moments, since Theorem I.1.4 and Remark I.1.6 indicate to us that each $\Psi^\epsilon(\cdot,u,\cdot)$ is a solution to following equation, for each $u \in \mathrm{i}\mathbb{V}$.*

$$\forall\ x \in \mathbb{X}, \qquad \begin{cases} \dot{\Psi}^\epsilon(t,u,x) = \Lambda^\epsilon\big(\psi^\epsilon(t,u),x\big), & t \geq 0\\ \Psi^\epsilon(0,u,x) = \langle u, x\rangle \end{cases}$$

This parameterization also applies to the liftings $\Psi^\epsilon(\underline{t},\cdot,\cdot)$ of Ψ^ϵ to finite-dimensional projections on partitions $\underline{t} \vdash [0,\infty)$.

$$\mathrm{E}_{\mathrm{P}_x^\epsilon} \exp\langle \underline{u}, \epsilon X_{\underline{t}}^\epsilon\rangle =: \exp\Psi(\underline{t},\underline{u},x/\epsilon), \quad \underline{u} \in \mathbb{V}^{|\underline{t}|},\ x \in \mathbb{X}$$

Denoting $\Psi(\underline{t},\cdot,\cdot) \coloneqq \Psi^1(\underline{t},\cdot,\cdot)$, the below result shows just this.

Proposition II.1.5. *Assume $\mathbb{X}$ is a cone satisfying $\operatorname{span}\mathbb{X} = \mathbb{V}$. Fix $\underline{t} \vdash [0,\infty)$, $x_0 \in \mathbb{X}^\circ$, and $\epsilon > 0$ and define $U_{\underline{t}}$ as in Proposition I.3.5. Each $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ satisfying $\underline{u} = U_{\underline{t}}(\underline{\theta})$ for some $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$ satisfies*

$$\Psi^\epsilon\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x_0\big) = \frac{1}{\epsilon}\Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), \epsilon x_0\big) < \infty,$$

and if no such $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$ exists, both are infinite.

$$\Psi^\epsilon(\underline{t},\underline{u},x_0) = \frac{1}{\epsilon}\Psi(\underline{t},\underline{u},\epsilon x_0) = \infty.$$

Proof. We start by recognizing two facts. Firstly, from Proposition II.1.3, we have an identity of the following sets.

$$\mathcal{D}_{\Psi^\epsilon}(\underline{t}) = \prod_{k=1}^{|\underline{t}|}\mathcal{D}_{\Psi^\epsilon}(\Delta t_k) = \prod_{k=1}^{|\underline{t}|}\mathcal{D}_{\Psi}(\Delta t_k) = \mathcal{D}_\Psi(\underline{t}), \tag{II.1.6}$$

Secondly, Proposition II.1.3 also shows us that the $U_{\underline{t}}^\epsilon$ associated with X^ϵ is identical to that $U_{\underline{t}}$ of X, as $\psi^\epsilon = \psi$. We now show the desired identity by fixing $\underline{u} \in \mathbb{V}^{|\underline{t}|}$ and considering each case.

First suppose $\underline{u} = U_{\underline{t}}(\underline{\theta})$ for some $\underline{\theta} \in \mathcal{D}_\Psi(\underline{t})$. The identity of (II.1.6) tells us $\underline{\theta} \in \mathcal{D}_{\Psi^\epsilon}(\underline{t})$ and so Propositions I.3.5 and II.1.3 give us the following.

$$\Psi^\epsilon(\underline{t},\underline{u},x_0) = \Psi^\epsilon\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), x_0\big)$$

$$\begin{aligned}
&= \sum_{k=1}^{|\underline{t}|} \psi_0^\epsilon(\Delta t_k, \theta_k) + \big\langle \psi^\epsilon(\Delta t_1, \theta_1), x_0 \big\rangle \\
&= \frac{1}{\epsilon}\bigg(\sum_{k=1}^{|\underline{t}|} \psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_1, \theta_1), \epsilon x_0 \big\rangle \bigg) \\
&= \frac{1}{\epsilon}\Psi\big(\underline{t}, U_{\underline{t}}(\underline{\theta}), \epsilon x_0\big) \\
&= \frac{1}{\epsilon}\Psi(\underline{t}, \underline{u}, \epsilon x_0)
\end{aligned}$$

On the other hand, suppose $\underline{u}$ is not in the image of $\mathcal{D}_\Psi(\underline{t})$ under $U_{\underline{t}}$. Seeing as $x_0 \in \mathbb{X}^\circ$ and $\mathbb{X}$ is a cone, we have $\epsilon x_0 \in \mathbb{X}^\circ$. Applying Theorem I.3.11, we then have $\Psi(\underline{t}, \underline{u}, \epsilon x_0) = \infty$. The identity in (II.1.6) also tells us that $\underline{u}$ is not in the image of $\mathcal{D}_\Psi(\underline{t})$ under $U_{\underline{t}}$. Theorem I.3.11 now tell us $\Psi^\epsilon(\underline{t}, \underline{u}, x_0) = \infty$. We conclude our final identity.

$$\Psi^\epsilon(\underline{t}, \underline{u}, x_0) = \frac{1}{\epsilon}\Psi(\underline{t}, \underline{u}, \epsilon x_0) = \infty$$

Now that we have established parameterizations for just about every object that relates to an affine process, we establish some intuition on the relationship between these distributions $(\mathrm{P}_x^\epsilon)_{\epsilon>0}$. The first of which is immediate from our preceding result, but it only makes sense when we consider the countable sequence $\epsilon_m := 1/m$ for $m \in \mathbb{N}$.

Proposition II.1.7. *For a fixed $x \in \mathbb{X}$, the family $(\mathrm{P}_x^{\epsilon_m})_{\epsilon>0}$ corresponds to a mean-field regime. That is to say, if we fix a probability space $(\Omega, \Sigma, \mathrm{P})$ equipped with a sequence of independent quantities $(X^{(j)})_{j\in\mathbb{N}}$ each distributing according to P_x, then we may realize each $\epsilon_m X^{\epsilon_m}$ as follows.*

$$\epsilon_m X^{\epsilon_m} = \frac{1}{m}\sum_{j=1}^{m} X^{(j)}, \quad m \in \mathbb{N}$$

Proof. We will prove this by showing that the finite-dimensional distributions agree by identity of their characteristic functions. Fixing $\underline{t} \vdash [0,\infty)$ and $\underline{u} \in \mathbb{V}^{|\underline{t}|}$, we apply Proposition I.3.2 and Remark II.1.4.

$$\begin{aligned}
\log \mathrm{E}_{\mathrm{P}} \exp \Big\langle \mathrm{i}\underline{u}, \sum_{j=1}^{m} X_{\underline{t}}^{(j)} \Big\rangle &= \log \Big(\mathrm{E}_{\mathrm{P}_x} \exp \langle \mathrm{i}\underline{u}, X_{\underline{t}} \rangle \Big)^m \\
&= \log \Big(\exp \Psi\big(\underline{t}, \mathrm{i}\underline{u}, x\big) \Big)^m \\
&= m\Psi\big(\underline{t}, \mathrm{i}\underline{u}, x\big) \\
&= m\bigg(\sum_{k=1}^{|\underline{t}|} \psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_1, \theta), x \big\rangle \bigg) \\
&= \sum_{k=1}^{|\underline{t}|} \frac{1}{\epsilon_m}\psi_0(\Delta t_k, \theta_k) + \big\langle \psi(\Delta t_1, \theta), x/\epsilon_m \big\rangle \Big) \\
&= \sum_{k=1}^{|\underline{t}|} \psi_0^{\epsilon_m}(\Delta t_k, \theta_k) + \big\langle \psi^{\epsilon_m}(\Delta t_1, \theta), x/\epsilon_m \big\rangle
\end{aligned}$$

$$\begin{aligned}&= \Psi^{\epsilon_m}(\underline{t}, \underline{u}, x/\epsilon_m)\\ &= \log \mathrm{E}_{\mathrm{P}_x^{\epsilon_m}} \exp \langle \underline{u}, X_{\underline{t}} \rangle\end{aligned}$$

We may also intuitively understand the relationship of $(\mathrm{P}^\epsilon_x)_{x\in\mathbb{X}}$ from a dynamical system perspective. In Theorem A.1.14, we see how jump-diffusions X always correspond to a weak solution of some stochastic differential equation driven by standard Brownian motion W and Poisson random measure p.

$$\begin{aligned}
&X_t = X_0 + \beta(X)\bullet\ell_t + \sigma(X)\bullet W_t + c(X,\mathrm{id}_{\mathbb{V}}) * \tilde{p}_t\\
&\forall\ x\in\mathbb{X},\quad \begin{cases}\mu(x,\Gamma) = \displaystyle\int_{\mathbb{V}} 1_\Gamma\big(c(x,v)\big)\mathrm{d}v, \quad \Gamma\in\mathscr{B}(\mathbb{V}-\{0\})\\ \alpha(x) = \sigma\sigma^*(x)\end{cases}
\end{aligned} \tag{II.1.8}$$

The following proposition gives perspectives on how the processes ϵX^ϵ may relate through these objects in two different perturbed dynamical systems.

Proposition II.1.9. *Fix a probability space $(\Omega,\Sigma,\mathrm{P})$ equipped with standard Brownian motion W on $\mathbb{V}$ and Poisson random measure p on $\mathscr{B}(\mathbb{R}_+\times\mathbb{V})$. Let $\sigma:\mathbb{X}\to\mathbb{L}(\mathbb{V})$ and $c:\mathbb{X}\times\mathbb{V}\to\mathbb{V}$ satisfy (II.1.8) for the special differential characteristics (β,α,μ), as granted by Theorem A.1.14. For each $x\in\mathbb{X}$, the family $(\mathrm{P}^\epsilon_x)_{\epsilon>0}$ of distributions P^ϵ_x may be recognized as each P^ϵ_x being a weak solution to the respective* scaled *stochastic dynamical system,*

$$\epsilon X^\epsilon_t = x + \beta(\epsilon X^\epsilon)\bullet\ell_t + \sqrt{\epsilon}\sigma(X)\bullet W_t + \epsilon c\big(\epsilon X^\epsilon, \sqrt[d]{\epsilon}\cdot\mathrm{id}_{\mathbb{V}}\big) * \tilde{p}_t,$$

or the time-changed *stochastic dynamical system.*

$$\begin{aligned}\epsilon X^\epsilon_t &= x + \beta(\epsilon X^\epsilon)\bullet\ell_t + \epsilon\sigma(X)\bullet W^\epsilon_t + \epsilon c\big(\epsilon X^\epsilon,\mathrm{id}_{\mathbb{V}}\big) * \tilde{p}^\epsilon_t,\\ W^\epsilon_t &\coloneqq W_{t/\epsilon}\\ p^\epsilon([0,t]\times\Gamma) &\coloneqq p([0,t/\epsilon]\times\Gamma)\end{aligned}$$

Proof. By [JS03, III.2.26], it suffices to check if the characteristics match; let's check those for X^ϵ. Let us first address the first system. Note that for any $i,j=1,\dots,d$, we use [JS03, Theorem I.4.40(d)] to resolve the predictable quadratic covariation of the continuous local martingale part $\frac{1}{\sqrt{\epsilon}}\sigma(\epsilon X^\epsilon)\bullet W$ of X^ϵ.

$$\begin{aligned}
\Big\langle\big(\frac{1}{\sqrt{\epsilon}}\sigma(\epsilon X^\epsilon)\bullet W\big)^i, \big(\frac{1}{\sqrt{\epsilon}}\sigma(\epsilon X^\epsilon)\bullet W\big)^j\Big\rangle &= \Big\langle \sum_{l=1}^d \frac{1}{\sqrt{\epsilon}}\sigma_{i,l}(\epsilon X^\epsilon)\bullet W^l, \sum_{m=1}^d\frac{1}{\sqrt{\epsilon}}\sigma_{j,m}(\epsilon X^\epsilon)\bullet W^m\Big\rangle\\
&= \frac{1}{\epsilon}\sum_{l,m=1}^d \sigma_{i,l}(\epsilon X^\epsilon)\sigma_{j,m}(\epsilon X^\epsilon)\bullet\big\langle W^l, W^m\big\rangle\\
&= \frac{1}{\epsilon}\sum_{l=1}^d \sigma_{i,l}(\epsilon X^\epsilon)\sigma_{j,l}(\epsilon X^\epsilon)\bullet\ell\\
&= \frac{1}{\epsilon}\alpha_{i,j}(\epsilon X^\epsilon)\bullet\ell\\
&= \langle X^{\epsilon,\mathrm{c},i}, X^{\epsilon,\mathrm{c},j}\rangle
\end{aligned}$$

Note that the accumulated jump process associated from X^ϵ in these dynamics is the following process.

$$t \mapsto \sum_{0<s\le t} c\big(\epsilon X^\epsilon_{s-},\, \sqrt[d]{\epsilon}\cdot \epsilon \Delta X^\epsilon_s\big)$$

This allows us to see that, for a non-negative predictable process $H : \Omega \times \mathbb{R}_+ \times \mathbb{V} \to \mathbb{R}$, we have the following identities, via changing coordinates.

$$\begin{aligned}
\mathrm{E_P}\Big(H * q^{X^\epsilon}_\infty\Big) &= \mathrm{E_P}\int_{\mathbb{R}_+\times\mathbb{V}} H\Big(\cdot, s, c\big(\epsilon X^\epsilon_{s-},\, \sqrt[d]{\epsilon}\cdot v\big)\Big) p(\mathrm{d}s, \mathrm{d}v) \\
&= \mathrm{E_P}\int_0^\infty \int_{\mathbb{V}} H\Big(\cdot, s, c\big(\epsilon X^\epsilon_{s-},\, \sqrt[d]{\epsilon}\cdot v\big)\Big) \mathrm{d}v \mathrm{d}s \\
&= \mathrm{E_P}\int_0^\infty \int_{\mathbb{V}} H\Big(\cdot, s, c\big(\epsilon X^\epsilon_{s-}, v\big)\Big) \frac{1}{\epsilon} \mathrm{d}v \mathrm{d}s && v \leftarrow \sqrt[d]{\epsilon}\cdot v \\
&= \mathrm{E_P}\int_0^\infty \int_{\mathbb{V}} H(\cdot, s, v) \frac{1}{\epsilon} \mu(\epsilon X^\epsilon_{s-}, \mathrm{d}v) \mathrm{d}s && v \leftarrow c(X_{s-}, v)
\end{aligned}$$

Now we address the second system using the same calculations; the continuous local martingale term of X^ϵ in this case is $\sigma(\epsilon X^\epsilon)\bullet W$, and we have the following identity.

$$\begin{aligned}
\Big\langle \big(\sigma(\epsilon X^\epsilon)\bullet W^\epsilon\big)^i, \big(\sigma(\epsilon X^\epsilon)\bullet W^\epsilon\big)^j \Big\rangle &= \Big\langle \sum_{l=1}^d \sigma_{i,l}(\epsilon X^\epsilon)\bullet W^{\epsilon,l}, \sum_{m=1}^d \sigma_{j,m}(\epsilon X^\epsilon)\bullet W^{\epsilon,m} \Big\rangle \\
&= \sum_{l,m=1}^d \sigma_{i,l}(\epsilon X^\epsilon)\sigma_{j,m}(\epsilon X^\epsilon)\bullet \big\langle W^{\epsilon,l}, W^{\epsilon,m}\big\rangle \\
&= \sum_{l=1}^d \sigma_{i,l}(\epsilon X^\epsilon)\sigma_{j,l}(\epsilon X^\epsilon)\bullet \big(\epsilon^{-1}\ell\big) \\
&= \frac{1}{\epsilon}\alpha_{i,j}(\epsilon X^\epsilon)\bullet \ell \\
&= \langle X^{\epsilon,\mathrm{c},i}, X^{\epsilon,\mathrm{c},j}\rangle
\end{aligned}$$

Meanwhile, our time-change of the Poisson random measure immediately gives us our desired characteristic.

$$\begin{aligned}
\mathrm{E_P}\Big(H * q^{X^\epsilon}_\infty\Big) &= \mathrm{E_P}\int_{\mathbb{R}_+\times\mathbb{V}} H\big(\cdot, s, c(\epsilon X^\epsilon_{s-}, v)\big) p^\epsilon(\mathrm{d}s, \mathrm{d}v) \\
&= \mathrm{E_P}\int_0^\infty \int_{\mathbb{V}} H\big(\cdot, s, c(\epsilon X^\epsilon_{s-}, v)\big) \mathrm{d}v \cdot \frac{1}{\epsilon}\mathrm{d}s \\
&= \mathrm{E_P}\int_0^\infty \int_{\mathbb{V}} H(\cdot, s, v) \frac{1}{\epsilon} \mu(\epsilon X^\epsilon_{s-}, \mathrm{d}v) \mathrm{d}s && v \leftarrow c(X_{s-}, v)
\end{aligned}$$

II.2 Assumptions

II.3 Dawson-Gärtner

II.4 Exponential martingales

II.5 Integral representation of rate function

Chapter III

Large deviation rate functions

III.1 Mogulskii's theorem

A surprisingly powerful theorem in the theory of large deviations of stochastic processes is that of Mogulskii (see [DZ10, Theorems 5.1.2 and 5.1.19 and Exercise 5.122]). Fixing a family $(V_j)_{j\in\mathbb{N}}$ of independent quantities distributing with common distribution κ with light tails,

$$\Lambda_\kappa(u) \coloneqq \log \int_{\mathbb{V}} e^{\langle u,v\rangle}\kappa(\mathrm{d}v) < \infty, \quad u \in \mathbb{V} \tag{III.1.1}$$

this theorem provides a large deviation principle for the laws associated to quantities Y^ϵ as below.

$$Y_t^\epsilon = \epsilon \sum_{j=1}^{[t/\epsilon]} V_j, \quad t \in [0,\tau]$$

It states that the associated laws $(\mathrm{P}^\epsilon)_{\epsilon>0}$ satisfy a large deviation principle on the space $\mathbb{L}^\infty[0,\tau]$ of bounded functions $[0,\tau] \to \mathbb{V}$, equipped with the supremum norm. The rate function, like ours, is an integral of the Fenchel-Legendre transform of Λ_κ.

$$\xi \mapsto \begin{cases} \displaystyle\int_0^\tau \Lambda_\kappa^*\big(\dot\xi(t)\big)\mathrm{d}t & \xi(0)=0,\ \xi \in \mathbb{A}([0,\tau],\mathbb{V}) \\ \infty & \text{otherwise} \end{cases}$$

Very minor adjustments can actually make this theorem similar to the context of our principle. Firstly, the principle may be lifted to the space $\mathbb{L}^\infty_{\mathrm{loc}}[0,\infty)$ of locally bounded functions $[0,\infty) \to \mathbb{V}$, equipped with the weighted supremum norm,

$$(\xi,\xi') \mapsto \sup_{t\in[0,\infty)} e^{-t}|\xi(t)-\xi'(t)|,$$

for this metric is consistent with $\xi_n \to \xi$ if and only if $\xi_n|_{[0,\tau]} \to \xi|_{[0,\tau]}$ uniformly for all $\tau \geq 0$, which is the same as the projective limit space induced by the restriction maps.

$$(\xi_\tau)_{\tau>0} \in \varprojlim_{\tau} \mathbb{L}^\infty[0,\tau] \quad \xleftrightarrow{\xi_\tau=\xi|_{[0,\tau]}} \quad \xi \in \mathbb{L}^\infty_{\mathrm{loc}}[0,\infty)$$

Applying Dawson-Gärtner [DZ10, Theorem 4.6.1], the rate function over this space is as follows.

$$\xi \mapsto \begin{cases} \sup_{\tau>0} \int_0^\tau \Lambda_\kappa^*\big(\dot\xi(t)\big)\mathrm{d}t & \xi(0)=0,\ \xi \in \mathbb{A}([0,\tau),\mathbb{V}) \text{ for all } \tau>0 \\ \infty & \text{otherwise} \end{cases}$$

From here, we recognize that each process Y^ϵ is càdlàg; if ν is supported on $\mathbb{X}$, the process takes values in $\mathbb{D}([0,\infty),\mathbb{X})$, and so we may restrict our principle (see [DZ10, Lemma 4.1.5(b)]). Our rate function then takes the same form (recall the local definition of absolute continuity $\mathbb{A}([0,\infty),\mathbb{X})$).

$$\text{(III.1.2)} \qquad \xi \mapsto \begin{cases} \int_0^\infty \Lambda_\kappa^*\big(\dot\xi(t)\big)\mathrm{d}t & \xi(0)=0,\ \xi \in \mathbb{A}([0,\infty),\mathbb{X}) \\ \infty & \text{otherwise} \end{cases}$$

Example III.1.3 (Brownian motion). *Applying Mogulskii's theorem when our increment distribution κ is* Normal$(0, \mathrm{id}_\mathbb{V})$*, the integral in our rate function in (III.1.2) becomes the following.*

$$\text{(III.1.4)} \qquad \int_0^\infty \frac{1}{2}\big|\dot\xi(t)\big|^2 \mathrm{d}t$$

Furthermore, for a Brownian motion W, the process $\sqrt{\epsilon}W$ ends up being exponentially equivalent to Y^ϵ,

$$\limsup_{\epsilon\to 0} \epsilon \log \mathrm{P}\big(|\sqrt{\epsilon}W - Y^\epsilon| \geq \delta\big) = -\infty,$$

which makes the family $\sqrt{\epsilon}W$ satisfy the large deviation principle with rate function (III.1.4); this result is known as Schilder's theorem (see [DZ10, Theorem 5.2.3]).

*Note that $(\sqrt{\epsilon}W)_{\epsilon>0}$ is a family of affine processes covered Theorem **??**. We have $\epsilon X^\epsilon = \sqrt{\epsilon}W$, where the base process X has special differential characteristics $(0, \mathrm{id}_\mathbb{V}, 0)$. The easiest way to see this is by considering Theorem **??** with initial state $x = 0$. Our theorem also immediately resolves (**??**) the same rate function.*

$$\Lambda^*(\dot x, x) = \sup_{u\in\mathbb{V}} \left(\langle u, \dot x\rangle - \frac{1}{2}\langle u, \mathrm{id}_\mathbb{V} \cdot u\rangle \right) = \frac{1}{2}|\dot x|$$

Example III.1.5 (Poisson). *One may apply a very similar argument for when our increment distribution κ is* Poisson(1)*. In this case, the integral in the rate function in (III.1.2) evaluates to*

$$\text{(III.1.6)} \qquad \int_0^\infty \Big(\dot\xi(t)\log\big(\dot\xi(t)\big) - \dot\xi(t) + 1\Big)\mathrm{d}t,$$

so long as $\xi(t) \geq 0$ for Lebesgue-almost-every $t \geq 0$ (otherwise, it is infinite). In the case that $\xi(t) = 0$, we are taking the continuous extension of the integrand, i.e. $0\log(0) := 0$. Similar to the work of Schilder's theorem, we may show, for a standard Poisson process N, $\epsilon N_{\cdot/\epsilon}$ is exponentially equivalent to this Y^ϵ, which makes the family satisfy a large deviation principle with rate function (III.1.6). In fact and exercise of our reference text, [DZ10, Exercise 5.2.12], suggests the reader to show just this.

*Again, such a family $(\epsilon N_{\cdot/\epsilon})_{\epsilon>0}$ is covered by Theorem **??**. To see this, consider a base affine process X on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ with special differential characteristics as below, where δ_1 denotes the degenerate distribution at $1 \in \mathbb{R}$.*

$$\beta(x) = 1, \quad \alpha(x) = 0, \quad \mu(x, \mathrm{d}v) = \delta_1(\mathrm{d}v)$$

*Setting the initial state $x = 0$ and looking at Theorem **??**, we may say that ϵX^ϵ can be realized as follows.*

$$\begin{aligned}
\epsilon X^\epsilon_t &= t + \epsilon 1_{[0,1]}(\mathrm{id}_{\mathbb{R}}) * \tilde{p}^\epsilon_t \\
&= t + \epsilon 1_{[0,1]}(\mathrm{id}_{\mathbb{R}}) * p^\epsilon_t - \epsilon 1_{[0,1]}(\mathrm{id}_{\mathbb{R}}) * \hat{p}^\epsilon_t \\
&= t + \epsilon p([0, t/\epsilon] \times [0,1]) - \int_0^{t/\epsilon} \int_{\mathbb{R}} \epsilon 1_{[0,1]}(v) \mathrm{d}v \mathrm{d}s \\
&= \epsilon p([0, t/\epsilon] \times [0,1])
\end{aligned}$$

As stated in [JS03, Theorem II.4.8], this Poisson random measure p is a Poisson point process with Lebesgue intensity. This means that, for each $t \geq 0$, $N_t \coloneqq p([0,t] \times [0,1])$ distributes $\mathrm{Poisson}(t)$*, and $N_t - N_s = p((s,t] \times [0,1])$ is independent of $N_s = p([0,s],[0,1])$ for each $0 \leq s < t$. In other words, N is a standard Poisson process and*

$$\epsilon X^\epsilon_t = \epsilon p([0, t/\epsilon] \times [0,1]) = \epsilon N_{t/\epsilon}.$$

*As with the normal increments, our rate function (**??**) resolves this immediately.*

$$\begin{aligned}
\Lambda^*(\dot{x}, x) = \sup_{u \in \mathbb{V}} \left(u\dot{x} - u - \int_{\mathbb{R}} \left(e^{uv} - 1 - uv\right) \delta_1(\mathrm{d}v) \right) &= \sup_{u \in \mathbb{V}} \left(u\dot{x} - e^u + 1 \right) \\
&= \begin{cases} \dot{x} \log \dot{x} - \dot{x} + 1 & \dot{x} \geq 0 \\ \infty & \textit{otherwise} \end{cases}
\end{aligned}$$

III.2 Transformations

While Mogulskii's theorem specifies that the processes $(Y^\epsilon)_{\epsilon>0}$—by design—have independent increments, we may use the contraction principle [DZ10, Theorem 4.2.1] to develop a large deviation principle for families of processes with state-dependent increments. Indeed, by introducing a continuous map F, we will have a large deviation principle for the family $(F_\# \mathrm{P}^\epsilon)_{\epsilon>0}$ of measures $F_\# \mathrm{P}^\epsilon$ associated with respective quantities $F(Y^\epsilon)$, and F can be chosen so that each $F(Y^\epsilon)$ is a process with state-dependent increments. Seeing as this section serves as a survey for intuition on rate functions, we will digress from discussing the specifics of continuity of F on restricted spaces and/or exponentially equivalent families in our examples below.

Example III.2.1 (Diffusions). *We can leverage Example III.1.3 to a family of processes $(\epsilon X^\epsilon)_{\epsilon>0}$,*

$$\epsilon X^\epsilon = \beta(\epsilon X^\epsilon) \bullet \ell + \sqrt{\epsilon} \sigma(\epsilon X^\epsilon) \bullet W, \tag{III.2.2}$$

where the drift $\beta : \mathbb{V} \to \mathbb{V}$ and diffusion $\alpha = \sigma\sigma^* : \mathbb{V} \to \mathbb{L}(\mathbb{V})$ are bounded and Lipschitz and α is invertible. The details of this result, attributed to Freidlin-Wentzel [DZ10, Theorems

5.6.3 and 5.6.7], are rather complicated, so we will explain a heuristic. Having a map $F_{\beta,\alpha}$ *which implicitly solves the equation,*

$$F_{\beta,\alpha}(\omega) = \xi, \quad \xi(t) = \beta(\xi) \bullet \ell_t + \sigma(\xi) \bullet \omega_t,$$

will make $F_\beta(\sqrt{\epsilon} W) = \epsilon X^\epsilon$ *for each* $\epsilon > 0$*, so the contraction principle states that the distributions of* $(\epsilon X^\epsilon)_{\epsilon>0}$ *satisfy a large deviation principle in which the rate function* I_X *is derived from that* I_W *from Example III.1.3.*

$$I_X(\xi) \coloneqq \inf \Big\{ I_W(\omega) : F_\beta(\omega) = \xi \Big\},$$
$$I_W(\omega) \coloneqq \begin{cases} \int_0^\infty \frac{1}{2} \big|\dot{\omega}(t)\big|^2 \mathrm{d}t & \omega(0) = 0, \ \omega \in \mathbb{A}([0,\infty), \mathbb{V}), \\ \infty & \textit{otherwise} \end{cases}$$

When $F_\beta(\omega) = \xi$*, equation (III.2.2) tells us* $\dot{\omega} = \sigma(\xi)^{-1}\big(\dot{\xi} - \beta(\xi)\big)$*, and so on the*

$$\begin{aligned} I_X(\xi) &= \int_0^\infty \frac{1}{2} \Big| \sigma\big(\xi(t)\big)^{-1} \Big(\dot{\xi}(t) - \beta\big(\xi(t)\big)\Big) \Big| \mathrm{d}t \\ &= \int_0^\infty \frac{1}{2} \Big\langle \Big(\dot{\xi}(t) - \beta\big(\xi(t)\big)\Big), \alpha\big(\xi(t)\big)^{-1} \Big(\dot{\xi}(t) - \beta\big(\xi(t)\big)\Big) \Big\rangle \mathrm{d}t \end{aligned}$$

Note that this result does not apply to the general class of affine diffusions, for β, α *are generally not bounded or Lipschitz, and* α *need not be invertible. However, [?]—a paper which inspires parts of our proof—first proved that affine diffusions with special characteristics* (β, α) *satisfy a large deviation principle with rate function similar to that above. Our rate function (??) from Theorem ?? immediately resolves an identical representation.*

$$\begin{aligned} \Lambda^*(\dot{x}, x) &= \sup_{u \in \mathbb{V}} \Big(\langle u, \dot{x} \rangle - \big\langle u, \beta(x) \big\rangle - \frac{1}{2} \big\langle u, \alpha(x) u \big\rangle \Big) \\ &= \begin{cases} \frac{1}{2} \Big\langle \big(\dot{x} - \beta(x)\big), \alpha(x)^\dagger \big(\dot{x} - \beta(x)\big) \Big\rangle & \dot{x} - \beta(x) \in \operatorname{image} \alpha(x) \\ \infty & \textit{otherwise} \end{cases} \end{aligned}$$

Above, $a^\dagger \in \mathbb{L}(\mathbb{V})$ *denotes the pseudoinverse of* $a \in \mathbb{L}(\mathbb{V})$*.*

III.3 Coupling

1. **Mogulskii's theorem.** Purpose of section is to familiarize with rate functions we already have and set the stage for how we operate with our theorem.
   (a) Cite Mogulskii's theorem.
   (b) Brownian motion of our regime: achieved with Mogulskii's theorem with Normal increments and exponential tightness (Schilder).
   (c) Poisson process of our regime: achieved with Mogulskii's theorem with Poisson increments and use exponential tightness.

2. **Simple contraction maps.**

   (a) Friedlin Wentzel (not so much our regime): use contraction mapping principles
   (b) Birth rate process: Can we similarly contraction map Poisson to get this?
   (c) Extension of Freidlin Wentzel to affine: KK

3. **Coupling states.** Indicate that when contraction mappings are not sufficient, we may *couple correlated states*, in the sense of looking at LDPs of joint processes.
   (a) Compound Poisson: Two *sources* of randomness; the arrivals and the jump sizes. Appeal to Duffy results for heuristical calculations.
   (b) Compound linear Hawkes: Similarly two *sources* of randomness. Duffy also gives us the calculations. Note on Zhu paper for *sidestep*; less general jumps, more general nonlinear relationship of arrivals.
   (c) The jumps of a general jump-diffusion do not have a well-posed notion of arrivals and jump sizes; we thus turn our focus to locally countable jump-diffusions, in which the three *sources* of randomness are the continuous local martingale, the arrival times, and the jump sizes. Note how this is discussed in next section.

4. **Locally countable affine processes.** Perform the necessary calculus and proceed to show our general formulation.
   (a) State result in numerous flavors, depending on which *base* quantities in which we choose to focus the large deviations.

$$\begin{aligned} X &= \beta(X) \bullet \ell + X^{\mathrm{c}} + \mathrm{id}_{\mathbb{V}} * \tilde{q}^X \\ N^X &= 1 * q^X \\ V^X &= \mathrm{id}_{\mathbb{V}} * q^X \end{aligned}$$

      i. **overdetermined flavor.** $(X, X^{\mathrm{c}}, N^X, V^X)$ produces an overdetermined system which requires another condition for $I(\xi, \omega, \eta, \gamma)$ to be finite.

$$\dot{\xi}(t) = \beta\big(\xi(t)\big) + \dot{\omega}(t) + \dot{\eta}(t)\dot{\gamma}(t)$$

      However, the rate function is very simple to understand.
      ii. **determine-continuous-noise flavor.** Normal term gets messy
      iii. **determine-arrivals flavor.** Poisson and jump-term-denominator gets messy
      iv. **determine-jumps flavor.** This is the one we have already presented; the jump-term-numerator gets messy.
   (b) Discuss how the deviations of X from the dynamical system $X = \beta(X) \bullet \ell$ are imposed from *continuous deviations* X^{c} and *discontinuous deviations* $\mathrm{id}_{\mathbb{V}} * \tilde{q}^X$.
   (c) Discuss how four quantities $X, X^{\mathrm{c}}, N^X, V^X$ heuristically relate in simple infinitesimal equality $(X, X^{\mathrm{c}}, N^X, V^X) \approx (\xi, \omega, \eta, \gamma)$.

$$\dot{\xi}(t) = \beta\big(\xi(t)\big) + \dot{\omega}(t) + \dot{\eta}(t) \cdot \dot{\gamma}(t)$$

   (d) Each of the primitive deviations have a simple analogy, when we think of infinitesimals.

$\dot{\omega}(t)$ *normal deviations of covariance* $\alpha\big(\xi(t)\big)$

$\dot{\eta}(t)$	*Poisson deviations of rate* $\lambda\big(\xi(t)\big)$
$\dot{\gamma}(t)$	*jump deviations from* $\kappa\big(\xi(t), \mathrm{d}v\big)$
$\xi(t) = \beta\big(\xi(t)\big) + \dot{\omega}(t) + \dot{\eta}(t) \cdot \dot{\gamma}(t)$	*all combined deviations*

(e) Think of results from first section in this regard.

  i. birth is $\dot{\xi}(t) = \dot{\eta}(t)$, so we only need $\xi \approx X$.

  ii. diffusion is $\dot{\xi}(t) = \beta\big(\xi(t)\big) + \dot{\omega}(t)$, and so we only need $\xi \approx X$ and rate function includes $\xi(t) - \beta\big(\xi(t)\big)$ where $\dot{\omega}$ is.

  iii. compound Poisson is $\dot{\xi}(t) = \dot{\eta}(t) \cdot \dot{\gamma}(t)$, so we choose one of the following pairs $(\xi, \eta) \approx (X, N^X)$, $(\xi, \gamma) \approx (X, V^X)$, or $(\eta, \gamma) \approx (N^X, V^X)$.

  iv. compound linear Hawkes is $\dot{\xi}(t) = \beta\big(\xi(t)\big) + \dot{\eta}(t) \cdot \dot{\gamma}(t)$, so we can choose $(\xi, \eta) \approx (X, N^X)$ or $(\xi, \dot{\gamma}) \approx (X, V^X)$.

Appendix A

Jump-diffusions

TODO:

- Motivate why I chose to put this in the appendix. Big point: I want to resolve abstractions and rigor of [JS03] to the digestible notions of special jump-diffusions.
- Point to the various papers we use that do not consolidate a similar set of assumptions.

In order to discuss jump-diffusions on a finite-dimensional real vector space, one must have a decent understanding of semimartingales. A great text for a comprehensive study of this is [JS03], which we will refer to in our proofs. In terms of notational differences, we choose our probability space $(\Omega, \Sigma, \mathrm{P})$ and filtration $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$, where $\mathscr{F}_\infty \subseteq \Sigma$ denotes the joined space. Furthermore, we do not explicitly write processes to take values in $\mathbb{R}^d$, but rather some vector space $\mathbb{V}$ with dimension $d \coloneqq \dim \mathbb{V}$ and inner-product $\langle \cdot, \cdot \rangle$. Surely—due to our isometric isomorphism $\mathbb{V} \equiv \mathbb{R}^d$—any componentwise or linear notion, such as integration or differentiation may be taken as equivalent. Furthermore, we sometimes specify that a stochastic process X has a Borel state space $\mathbb{X} \subseteq \mathbb{V}$, as this is the case when studying affine processes. We find it important to highlight the following important notation of objects introduced in [JS03, Chapters I-II].

- Given $(\mathrm{P}, \mathscr{F})$ locally square-integrable martingales $M, N : \Omega \times \mathbb{R}_+ \to \mathbb{R}$, denote $\langle M, N \rangle$ the predictable quadratic covariation.
- Given $H, X : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ with H being $\mathscr{F}$ predictable and $(\mathrm{P}, \mathscr{F})$ locally bounded and X a $(\mathrm{P}, \mathscr{F})$ semimartingale, denote the stochastic integral process as follows.

  $$H \bullet X_t = \int_0^t H_s \mathrm{d}X_s$$

  We may lift this concept componentwise and linearly. This allows us to choose the codomains of H, X to various combinations of $\mathbb{V}$ and $\mathbb{L}(\mathbb{V}, \mathbb{W})$ when evaluating $H \bullet X$, so long as such a combination allows for $H_t \cdot X_t$ to make sense.
- Denote $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ the identity map to allow a concise notation for Lebesgue integration.

  $$H \bullet \ell_t = \int_0^t H_s \mathrm{d}s$$

- Given a random measure $q : \Omega \times \mathscr{B}(\mathbb{R}_+ \times \mathbb{V}) \to [0, \infty]$, denote the stochastic integral process against some suitably integrable process $H : \Omega \times \mathbb{R}_+ \times \mathbb{V} \to \mathbb{R}$ as follows.

$$H * q_t = \int_{[0,t]\times\mathbb{V}} H_s(v) q(\mathrm{d}s, \mathrm{d}v)$$

Denote its $(\mathrm{P}, \mathscr{F})$ predictable projection by $\hat{q}$ and the compensated measure $\tilde{q} = q - \hat{q}$. Also denote $H * \tilde{q}$ the compensated local martingale process for suitable $H \in G_{\mathrm{loc}}(q)$, as constructed in [JS03, Definition II.1.27]. Lift these integration notions to vector-valued H componentwise. Instead of choosing a canonical variable for integrating expressions in this form, we use the identity maps $\mathrm{id}_{\mathbb{V}}$ or ℓ.

$$f(\ell, \mathrm{id}_{\mathbb{V}}) * q_t = \int_{[0,t]\times\mathbb{V}} f(s, v) q(\mathrm{d}s, \mathrm{d}v)$$

- Given $(\mathrm{P}, \mathscr{F})$ semimartingales $X, Y : \Omega \times \mathbb{R}_+ \to \mathbb{R}$, denote $[X, Y]$ the quadratic covariation.

- Given a semimartingale X, denote X^{c} its continuous local martingale component and q^X its jump measure.

A.1 Formulation

As in [JS03, Definition III.2.18], a $(\mathrm{P}, \mathscr{F})$ jump-diffusion X on state space $(\mathbb{X}, \mathscr{B}(\mathbb{X}))$ is a $(\mathrm{P}, \mathscr{F})$ semimartingale in which the χ-characteristics $(B^\chi, A, \hat{q}^X)$ have the following decompositions.

$$B^\chi_t = \int_0^t \beta^\chi(X_s)\mathrm{d}s, \quad A_t = \int_0^t \alpha(X_s)\mathrm{d}s, \quad \hat{q}^X(\mathrm{d}s, \mathrm{d}v) = \mu(X_s, \mathrm{d}v)\mathrm{d}s, \tag{A.1.1}$$

where the functions have the following properties.

- $\beta^\chi : \mathbb{X} \to \mathbb{V}$ is Borel measurable, $\beta^\chi \in \mathscr{B}(\mathbb{X})/\mathscr{B}(\mathbb{V})$.

- $\alpha : \mathbb{X} \to \mathbb{L}(\mathbb{V})$ is Borel measurable, $\alpha \in \mathscr{B}(\mathbb{X})/\mathscr{B}(\mathbb{L}(\mathbb{V}))$, and $\alpha(x)$ is self-adjoint and non-negative for each $x \in \mathbb{X}$.

- $\mu : \mathbb{X} \times \mathscr{B}(\mathbb{V}) \to [0, \infty]$ is a transition kernel from $\mathbb{X}$ to $\mathbb{V}$, and it satisfies the following properties for each $x \in \mathbb{X}$.

$$\mu\big(x, \{0\}\big) = 0, \quad \int_{\mathbb{V}} \big(1 \wedge |v|^2\big)\mu(x, \mathrm{d}v) < \infty \tag{A.1.2}$$

In other words, our jump-diffusion X has the following canonical semimartingale representation (see [JS03, Theorem II.2.34] for definition).

$$\begin{aligned} X &= X_0 + \beta^\chi(X) \bullet \ell + X^{\mathrm{c}} + \chi * \tilde{q}^X + (\mathrm{id}_{\mathbb{V}} - \chi) * q^X \\ \langle X^{\mathrm{c},i}, X^{\mathrm{c},j} \rangle &= \alpha_{ij}(X) \bullet \ell \\ \hat{q}^X(\mathrm{d}s, \mathrm{d}v) &= \mu(X_s, \mathrm{d}v)\mathrm{d}s \end{aligned} \tag{A.1.3}$$

Remark A.1.4. *(a) Note that we differ slightly from the definition we reference by imposing a time-homogeneity formulation. There is no loss of generality in doing so, because we may always extend the state to $\mathbb{R}_+ \times \mathbb{X}$ via $\hat{X}_t = (t, X_t)$.*

(b) Note that (A.1.1) can be written concisely by using the identity ℓ on $\mathbb{R}_+$.

$$B_t^\chi = \beta^\chi(X) \bullet \ell_t, \quad A_t = \alpha(X) \bullet \ell_t, \quad \hat{q}^X([0,t], \mathrm{d}v) = \mu(X, \mathrm{d}v) \bullet \ell_t$$

(c) If we have a jump-diffusion with χ-characteristics in (A.1.1), we call $(\beta^\chi, \alpha, \mu)$ the differential χ-characteristics. *We see from (A.1.3) that β^χ and $\beta^{\hat{\chi}}$ relate between different truncation functions $\chi, \hat{\chi}$ with the simple identity.*

$$\beta^{\hat{\chi}}(x) = \beta^\chi(x) + \int_{\mathbb{V}} \big(\hat{\chi}(v) - \chi(v)\big)\mu(x, \mathrm{d}v) \tag{A.1.5}$$

(d) The conditions on $\alpha(x)$ and $\mu(x, \mathrm{d}v)$ are immediate consequences of (A.1.1). For the most general setting, see the corresponding result for any *semimartingale, in [JS03, Proposition II.2.9].*

Example A.1.6. *Fix a probability space $(\Omega, \Sigma, \mathrm{P})$ and filtration $\mathscr{F} = (\mathscr{F})_{t\geq 0}$.*

Just as with $(\mathbb{R}^d, \mathscr{B}(\mathbb{V}))$, we say that W is a standard $(\mathrm{P}, \mathscr{F})$ Brownian motion on $(\mathbb{V}, \mathscr{B}(\mathbb{V}))$ if it is a continuous $(\mathrm{P}, \mathscr{F})$ martingale with predictable quadratic covariation as follows.

$$\langle W^i, W^j \rangle_t = \begin{cases} t & i = j \\ 0 & \textit{otherwise} \end{cases}$$

It is clear that W is a $(\mathrm{P}, \mathscr{F})$ jump-diffusion with differential χ-characteristics $(0, \alpha, 0)$, where $\alpha(x) = \mathrm{id}_{\mathbb{V}}$ for all $x \in \mathbb{X}$.

Similarly, we say that p is a standard $(\mathrm{P}, \mathscr{F})$ Poisson random measure on $\mathscr{B}(\mathbb{R}_+ \times \mathbb{V})$ if its $(\mathrm{P}, \mathscr{F})$ predictable projection is the Lebesgue measure $\hat{p}(\mathrm{d}s, \mathrm{d}v) = \mathrm{d}s \otimes \mathrm{d}v$ (identifying measures on $\mathscr{B}(\mathbb{R}^d)$ as those on $\mathscr{B}(\mathbb{V})$). By [JS03, Theorem II.4.8], this p is the same as a Poisson point process with Lebesgue intensity, which has infinitely many jumps on any nonempty interval of time. The accumulated jumps $\mathrm{id}_{\mathbb{V}} * p$ form a $(\mathrm{P}, \mathscr{F})$ jump-diffusion with parameters as follows.

$$\beta^\chi(x) = \int_{\mathbb{V}} \chi(v)\mathrm{d}v, \quad \alpha(x) = 0, \quad \mu(x, \mathrm{d}v) = \mathrm{d}v,$$

because we have the following decomposition.

$$\begin{aligned} \mathrm{id}_{\mathbb{V}} * p &= \chi * p + (\mathrm{id}_{\mathbb{V}} - \chi) * p \\ &= \chi * \hat{p} + \chi * \tilde{p} + (\mathrm{id}_{\mathbb{V}} - \chi) * p \\ &= \beta^\chi \bullet \ell + \chi * \tilde{p} + (\mathrm{id}_{\mathbb{V}} - \chi) * p \end{aligned}$$

Note that the infinite activity of p means that the last term cannot be compensated.

We will see at the end of this section that these two objects W and p are the fundamental building blocks of all jump-diffusions.

The following Lemma will be repeatedly used as a shortcut of Itô's formula and various identities that always apply with jump-diffusions.

Lemma A.1.7. *Let X be a jump-diffusion with differential χ-characteristics $(\beta^\chi, \alpha, \mu)$ and $f \in \mathbb{C}^2(\mathbb{V}, \mathbb{R})$. The composition $f(X)$ has the following semimartingale representation.*

$$\begin{aligned} f(X_t) = f(X_0) &+ \Big(\mathrm{D}f(X)\cdot\beta^\chi(X)\Big)\bullet\ell_t + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f(X)\circ\alpha(X)\Big)\bullet\ell_t + \mathrm{D}f(X_-)\bullet X^{\mathrm{c}} \\ &+ \Big(\mathrm{D}f(X_-)\cdot\chi\Big) * \tilde{q}^X_t + \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\chi\Big) * q^X_t \end{aligned}$$

Proof. Apply Itô's formula [JS03, Theorem I.4.57] and use the predictable covariation identity in (A.1.3) to get the following.

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{i=1}^d \mathrm{D}_i f(X_-)\bullet X^i_t + \frac{1}{2}\sum_{i,j=1}^d \mathrm{D}_{ij} f(X_-)\bullet\langle X^{\mathrm{c},i}, X^{\mathrm{c},j}\rangle_t \\ &\quad + \sum_{0<s\le t}\Big(f(X_s) - f(X_{s-}) - \sum_{i=1}^d \mathrm{D}f_i(X_{s-})\Delta X_s\Big) \\ &= f(X_0) + \mathrm{D}f(X_-)\bullet X_t + \frac{1}{2}\sum_{i,j=1}^d \mathrm{D}_{ij} f(X_-)\bullet\big(\alpha_{ij}(X)\bullet\ell\big)_t \\ &\quad + \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\mathrm{id}_{\mathbb{V}}\Big) * q^X_t \end{aligned}$$

Using the iterated stochastic integral formula [JS03, Remark I.4.37], we may simplify the above equation to the following.

$$\begin{aligned} f(X_t) = f(X_0) &+ \mathrm{D}f(X_-)\bullet X_t + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}_{ij} f(X_-)\circ\alpha(X)\Big)\bullet\ell_t \\ &+ \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\mathrm{id}_{\mathbb{V}}\Big) * q^X_t \end{aligned}$$

Now substitute our representation of (A.1.3) and repeat the iterated stochastic integral to get the following.

$$\begin{aligned} f(X_t) &= f(X_0) + \mathrm{D}f(X_-)\bullet\big(X_0 + \beta^\chi(X)\bullet\ell + X^{\mathrm{c}} + \chi * \tilde{q}^X + (\mathrm{id}_{\mathbb{V}} - \chi) * q^X\big)_t \\ &\quad + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f(X_-)\circ\alpha(X)\Big)\bullet\ell_t + \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\mathrm{id}_{\mathbb{V}}\Big) * q^X_t \\ &= f(X_0) + \Big(\mathrm{D}f(X_-)\cdot\beta^\chi(X)\Big)\bullet\ell_t + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f(X_-)\circ\alpha(X)\Big)\bullet\ell_t + \mathrm{D}f(X_-)\bullet X^{\mathrm{c}} \\ &\quad + \mathrm{D}f(X_-)\bullet\big(\chi * \tilde{q}^X\big)_t + \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\chi\Big) * q^X_t \end{aligned}$$

Furthermore, since $X_- = X$ on all but a countable amount of jumps, we may rewrite the Lebesgue integrals.

(A.1.8)

$$\begin{aligned} f(X_t) = f(X_0) &+ \Big(\mathrm{D}f(X)\cdot\beta^\chi(X)\Big)\bullet\ell_t + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f(X)\circ\alpha(X)\Big)\bullet\ell_t + \mathrm{D}f(X_-)\bullet X^{\mathrm{c}} \\ &+ \mathrm{D}f(X_-)\bullet\big(\chi * \tilde{q}^X\big)_t + \Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-)\cdot\chi\Big) * q^X_t \end{aligned}$$

For the remaining equality, we construct localizing sequence $(T_n)_{n\in\mathbb{N}}$ of $\mathscr{F}$ stopping times,

(A.1.9) $$T_n(\omega) \coloneqq \inf\big\{t > 0 : X_t(\omega) > n\big\} \wedge n, \quad \omega\in\Omega,\ n\in\mathbb{N},$$

to see that $\mathrm{D}f(X_-)$ is $(\mathrm{P}, \mathscr{F})$ locally bounded.

$$\left|\mathrm{D}f(X^{T_n}_{s-})\right| \leq \sup_{|x|\leq n} \left|\mathrm{D}f(x)\right|$$

Thus, by [JS03, Proposition II.1.30], we may rewrite the following.

$$\mathrm{D}f(X_-) \bullet \left(\chi * \tilde{q}^X\right)_t = \left(\mathrm{D}f(X_-) \cdot \chi\right) * \tilde{q}^X_t,$$

which when substituted into (A.1.8) gives us our desired identity.

In the above lemma, the final term in the semimartingale decomposition of $f(X)$ is typically not able to be compensated into a local martingale. If we did have local integrability of the following quantity,

$$\left| f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) + \mathrm{D}f(X_-) \cdot \chi \right| * \hat{q}^X,$$

then by [JS03, Proposition II.1.28] we could rewrite $f(X)$ into a canonical special semimartingale decomposition.

$$\begin{aligned} &f(X_t) = f(X_0) + \mathcal{L}f(X) \bullet \ell_t + \mathrm{D}f(X_-) \bullet X^{\mathrm{c}} + \left(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-)\right) * \tilde{q}^X_t \\ &\mathcal{L}f(x) \coloneqq \mathrm{D}f(x) \cdot \beta^\chi(x) + \frac{1}{2} \operatorname{tr}\left(\mathrm{D}^2 f(x) \circ \alpha(x)\right) \\ &\qquad\qquad + \int_{\mathbb{V}} \Big(f(x+v) - f(x) - \mathrm{D}f(x) \cdot \chi(v)\Big)\mu(x, \mathrm{d}v) \end{aligned} \tag{A.1.10}$$

So long as f is bounded, we can guarantee this special semimartingale property.

Proposition A.1.11. *Let X and f as in Lemma A.1.7, and further impose f is bounded. Then the composition $f(X)$ is a special semimartingale with the decomposition as in (A.1.10).*

Proof. Seeing as f is bounded, [JS03, Lemma I.4.24] tells us that $f(X)$ is a special semimartingale. By [JS03, Proposition I.4.23], it is then the case that the following term is locally integrable.

$$\Big(f(X_- + \mathrm{id}_{\mathbb{V}}) - f(X_-) - \mathrm{D}f(X_-) \cdot \chi\Big) * q^X_t$$

By our discussion above, this suffices to conclude (A.1.10).

This operator $\mathcal{L}$ in (A.1.10) gives a nice closed from for suitable $f(X)$, and so we reserve it the term of *generator* associated with X. Note that we do not mark dependence on χ, as any other truncation function $\hat{\chi}$ will produce the same operator; see Remark A.1.4(c) and note that the displacement from β^χ and $\beta^{\hat{\chi}}$ would be the same as that in the integral term. One particular setting in which this result is useful is establishing a Lévy-Khintchine formula for jump-diffusions.

Proposition A.1.12. *Fix a jump-diffusion X with differential χ-characteristics $(\beta^\chi, \alpha, \mu)$. Then, for each $u \in \mathrm{i}\mathbb{V}$, the process $\exp\left(\langle u, X\rangle - \Lambda(u, X) \bullet \ell\right)$ is a complex-valued $(\mathrm{P}, \mathscr{F})$ local martingale, where $\Lambda : \mathrm{i}\mathbb{V} \times \mathbb{X} \to \mathbb{R}$ is the associated* Lévy-Khintchine map*.*

$$\Lambda(u, x) = \left\langle u, \beta^\chi(x)\right\rangle + \frac{1}{2}\left\langle u, \alpha(x)\right\rangle + \int_{\mathbb{V}} \left(e^{\langle u, v\rangle} - 1 - \langle u, \chi(v)\rangle\right)\mu(x, \mathrm{d}v),$$

Proof. For a fixed $u \in \mathrm{i}\mathbb{V}$, note that the map f_u, defined by $f_u(v) = \exp\langle u, v\rangle$ is bounded. Thus, by Proposition A.1.11, we have

$$f_u(X_t) = f_u(X_0) + \mathcal{L}f_u(X) \bullet \ell_t + M_t,$$

where M is a $(\mathrm{P}, \mathscr{F})$ local martingale. Observe that the partial derivatives of f are as follows,

$$\mathrm{D}_i f_u(x) = f_u(x) u_i, \qquad \mathrm{D}_{ij} f_u(x) = f_u(x) u_i u_j, \tag{A.1.13}$$

so we have the following equation.

$$\begin{aligned}
\mathcal{L}f_u(x) &= \mathrm{D}f_u(x) \cdot \beta^\chi(x) + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f_u(x) \circ \alpha(x)\Big) \\
&\quad + \int_{\mathbb{V}} \Big(f_u(x+v) - f_u(x) - \mathrm{D}f(x) \cdot \chi(v)\Big)\mu(x, \mathrm{d}v) \\
&= f_u(x)\big\langle u, \beta^\chi(x)\big\rangle + \frac{1}{2} f_u(x)\big\langle u, \alpha(x)u\big\rangle + f_u(x)\int_{\mathbb{V}} \Big(f_u(v) - 1 - \big\langle u, \chi(v)\big\rangle\Big)\mu(x, \mathrm{d}v) \\
&= f_u(x) \cdot \Lambda(u, x)
\end{aligned}$$

Denoting $A = f_u(X) = \exp\langle u, X\rangle$ and $B = \exp\big(-\Lambda(u, X)\bullet\ell\big)$, we now use the fact that B is $\mathscr{F}$ predictable and of finite-variation, so [JS03, Proposition I.4.49(b)] gives us the following.

$$\begin{aligned}
&\exp\Big(\langle u, X\rangle - \Lambda(u, X) \bullet \ell\Big) \\
&= A_t B_t \\
&= A_0 B_0 + A_- \bullet B_t + B \bullet A_t \\
&= \exp\langle u, X_0\rangle + A_- \bullet \Big(\big(-B \cdot \Lambda(u, X)\big) \bullet \ell\Big)_t + B \bullet \Big(f_u(X_0) + \mathcal{L}f_u(X) \bullet \ell + M\Big)_t \\
&= \exp\langle u, X_0\rangle - \Big(A \cdot B \cdot \Lambda(u, X)\Big) \bullet \ell_t + \Big(B \cdot f_u(X) \cdot \Lambda(u, X)\Big) \bullet \ell_t + B \bullet M_t \\
&= \exp\langle u, X_0\rangle + B \bullet M_t
\end{aligned}$$

This identity and [JS03, Remark I.4.34(b)] concludes the proof.

It turns out that each of the preceding results is sufficient in characterizing a semimartingale X as a jump-diffusion.

Theorem A.1.14. *The following statements are equivalent for a stochastic process X on state space $(\mathbb{X}, \mathscr{B}(\mathbb{X}))$.*

(a) X is a $(\mathrm{P}, \mathscr{F})$ jump-diffusion with differential χ-characteristics $(\beta^\chi, \alpha, \mu)$.

(b) For each bounded $f \in \mathrm{C}^2(\mathbb{V}, \mathbb{R})$, the process $f(X_t) - \mathcal{L}f(X_t) \bullet \ell_t$ is a $(\mathrm{P}, \mathscr{F})$ local martingale, where

$$\mathcal{L}f(x) \coloneqq \mathrm{D}f(x) \cdot \beta^\chi(x) + \frac{1}{2}\operatorname{tr}\Big(\mathrm{D}^2 f(x) \circ \alpha(x)\Big) + \int_{\mathbb{V}} \Big(f(x+v) - f(x) - \mathrm{D}f(x) \cdot \chi(v)\Big)\mu(x, \mathrm{d}v)$$

(c) For each $u \in \mathrm{i}\mathbb{V}$, the process $\exp\Big(\langle u, X\rangle - \Lambda(u, X) \bullet \ell\Big)$ is a $(\mathrm{P}, \mathscr{F})$ local martingale, where Λ is our Lévy-Khintchine map.

$$\Lambda(u, x) = \big\langle u, \beta^\chi(x)\big\rangle + \frac{1}{2}\big\langle u, \alpha(x)\big\rangle + \int_{\mathbb{V}} \big(e^{\langle u, v\rangle} - 1 - \langle u, \chi(v)\rangle\big)\mu(x, \mathrm{d}v),$$

(d) Denoting $(\mathrm{P}_x)_{x\in\mathbb{X}}$ *the* P*-conditional distributions of* X *factored through the initial state* X_0 *and selecting Borel functions* σ, c *to satisfy,*

$$\begin{aligned} &\sigma : \mathbb{X} \to \mathbb{L}(\mathbb{V}) \quad \sigma\sigma^*(x) = \alpha(x) \\ &c : \mathbb{X}\times\mathbb{V} \to \mathbb{V} \quad \mu(x,\Gamma) = \int_{\mathbb{V}} 1_\Gamma\big(c(x,v)\big)\mathrm{d}v \end{aligned} \tag{A.1.15}$$

each P_x *is a solution to the equation associated with a standard Brownian motion* W *and Poisson random measure* p*, where* $\chi' = \mathrm{id}_{\mathbb{V}} - \chi$*.*

$$X_t = x + \beta^\chi(X)\bullet\ell_t + \sigma(X_-)\bullet W_t + \big(\chi\circ c(X_-,\mathrm{id}_{\mathbb{V}})\big) * \tilde{p}_t + \big(\chi'\circ c(X_-,\mathrm{id}_{\mathbb{V}})\big) * p_t$$

Proof. This is simply restating [JS03, Theorems II.2.42, II.2.49, and III.2.26] in terms of our identities from the previous propositions and lemmas. The choice of standard intensity $\mathrm{d}t\otimes\mathrm{d}v$ for the Poisson random measure is such that the jump factor $\mathrm{d}v$ satisfies the atom-free and infinite properties in [JS03, Remark III.2.28(3)].

Remark A.1.16. *In the final part above, the push-forward map* c *may put mass on* 0,

$$\int_{\mathbb{V}} 1_{\{0\}}\big(c(x,v)\big)\mathrm{d}v > 0,$$

to thin *or* delete *jumps coming from* p *(of which there are infinitely many). However, this contradicts the condition (A.1.2) that* $\mu(x,\{0\}) = 0$ *for all* $x\in\mathbb{X}$*. Explicitly, the push-forward in (A.1.15) happens on the space* $\mathbb{V}_0 \coloneqq \mathbb{V} - \{0\}$*,*

$$\mu(x,\Gamma) = \int_{\mathbb{V}} 1_\Gamma\big(c(x,v)\big)\mathrm{d}v, \quad \Gamma\in\mathscr{B}(\mathbb{V}_0)$$

to allow for such thinning.

A.2 Special jump-diffusions

We now turn our focus to $(\mathrm{P},\mathscr{F})$ jump-diffusions which are additionally *special* in the sense of them having a semimartingale decomposition in which the finite-variation term is predictable. When looking at the canonical representation of a jump-diffusion X with χ-characteristics (β^χ,α,μ), it is clear how to make this predictable.

$$\begin{aligned} X_t &= X_0 + \beta^\chi(X)\bullet\ell_t + X^{\mathrm{c}}_t + \chi * \tilde{q}^X + (\mathrm{id}_{\mathbb{V}} - \chi) * q^X \\ &= X_0 + \beta^\chi(X)\bullet\ell_t + (\mathrm{id}_{\mathbb{V}} - \chi) * \hat{q}^X + X^{\mathrm{c}}_t + \mathrm{id}_{\mathbb{V}} * \tilde{q}^X \\ &= X_0 + \left(\beta^\chi(X) + \int_{\mathbb{V}} \big(v - \chi(v)\big)\mu(X,\mathrm{d}v)\right)\bullet\ell_t + X^{\mathrm{c}}_t + \mathrm{id}_{\mathbb{V}} * \tilde{q}^X \end{aligned} \tag{A.2.1}$$

In such a case, it is nice to define the function $\beta : \mathbb{X}\to\mathbb{V}$,

$$\beta(x) \coloneqq \beta^\chi(x) + \int_{\mathbb{V}} \big(v - \chi(v)\big)\mu(x,\mathrm{d}v), \tag{A.2.2}$$

so that (A.2.1) may be simplified to a concise special semimartingale decomposition.

$$X_t = X_0 + \beta(X)\bullet\ell + X^{\mathrm{c}} + \mathrm{id}_{\mathbb{V}} * \tilde{q}^X_t$$

We call the triplet (β, α, μ) that results from (A.2.2) the *special differential characteristics* and its components β, α, μ the *drift*, *diffusion*, and *jump kernel*, respectively.

The calculus of (A.2.1) begs the question that $(\mathrm{id}_{\mathbb{V}} - \chi) * q^X$ can be compensated which is not generally the case—otherwise, the term *special* would be a misnomer! The next result specifies conditions on which we may perform the above calculus.

Lemma A.2.3. *Let X be a $(\mathrm{P}, \mathscr{F})$ jump-diffusion with differential χ-characteristics $(\beta^\chi, \alpha, \mu)$, such that μ satisfies the following condition.*

$$x \mapsto \int_{\mathbb{V}} \big|v - \chi(v)\big| \mu(x, \mathrm{d}v) \text{ is bounded on compact subsets}$$

Then, X is special with drift β as in (A.2.2).

Proof. By choosing a $\mathscr{F}$ localizing sequence $(T_n)_{n \in \mathbb{N}}$ as in (A.1.9), our hypothesis gives us the following integrability.

$$\mathrm{E}_{\mathrm{P}} \big|\mathrm{id}_{\mathbb{V}} - \chi\big| * \hat{q}^X_{T_n} = \mathrm{E}_{\mathrm{P}} \int_0^{T_n} \int_{\mathbb{V}} \big|v - \chi(v)\big| \mu(X_t, \mathrm{d}v) \mathrm{d}t \le n \cdot \sup_{|x| \le n} \int_{\mathbb{V}} \big|v - \chi(v)\big| \mu(x, \mathrm{d}v) < \infty$$

Now, [JS03, Proposition II.1.28] allows us to compensate as we did in (A.2.1)

Seeing as $(\mathrm{id}_{\mathbb{V}} - \chi) * q^X$ may be compensated for special jump-diffusions X, all the characterizing objects of Theorem A.1.14 may be rewritten in terms of our drift β—effectively, χ becomes the identity.

$$\mathcal{L}f(x) \coloneqq \mathrm{D}f(x) \cdot \beta(x) + \frac{1}{2} \operatorname{tr}\Big(\mathrm{D}^2 f(x) \circ \alpha(x)\Big) + \int_{\mathbb{V}} \Big(f(x+v) - f(x) - \mathrm{D}f(x) \cdot v\Big) \mu(x, \mathrm{d}v)$$
$$\Lambda(u, x) = \big\langle u, \beta(x)\big\rangle + \frac{1}{2}\big\langle u, \alpha(x)\big\rangle + \int_{\mathbb{V}} \big(e^{\langle u, v\rangle} - 1 - \langle u, v\rangle\big) \mu(x, \mathrm{d}v),$$
$$X_t = x + \beta(X) \bullet \ell_t + \sigma(X_-) \bullet W_t + c(X_-, \mathrm{id}_{\mathbb{V}}) * \tilde{p}_t$$

A.3 Locally countable jump-diffusions

We see that a $(\mathrm{P}, \mathscr{F})$ jump-diffusion X is special if the accumulated large jumps $(\mathrm{id}_{\mathbb{V}} - \chi) * q^X$ may be compensated. To this end, being special is a condition on the jumps *away* from the origin. We now turn our focus to jump-diffusions X in which the jumps *near* the origin behave nicely. For any jump-diffusion X, we may count the jumps with the jump process N^X.

$$N^X_t \coloneqq \sum_{0 < s \le t} 1_{\Delta X_s \neq 0} = 1 * q^X_t \tag{A.3.1}$$

For many jump diffusions, it may be the case that we P-almost-surely have $N^X_t = \infty$ for all $t > 0$. We way that X has $(\mathrm{P}, \mathscr{F})$ *locally countable*, so long as N^X is $(\mathrm{P}, \mathscr{F})$ locally integrable. Below, we state how to verify this using the differential characteristics.

Lemma A.3.2. *Fix a $(\mathrm{P}, \mathscr{F})$ jump-diffusion X with differential χ-characteristics $(\beta^\chi, \alpha, \mu)$ satisfying*

$$x \mapsto \mu\big(x, \mathbb{V}\big) \text{ is bounded on compact sets,}$$

then X is locally countable. Moreover, we may define $\lambda : \mathbb{X} \to \mathbb{R}_+$ and probability kernel $\kappa : \mathbb{X} \times \mathscr{B}(\mathbb{V}) \to [0,1]$ by the following factoring.

$$\lambda(x) \coloneqq \mu(x, \mathbb{V}), \quad \mu(x, \mathrm{d}v) =: \lambda(x)\kappa(x, \mathrm{d}v)$$

Also, N has $(\mathrm{P}, \mathscr{F})$ intensity $\lambda(X)$.

Proof. Select the sequence $(T_n)_{n\in\mathbb{N}}$ as in (A.1.9). Note now that, since the constant function 1 is predictable,

$$\mathrm{E_P} N^X_{T_n} = \mathrm{E_P} 1 * q^X_{T_n} = \mathrm{E_P} 1 * \hat{q}^X_{T_n} = \mathrm{E_P} \int_0^{T_n} \mu(X_t, \mathbb{V})\mathrm{d}t \leq n \cdot \sup_{|x|\leq n} \mu(x, \mathbb{V}) < \infty$$

This means that N^X is locally integrable, making X locally countable. Moreover, by [JS03, Theorem II.1.8],

$$N^X - \int_0^t \lambda(X_s)\mathrm{d}s = 1 * q^X - \int_0^t \int_{\mathbb{V}} \mu(X_s, \mathrm{d}v)\mathrm{d}s = 1 * q^X - 1 * \hat{q}^X$$

is a $(\mathrm{P}, \mathscr{F})$ local martingale, which finishes the proof.

Remark A.3.3. *(a) Such objects λ, κ always exist with our assumption of the Lemma. Seeing as μ is a transition kernel from $(\mathbb{X}, \mathscr{B}(\mathbb{X}))$ to $(\mathbb{V}, \mathscr{B}(\mathbb{V}))$, we have our desired measurability.*

$$\lambda \coloneqq \mu(\cdot, \mathbb{V}) \in \mathscr{B}(\mathbb{X})/\mathscr{B}(\mathbb{R}_+)$$

Constructing κ should be obvious algebra, so long as we have no zero measures; otherwise, we may define

$$\kappa(x, \Gamma) \coloneqq \delta_{e_1}(\Gamma) \cdot 1_{\lambda^{-1}\{0\}}(x) + \frac{\mu(x, \Gamma)}{\lambda(x)} 1_{\mathbb{X}-\lambda^{-1}\{0\}}(x),$$

where δ_{e_1} is the degenerate measure at $e_1 \in \mathbb{V}$. This ensures that any $\kappa(\cdot, \Gamma) \in \mathscr{B}(\mathbb{X})/\mathscr{B}([0,1])$ and any $\kappa(x, \cdot)$ a probability measure on $\mathscr{B}(\mathbb{V})$. Also, when $\mu(x, \cdot)$ is the zero measure,

$$\mu(x, \mathrm{d}v) = 0 = \lambda(x) \cdot \delta_{e_1}(\mathrm{d}v) = \lambda(x)\kappa(x, \mathrm{d}v),$$

and otherwise,

$$\mu(x, \mathrm{d}v) = \mu(x, \mathbb{V}) \frac{\mu(x, \mathrm{d}v)}{\mu(x, \mathbb{V})} = \lambda(x)\kappa(x, \mathrm{d}v).$$

(b) We call λ the intensity map *and κ the* (conditional) jump distribution

(c) As far as we know, there is no widely accepted source which explores jump-diffusions to the extent of declaring a notion like locally countable*, as we have. This means that there is likely some clash of terminology, should such a concept already exist.*

A.4 Real moments of jump-diffusions

We now turn our focus to the real moments of $(\mathrm{P}, \mathscr{F})$ jump-diffusions and the extension of our Lévy-Khintchine map Λ to real moments.

$$\Lambda(u,x) = \big\langle u, \beta^{\chi}(x)\big\rangle + \frac{1}{2}\big\langle u, \alpha(x)\big\rangle + \int_{\mathbb{V}} \big(e^{\langle u,v\rangle} - 1 - \langle u, \chi(v)\rangle\big)\mu(x, \mathrm{d}v), \quad u \in \mathbb{V},\ x \in \mathbb{X}$$

The above expression may be infinite, as the final term includes an unbounded integral over a possibly infinite measure. That said, we find it imperative to denote the following sets of finiteness.

$$\mathcal{D}_\Lambda(x) \coloneqq \Big\{u \in \mathbb{V} : \Lambda(u,x) < \infty\Big\}, \qquad \mathcal{D}_\Lambda \coloneqq \bigcap_{x\in\mathbb{X}} \mathcal{D}_\Lambda(x) \tag{A.4.1}$$

The following results will explore the nature of the maps $\Lambda(\cdot, x) : \mathcal{D}_\Lambda(x) \to \mathbb{R}$ for fixed differentiable χ-characteristics $(\beta^{\chi}, \alpha, \mu)$, where our truncation function χ is defined by $\chi(v) = v 1_{|v|\le 1}$. Note that there is no loss of generality in selecting this truncation function, since they all evaluate Λ identically.

Lemma A.4.2. *For any $x \in \mathbb{X}$, we have $u \in \mathcal{D}_\Lambda(x)$ if and only if $\int_{|v|>1} e^{\langle u,v\rangle}\mu(x, \mathrm{d}v) < \infty$.*

Proof. To each $u, v \in \mathbb{V}$, Taylor's theorem gives us $\gamma_{u,v} \in [0,1]$ such that

$$e^{\langle u,v\rangle} = 1 + \langle u, v\rangle + \frac{1}{2} e^{\gamma_{u,v}\langle u,v\rangle}\langle u, v\rangle^2.$$

This allows us to see that, for each $x \in \mathbb{X}$, $\Lambda(u,x)$ and $\int_{|v|>1} e^{\langle u,v\rangle}\mu(x, \mathrm{d}v)$ differ by finite expressions.

$$\begin{aligned}
&\left|\Lambda(u,x) - \int_{|v|>1} e^{\langle u,v\rangle}\mu(x,\mathrm{d}v)\right| \\
&= \left|\langle u, \beta^{\chi}(x)\rangle + \frac{1}{2}\langle u, \alpha(x)u\rangle + \int_{|v|\le 1}\big(e^{\langle u,v\rangle} - 1 - \langle u,v\rangle\big)\mu(x,\mathrm{d}v) - \int_{|v|>1}\mu(x,\mathrm{d}v)\right| \\
&\le \left|\langle u, \beta^{\chi}(x)\rangle + \frac{1}{2}\langle u, \alpha(x)u\rangle\right| + \left|\int_{|v|\le 1}\frac{1}{2} e^{\gamma_{u,v}\langle u,v\rangle}\langle u,v\rangle^2\mu(x,\mathrm{d}v)\right| + \int_{|v|>1}\mu(x,\mathrm{d}v) \\
&\le \left|\langle u, \beta^{\chi}(x)\rangle + \frac{1}{2}\langle u, \alpha(x)u\rangle\right| + \left(\frac{1}{2}e^{|u|} + 1\right)\int_{\mathbb{V}}\big(1 \wedge |v|^2\big)\mu(x,\mathrm{d}v)
\end{aligned}$$

Thus, one can be defined as a finite displacement of the other.

Lemma A.4.3. *For each $x \in \mathbb{X}$, $\mathcal{D}_\Lambda(x)$ is convex.*

Proof. We use our characterization of $\mathcal{D}_\Lambda(x)$ from Lemma A.4.2. Let $u, u' \in \mathcal{D}_\Lambda(x)$, $\gamma \in (0,1)$, and use Hölder's inequality to see the following.

$$\begin{aligned}
&\int_{|v|>1} e^{\langle u' + \gamma(u-u'), v\rangle}\mu(x,\mathrm{d}v) \\
&\quad = \int_{|v|>1}\big|(e^{\langle u,v\rangle})^{\gamma}\cdot(e^{\langle u',v\rangle})^{1-\gamma}\big|\mu(x,\mathrm{d}v)
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(\int_{|v|>1} \left| (e^{\langle u,v\rangle})^{\gamma} \right|^{\frac{1}{\gamma}} \mu(x, \mathrm{d}v) \right)^{\gamma} \left(\int_{|v|>1} \left| (e^{\langle u',v\rangle})^{1-\gamma} \right|^{\frac{1}{1-\gamma}} \mu(x, \mathrm{d}v) \right)^{1-\gamma} \\
&= \left(\int_{|v|>1} e^{\langle u,v\rangle} \mu(x, \mathrm{d}v) \right)^{\gamma} \left(\int_{|v|>1} e^{\langle u',v\rangle} \mu(x, \mathrm{d}v) \right)^{1-\gamma} \\
&< \infty
\end{aligned}
$$

An arbitrary convex combination now satisfies $\gamma u + (1-\gamma)u' \in \mathcal{D}_\Lambda(x)$.

Lemma A.4.4. *For each $x \in \mathbb{X}$, the map $\Lambda(\cdot, x)$ is continuously differentiable on $\mathcal{D}_\Lambda(x)^\circ$, with derivative $\mathrm{D}\Lambda(\cdot, x) : \mathcal{D}_\Lambda(x)^\circ \to \mathbb{L}(\mathbb{V}, \mathbb{R})$ as follows.*

$$
\mathrm{D}\Lambda(u,x)w = \left\langle \beta^{\chi}(x) + \alpha(x)u + \int_{\mathbb{V}} \left(e^{\langle u,v\rangle} v - \chi(v)\right) \mu(x, \mathrm{d}v), w \right\rangle, \quad u \in \mathcal{D}_\Lambda(x)^\circ \tag{A.4.5}
$$

Proof. Fix $x \in \mathbb{X}$, $u \in \mathcal{D}_\Lambda(x)^\circ$. Let $\epsilon > 0$ such that $B(u, \epsilon) \subseteq \mathcal{D}_\Lambda(x)$. For all $0 < \delta < \epsilon$ and $i = 1, \dots, d$, we now have the following identity

$$
\begin{aligned}
\frac{\Lambda(u + \delta e_i, x) - \Lambda(u, x)}{\delta} &= \langle e_i, \beta^{\chi}(x)\rangle + \langle e_i, \alpha(x)u\rangle + \frac{1}{2}\langle \delta e_i, \alpha(x)u\rangle \\
&\quad + \int_{|v|\leq 1} \frac{1}{\delta}\left(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle} - \langle \delta e_i, v\rangle \right) \mu(x, \mathrm{d}v) \\
&\quad + \int_{|v|>1} \frac{1}{\delta}\left(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle} \right) \mu(x, \mathrm{d}v)
\end{aligned} \tag{A.4.6}
$$

Evaluating the limit of (A.4.6) as $\delta \to 0$ is now a matter of exchanging the limit with integration; we will do this by using the dominated convergence theorem.

For the first integral, Taylor's theorem provides us $\gamma_0, \gamma_1 \in [0, 1]$ such that the following hold.

$$
\begin{aligned}
e^{\langle u+\delta e_i, v\rangle} &= 1 + \langle u + \delta e_i, v\rangle + \frac{1}{2}\langle u + \delta e_i, v\rangle^2 e^{\gamma_0 \langle u+\delta e_i, v\rangle} \\
e^{\langle u,v\rangle} &= 1 + \langle u, v\rangle + \frac{1}{2}\langle u, v\rangle^2 e^{\gamma_1 \langle u,v\rangle}
\end{aligned}
$$

This shows us that, for all $0 < \delta < \epsilon$ and $|v| \leq 1$,

$$
\begin{aligned}
\left| \frac{1}{\delta}\left(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle} - \langle \delta e_i, v\rangle \right) \right| &= \left| \frac{1}{2}\langle u + \delta e_i, v\rangle^2 e^{\gamma_0 \langle u+\delta e_i, v\rangle} + \frac{1}{2}\langle u, v\rangle^2 e^{\gamma_1 \langle u,v\rangle} \right| \\
&\leq \left(\left(|u| + \epsilon\right)^2 e^{|u|+\epsilon} \right) |v|^2.
\end{aligned}
$$

This dominating function is integrable,

$$
\int_{|v|\leq 1} \left(\left(|u| + \epsilon\right)^2 e^{|u|+\epsilon} \right) |v|^2 \mu(x, \mathrm{d}v) \leq \left(\left(|u| + \epsilon\right)^2 e^{|u|+\epsilon} \right) \int_{\mathbb{V}} (1 \wedge |v|^2) \mu(x, \mathrm{d}v) < \infty,
$$

so we may apply the dominated convergence theorem.

$$
\lim_{\delta \to 0} \int_{|v|\leq 1} \frac{1}{\delta}\left(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle} - \langle \delta e_i, v\rangle \right) \mu(x, \mathrm{d}v)
$$

$$= \int_{|v|\leq 1} \lim_{\delta\to 0} \frac{1}{\delta}\Big(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle} - \langle \delta e_i, v\rangle\Big)\mu(x,\mathrm{d}v)$$

$$= \int_{|v|\leq 1} \Big(e^{\langle u,v\rangle} v_i - v_i\Big)\mu(x,\mathrm{d}v) \tag{A.4.7}$$

For the second integral, we again use Taylor's theorem to establish for each $0 < \delta < \epsilon/2$, some $\gamma_\delta \in [0,\delta]$ such that

$$e^{\langle u+\delta e_i, v\rangle} = e^{\langle u,v\rangle} + \langle \delta e_i, v\rangle e^{\langle u+\gamma_\delta e_i, v\rangle}$$

This way, we have the following dominating function.

$$\left|\frac{1}{\delta}\Big(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle}\Big)\right| \leq \left|\langle e_i, v\rangle e^{\langle u+\gamma_\delta e_i, v\rangle}\right| \leq |v_i| e^{\langle u,v\rangle + \epsilon|v_i|/2}$$

The claim is that this dominating function is integrable. To see this, first note that because we have the following limit,

$$\lim_{|v|\to\infty} \frac{|v_i| e^{\langle u,v\rangle+\epsilon|v_i|/2}}{e^{\langle u,v\rangle + 2\epsilon|v_i|/3}} = \lim_{|v|\to\infty} \frac{|v_i|}{e^{\epsilon|v_i|/6}} = 0$$

There exists $M > 0$ such that for all $|v| > M$,

$$|v_i| e^{\langle u,v\rangle+\epsilon|v_i|/2} < e^{\langle u,v\rangle+2\epsilon|v_i|/3}.$$

We now see that

$$\begin{aligned}
&\int_{|v|>1} |v_i| e^{\langle u,v\rangle+\epsilon|v_i|/2}\mu(x,\mathrm{d}v) \\
&\quad= \int_{1<|v|\leq M} |v_i| e^{\langle u,v\rangle+\epsilon|v_i|/2}\mu(x,\mathrm{d}v) + \int_{|v|>M} |v_i| e^{\langle u,v\rangle+\epsilon|v_i|/2}\mu(x,\mathrm{d}v) \\
&\quad\leq \int_{1<|v|\leq M} M e^{(|u|+\epsilon/2)M}\mu(x,\mathrm{d}v) + \int_{|v|>M} e^{\langle u,v\rangle+2\epsilon|v_i|/3}\mu(x,\mathrm{d}v) \\
&\quad\leq M e^{(|u|+\epsilon/2)M} \int_{\mathbb{V}} (1\wedge |v|^2)\mu(x,\mathrm{d}v) + \sum_{\ell=0}^{1} \int_{|v|>1} e^{\langle u+2\epsilon e_i/3, v\rangle}\mu(x,\mathrm{d}v) \\
&\quad< \infty.
\end{aligned}$$

We again use the dominated convergence theorem to deduce the following.

$$\lim_{\delta\to 0}\int_{|v|>1} \frac{1}{\delta}\Big(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle}\Big)\mu(x,\mathrm{d}v)$$

$$= \int_{|v|>1} \lim_{\delta\to 0}\frac{1}{\delta}\Big(e^{\langle u+\delta e_i, v\rangle} - e^{\langle u,v\rangle}\Big)\mu(x,\mathrm{d}v)$$

$$= \int_{|v|>1} e^{\langle u,v\rangle} v_i \mu(x,\mathrm{d}v) \tag{A.4.8}$$

Combining equations (A.4.6), (A.4.7), and (A.4.8) now yields our desired identity.

$$\mathrm{D}_i\Lambda(u,x) = \Big\langle e_i, \beta^\chi(x) + \alpha(x)u + \int_{\mathbb{V}} \Big(e^{\langle u,v\rangle} v - \chi(v)\Big)\mu(x,\mathrm{d}v)\Big\rangle$$

Continuity of $\mathrm{D}_i\Lambda(u,x)$ for $u \in \mathcal{D}_\Lambda(x)^\circ$ involves very similar dominated convergence theorem arguments as above. From here, it is clear that Λ is continuously differentiable with the form in (A.4.5).

As we have seen in Lemmas A.2.3 and A.3.2, if we have local boundedness of certain integrals of a jump kernel μ, we can leverage these to $(\mathrm{P},\mathscr{F})$ local conditions of the associated jump-diffusion X. Throughout the remainder of this section, we impose the following uniform-boundedness principle for the kernel μ.

$$\begin{aligned} &f \in \mathscr{B}(\mathbb{V})/\mathscr{B}(\mathbb{R}), \quad \int_{\mathbb{V}} |f(v)|\mu(x,\mathrm{d}v) < \infty \text{ for all } x \in \mathbb{X} \\ &\implies \quad x \mapsto \int_{\mathbb{V}} |f(v)|\mu(x,\mathrm{d}v) \text{ bounded on compact sets} \end{aligned} \tag{A.4.9}$$

With this assumption, we get some nice results on finite exponential moments of X.

Proposition A.4.10. *Fix a $(\mathrm{P},\mathscr{F})$ jump-diffusion X with differential χ-characteristics (β^χ,α,μ). Suppose we have the regularity condition (A.4.9) above. If $0 \in \mathcal{D}_\Lambda^\circ$, then X is special.*

Proof. If $0 \in \mathcal{D}_\Lambda^\circ$, then there exists some $\delta > 0$ such that $\overline{B}(0,\delta) \subseteq \mathcal{D}_\Lambda$. Observe the following implication of this fact, for each $x \in \mathbb{X}$.

$$\begin{aligned} \int_{\mathbb{V}} \big|v-\chi(v)\big|\mu(x,\mathrm{d}v) &= \int_{|v|>1} |v|\mu(x,\mathrm{d}v) \\ &\le \int_{|v|>1} \frac{\sqrt{d}}{\delta}\exp\Big(\frac{\delta|v|}{\sqrt{d}}\Big)\mu(x,\mathrm{d}v) \\ &\le \frac{\sqrt{d}}{\delta}\int_{|v|>1} \exp\Big(\max_{i=1}^{d}\max_{\ell=0}^{1}\big\langle(-1)^\ell\delta e^i,v\big\rangle\Big)\mu(x,\mathrm{d}v) \\ &\le \frac{\sqrt{d}}{\delta}\sum_{i=1}^{d}\sum_{\ell=0}^{1}\int_{|v|>1}\exp\big\langle(-1)^\ell\delta e^i,v\big\rangle\mu(x,\mathrm{d}v) \\ &< \infty \end{aligned}$$

Our regularity condition (A.4.9) now allows us to apply Lemma A.2.3 to conclude X is special.

Proposition A.4.11. *Fix a $(\mathrm{P},\mathscr{F})$ jump-diffusion X with differential χ-characteristics (β^χ,α,μ). Suppose we have the regularity condition (A.4.9) above. If $u \in \mathcal{D}_\Lambda$, then $\exp\langle u,X\rangle$ is special, and $\exp\big(\langle u,X\rangle - \Lambda(u,X)\bullet\ell\big)$ is a $(\mathrm{P},\mathscr{F})$ local martingale.*

Proof. Using Lemma A.1.7 for the function $f_u(v) = \exp\langle u,v\rangle$ and its derivative identities as in (A.1.13), we get the following.

$$\begin{aligned} \exp\langle u,X_t\rangle = {} &\exp\langle u,X_0\rangle + \exp\langle u,X_t\rangle\Big(\langle u,\beta^\chi(X)\rangle + \frac{1}{2}\langle u,\alpha(X)u\rangle\Big)\bullet\ell_t \\ &+ \mathrm{D}f_u(X_-)\bullet X^{\mathrm{c}} + \Big(\exp\langle u,X_-\rangle\langle u,\chi\rangle\Big) * \tilde{q}_t^X \\ &+ \exp\langle u,X_-\rangle\cdot\Big(\exp\langle u,\mathrm{id}_{\mathbb{V}}\rangle - 1 - \langle u,\chi\rangle\Big) * q^X \end{aligned} \tag{A.4.12}$$

Note that localizing our final term on the sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times in (A.1.9), we get the following.

$$\begin{aligned}
&\mathrm{E}_{\mathrm{P}}\Big|\exp\langle u, X_-\rangle\Big(\exp\langle u, \mathrm{id}_{\mathbb{V}}\rangle - 1 - \langle u,\chi\rangle\Big)\Big| * \hat{q}^X_{T_n}\\
&= \mathrm{E}_{\mathrm{P}}\int_0^{T_n}\int_{\mathbb{V}}\Big|\exp\langle u, X_s\rangle\Big(\exp\langle u, v\rangle - 1 - \langle u,\chi(v)\rangle\Big)\Big|\mu(X_s,\mathrm{d}v)\mathrm{d}s\\
&\le n\cdot\sup_{|x|\le n}\Big(e^{\langle u,x\rangle}\int_{\mathbb{V}}\big|e^{\langle u,v\rangle} - 1 - \langle u,\chi(v)\rangle\big|\mu(x,\mathrm{d}v)\Big)
\end{aligned}$$

Seeing as $u \in \mathcal{D}_\Lambda$, the integral in the above quantity is finite, and so (A.4.9) gives us finiteness of the supremum. Using [JS03, Proposition II.1.28] now allows us to compensate the jump term in (A.4.12).

$$\exp\langle u, X_t\rangle = \exp\langle u, X_0\rangle + \Big(\exp\langle u, X_t\rangle\cdot\Lambda(u,X)\Big)\bullet\ell_t + \mathrm{D}f_u(X_-)\bullet X^{\mathrm{c}} + \Big(\exp\langle u, X_-\rangle\langle u,\chi\rangle\Big)*\tilde{q}^X_t$$

This is a representation of $\exp\langle u, X\rangle$ as an initial term, predictable term of finite variation, and a local martingale. Thus, it is a special semimartingale. From here, we may perform the product rule on $\exp\big(\langle u, X\rangle - \Lambda(u,X)\bullet\ell\big)$ as we did in Proposition A.1.12 to show that the process is a local martingale.

Theorem A.4.13. *Fix a* $(\mathrm{P},\mathscr{F})$ *jump-diffusion* X *with differential* χ*-characteristics* (β^χ,α,μ)*. Suppose we have the regularity condition (A.4.9) above and that* $0\in\mathcal{D}^\circ_\Lambda$*. For each* $h\in\mathbb{D}([0,\infty),\mathbb{V})$ *of finite-variation with image contained in* $\mathcal{D}^\circ_\Lambda$*, the process* $\exp(h\bullet X)$ *is special and*

$$\exp\Big(h\bullet X - \Lambda(h,X)\bullet\ell\Big)$$

is a $(\mathrm{P},\mathscr{F})$ *local martingale.*

Proof. We first note that Proposition A.4.10 allows us to conclude X is special. Perform Itô's formula [JS03, Theoerem I.4.57] in addition to its jump-diffusion variant in Lemma A.1.7 and various stochastic integral identities [JS03, Remarks I.4.36, I.4.37, Theorem I.4.40(d), Proposition II.1.30(b)].

$$\begin{aligned}
&\exp\big(h\bullet X_t\big)\\
&= \exp\big(h\bullet X_-\big)\bullet\big(h\bullet X\big)_t + \frac{1}{2}\exp\big(h\bullet X_-\big)\bullet\big\langle(h\bullet X)^{\mathrm{c}},(h\bullet X)^{\mathrm{c}}\big\rangle_t\\
&\quad+\sum_{0<s\le t}\Big(\exp\big(h\bullet X_{s-} + \Delta\big(h\bullet X\big)_s\big) - \exp\big(h\bullet X_{s-}\big) - \exp\big(h\bullet X_{s-}\big)\Delta\big(h\bullet X\big)_s\Big)\\
&= \Big(\exp\big(h\bullet X_-\big)\cdot h\Big)\bullet X_t + \frac{1}{2}\exp\big(h\bullet X\big)\big\langle h,\alpha(X)h\big\rangle\bullet\ell_t\\
&\quad+\exp\big(h\bullet X_-\big)\Big(e^{\langle h,\mathrm{id}_{\mathbb{V}}\rangle} - 1 - \langle h,\mathrm{id}_{\mathbb{V}}\rangle\Big)*q^X_t
\end{aligned}$$

$$\begin{aligned}
&= \Big(\exp\big(h\bullet X\big)\cdot\langle h,\beta\rangle + \frac{1}{2}\exp\big(h\bullet X\big)\big\langle h,\alpha(X)h\big\rangle\Big)\bullet\ell_t + \Big(\exp\big(h\bullet X_-\big)\cdot h\Big)\bullet X^{\mathrm{c}}_t\\
&\quad+\exp\big(h\bullet X_-\big)\langle h,\mathrm{id}_{\mathbb{V}}\rangle * \tilde{q}^X_t\\
&\quad+\exp\big(h\bullet X_-\big)\Big(e^{\langle h,\mathrm{id}_{\mathbb{V}}\rangle} - 1 - \langle h,\mathrm{id}_{\mathbb{V}}\rangle\Big)*q^X_t
\end{aligned}\tag{A.4.14}$$

Now, choosing our $(\mathrm{P},\mathscr{F})$ localizing sequence $(T_n)_{n\in\mathbb{N}}$ as in A.1.9, we have the following bound.

$$\begin{aligned}
&\mathrm{E}_{\mathrm{P}}\Big|\exp\big(h\bullet X_-\big)\Big(e^{\langle h,\mathrm{id}_{\mathbb{V}}\rangle}-1-\langle h,\mathrm{id}_{\mathbb{V}}\rangle\Big)*\hat{q}^X_{T_n}\Big|\\
&=\mathrm{E}_{\mathrm{P}}\int_0^{T_n}\int_{\mathbb{V}}\Big|\exp\big(h\bullet X_s\big)\Big(e^{\langle h(s),v\rangle}-1-\langle h(s),v\rangle\Big)\Big|\mu(X_s,\mathrm{d}v)\mathrm{d}s\\
&\le n\cdot\sup_{|x|\le n}\sup_{s\in[0,n]}e^{|x|\cdot|h(s)|}\int_{\mathbb{V}}\big|e^{\langle h(s),v\rangle}-1-\langle h(s),v\rangle\big|\mu(x,\mathrm{d}v)
\end{aligned}$$

Seeing as $\Lambda(\cdot,x)$ is continuously differentiable, it is uniformly bounded on $\mathcal{D}^\circ_\Lambda$. This, along with assumption (A.4.9) allow us to conclude that the preceding expression is finite. Thus, we may compensate the final jump integral in (A.4.14).

$$\begin{aligned}
\exp\big(h\bullet X_t\big)=&\Big(\exp\big(h\bullet X\big)\cdot\Lambda(h,X)\Big)\bullet\ell_t+\Big(\exp\big(h\bullet X_-\big)\cdot h\Big)\bullet X^{\mathrm{c}}_t\\
&+\exp\big(h\bullet X_-\big)\Big(e^{\langle h,\mathrm{id}_{\mathbb{V}}\rangle}-1\Big)*q^X_t
\end{aligned}$$

The decomposition of $\exp(h\bullet X)$ into a predictable finite-variation process and a local martingale implies that it is special. Now, we write M as the local martingale term above, $A=\exp(h\bullet X)$, and $B=\exp(-\Lambda(h,X)\bullet\ell)$. We now recognize that B is predictable and finite-variation and use [JS03, Proposition I.4.49(b)] to conclude our proof.

$$\begin{aligned}
\exp\big(h\bullet X_t-\Lambda(h,X)\bullet\ell_t\big)&=A_tB_t\\
&=A_-\bullet B_t+B\bullet A_t\\
&=\big(A\cdot B\cdot-\Lambda(h,X)\big)\bullet\ell_t+B\bullet\Big(\big(\exp(h\bullet X)\cdot\Lambda(h,X)\big)\bullet\ell+M\Big)_t\\
&=\big(A\cdot B\cdot-\Lambda(h,X)\big)\bullet\ell_t+\big(B\cdot A\cdot\Lambda(h,X)\big)\bullet\ell_t+B\bullet M_t\\
&=B\bullet M_t
\end{aligned}$$

Bibliography

[Cuc11] Christina Cuchiero. Affine and polynomial processes, 2011.

[DZ10] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*, volume 38. Springer Berlin Heidelberg, 2 edition, 2010.

[JS03] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288. Springer Berlin Heidelberg, 2003.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability*. Springer New York, 2002.

[KRM15] Martin Keller-Ressel and Eberhard Mayerhofer. Exponential moments of affine processes. *The Annals of Applied Probability*, 25(2):714–752, apr 2015.

[SV10] Peter Spreij and Enno Veerman. The affine transform formula for affine jump-diffusions with a general closed convex state space. *arXiv preprint arXiv:1005.1099*, 2010.

[Wal98] Wolfgang Walter. *Ordinary Differential Equations*, volume 182. Springer New York, 1998.