Markov chain approximations to, and some fluctuation results for, Lévy processes

by

Matija Vidmar

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Statistics

October 2014

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments

It is pleasure to thank Aleksandar Mijatović and Saul Jacka for their apprenticeship in Mathematics and continued guidance during my years as a PhD student.

The support of the Slovene Human Resources Development and Scholarship Fund under contract number 11010-543/2011 is acknowledged.

Declarations

Chapter 2 is joint work with Aleksandar Mijatović and Saul Jacka, ‘Markov chain approximations for transition densities of Lévy processes’, Electronic Journal of Probability **19** (2014), no. 7, pp. 1-37. Chapter 3 is based on ‘Non-random overshoots of Lévy processes’, arXiv:1301.4463, and on ‘Fluctuation theory for upwards skip-free Lévy chains’, arXiv:1309.5328. Chapter 4 is joint work with Aleksandar Mijatović and Saul Jacka, ‘Markov chain approximations to scale functions of Lévy processes’, arXiv:1310.1737.

Abstract

We introduce, and analyze in terms of convergence rates of transition kernels, a continuous-time Markov chain approximation to Lévy processes. A full fluctuation theory for what are right-continuous random walks embedded into continuous-time as compound Poisson processes, is provided. These results are applied to obtaining a general algorithm for the calculation of the scale functions of a spectrally negative Lévy process. In a related result, the class of Lévy processes having non-random overshoots is precisely characterized.

Chapter 1

Introduction

In this chapter we (i) review the relevant theory and terminology; in the process, or otherwise, (ii) fix notation (Sections 1.1 and 1.2); and (iii) explain the structure of the remainder of the thesis (Section 1.3). Further (more specific) concepts and notation will be introduced in subsequent chapters, as and where appropriate.

1.1 Miscellaneous general notation

Notation 1.1 (Number sets).

(1) For $h > 0$, $d \in \mathbb{N}$: $\mathbb{Z}_h := h\mathbb{Z} := \{hk : k \in \mathbb{Z}\}$ and $\mathbb{Z}_h^d := (\mathbb{Z}_h)^d = \{hk : k \in \mathbb{Z}^d\}$.

(2) The nonnegative, nonpositive, positive and negative real numbers are denoted by $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$, $\mathbb{R}_- := \{x \in \mathbb{R} : x \leq 0\}$, $\mathbb{R}^+ := \mathbb{R}_+ \backslash \{0\}$ and $\mathbb{R}^- := \mathbb{R}_- \backslash \{0\}$, respectively. Then $\mathbb{Z}_+ := \mathbb{R}_+ \cap \mathbb{Z}$, $\mathbb{Z}_- := \mathbb{R}_- \cap \mathbb{Z}$, $\mathbb{Z}^+ := \mathbb{R}^+ \cap \mathbb{Z}$ and $\mathbb{Z}^- := \mathbb{R}^- \cap \mathbb{Z}$ are the nonnegative, nonpositive, positive and negative integers, respectively.

(3) Similarly, for $h > 0$, $\mathbb{Z}_h^+ := \mathbb{Z}_h \cap \mathbb{R}_+$, $\mathbb{Z}_h^- := \mathbb{Z}_h \cap \mathbb{R}_-$, $\mathbb{Z}_h^{++} := \mathbb{Z}_h \cap \mathbb{R}^+$ and $\mathbb{Z}_h^{--} := \mathbb{Z}_h \cap \mathbb{R}^-$, are the apposite elements of $\mathbb{Z}_h$.

(4) The following introduces notation for the relevant half-planes of $\mathbb{C}$; the arrow notation is meant to be suggestive of which half-plane is being considered: $\mathbb{C}^{\rightarrow} := \{z \in \mathbb{C} : \Re z > 0\}$, $\mathbb{C}^{\leftarrow} := \{z \in \mathbb{C} : \Re z < 0\}$, $\mathbb{C}^{\downarrow} := \{z \in \mathbb{C} : \Im z < 0\}$ and $\mathbb{C}^{\uparrow} := \{z \in \mathbb{C} : \Im z > 0\}$. $\overline{\mathbb{C}^{\rightarrow}}$, $\overline{\mathbb{C}^{\leftarrow}}$, $\overline{\mathbb{C}^{\downarrow}}$ and $\overline{\mathbb{C}^{\uparrow}}$ are then the respective closures of these sets.

(5) $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ are the positive and nonnegative integers, respectively.

Notation 1.2 (Balls and spheres). $B(x, \delta)$ (respectively $\overline{B}(x, \delta)$, $S(x, \delta)$) is the *open ball* (respectively *closed ball*, *sphere*), centre $x \in \mathbb{R}^d$, radius $\delta > 0$ ($d \in \mathbb{N}$).

Notation 1.3 (Ceiling and min/max functions). $\lceil x \rceil := \min\{k \in \mathbb{Z} : k \geq x\}$ ($x \in \mathbb{R}$) is the *ceiling* function. For $\{a, b\} \subset [-\infty, +\infty]$: $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$.

Convention 1.4 (Usage of increasing and decreasing, positive and negative). *Increasing* will always mean strictly increasing and *positive* will mean strictly positive (similarly *decreasing* and *negative*). *Exceeding* will mean strictly exceeding. $x_n \downarrow x$ (respectively $x_n \uparrow x$) will mean $x_n > x$ (respectively $x_n < x$) and x_n nonincreasing (respectively nondecreasing) to x (as $n \to \infty$).

Notation 1.5 (Big O and little o notation; usage of $\sim$). For functions f and $g > 0$, defined on some right neighborhood of 0, we shall write $f = O(g)$ (respectively $f = o(g)$, $f \sim g$) for $\limsup_{h \downarrow 0} |f|(h)/g(h) < \infty$ (respectively $\lim_{h \downarrow 0} |f|(h)/g(h) = 0$, $\lim_{h \downarrow 0} |f|(h)/g(h) \in (0, \infty)$) — if further g converges to 0, then we will say f *decays no slower than* (respectively *faster than*, *at the same rate as*) g. Analogous notation obtains for the behavior of functions at $+\infty$ or $-\infty$.

Notation 1.6 (The geometric and exponential laws). The *geometric law* $\mathrm{geom}(p)$ on $\mathbb{N}_0$ with *success* parameter $p \in (0, 1]$ has $\mathrm{geom}(p)(\{k\}) = p(1-p)^k$ ($k \in \mathbb{N}_0$). $1 - p$ is called the *failure* parameter. The *exponential law* $\mathrm{Exp}(\lambda)$ on $\mathbb{R}^+$ with parameter $\lambda \in (0, \infty)$ is specified by the density $\mathrm{Exp}(\lambda)(dt) = \lambda e^{-\lambda t} dt$. Additionally, we will understand any random element, which is equal to $+\infty$, a.s., to have the $\mathrm{Exp}(0)$-law.

Notation 1.7 (Image measure; random elements). If μ is a measure on some measurable space $(X, \mathcal{A})$ and f is a measurable mapping between the measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, then the *push-forward* or *image* measure $f_\star \mu = \mu \circ f^{-1}$ on $(Y, \mathcal{B})$ is given by $f_\star \mu(B) = \mu(f^{-1}(B))$ ($B \in \mathcal{B}$).

If $(\Omega, \mathcal{F}, \mathsf{P})$ is a probability space, then by a *random element* X thereon, we mean a measurable mapping from $(\Omega, \mathcal{F})$ into some measurable space $(S, \mathcal{S})$. In this instance we shall use the notation $\mathsf{P}_X = X_\star \mathsf{P}$ for the *law* of X (on $(S, \mathcal{S})$ with respect to P) [Kallenberg, 1997, p. 24].

Notation 1.8 (Borel σ-fields and supports; Dirac measures). $\mathcal{B}(S)$ will always denote the Borel σ-field of a topological space S; $\mathrm{supp}(m)$ the support of a measure m thereon [Kallenberg, 1997, p. 9]; we shall say m is *carried* by $A \in \mathcal{B}(S)$, if $m(S \backslash A) = 0$; $\delta_x := (A \mapsto \mathbb{1}_A(x))$, mapping $\mathcal{B}(S)$ into $[0, 1]$, is the *Dirac measure* at $x \in S$.

Notation 1.9 (Completions). For a measure μ on a σ-field $\mathcal{A}$, $\overline{\mathcal{A}}^{\mu}$ denotes the *completion* of $\mathcal{A}$ with respect to μ, while $\overline{\mu}$ is the unique extension of μ to $\overline{\mathcal{A}}^{\mu}$ [Kallenberg, 1997, p. 13]. $A \subset X$ is said to be *μ-negligible* (or *μ-null*), if A is measurable and of μ-measure 0.

The *universal completion* of a σ-field $\mathcal{F}$ will be denoted by $\mathcal{F}^{\star}$.

Convention 1.10 (Abbreviations). DCT (MCT) is shorthand for *dominated* (*monotone*) *convergence theorem*. CP stands for *compound Poisson*.

Convention 1.11 (Usage of $0\cdot\infty$, $a/0$). By convention, $0\cdot\infty = 0$. We will understand $a/0 = \pm\infty$ for $a \in \pm(0,\infty)$.

Convention 1.12 (Usage of $\perp$). The symbol $\perp$ will sometimes be used to indicate *stochastic independence* (relative to the probability measure P, or some conditional measure $\mathsf{P}(\cdot|A)$ (with $A \in \mathcal{F}$ and $\mathsf{P}(A) > 0$) derived therefrom, depending on the context).

Notation 1.13 (Identity function). For a set A, id_A shall denote the *identity function* on A.

Notation 1.14 (Laplace transforms and Lebesgue-Stieltjes measures). The *Laplace transform* of a measure μ on $\mathbb{R}$, concentrated on $[0,\infty)$, is denoted $\hat{\mu}$: $\hat{\mu}(\beta) = \int_{[0,\infty)} e^{-\beta x}\mu(dx)$ (for all $\beta \geq 0$ such that this integral is finite). To a nondecreasing right-continuous function $F : \mathbb{R} \to \mathbb{R}$, a measure dF may be associated in the Lebesgue-Stieltjes sense.

Definition 1.15 (Functions of exponential order; limits at infinity)**.** A function $f : [0,\infty) \to [0,\infty)$ is said to be of *exponential order*, if there are $\{\alpha, A\} \subset \mathbb{R}_+$, such that $f(x) \leq Ae^{\alpha x}$ ($x \geq 0$); $f(+\infty) := \lim_{x\to\infty} f(x)$, when this limit exists.

Definition 1.16 (Usual assumptions/conditions; augmentations)**.** A filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ is *complete*, if $\mathcal{F}$ is complete relative to the measure P, and $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$. If in addition $\mathbb{F}$ is right-continuous, we say that the filtered probability space satisfies the *usual assumptions/conditions*. There is a smallest augmentation of the filtration (upon completion of $(\mathcal{F}, \mathsf{P})$) which achieves this, when it is not already so, see e.g. [Kallenberg, 1997, p. 101, Lemma 6.8]. We refer to the latter as the *usual augmentation*.

1.2 Lévy processes and continuous-time Markov chains

1.2.1 Lévy processes

Throughout this subsection we fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ and a dimension $d \in \mathbb{N}$. For Lévy processes canonical references are [Bertoin, 1996;

Sato, 1999].

Definition 1.17 (Lévy process)**.** A continuous-time stochastic process $X = (X_t)_{t\geq 0}$, with state space $\mathbb{R}^d$, is a *Lévy process*, if it starts at 0, P-a.s., is continuous in probability, has independent and stationary increments and is càdlàg off a P-negligible set [Sato, 1999, p. 3, Definition 1.6]. It is a Lévy process *in law* if we do not insist on the càdlàg property. Finally it is so with respect to the filtration $\mathbb{F}$, if it is $\mathbb{F}$-adapted and $\mathcal{F}_s \perp (X_t - X_s)$, whenever $t \geq s \geq 0$.

A Lévy process in law X is uniquely characterized by its characteristic triplet $(\Sigma, \lambda, \mu)_{\tilde{c}}$. Here $\Sigma \in \mathbb{R}^{d\times d}$ is a symmetric nonnegative definite matrix, called the *diffusion matrix* (reducing to the scalar, *diffusion coefficient*, when $d = 1$), Λ is a measure on $\mathbb{R}^d$ satisfying $\Lambda(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(|x|^2 \wedge 1)\Lambda(dx) < \infty$ (i.e. Λ does not charge $\{0\}$, integrates 1 outside every neighborhood of 0 and $|\mathrm{id}_{\mathbb{R}}|^2$ on every compact neighborhood of 0 — we say it is a *Lévy measure*), and $\mu \in \mathbb{R}^d$ is the *drift coefficient* relative to some *cut-off function* (also called *truncation function*) $\tilde{c}$ [Sato, 1999, p. 39].

The characteristic function of the law $\mu^t := X_{t*}\mathsf{P}$ $(t \geq 0)$ is then given by the celebrated *Lévy-Khintchine formula* [Sato, 1999, p. 38, Corollary 8.3]:

$$\phi_{X_t}(p) := \int_{\mathbb{R}} e^{i\langle p,x\rangle}\mu^t(dx) = \exp\{t\Psi(p)\} \qquad (t \geq 0,\ p \in \mathbb{R}^d) \tag{1.1}$$

where:

$$\Psi(p) := -\frac{1}{2}\langle p, \Sigma p\rangle + i\langle \mu, p\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle p,x\rangle} - 1 - i\langle p, x\rangle\tilde{c}(x)\right)\Lambda(dx) \quad (p \in \mathbb{R}^d) \tag{1.2}$$

is the *characteristic exponent.* Note that X is a Markov process admitting a temporally and spatially homogeneous transition function $P_{t,T}(x, B) := \mu^{T-t}(B - x)$, where $0 \leq t \leq T$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$ (see e.g. [Sato, 1999, pp. 54-58]).

Definition 1.18 (Compound Poisson processes)**.** X is called a *compound Poisson* (CP) *process in law*, if Λ is finite and, with $\tilde{c} = 0$, $\Sigma = 0$ and $\mu = 0$. In case X is a Lévy process, the qualification "in law" is of course dropped. Remark that we allow in this definition for the case when $X = 0$ identically (P-a.s.).

1.2.2 Fluctuation theory of Lévy processes

Fluctuation theory of Lévy processes studies the first passage times (above, or below a certain level), the running supremum and infimum processes, the two-sided exit problem, and related concepts — with the Wiener-Hopf factorization being one of its most important results. [Kyprianou, 2006] is a book dedicated entirely to these ideas and their applications.

In this subsection we fix a Lévy process X on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$. We shall assume without loss of generality that the latter satisfies the usual assumptions (see Definition 1.16).[1]

We also let e_1 stand for an Exp(1)-distributed random variable independent of X, then define $e_p := e_1/p$ ($p \in (0,\infty)\backslash\{1\}$). We insist (harmlessly) that our probability space is already rich enough to support e_1 (if not, it can be made as such, by taking a suitable tensor product).

Definition 1.19 (Subordinators)**.** X is called a *subordinator*, if it is nondecreasing off a P-null set.

Definition 1.20 (First passage times and overshoots)**.** For $x \in \mathbb{R}$ introduce $T_x := \inf\{t \geq 0 : X_t \geq x\}$ (respectively $\hat{T}_x := \inf\{t \geq 0 : X_t > x\}$, $T_x^- := \inf\{t \geq 0 : X_t < -x\}$), the first entrance time of X to $[x, \infty)$ (respectively (x, ∞), $(-\infty, -x)$). We will informally refer to T_x and $\hat{T}_x$ (respectively T_x^-) as the *times of first passage* above (respectively, below) the level x (respectively $-x$). $R_x := X(\hat{T}_x) - x$ is the *overshoot* at the level x, $x \geq 0$ [Sato, 1999, p. 369].

Remark 1.21. By the Début Theorem [Kallenberg, 1997, p. 101, Theorem 6.7], times of first passage, as in Definition 1.20, are stopping times.

Definition 1.22 (Supremum and infimum pocesses)**.** We define:

(a) $\overline{X}_t := \sup\{X_s : s \in [0,t]\}$ ($t \geq 0$), the *running supremum* or *maximum* process.

(b) $\underline{X} := -\overline{-X}$, the *running infimum* or *minimum* process.

(c) $\overline{G}_t := \sup\{s \in [0,t] : X_s = \overline{X}_s\}$, the last time on $[0,t]$ of attaining the running supremum ($t \geq 0$).

(d) $\underline{G}_t := \sup\{s \in [0,t] : X_s = \underline{X}_s\}$, the last time on $[0,t]$ of attaining the running infimum ($t \geq 0$).

[1] We can always achieve this by first completing $(\Omega, \mathcal{F}, \mathsf{P})$ (clearly X remains a Lévy process on this space as well); then (harmlessly) discarding the $\overline{\mathsf{P}}$-negligible set on which X is not càdlàg; and finally performing the usual augmentation (see Definition 1.16) of the filtration $\mathbb{F}$, by (i) making it first right-continuous, and then (ii) adding the P-null sets. In both steps (i) and (ii), the property of X being a Lévy process with respect to the filtration $\mathbb{F}$ is preserved. For, in step (i), if $0 \leq s < t$, then for any $s < s' < t$ one has $(X_t - X_{s''})_{s' \leq s'' \leq t} \perp \mathcal{F}_{s'} \supset \mathcal{F}_{s+}$. Therefore (by a π/λ-argument) $(X_t - X_{s'})_{s < s' \leq t} \perp \mathcal{F}_{s+}$. By right-continuity of the sample paths, this implies $(X_t - X_s) \perp \mathcal{F}_{s+}$. In step (ii), we have as follows. If $\mathcal{N}$ is the set of $\overline{\mathsf{P}}$-negligible sets, then for every $0 \leq s < t$, $\mathcal{F}_{s+} \cup \mathcal{N}$ is a π-system, independent of $X_t - X_s$. Since one is able to raise independence from a π-system to the σ-algebra generated by it, this establishes the property.

When X is CP we additionally set:

(e) $\overline{G}^*_t := \inf\{s \in [0,t] : X_s = \overline{X}_t\}$, i.e., P-a.s., $\overline{G}^*_t$ is the last time in the interval $[0,t]$ that X attains a new maximum ($t \geq 0$).

(e) $\underline{G}^*_t := \inf\{s \in [0,t] : X_s = \underline{X}_t\}$, i.e., P-a.s., $\underline{G}^*_t$ is the last time in the interval $[0,t]$ that X attains a new minimum ($t \geq 0$).

Assume henceforth (for convenience, cf. Footnote 1) that X is càdlàg with certainty (rather than just P-a.s.).

Remark 1.23. In the above we have taken right continuous versions of the nondecreasing processes $\overline{G}$ and $\underline{G}$. Since in the sequel they enter the results only after they have been evaluated at e_p, working with their left continuous versions instead (they are: $\overline{G}'_0 := 0$, $\overline{G}'_t := \sup\{s \in [0,t) : X_s = \overline{X}_s\} = \lim_{s\uparrow t} \overline{G}_s$ ($t > 0$) and $\underline{G}'_0 := 0$, $\underline{G}'_t := \sup\{s \in [0,t) : X_s = \underline{X}_s\} = \lim_{s\uparrow t} \underline{G}_s$ ($t > 0$)) would not change any of the results (thus, $\overline{G}'_{e_p} = \overline{G}_{e_p}$ and $\underline{G}'_{e_p} = \underline{G}_{e_p}$ P-a.s.).[2] Moreover, by the remark in the introduction to [Bertoin, 1996, Section VI.2], it follows that we could just as easily also work with the definitions $\overline{G}''_t := \sup\{s \in [0,t] : \overline{X}_t \in \{X_s, X_{s-}\}\}$ and $\underline{G}''_t := \sup\{s \in [0,t] : \underline{X}_t \in \{X_s, X_{s-}\}\}$ (thus, $\overline{G}''_{e_p} = \overline{G}_{e_p}$ and $\underline{G}''_{e_p} = \underline{G}_{e_p}$ P-a.s.).

An almost indispensable tool in studying fluctuation theory are also the notions of local time and ladder processes. We take the definition from [Kyprianou, 2006].

Definition 1.24 (Local time, ladder and reflected processes)**.** A (continuous, unless 0 is irregular for $[0,\infty)$ see [Kyprianou, 2006, p. 144, Theorem 6.7]), nondecreasing, $\mathbb{R}_+$-valued, $\mathbb{F}$-adapted process $L = (L_t)_{t\geq 0}$ is called a *local time at the maximum* (or just *local time* for short) if the following hold:

1. The support of the Stieltjes measure dL is the closure of the (random) set of times $\{t \geq 0 : X_t = \overline{X}_t\}$ for each $t \geq 0$.

2. For every $\mathbb{F}$-stopping time T such that $X_T = \overline{X}_T$ on $\{T < \infty\}$, P-a.s., the shifted trivariate process

$$(X_{T+t} - X_T, \overline{X}_{T+t} - X_T, L_{T+t} - L_T))_{t\geq 0}$$

is independent of $\mathcal{F}_T$ conditionally on $\{T < \infty\}$ and has the same law under $\mathsf{P}(\cdot|\{T < \infty\})$ as does $(X, \overline{X} - X, L)$ under P.

[2] Since a nondecreasing function has only countably many points of discontinuity (jumps), then G and G' (generically for both of the cases) disagree at most on a countable set. Thus, as $e_p \perp X$, P-a.s., e_p will not equal a point of disagreement of G and G'.

The process which is, P-a.s., identically equal to zero, is excluded. By applying this definition to $-X$, one gets also the notion of a *local time at the minimum*, denoted $\hat{L}$. $\overline{X} - X$ is called the *reflected process at the maximum* (similarly $X - \underline{X}$ the *reflected process at the minimum*).

We let $L_t^{-1} := \inf\{s \geq 0 : L_s > t\}$ be the right-continuous *inverse local time at the maximum* and $H_t = X(L_t^{-1})$ be the *ascending ladder heights* (for $0 \leq t < L_{+\infty}$; $L_t^{-1} := +\infty$ and $H_t := +\infty$ for $t \geq L_{+\infty}$). (L^{-1}, H) is the *ascending ladder process*. By applying the same procedure to $\hat{L}$ we get also the *descending ladder process* $(\hat{L}^{-1}, \hat{H})$.

Remark 1.25. Local times, as described above, always exist (and are unique up to a multiplicative constant, unless 0 is irregular for $[0, \infty)$). The ascending and descending ladder processes are (possibly exponentially killed [Kallenberg, 1997, p. 242]) bivariate subordinators. To them there correspond the bivariate Laplace exponents κ and $\hat{\kappa}$ with

$$\mathsf{E}[\exp\{-\alpha L_1^{-1} - \beta H_1\}\mathbb{1}_{\{1 < L_\infty\}}] = \exp\{-\kappa(\alpha, \beta)\}$$

and

$$\mathsf{E}[\exp\{-\alpha \hat{L}_1^{-1} - \beta \hat{H}_1\}\mathbb{1}_{\{1 < \hat{L}_\infty\}}] = \exp\{-\hat{\kappa}(\alpha, \beta)\}$$

($\{\alpha, \beta\} \subset \overline{\mathbb{C}^{\rightarrow}}$) [Kyprianou, 2006, pp. 149 & 157]. Indeed, κ and $\hat{\kappa}$ are non-zero whenever $\alpha \in \mathbb{C}^{\rightarrow}$, continuous (by the DCT) and they are analytic in the interior of their domains (use e.g. the theorems of Cauchy [Rudin, 1970, p. 206, 10.13 Cauchy's theorem for triangle], Morera [Rudin, 1970, p. 209, 10.17 Morera's theorem] and Fubini).

We do not offer any more details here but refer the reader to, say, [Kyprianou, 2006, Chapter 6].

Next, while e.g. [Kyprianou, 2006, p. 158, Theorem 6.16] is explicit regarding the Wiener-Hopf factorization in the case when X is *not* compound Poisson, we shall actually find use in the sequel of the following result:

Proposition 1.26 (Wiener-Hopf factorization for CP processes)**.** *Let X be compound Poisson and $p > 0$. Then:*

(i) The pairs $(\overline{G}^*_{e_p}, \overline{X}_{e_p})$ and $(e_p - \overline{G}^*_{e_p}, \overline{X}_{e_p} - X_{e_p})$ are independent and infinitely divisible, yielding the factorisation:

$$\frac{p}{p - i\eta - \Psi(\theta)} = \Psi_p^+(\eta, \theta)\Psi_p^-(\eta, \theta),$$

where for $\{\theta, \eta\} \subset \mathbb{R}$,

$$\Psi_p^+(\eta, \theta) := \mathsf{E}[\exp\{i\eta \overline{G}^*_{e_p} + i\theta \overline{X}_{e_p}\}] \text{ and } \Psi_p^-(\eta, \theta) := \mathsf{E}[\exp\{i\eta \underline{G}_{e_p} + i\theta \underline{X}_{e_p}\}].$$

Duality: $(e_p - \overline{G}^*_{e_p}, \overline{X}_{e_p} - X_{e_p})$ *is equal in distribution to* $(\underline{G}_{e_p}, -\underline{X}_{e_p})$. Ψ_p^+ *and* Ψ_p^- *are the Wiener-Hopf factors.*

(ii) The Wiener-Hopf factors may be identified as follows:

$$\mathsf{E}[\exp\{-\alpha \overline{G}^*_{e_p} - \beta \overline{X}_{e_p}\}] = \frac{\kappa^*(p, 0)}{\kappa^*(p + \alpha, \beta)}$$

and

$$\mathsf{E}[\exp\{-\alpha \underline{G}_{e_p} + \beta \underline{X}_{e_p}\}] = \frac{\hat{\kappa}(p, 0)}{\hat{\kappa}(p + \alpha, \beta)}$$

for $\{\alpha, \beta\} \subset \overline{\mathbb{C}^{\rightarrow}}$.

(iii) Here, in terms of the law of X,

$$\kappa^*(\alpha, \beta) := k^* \exp\left(\int_0^\infty \int_{(0,\infty)} (e^{-t} - e^{-\alpha t - \beta x}) \frac{1}{t} \mathsf{P}(X_t \in dx) dt\right)$$

and

$$\hat{\kappa}(\alpha, \beta) = \hat{k} \exp\left(\int_0^\infty \int_{(-\infty,0]} (e^{-t} - e^{-\alpha t + \beta x}) \frac{1}{t} \mathsf{P}(X_t \in dx) dt\right)$$

for $\alpha \in \mathbb{C}^{\rightarrow}$, $\beta \in \overline{\mathbb{C}^{\rightarrow}}$ *and some constants* $\{k^*, \hat{k}\} \subset \mathbb{R}^+$.

(iv) For some constant $k' < 0$ *and then all* $\theta \in \mathbb{R}$*:*

$$k' \Psi(\theta) = \kappa^*(0, -i\theta) \hat{\kappa}(0, i\theta).$$

Alternatively:

(i) The pairs $(\overline{G}_{e_p}, \overline{X}_{e_p})$ *and* $(e_p - \overline{G}_{e_p}, \overline{X}_{e_p} - X_{e_p})$ *are independent and infinitely divisible, yielding the factorisation:*

$$\frac{p}{p - i\eta - \Psi(\theta)} = \Psi_p^+(\eta, \theta) \Psi_p^-(\eta, \theta),$$

where for $\{\theta, \eta\} \subset \mathbb{R}$,

$$\Psi_p^+(\eta, \theta) := \mathsf{E}[\exp\{i\eta \overline{G}_{e_p} + i\theta \overline{X}_{e_p}\}] \text{ and } \Psi_p^-(\eta, \theta) := \mathsf{E}[\exp\{i\eta \underline{G}^*_{e_p} + i\theta \underline{X}_{e_p}\}].$$

Duality: $(e_p - \overline{G}_{e_p}, \overline{X}_{e_p} - X_{e_p})$ *is equal in distribution to* $(\underline{G}^*_{e_p}, -\underline{X}_{e_p})$*.* Ψ^+_p *and* Ψ^-_p *are the Wiener-Hopf factors.*

(ii) The Wiener-Hopf factors may be identified as follows:

$$\mathsf{E}[\exp\{-\alpha\overline{G}_{e_p} - \beta\overline{X}_{e_p}\}] = \frac{\kappa(p,0)}{\kappa(p+\alpha,\beta)}$$

and

$$\mathsf{E}[\exp\{-\alpha\underline{G}^*_{e_p} + \beta\underline{X}_{e_p}\}] = \frac{\hat{\kappa}^*(p,0)}{\hat{\kappa}^*(p+\alpha,\beta)}$$

for $\{\alpha,\beta\} \subset \overline{\mathbb{C}^{\rightarrow}}$*.*

(iii) Here, in terms of the law of X*,*

$$\kappa(\alpha,\beta) = k\exp\left(\int_0^\infty \int_{[0,\infty)} (e^{-t} - e^{-\alpha t - \beta x})\frac{1}{t}\mathsf{P}(X_t \in dx)dt\right)$$

and

$$\hat{\kappa}^*(\alpha,\beta) := \hat{k}^*\exp\left(\int_0^\infty \int_{(-\infty,0)} (e^{-t} - e^{-\alpha t + \beta x})\frac{1}{t}\mathsf{P}(X_t \in dx)dt\right)$$

for $\alpha \in \mathbb{C}^{\rightarrow}$*,* $\beta \in \overline{\mathbb{C}^{\rightarrow}}$ *and some constants* $\{k, \hat{k}^*\} \subset \mathbb{R}^+$*.*

(iv) For some constant $k' < 0$ *and then all* $\theta \in \mathbb{R}$*:*

$$k'\Psi(\theta) = \kappa(0,-i\theta)\hat{\kappa}^*(0,i\theta).$$

Proof. These claims are contained in the remarks regarding compound Poisson processes in [Kyprianou, 2006, p. 167] pursuant to the proof of Theorem 6.16 therein. Analytic continuations have been effected in both parts (iii) using properties of zeros of holomorphic functions [Rudin, 1970, p. 209, Theorem 10.18], the theorems of Cauchy, Morera and Fubini, and finally the finiteness/integrability properties of potential measures [Sato, 1999, p. 203, Theorem 30.10(ii)]. □

Finally we consider:

Definition 1.27 (Spectrally negative Lévy processes)**.** X is said to be *spectrally negative*, if it has no positive jumps, a.s., and does not have a.s. monotone paths.

For the remainder of this subsection, assume X is spectrally negative. In this case, further particulars of the fluctuation theory (e.g. of the Wiener-Hopf factorization) for X can be established, see [Bertoin, 1996, Chapter VII] [Sato, 1999, Section 9.46] and especially [Kyprianou, 2006, Chapter 8]. In particular, X admits a Laplace exponent ψ, defined via $\psi(\beta) := \log \mathsf{E}[e^{\beta X_1}]$ ($\beta \in \overline{\mathbb{C}^{\rightarrow}}$) and for $q \geq 0$, we may let $\Phi(q)$ denote the largest root of $\psi - q$ on $[0, \infty)$ [Kyprianou, 2006, p. 211].

In addition, the *two-sided exit problem* admits a semi-explicit solution in terms of two families of *scale functions*, $(W^{(q)})_{q\in[0,\infty)}$ and $(Z^{(q)})_{q\in[0,\infty)}$ [Kyprianou, 2006, Section 8.2]. Indeed, for each $q \geq 0$, we have $W^{(q)}(x) = 0$ for $x < 0$ and on $[0, \infty)$, $W^{(q)}$ is characterized as the unique continuous and strictly increasing function whose Laplace transform satisfies:

$$\int_0^\infty e^{-\beta x} W^{(q)}(x)dx = \frac{1}{\psi(\beta) - q} \quad (\beta > \Phi(q)),$$

whereas:

$$Z^{(q)}(x) = 1 + q\int_0^x W^{(q)}(y)dy \quad (x \in \mathbb{R}).$$

Then, with $\{x, y\} \subset (0, \infty)$:

$$\mathsf{E}[e^{-qT_y}\mathbb{1}_{\{T_y<T_x^-\}}] = \frac{W^{(q)}(x)}{W^{(q)}(x+y)},$$

while:

$$\mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<T_y\}}] = Z^{(q)}(x) - Z^{(q)}(x+y)\frac{W^{(q)}(x)}{W^{(q)}(x+y)}$$

(note that from the regularity of 0 for $(0, \infty)$ [Kyprianou, 2006, p. 212] and by the strong Markov property of Lévy processes [Sato, 1999, p. 278, Theorem 40.10] applied at the time T_y, $\hat{T}_y = T_y$, P-a.s.).

1.2.3 Continuous-time Markov chains

For the general theory of continuous-time Markov chains (CTMCs) see e.g. [Chung, 1960; Grimmett and Stirzaker, 2001; Norris, 1997].

For our part, and for the reader's convenience, we provide below a rigorous exposition of the construction of a (non-explosive) CTMC starting with a regular Q-matrix Q (see Definition 1.28 below) as its basic datum (we will see Q in just such a rôle in Chapter 2 below). This is as much to introduce the relevant concepts, as it is to demonstrate that CTMCs are (comparatively speaking) very simple stochastic objects indeed. Moreover, given that CTMCs feature (together with Lévy processes) centrally in this thesis, such an exposition seems a small sacrifice of space with the benefit that the thesis is more self-contained. The reader already familiar with these classical results will of course harmlessly skip their proofs (but not their formulation).

In addition to this, we also establish some more specific properties of CTMCs, which will prove useful in the sequel (see Proposition 1.35 and Theorem 1.39).

Throughout we fix a countable non-empty set S endowed (where necessary) with the discrete σ-algebra making it into a standard measurable space [Dudley, 2004, p. 440].

Let us first recapitulate the definition of a Q-matrix [Norris, 1997, p. 60]:

Definition 1.28 (Q-matrix)**.** A *Q-matrix* is a mapping from $S \times S$ into the reals, entries denoted $Q_{su} := Q(s, u) := Q((s, u))$ (for $\{s, u\} \subset S$), and having the following properties:

(i) nonnegative off-diagonal entries: $Q_{su} \geq 0$, whenever $\{s, u\} \subset S$ and $s \neq u$.

(ii) nonpositive diagonal entries: $Q_{ss} \leq 0$ for $s \in S$.

(iii) rows summing to 0: for each $s \in S$, $-Q_{ss} = \sum_{u \in S \backslash \{s\}} Q_{su}$.

It is called *regular* if $\sup\{-Q_{ss} : s \in S\} < \infty$, i.e. if the entries of Q are uniformly bounded in absolute value.

One considers the Q-matrix as furnishing the infinitesimal generator of the continuous-time Markov chain. Hence the following definition:

Definition 1.29 (Infinitesimal generator)**.** Let Q be a regular Q-matrix on S. We define the *infinitesimal generator* corresponding to Q, as the mapping $L := L_Q$ on $l^\infty(S)$, the set of all real- (or complex-) valued bounded functions on S, as follows ($f \in l^\infty(S)$, $s \in S$):

$$Lf(s) := \sum_{u \in S} Q_{su} f(u). \tag{1.3}$$

Lemma 1.30. *If Q is a regular Q-matrix, then the corresponding infinitesimal generator $L : l^\infty(S) \to l^\infty(S)$ is a bounded linear mapping on the Banach space $l^\infty(S)$ with the supremum norm.*

Proof. Follows at once from the regularity of Q. □

We may hence fully exploit the theory of Banach spaces. Recall that the space of bounded linear operators on $l^\infty(S)$, denoted $\mathcal{L}(l^\infty(S))$, is in turn a Banach space, and if a sequence $(A_n)_{n \in \mathbb{N}}$ in this space converges to A, then for any $f \in l^\infty(S)$,

$(A_n f)_{n\in\mathbb{N}}$ converges in $l^\infty(S)$ and its limit is $\lim_{n\to\infty} A_n f = Af$ (see e.g. [Reed and Simon, 1980, p. 70, Theorem III.2]). Furthermore, for a sequence $(f_n)_{n\in\mathbb{N}}$ converging to f in $l^\infty(S)$, $(f_n(s))_{n\in\mathbb{N}}$ converges to $f(s)$ (uniformly in $s \in S$).

Definition 1.31 (Transition semigroup)**.** Let Q be a regular Q-matrix on the state space S with corresponding infinitesimal generator L. We define the *transition semigroup* $(P_t)_{t\geq 0}$ associated to Q by

$$P_t := \exp(tL) := \sum_{k=0}^{\infty} \frac{(tL)^k}{k!} \quad (t \geq 0) \tag{1.4}$$

with the series converging in $\mathcal{L}(l^\infty(S))$ as it is absolutely summable (see e.g. [Reed and Simon, 1980, p. 71, Theorem III.3]).

We have the following key result (cf. [Norris, 1997, pp. 62 & 63, Theorems 2.1.1 & 2.1.2] for the case of S finite and [Grimmett and Stirzaker, 2001, p. 267, Theorem 10] for the general case):

Theorem 1.32 (Transition semigroup)**.** *$(P_t)_{t\geq 0}$ is a family of bounded linear operators on $l^\infty(S)$ satisfying:*

(i) $P_t P_s = P_{t+s}$*, whenever* $\{t, s\} \subset [0, \infty)$*.*

(ii) $P_0 = I$*, the identity on* $l^\infty(S)$*.*

(iii) $\lim_{t\downarrow 0} P_t = I$*, while* $L = \frac{\mathrm{d}P_t}{\mathrm{d}t}|_{t=0+}$ *(in* $\mathcal{L}(l^\infty(S))$*).*

(iv) $\|P_t\| \leq 1$ *and* $P_t \mathbb{1}_S = \mathbb{1}_S$ *for all* $t \in [0, \infty)$*.*

(v) If $t \geq 0$*,* $f \in l^\infty(S)$ *and* $f \geq 0$*, then* $P_t f \geq 0$*.*

Moreover, if we define $P_{su}(t) = (P_t \mathbb{1}_{\{u\}})(s)$ *(*$t \in [0, \infty)$*,* $\{s, u\} \subset S$*) then for each* $t \in [0, \infty)$*,* $(P_{su}(t))_{(s,u)\in S\times S}$ *is a stochastic matrix, i.e.:*

(a) $P_{su}(t) \geq 0$*, whenever* $\{s, u\} \subset S$*.*

(b) $\sum_{u\in S} P_{su}(t) = 1$ *for any* $s \in S$ *and* $t \geq 0$*.*

and the matrices satisfy the Chapman-Kolmogorov equations:

(c) $\sum_{w\in S} P_{uw}(t)P_{wv}(s) = P_{uv}(t + s)$*, whenever* $\{s, t\} \subset [0, \infty)$ *and* $\{u, v\} \subset S$*.*

Finally, letting $P_t(s, A) := \sum_{u\in A} P_{su}(t)$ *(for* $t \in [0, \infty)$*,* $A \subset S$ *and* $s \in S$*), the latter constitute a transition function on* S*, to wit:*

(I) for a fixed $s \in S$ and $t \in [0, \infty)$, $(A \mapsto P_t(s, A))$ is a probability measure on $(S, 2^S)$.

(II) for a fixed $A \subset S$ and $t \in [0, \infty)$, $(s \mapsto P_t(s, A))$ is measurable (this is trivial, since $(S, 2^S)$ is discrete).

(III) for $s \in S$ and $A \subset S$, we have $P_0(s, A) = \delta_s(A) = \mathbb{1}_A(s)$.

(IV) $\int_S P_t(x, dy) P_s(y, A) = P_{t+s}(x, A)$, for any $x \in S$ and $A \subset S$.

Proof. (ii) is clear. Consider now (i). Let $f \in l^\infty(S)$ and $u \in S$. Then:

$$\begin{aligned}
P_{t+s} f(u) &= \left(\sum_{k=0}^{\infty} \frac{(t+s)^k}{k!} L^k f \right)(u) = \sum_{k=0}^{\infty} \frac{(t+s)^k}{k!} L^k f(u) \\
&= \sum_{k=0}^{\infty} \sum_{n=0}^{k} \frac{1}{n!(k-n)!} t^n s^{k-n} L^k f(u) \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{k=n}^{\infty} \frac{s^{k-n}}{(k-n)!} (L^n (L^{k-n} f))(u) \\
&= \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{k=n}^{\infty} \frac{s^{k-n}}{(k-n)!} L^n L^{k-n} f \right)(u) \\
&= \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} L^n \sum_{k=n}^{\infty} \frac{s^{k-n}}{(k-n)!} L^{k-n} f \right)(u) = (P_t(P_s f))(u),
\end{aligned}$$

where the interchange of the order of summation in line three is justified via Fubini by the fact that the series is absolutely convergent, whilst the "taking out of L^n" in the last line uses continuity and linearity of L^n.

We next establish a key version of bounded convergence;

Lemma 1.33 (Lemma on bounded convergence). *We have for any $t \geq 0$, $f \in l^\infty(S)$ and $s \in S$:*

$$(P_t f)(s) = \sum_{u \in S} f(u)(P_t \mathbb{1}_{\{u\}})(s), \tag{1.5}$$

where the series on the right converges absolutely.

Proof. To see this, note that for all $f \in l^\infty(S)$, $s \in S$,

$$(Lf)(s) = \sum_{u_1 \in S} Q_{su_1} f(u_1)$$

and the series is absolutely convergent, since $\sum_{u_1 \in S} |Q_{su_1} f(u_1)| \leq 2q\|f\|$, where $q := \sup\{-Q_{ss} : s \in S\}$. We claim furthermore that for each $k \geq 1$, for all $f \in l^\infty(S)$, $s \in S$:

$$(L^k f)(s) = \sum_{u_1 \in S} \cdots \sum_{u_k \in S} Q_{su_1} \cdots Q_{u_{k-1}u_k} f(u_k) \tag{1.6}$$

where the iterated series $\sum_{u_1} \cdots \sum_{u_k}$ converges absolutely, moreover:

$$\sum_{u_1 \in S} \cdots \sum_{u_k \in S} |Q_{su_1} \cdots Q_{u_{k-1}u_k} f(u_k)| \leq (2q)^k \|f\|. \tag{1.7}$$

We prove by induction. Supposing the claim for k, we have for every $f \in l^\infty(S)$ and each $s \in S$:

$$\begin{aligned}
(L^{k+1}f)(s) &= (LL^k f)(s) = \sum_{u_1 \in S} Q_{su_1}(L^k f)(u_1) \\
&= \sum_{u_1 \in S} Q_{su_1} \left(\sum_{u_2 \in S} \cdots \sum_{u_{k+1} \in S} Q_{u_1 u_2} \cdots Q_{u_k u_{k+1}} f(u_{k+1}) \right) \\
&= \sum_{u_1 \in S} \cdots \sum_{u_{k+1} \in S} Q_{su_1} \cdots Q_{u_k u_{k+1}} f(u_{k+1}),
\end{aligned}$$

where the last equality follows from $\sum_{u_1 \in S} \cdots \sum_{u_{k+1} \in S} |Q_{su_1} \cdots Q_{u_k u_{k+1}} f(u_{k+1})| \leq \sum_{u_1 \in S} |Q_{su_1}| (2q)^k \|f\| \leq (2q)^{k+1} \|f\|$ so the iterated series $\sum_{u_1} \cdots \sum_{u_{k+1}}$ is again absolutely convergent, and thus (1.6) and (1.7) obtain at once for $k+1$, whence by induction we are done.

Thus:

$$\begin{aligned}
(P_t f)(s) &= \sum_{k=0}^{\infty} \frac{t^k}{k!} (L^k f)(s) \\
&= f(s) + \sum_{k=1}^{\infty} \frac{t^k}{k!} \left(\sum_{u_1 \in S} \cdots \sum_{u_k \in S} Q_{su_1} \cdots Q_{u_{k-1}u_k} f(u_k) \right) \\
&= f(s) + \sum_{k=1}^{\infty} \left(\sum_{u \in S} \frac{t^k}{k!} f(u) \sum_{u_1 \in S} \cdots \sum_{u_{k-1} \in S} Q_{su_1} \cdots Q_{u_{k-1}u} \right) \\
&= f(s) + \sum_{u \in S} f(u) \left(\sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{u_1 \in S} \cdots \sum_{u_{k-1} \in S} Q_{su_1} \cdots Q_{u_{k-1}u} \right) \\
&= \sum_{u \in S} f(u) \left(\delta_{su} + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{u_1 \in S} \cdots \sum_{u_{k-1} \in S} Q_{su_1} \cdots Q_{u_{k-1}u} \right) \\
&= \sum_{u \in S} f(u) (P_t \mathbb{1}_{\{u\}})(s),
\end{aligned}$$

where $\sum_{u_1 \in S} \cdots \sum_{u_{k-1} \in S} Q_{su_1} \cdots Q_{u_{k-1}u}$ is understood to mean Q_{su}, when $k = 1$; whereas by Fubini we were allowed to interchange the order of summation in line four because the series is absolutely convergent:

$$\sum_{k=0}^{\infty}\sum_{u_1\in S}\cdots\sum_{u_k\in S}\frac{t^k}{k!}|Q_{su_1}\cdots Q_{u_{k-1}u_k}f(u_k)|\le e^{2qt}\|f\|$$

(the term with $k=0$ being understood to mean $f(s)$). In particular, we see that the last series in the evaluation of $(P_tf)(s)$ above is absolutely convergent. Thus we obtain (1.5). $\square$

Now, with regard to (v), fix $t\ge 0$, and write for $n\in\mathbb{N}$, $P_t=(P_{t/n})^n=(I+\frac{t}{n}L+B_n)^n$, where $B_n:=\sum_{k=2}^{\infty}\frac{(t/n)^k}{k!}L^k$ and hence $\|B_n\|\le\frac{t^2\|L\|^2}{n^2}e^{\|L\|t/n}$. Now $((I+\frac{t}{n}L)f)(s)=f(s)+\sum_{u\in S}\frac{t}{n}Q_{su}f(u)$. Denote $A_n:=\frac{t}{n}L$. Then, since $q:=\sup\{-Q_{ss}:s\in S\}<\infty$, for all sufficiently large n, and then for all $f\in l^\infty(S)$, $(I+A_n)f\ge 0$ and hence $(I+A_n)^nf\ge 0$. Next $(I+A_n+B_n)^n=(I+A_n)^n+R_n$, where $R_n:=\sum_{k=0}^{n-1}\binom{n}{k}(I+A_n)^kB_n^{n-k}$. It follows that (for $f\in l^\infty(S)$): $P_tf=(I+A_n)^nf+R_nf$, where $\|R_nf\|\le\|R_n\|\|f\|\le\|f\|\sum_{k=0}^{n-1}\binom{n}{k}(1+\|A\|)^k\|B_n\|^{n-k}=\|f\|((1+\|A_n\|+\|B_n\|)^n-(1+\|A_n\|)^n)$. Letting $a:=t\|L\|$ and $b_n:=\|B_n\|$, we have, for all sufficiently large n, $P_tf\ge-\mathbb{1}_S\|f\|\left[(1+\frac{a}{n}+b_n)^n-(1+\frac{a}{n})^n\right]\to 0$ pointwise, as $n\to\infty$.[3] (a) is then a simple corollary, whereas for (b) note as follows (for $t\ge 0$, $s\in S$):

$$\sum_{u\in S}P_{su}(t)\quad=\quad\sum_{u\in S}(P_t\mathbb{1}_{\{u\}})(s)=(P_t\mathbb{1}_S)(s)=1,$$

by the Lemma on bounded convergence (with $f=\mathbb{1}_S$) and the property that rows of Q sum to zero (so that $L^k\mathbb{1}_S=L^{k-1}(L\mathbb{1}_S)=0$ for all $k\ge 1$, while $L^0\mathbb{1}_S=\mathbb{1}_S$). Then (iv) follows using (a), (b) and the Lemma on bounded convergence.

With regard to (c), we have (the first sum is pointwise, then use the Lemma on bounded convergence):

$$\begin{aligned}P_{uv}(t+s)&=(P_{t+s}\mathbb{1}_{\{v\}})(u)=(P_tP_s\mathbb{1}_{\{v\}})(u)\\&=\left(P_t\left(\sum_{w\in S}\mathbb{1}_{\{w\}}P_{wv}(s)\right)\right)(u)=\sum_{w\in S}P_{wv}(s)(P_t\mathbb{1}_{\{w\}})(u)\end{aligned}$$

[3] It is a standard result that for any $a\in\mathbb{R}$ and any sequence of real numbers b_n with $b_nn\to 0$ as $n\to\infty$, one has $(1+\frac{a}{n}+b_n)^n\to e^a$ as $n\to\infty$. Indeed, let $\epsilon>0$. Then for all sufficiently large n, $|b_n|\le\epsilon/n$, hence for all sufficiently large n:

$$(1+\frac{a-\epsilon}{n})^n\le(1+\frac{a}{n}+b_n)^n\le(1+\frac{a+\epsilon}{n})^n.$$

Now take the limits inferior and superior as $n\to\infty$, and let finally $\epsilon\downarrow 0$ to get the desired claim (via the continuity of the exponential function).

$$= \sum_{w \in S} P_{uw}(t) P_{wv}(s).$$

For (iii), simply note that $\|P_t - I\| = \|\sum_{k=1}^{\infty} \frac{t^k}{k!} L^k\| \leq t\|L\|e^{\|L\|t} \downarrow 0$, while $\|t^{-1}(P_t - I) - L\| = \|\sum_{k=2}^{\infty} \frac{t^{k-1}L^k}{k!}\| \leq t \sum_{k=2}^{\infty} \frac{t^{k-2}\|L\|^k}{k!} \leq t\|L\|^2 e^{\|L\|t} \downarrow 0$, as $t \downarrow 0$.

Finally, taking all of the above into account, (I)—(IV) are immediate. □

We have seen in this key result, then, that a regular Q-matrix gives rise, in a natural way, to a transition function on S. We shall use the latter to show the existence of our continuous-time Markov chain. Let then Q be a regular Q-matrix on S, as above. Let δ be a distribution on S. Define for any $n \geq 0$, $0 < t_1 < \cdots < t_n$ and $B \subset S^{n+1}$:

$$\begin{aligned} \mu^{\delta}_{0,t_1,\ldots,t_n}(B) &:= \int \delta(dx_0) \int P_{t_1}(x_0, dx_1) \int P_{t_2 - t_1}(x_1, dx_2) \\ &\quad \times \cdots \times \int P_{t_n - t_{n-1}}(x_{n-1}, dx_n) \mathbb{1}_B(x_0, \ldots, x_n). \end{aligned} \tag{1.8}$$

We obtain $\mu^{\delta}_{t_1,\ldots,t_n}$ by taking $B = S \times H$, $H \subset S^n$. Thus we have defined a consistent family of laws. Indeed, by the Chapman-Kolmogorov identity (IV) and a monotone class argument:

$$\int_S P_t(x, dy) \int_S P_s(y, dz) f(z) = \int_S P_{t+s}(x, dz) f(z) \tag{1.9}$$

(for all $f \in l^{\infty}(S)$, whenever $\{s, t\} \subset [0, \infty)$ and $x \in S$).

Hence, by the Kolmogorov extension theorem (see e.g. [Dudley, 2004, p. 441]) there exists a unique measure P^{δ} on $S^{[0,\infty)}$ (with the product σ-algebra) extending this family.

Definition 1.34 (Continuous time Markov chain in law)**.** An S-valued process $X := (X_t)_{t \geq 0}$ on a probability space $(\Omega, \mathcal{F}, P)$ is called a *continuous-time Markov chain (in law)*, state space S, initial distribution δ, Q-matrix Q, if its law on the space $S^{[0,\infty)}$ under the measure P is P^{δ}. We shall usually drop the qualification "in law" altogether.

Proposition 1.35. *Suppose X is, as in the above definition, a continuous-time Markov chain in law with a regular Q-matrix. It is then a time-homogeneous Markov process with transition functions $(P_t)_{t \geq 0}$ and for any $0 \leq t \leq T$, $u \in S$ one has, in particular:*

$$\mathsf{P}(X_T = u | X_t = v) = P_{vu}(T - t) = (e^{(T-t)L} \mathbb{1}_{\{u\}})(v), \tag{1.10}$$

P_{X_t}*-a.s. in* $v \in S$*. Moreover, endowing* S *with the discrete topology,* X *is a Feller process, if and only if the following "Feller condition" on* Q *is verified:* $(s \mapsto Q_{ss'}) \in c_0(S)$ *for every* $s' \in S$*.*[4]

Proof. The first claim is immediate from the definition and the fact that Identity (1.8) extends to all nonnegative functions by the usual argument; then one can apply [Revuz and Yor, 1999, p. 81, Proposition 1.4]. For the Feller property, we have as follows. First, if for some $s' \in S$, $(s \mapsto Q_{ss'}) \notin c_0(S)$, then $L\mathbb{1}_{\{s'\}} \notin c_0(S)$, while $P_t\mathbb{1}_{\{s'\}} = \mathbb{1}_{\{s'\}} + tL\mathbb{1}_{\{s'\}} + g_t$, for some $g_t \in l^\infty(S)$ with $\|g_t\|/t \to 0$ as $t \downarrow 0$. But then X cannot be a Feller process [Revuz and Yor, 1999, Section III.2]. Conversely, suppose $(s \mapsto Q_{ss'}) \in c_0(S)$ for every $s' \in S$. We show $L(c_0(S)) \subset c_0(S)$. Let then $f \in c_0(S)$, $\epsilon > 0$. Put $S' := \{|f| \geq \epsilon/(4q)\}$, where $q := \sup\{-Q_{ss} : s \in S\}$. Then $\{|Lf| \geq \epsilon\} \subset \{s \in S : |\sum_{s' \in S'} Q_{ss'} f(s')| \geq \epsilon/2\} \subset \cup_{s' \in S'}\{s \in S : |Q_{ss'}| \geq \frac{\epsilon}{2\|f\||S'|}\}$ and this is a finite set. Then we can show $P_t(c_0(S)) \subset c_0(S)$, $t \geq 0$, whence it will be established (together with the findings of Theorem 1.32) that $(P_t)_{t \geq 0}$ is a Feller semigroup [Revuz and Yor, 1999, Section III.2]. Let then, for the last time, $t \geq 0$, $f \in c_0(S)$ and $\epsilon > 0$. Note that for a sufficiently large, but finite, $K \in \mathbb{N}$, with $P_{Kt} := \sum_{k=0}^{K} \frac{t^k L^k}{k!}$, $\|(P_{Kt} - P_t)f\| \leq \epsilon/2$, so that $\{|P_t f| \geq \epsilon\} \subset \{|P_{Kt} f| \geq \epsilon/2\}$. But clearly $P_{Kt}(c_0(S)) \subset c_0(S)$, since $L(c_0(S)) \subset c_0(S)$, so this concludes the argument. □

Remark 1.36. In case X from Proposition 1.35 has the Feller property, it admits a càdlàg modification [Kallenberg, 1997, p. 325, Theorem 17.15]. For such a version, then, the number of jumps on every sample path of X is locally finite, so that the sequence of jump times $(J_j)_{j \geq 1}$ is increasing to, and possibly reaching, $+\infty$.

The construction of a (non-explosive) CTMC via the Kolmogorov extension theorem, as above, is certainly very straightforward and analytically appealing, but also quite abstract. Probabilistically, we like to think of CTMCs in the following terms:

Theorem 1.37. *On a probability space* $(\Omega, \mathcal{F}, \mathsf{P})$*, let* X *be a sample-path right-continuous CTMC with a regular Q-matrix* Q*, infinitesimal generator* L*, state space* S*. Denote by* $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ *its natural filtration, and by* $H_t := \inf\{s \geq t : X_s \neq X_t\}$ *the time of the first jump of* X *after* t *(*$t \geq 0$*). Let furthermore* τ *be any stopping time of* $\mathbb{F}$*, such that* $\mathsf{P}(\tau < \infty) > 0$*. Then, under the conditional measure* $\mathsf{P}(\cdot|\tau < \infty)$*, conditionally on* X_τ*;*

1. $\mathcal{F}_\tau$ *is independent of* $(X_{\tau+t})_{t \geq 0}$*, and the law of* $(X_{\tau+t})_{t \geq 0}$ *is* $P^{\mathsf{P} \circ X_\tau^{-1}}$*,* $\mathsf{P}(\cdot|\tau < \infty)$*-a.s. (strong Markov property);*

[4]Note that $f \in c_0(S)$, if and only if $\{|f| \geq \epsilon\}$ is finite for every $\epsilon > 0$.

2. *the law of $H_\tau - \tau$ is $\mathrm{Exp}(-Q_{X_\tau X_\tau})$, $\mathsf{P}(\cdot|\tau < \infty)$-a.s.; $H_\tau - \tau$ is independent of $X(H_\tau)$ (conditionally on $\{-Q_{X_\tau X_\tau} > 0\}$), whilst the probability mass function of $X(H_\tau)$ on $S\backslash\{X_\tau\}$ is $(u \mapsto Q_{X_\tau u}/(-Q_{X_\tau X_\tau}))$, $\mathsf{P}(\cdot|\tau < \infty, -Q_{X_\tau X_\tau} > 0)$-a.s.*

So (assuming $-Q_{ss} > 0$ for all $s \in S$), conditionally on X_τ, the history of X up to τ, $\mathcal{F}_\tau$, the time to the next jump of X after τ, $H_\tau - \tau$, as well as the position of X at the next jump after τ, $X(H_\tau)$, are all jointly independent (under $\mathsf{P}(\cdot|\tau < \infty)$).

Remark 1.38. For a right-continuous CTMC (with sequence of jump times $(J_j)_{j\geq 1}$) it follows then that $J_j < +\infty$ for all $j \geq 1$, P-a.s., provided $-Q_{ss} > 0$ for all $s \in S$.

Proof. For the first part, we need only establish for $A \in \mathcal{F}_\tau$, $n \in \mathbb{N}_0$, $0 < t_1 < \cdots t_n$, $\{u_1, \ldots, u_n\} \subset S$, and $u_0 \in S$ with $\mathsf{P}(X_\tau = u_0, \tau < \infty) > 0$, that

$$\begin{aligned}&\mathsf{P}(A, \{X_{\tau+t_1} = u_1, \ldots, X_{\tau+t_n} = u_n\}, X_\tau = u_0, \tau < \infty) = \\ &\mathsf{P}(A, X_\tau = u_0, \tau < \infty)(e^{Lt_1}\mathbb{1}_{u_1})(u_0)\cdots(e^{L(t_n-t_{n-1})}\mathbb{1}_{u_n})(u_{n-1}).\end{aligned}$$

Since one can approximate τ by a nonincreasing sequence of $\mathbb{F}$-stopping times $(\tau_n)_{n\geq 1} \to \tau$, each assuming only finitely many values and with $\{\tau_n < \infty\} = \{\tau < \infty\}$ for all $n \in \mathbb{N}$, and pass to the limit using the DCT and right-continuity of the sample paths, it will be assumed without loss of generality τ assumes only denumerably many values. Then by additivity of probability measures, it is sufficient to verify the above equality with $\tau = t$ in place of $\tau < \infty$, where t is any, but fixed, finite element of the range of τ, satisfying $\mathsf{P}(X_t = u_0) > 0$. But the latter is then immediately made clear by the Markov property of X.

To establish the second part of the theorem let us show for $l \in \cup_{n\in\mathbb{N}_0}\{\frac{k}{2^n} : k \in \mathbb{N}_0\}$ and $\{u, u_0\} \subset S$, satisfying $\mathsf{P}(X_\tau = u_0, \tau < +\infty, H_\tau < \infty) > 0$, $u \neq u_0$, that:

$$\begin{aligned}&\mathsf{P}(X(H_\tau) = u, H_\tau - \tau > l, H_\tau < \infty, X_\tau = u_0, \tau < \infty) = \\ &\mathsf{P}(H_\tau < \infty, X_\tau = u_0, \tau < \infty)e^{Q_{u_0u_0}l}\frac{Q_{u_0u}}{-Q_{u_0u_0}}.\end{aligned}$$

Again by approximation, we may assume, without loss of generality, that τ assumes only denumerably many values; and further by additivity of probability measures it is just as well if $\tau < \infty$ in the above is replaced by $\tau = t$, where t is any finite element of the range of τ with $\mathsf{P}(H_t < \infty, X_t = u_0, \tau < \infty) > 0$. Then thanks to right-continuity of the sample paths and the DCT:

$$\mathsf{P}(X(H_t) = u, H_t - t > l, H_t < \infty, X_t = u_0) =$$
$$\lim_{n\to\infty} \sum_{m=0}^{\infty} \mathsf{P}(X_t = u_0, X_{t+1/2^n} = u_0, \dots, X_{t+l+m/2^n} = u_0, X_{t+l+(m+1)/2^n} = u) =$$
$$\mathsf{P}(X_t = u_0) \lim_{n\to\infty} ((e^{L\frac{1}{2^n}} \mathbb{1}_{u_0})(u_0))^{l/(1/2^n)} \sum_{m=0}^{\infty} ((e^{L\frac{1}{2^n}} \mathbb{1}_{u_0})(u_0))^m (e^{L\frac{1}{2^n}} \mathbb{1}_u)(u_0).$$

Thus, in order to establish the second part of the theorem it will be sufficient to demonstrate that:

$$\mathsf{P}(H_t - t > l, X_t = u_0) = \mathsf{P}(X_t = u_0) \lim_{n\to\infty} ((e^{L\frac{1}{2^n}} \mathbb{1}_{u_0})(u_0))^{l/(1/2^n)} = \mathsf{P}(X_t = u_0) e^{Q_{u_0 u_0} l}$$

and

$$\lim_{n\to\infty} \sum_{m=0}^{\infty} ((e^{L\frac{1}{2^n}} \mathbb{1}_{u_0})(u_0))^m (e^{L\frac{1}{2^n}} \mathbb{1}_u)(u_0) = \frac{Q_{u_0 u}}{-Q_{u_0 u_0}},$$

the latter only provided $-Q_{u_0 u_0} > 0$. But this is clear. □

Lastly, we note for future reference, the result:

Theorem 1.39. *Let X be a continuous-time Markov chain, state space S, initial distribution δ, Q-matrix Q (assumed regular); $d \in \mathbb{N}$. Then:*

(a) Suppose $S = \mathbb{Z}_h^d$ and Q is spatially homogeneous, i.e. $Q_{ss'}$ depends only on $s - s'$ ($\{s, s'\} \subset S$). Then so is the infinitesimal generator L (respectively the transition semigroup $(P_t)_{t\geq 0}$) associated to Q (meaning, $(L\mathbb{1}_{\{s\}})(s')$ (respectively, for every $t \geq 0$, $(P_t\mathbb{1}_{\{s\}})(s')$) depends only on $s - s'$). Furthermore, for $t \geq 0$, $\{x, s\} \subset S$ and any $[-\infty, +\infty]$-valued measurable function f on S, one has:

$$\int P_t(x, dy) f(y) = \int P_t(x + s, dy) f(y - s), \tag{1.11}$$

in the sense that the left-hand side is well-defined, precisely when the right-hand side is so, in which case they are equal. Finally, X has stationary and independent increments and for $0 \leq s \leq t$, the law of $X_t - X_s$ is $P_{t-s}(0, \cdot)$.

(b) For any $s \geq 0$, $\mathsf{P}(X_t \neq X_s) \to 0$, as $t \to s$. In particular, if S is endowed with any metric d, then X is stochastically continuous, i.e. for every $\epsilon > 0$ and $s \geq 0$, one has $\mathsf{P}(d(X_t, X_s) > \epsilon) \to 0$ as $t \to s$.

Proof. We consider first (a). That $(P_t)_{t\geq 0}$ is spatially homogeneous, when Q is, follows at once from the equality $(P_t\mathbb{1}_{\{s\}})(s') = \sum_{k=0}^{\infty} \frac{t^k}{k!} (L^k \mathbb{1}_{\{s\}})(s')$, upon noting

(1.6), which gives spatial homogeneity of L^k for every $k \geq 0$. Equation (1.11) follows from the the spatial homogeneity of $(P_t)_{t\geq 0}$ and a monotone class argument. Using this relation in (1.8), from the innermost to the outermost integral, in consecutive order, one obtains for every $n \in \mathbb{N}_0$, $0 < t_1 < \cdots < t_n$ and $B \subset S^{n+1}$:

$$\mathsf{P}_{(X_0, X_{t_1} \cdots, X_{t_n})}(B) = \int \delta(dx_0) \int P_{t_1}(0, dx_1) \cdots \int P_{t_n - t_{n-1}}(0, dx_n) \mathbb{1}_B \circ A(x_0, \ldots, x_n),$$

where $A : S^{n+1} \to S^{n+1}$ is given by $A(x_0, x_1, \ldots, x_n) = (x_0, x_1 + x_0, \ldots, x_n + x_{n-1} + \cdots + x_0)$, $x \in A$. Then use "change of variables" theorem for the bijection A (see e.g. [Dudley, 2004, p. 121, Theorem 4.1.11]) to obtain:

$$\mathsf{P}_{(X_0, X_{t_1}, \ldots, X_{t_n})} = (\delta \times P_{t_1}(0, \cdot) \times \cdots \times P_{t_n - t_{n-1}}(0, \cdot)) \circ A^{-1},$$

or, since A is a bijection,

$$\mathsf{P}_{(X_0, X_{t_1}, \ldots, X_{t_n})} \circ (A^{-1})^{-1} = \delta \times P_{t_1}(0, \cdot) \times \cdots \times P_{t_n - t_{n-1}}(0, \cdot).$$

But

$$\mathsf{P}_{(X_0, X_{t_1}, \ldots, X_{t_n})} \circ (A^{-1})^{-1} = \mathsf{P}_{(X_0, X_{t_1} - X_0, \ldots, X_{t_n} - X_{t_{n-1}})},$$

from which the stationarity and the independence of increments follows at once.

Now consider (b). We have (assume first $t > s$; let $S' := \{x \in S : \mathsf{P}(X_s = x) > 0\}$):

$$\begin{aligned}
\mathsf{P}(X_t \neq X_s) = 1 - \mathsf{P}(X_t = X_s) &= \sum_{x \in S'} \mathsf{P}(X_s = x)(1 - \mathsf{P}(X_t = x | X_s = x)) \\
&= \sum_{s \in S'} \mathsf{P}(X_s = x)(1 - P_{t-s}\mathbb{1}_{\{x\}})(x) = \sum_{x \in S'} \mathsf{P}(X_s = x)((I - P_{t-s})\mathbb{1}_{\{x\}})(x) \\
&\leq \sum_{x \in S'} \mathsf{P}(X_s = x) \|I - P_{t-s}\| = \|I - P_{t-s}\|.
\end{aligned}$$

Then, for any $t \neq s$, $\mathsf{P}(X_t \neq X_s) \leq \|I - P_{|t-s|}\| \to 0$ as $t \to s$ by Theorem 1.32(iii). The final claim is an immediate corollary. $\square$

Remark 1.40. It follows that a continuous-time Markov chain having state space $\mathbb{Z}_h^d$, a spatially homogeneous Q-matrix, and initial position 0, a.s., is a compound Poisson process in law [Sato, 1999, p. 135, Theorem 21.2] (once one respects the natural inclusion $\mathbb{Z}_h^d \hookrightarrow \mathbb{R}^d$).

1.3 Structure of the remainder of the thesis

In Chapter 2 we present and analyze a weak continuous-time Markov chain approximation to a Lévy process.

In Chapter 4 this approximation is applied to obtaining a general algorithm for the calculation of the scale functions of a spectrally negative Lévy process.

A spectrally negative Lévy process is thus approximated by what is a random walk, skip-free to the right, and embedded into continuous time as a compound Poisson process. Hence, the setting up of the algorithm necessitates a fluctuation theory (and, in particular, a theory of scale functions) for the latter type of Lévy processes (called 'upwards skip-free Lévy chains'), and this is the subject matter of Chapter 3, Section 3.2. These results are interesting in their own right.

Moreover, together with Lévy processes which have no positive jumps (a.s.), upwards skip-free Lévy chains exhaust the class of Lévy processes exhibiting the property of having (conditionally on the process going above the level in question) a.s. constant overshoots. We discuss this related finding in Chapter 3, Section 3.1.

Finally, note that the chapter abstracts (immediately following the chapter titles) provide somewhat more exhaustive summaries of their respective contents.

Chapter 2

A continuous-time Markov chain approximation to Lévy processes

We consider the convergence of a continuous-time Markov chain approximation X^h, $h > 0$, to an $\mathbb{R}^d$-valued Lévy process X. The state space of X^h is an equidistant lattice and its Q-matrix is chosen to approximate the generator of X. In dimension one ($d = 1$), and then under a general sufficient condition for the existence of transition densities of X, we establish sharp convergence rates of the normalised probability mass function of X^h to the probability density function of X. In higher dimensions ($d > 1$), rates of convergence are obtained under a technical condition, which is satisfied when the diffusion matrix is non-degenerate.

2.1 Introduction

Discretization schemes for stochastic processes are relevant both theoretically, as they shed light on the nature of the underlying stochasticity, and practically, since they lend themselves well to numerical methods. Lévy processes, in particular, constitute a rich and fundamental class with applications in diverse areas such as mathematical finance, risk management, insurance, queuing, storage and population genetics etc. (see e.g. [Kyprianou, 2006]).

2.1.1 Short statement of problem and results

We study the rate of convergence of a weak approximation of an $\mathbb{R}^d$-valued ($d \in \mathbb{N}$) Lévy process X by a continuous-time Markov chain (CTMC). Our main aim is to understand the rates of convergence of transition densities. These cannot be viewed as expectations of (sufficiently well-behaved, e.g. bounded continuous) real-valued functions against the marginals of the processes, and hence are in general hard to study.

Since the results are easier to describe in dimension one ($d = 1$), we focus first on this setting. Specifically, our main result in this case, Theorem 2.4, establishes the precise convergence rate of the normalised probability mass function of the approximating Markov chain to the transition density of the Lévy process for the two proposed discretisation schemes, one in the case where X has a non-trivial diffusion component and one when it does not. More precisely, in both cases we approximate X by a CTMC X^h with state space $\mathbb{Z}_h := h\mathbb{Z}$ and Q-matrix defined as a natural discretised version of the generator of X. This makes the CTMC X^h into a continuous-time random walk, which is skip-free (i.e. simple) if X is without jumps (i.e. Brownian motion with drift). The quantity:

$$\kappa(\delta) := \int_{[-1,1]\backslash[-\delta,\delta]} |x| d\lambda(x), \quad \delta \geq 0,$$

where λ is the Lévy measure of X, is related to the activity of the small jumps of X and plays a crucial role in the rate of convergence. We assume that either the diffusion component of X is present ($\sigma^2 > 0$) or the jump activity of X is sufficient (Orey's condition [Orey, 1968], see also Assumption 2.3 below) to ensure that X admits continuous transition densities $p_{t,T}(x, y)$ (from x at time t to y at time $T > t$), which are our main object of study.

Let $P^h_{t,T}(x, y) := \mathsf{P}(X^h_T = y | X^h_t = x)$ denote the corresponding transition probabilities of X^h and let

$$\Delta_{T-t}(h) := \sup_{x,y \in \mathbb{Z}_h} \left| p_{t,T}(x, y) - \frac{1}{h} P^h_{t,T}(x, y) \right|.$$

The following table summarizes our result (see Notation 1.5 for the usage of big O, and, later on, little o and the symbol $\sim$):

	$\sigma^2 > 0$	$\sigma^2 = 0$
$\lambda(\mathbb{R}) = 0$	$\Delta_{T-t}(h) = O(h^2)$	$\times$
$0 < \lambda(\mathbb{R}) < \infty$	$\Delta_{T-t}(h) = O(h)$	$\times$
$\lambda(\mathbb{R}) = \infty$	$\Delta_{T-t}(h) = O(h\kappa(h/2))$	

We also prove that the rates stated here are sharp in the sense that there exist Lévy processes for which convergence is no better than stated.

Note that the rate of convergence depends on the Lévy measure λ, it being best when $\lambda = 0$ (quadratic when $\sigma^2 > 0$), and linear otherwise, unless the pure jump part of X has infinite variation, in which case it depends on the quantity κ. This is due to the nature of the discretisation of the Brownian motion with drift (which gives a quadratic order of convergence, when $\sigma^2 > 0$), and then of the Lévy measure, which is aggregated over intervals of length h around each of the lattice points; see also (v) of Remark 2.21. In the infinite activity case, $\kappa(h) = o(1/h)$, indeed κ is bounded, if in addition $\kappa(0) < \infty$. However, the convergence of $h\kappa(h/2)$ to zero, as $h \downarrow 0$, can be arbitrarily slow. Finally, if X is a compound Poisson process (i.e. $\lambda(\mathbb{R}) \in (0, \infty)$) without a diffusion component, but possibly with a drift, there is always an atom present in the law of X at a fixed time, which is why the finite Lévy measure case is studied only when $\sigma^2 > 0$.

By way of example, note that if $\lambda([-1,1]\backslash[-h,h]) \sim 1/h^{1+\alpha}$ for some $\alpha \in (0,1)$, then $\kappa(h) \sim h^{-\alpha}$ and the convergence of the normalized probability mass function to the transition density is by Theorem 2.4 of order $h^{1-\alpha}$, since $\kappa(0) = \infty$ and Orey's condition is satisfied. In particular, in the case of the CGMY [Carr et al., 2002] (tempered stable) or β-stable [Sato, 1999, p. 80] processes with stability parameter $\beta \in (1,2)$, we have $\alpha = \beta - 1$ and hence convergence of order $h^{2-\beta}$. More generally, if $\beta := \inf\{p > 0 : \int_{[-1,1]} |x|^p d\lambda(x) < \infty\}$ is the *Blumenthal-Getoor index* of X, and $\beta \geq 1$, then for any $p > \beta$ we have $\kappa(h) = O(h^{1-p})$. Conversely, if for some $p \geq 1$, $\kappa(h) = O(h^{1-p})$, then $\beta \leq p$.

The proof of Theorem 2.4 is in two steps: we first establish the convergence rate of the characteristic exponent of X_t^h to that of X_t (Subsection 2.3.2). In the second step we apply this to the study of the convergence of transition densities (Section 2.4) via their spectral representations (established in Subsection 2.3.1). Note that in general the rates of convergence of the characteristic functions do not carry over directly to the distribution functions. We are able to follow through the above programme by exploiting the special structure of the infinitely divisible distributions in what amounts to a detailed comparison of the transition kernels $p_{t,T}(x,y)$ and $P^h_{t,T}(x,y)$.

This gives the overall picture in dimension one. In dimensions higher than one ($d > 1$), and then under a straightforward extension of the discretization described above, essentially the same rates of convergence are obtained as in the univariate case; this time under a technical condition (cf. Assumption 2.6), which is satisfied when the diffusion-matrix is non-degenerate. Our main result in this case is Theorem 2.8.

2.1.2 Literature overview

In general, there has been a plethora of publications devoted to the subject of discretization schemes for stochastic processes, see e.g. [Kloeden and Platen, 1992], and with regard to the pricing of financial derivatives [Glasserman, 2003] and the references therein. In particular, there exists a wealth of literature concerning approximations of Lévy processes in one form or another and a brief overview of simulation techniques is given by [Rosiński, 2008].

In continuous time, for example, [Kiessling and Tempone, 2011] approximates by replacing the small jumps part with a diffusion, and discusses also rates of convergence for $\mathsf{E}[g \circ X_T]$, where g is real-valued and satisfies certain integrability conditions, T is a fixed time and X the process under approximation; [Crosby et al., 2010] approximates by a combination of Brownian motion and sums of compound Poisson processes with two-sided exponential densities. In discrete time, Markov chains have been used to approximate the much larger class of Feller processes and [Böttcher and Schilling, 2009] proves convergence in law of such an approximation in the Skorokhod space of càdlàg paths, but does not discuss rates of convergence; [Szimayer and Maller, 2007] has a finite state space path approximation and applies this to option pricing together with a discussion of the rates of convergence for the prices. With respect to Lévy-process-driven SDEs, [Kohatsu-Higa et al., to appear] (respectively [Tanaka and Kohatsu-Higa, 2009]) approximates solutions Y thereto using a combination of a compound Poisson process and a high order scheme for the Brownian component (respectively discrete-time Markov chains and an operator approach) — rates of convergence are then discussed for expectations of sufficiently regular real-valued functions against the marginals of the solutions.

We remark that approximation/simulation of Lévy processes in dimensions higher than one is in general more difficult than in the univariate case, see, e.g. the discussion on this in [Cohen and Rosiński, 2007] (which has a Gaussian approximation and establishes convergence in the Skorokhod space [Cohen and Rosiński, 2007, p. 197, Theorem 2.2]). Observe also that in terms of pricing theory, the probability density function of a process can be viewed as the Arrow-Debreu state price, i.e. the current value of an option whose payoff equals the Dirac delta function. The singular nature of this payoff makes it hard, particularly in the presence of jumps, to study the convergence of the prices under the discretised process to their continuous counterparts.

Indeed, Theorem 2.8 can be viewed as a generalisation of such convergence results for the well-known discretisation of the multi-dimensional Black-Scholes model (see e.g. [Mijatović, 2007] for the case of Brownian motion with drift in dimension one). In addition, existing literature, as specific to approximations of densities of Lévy processes (or generalizations thereof), includes [Figueroa-López, 2010] (polynomial expansion for a bounded variation driftless pure-jump process) and [Filipovic et al., 2013] (density expansions for multivariate affine jump-diffusion processes). [Knopova and Schilling, 2012; Sztonyk, 2011] study upper estimates for the densities. On the other hand [Bally and Talay, 2009] has a result similar in spirit to ours, but for solutions to SDEs: for the case of the Euler approximation scheme, the authors there also study the rate of convergence of the transition densities.

Further, from the point of view of partial integro-differential equations (PIDEs), the density $p : (0, \infty) \times \mathbb{R}^d \to [0, \infty)$ of the Lévy process X is the classical fundamental solution of the Cauchy problem (in $u \in C_0^{1,2}((0, \infty), \mathbb{R}^d)$) $\frac{\partial u}{\partial t} = Lu$, L being the infinitesimal generator of X [Cont and Tankov, 2004, Chapter 12] [Garroni and Menaldi, 1992, Chapter IV]. Note that Assumption 2.3 in dimension one (respectively Assumption 2.6 in the multivariate case) guarantees $p \in C_0^{1,\infty}$. There are numerous numerical methods in dealing with such PIDEs (and PIDEs in general): fast Fourier transform, trees and discrete-time Markov chains, viscosity solutions, Galerkin methods, see, e.g. [Cont and Voltchkova, 2005, Subsection 1.1] [Cont and Tankov, 2004, Subsections 12.3-12.7] and the references therein. In particular, we mention the finite-difference method, which is in some sense the counterpart of the present article in the numerical analysis literature, discretising both in space and time, whereas we do so only in space. In general, this literature often restricts to finite activity processes, and either avoids a rigorous analysis of (the rates of) convergence, or, when it does, it does so for initial conditions $h = u(0, \cdot)$, which exclude the singular δ-distribution. For example, [Cont and Voltchkova, 2005, p. 1616, Assumption 6.1] requires h continuous, piecewise C^∞ with bounded derivatives of all orders; compare also Propositions 2.29 and 2.32 concerning convergence of expectations in our setting. Moreover, unlike in our case where the discretisation is made outright, the approximation in [Cont and Voltchkova, 2005] is sequential, as is typical of the literature: beyond the restriction to a bounded domain (with boundary conditions), there is a truncation of the integral term in L, and then a reduction to the finite activity case, at which point our results are in agreement with what one would expect from the linear order of convergence of [Cont and Voltchkova, 2005, p. 1616, Theorem 6.7].

The rest of the chapter is organised as follows. Section 2.2 introduces the setting by specifying the Markov generator of X^h and precisely states the main results. Then Section 2.3 provides integral expressions for the transition kernels by applying spectral theory to the generator of the approximating chain and studies the convergence of the characteristic exponents. In Section 2.4 this allows us to establish convergence rates for the transition densities. While Sections 2.3 and 2.4 restrict this analysis to the univariate case, explicit comments are made in both on how to extend the results to the multivariate setting (this extension being, for the most part, direct and trivial). Finally, Section 2.5 derives some results regarding convergence of expectations $\mathsf{E}[f \circ X_t^h] \to \mathsf{E}[f \circ X_t]$ for suitable test functions f; presents a numerical algorithm, under which computations are eventually done; discusses the corresponding truncation/localization error and gives some numerical experiments.

2.2 Definitions, notation and statement of results

2.2.1 Setting

Fix a dimension $d \in \mathbb{N}$ and let $(e_j)_{j=1}^d$ be the standard orthonormal basis of $\mathbb{R}^d$. Further, let X be an $\mathbb{R}^d$-valued Lévy process with characteristic exponent (cf. (1.1)-(1.2)):

$$\Psi(p) = -\frac{1}{2}\langle p, \Sigma p\rangle + i\langle \mu, p\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle p,x\rangle} - i\langle p, x\rangle \mathbb{1}_{[-V,V]^d}(x) - 1\right) d\lambda(x) \tag{2.1}$$

($p \in \mathbb{R}^d$). Here $(\Sigma, \lambda, \mu)_{\tilde{c}}$ is the characteristic triplet relative to the cut-off function $\tilde{c} = \mathbb{1}_{[-V,V]^d}$; V is 1 or 0, the latter only if $\int_{[-1,1]^d} |x| d\lambda(x) < \infty$. Recall that X is then a Markov process with transition function $P_{t,T}(x, B) := \mathsf{P}(X_{T-t} \in B - x)$ ($0 \leq t \leq T$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$) and (for $t \geq 0$, $p \in \mathbb{R}^d$) $\phi_{X_t}(p) := \mathsf{E}[e^{i\langle p, X_t\rangle}] = \exp\{t\Psi(p)\}$.

Since $\Sigma \in \mathbb{R}^{d\times d}$ is symmetric, nonnegative definite, it is assumed without loss of generality that $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_d^2)$ with $\sigma_1^2 \geq \cdots \geq \sigma_d^2$. We let $l := \max\{k \in \{1, \ldots, d\} : \sigma_k^2 > 0\}$ ($\max \emptyset := 0$). In the univariate case, $d = 1$, Σ reduces to the scalar $\sigma^2 := \sigma_1^2$.

Now fix $h > 0$. Consider a CTMC $X^h = (X_t^h)_{t\geq 0}$, approximating our Lévy process X (in law). We describe (see Subsection 1.2.3 for terminology and relevant results on CTMCs) X^h as having state space $\mathbb{Z}_h^d$ (recall Notation 1.1(1)), initial state $X_0^h = 0$, a.s., and an infinitesimal generator L^h given by a spatially homogeneous

Q-matrix Q^h (i.e. $Q^h_{ss'}$ depends only on $s-s'$, for $\{s,s'\} \subset \mathbb{Z}^d_h$). Thus L^h is a mapping defined on the set $l^\infty(\mathbb{Z}^d_h)$ of bounded functions f on $\mathbb{Z}^d_h$, and $L^h f(s) = \sum_{s' \in \mathbb{Z}^d_h} Q^h_{ss'} f(s')$.

It remains to specify Q^h. To this end we discretise on $\mathbb{Z}^d_h$ the infinitesimal generator L of the Lévy process X, thus obtaining L^h. Recall that [Sato, 1999, p. 208, Theorem 31.5]:

$$Lf(x) = \sum_{j=1}^{d} \left(\frac{\sigma_j^2}{2} \partial_{jj} f(x) + \mu_j \partial_j f(x) \right) + \int_{\mathbb{R}^d} \left(f(x+y) - f(x) - \sum_{j=1}^{d} y_j \partial_j f(x) \mathbb{1}_{[-V,V]^d}(y) \right) d\lambda(y)$$

($f \in C_0^2(\mathbb{R}^d)$, $x \in \mathbb{R}^d$). We specify L^h separately in the univariate, $d=1$, and in the general, multivariate, setting.

Univariate case

In the case when $d=1$, we introduce two schemes. Referred to as **discretization scheme 1** (respectively **2**), and given by (2.2) (respectively (2.4)) below, they differ in the discretization of the first derivative, as follows.

Under **discretisation scheme 1**, for $s \in \mathbb{Z}_h$ and $f : \mathbb{Z}_h \to \mathbb{R}$ vanishing at infinity (i.e. $f \in c_0(\mathbb{Z}_h)$):

$$\begin{aligned} L^h f(s) &= \frac{1}{2}\left(\sigma^2 + c_0^h\right) \frac{f(s+h) + f(s-h) - 2f(s)}{h^2} + \left(\mu - \mu^h\right) \frac{f(s+h) - f(s-h)}{2h} + \\ & \sum_{s' \in \mathbb{Z}_h \backslash \{0\}} \left[f(s+s') - f(s)\right] c^h_{s'} \end{aligned} \tag{2.2}$$

where the following notation has been introduced:

- for $s \in \mathbb{Z}_h$:
$$A^h_s := \begin{cases} [s-h/2, s+h/2), & \text{if } s<0 \\ [-h/2, h/2], & \text{if } s=0 \\ (s-h/2, s+h/2], & \text{if } s>0 \end{cases};$$
- for $s \in \mathbb{Z}_h \backslash \{0\}$: $c^h_s := \lambda(A^h_s)$;
- and finally:
$$c_0^h := \int_{A_0^h} y^2 \mathbb{1}_{[-V,V]}(y) d\lambda(y) \quad \text{and} \quad \mu^h := \sum_{s \in \mathbb{Z}_h \backslash \{0\}} s \int_{A^h_s} \mathbb{1}_{[-V,V]}(y) d\lambda(y).$$

Note that Q^h has nonnegative off-diagonal entries for all h for which:

$$\frac{\sigma^2 + c_0^h}{2h^2} + \frac{\mu - \mu^h}{2h} + c^h_h \geq 0 \qquad \text{and} \qquad \frac{\sigma^2 + c_0^h}{2h^2} - \frac{\mu - \mu^h}{2h} + c^h_{-h} \geq 0 \tag{2.3}$$

and in that case Q^h is a genuine Q-matrix. Moreover, due to spatial homogeneity, its entries are then also uniformly bounded in absolute value.

Further, when $\sigma^2 > 0$, it will be shown that (2.3) always holds, at least for all sufficiently small h (see Proposition 2.18). However, in general, (2.3) may fail. It is for this reason that we introduce scheme 2, under which the condition on the nonnegativity of off-diagonal entries of Q^h holds vacuously.

To wit, we use in **discretization scheme 2** the one-sided, rather than the two-sided discretisation of the first derivative, so that (2.2) reads:

$$\begin{aligned} L^h f(s) &= \frac{1}{2}\left(\sigma^2 + c_0^h\right)\frac{f(s+h)+f(s-h)-2f(s)}{h^2} + \sum_{s'\in\mathbb{Z}_h\backslash\{0\}}[f(s+s')-f(s)]c_{s'}^h + \\ & (\mu-\mu^h)\left(\frac{f(s+h)-f(s)}{h}\mathbb{1}_{[0,\infty)}(\mu-\mu^h) + \frac{f(s)-f(s-h)}{h}\mathbb{1}_{(-\infty,0]}(\mu-\mu^h)\right) \end{aligned} \tag{2.4}$$

Importantly, while scheme 2 is always well-defined, scheme 1 is not; and yet the two-sided discretization of the first derivative exhibits better convergence properties than the one-sided one (cf. Proposition 2.20). We therefore retain the treatment of both these schemes in the sequel.

For ease of reference we also summarize here the following notation which will be used from Subsection 2.3.2 onwards:

$$c := \lambda(\mathbb{R}), \quad b := \kappa(0), \quad d := \lambda(\mathbb{R}\backslash[-1,1])$$

and for $\delta \in (0,1]$:

$$\zeta(\delta) := \delta\int_{[-1,1]\backslash[-\delta,\delta]}|x|d\lambda(x) \text{ and } \gamma(\delta) := \delta^2\int_{[-1,1]\backslash[-\delta,\delta]}d\lambda(x).$$

Multivariate case

For the sake of simplicity we introduce only one discretisation scheme in this general setting. If necessary, and to avoid confusion, we shall refer to it as the **multivariate scheme**. We choose $V = 0$ or $V = 1$, according as to whether $\lambda(\mathbb{R}^d)$ is finite or infinite. L^h is then given by:

$$\begin{aligned} L^h f(s) = & \frac{1}{2}\sum_{j=1}^{d}\left(\sigma_j^2 + c_{0j}^h\right)\frac{f(s+he_j)+f(s-he_j)-2f(s)}{h^2} + \sum_{j=1}^{l}(\mu_j-\mu_j^h)\frac{f(s+he_j)-f(s-he_j)}{2h} \\ & + \sum_{j=l+1}^{d}(\mu_j-\mu_j^h)\left(\frac{f(s+he_j)-f(s)}{h}\mathbb{1}_{[0,\infty)}(\mu_j-\mu_j^h) + \frac{f(s)-f(s-he_j)}{h}\mathbb{1}_{(-\infty,0]}(\mu_j-\mu_j^h)\right) + \\ & \sum_{s'\in\mathbb{Z}_h^d}\left(f(s+s')-f(s)\right)c_{s'}^h \end{aligned}$$

($f \in c_0(\mathbb{Z}_h^d)$, $s \in \mathbb{Z}_h^d$; and we agree $\sum_\emptyset := 0$). Here the following notation has been introduced:

- for $s \in \mathbb{Z}_h^d$: $A_s^h := \prod_{j=1}^d I_{s_j}^h$, where for $s \in \mathbb{Z}_h$:

$$I_s^h := \begin{cases} [s-h/2, s+h/2), & \text{if } s < 0 \\ [-h/2, h/2], & \text{if } s = 0 \\ (s-h/2, s+h/2], & \text{if } s > 0 \end{cases}$$

so that $\{A_s^h : s \in \mathbb{Z}_h^d\}$ constitutes a partition of $\mathbb{R}^d$;

- for $s \in \mathbb{Z}_h^d \backslash \{0\}$: $c_s^h := \lambda(A_s^h)$;
- and finally for $j \in \{1, \ldots, d\}$:

$$c_{0j}^h := \int_{A_0^h} x_j^2 \mathbb{1}_{[-V,V]^d}(x) d\lambda(x) \qquad \text{and} \qquad \mu_j^h := \sum_{s \in \mathbb{Z}_h^d \backslash \{0\}} s_j \int_{A_s^h} \mathbb{1}_{[-V,V]^d}(y) d\lambda(y).$$

Notice that when $d = 1$, this scheme reduces to scheme 1 or scheme 2, according as to whether $\sigma^2 > 0$ or $\sigma^2 = 0$. Indeed, statements pertaining to the multivariate scheme will always be understood to include also the univariate case $d = 1$.

Remark 2.1. The complete analogue of c_0^h from the univariate case would be the matrix c_0^h, entries $(c_0^h)_{ij} := \int_{A_0^h} x_i x_j \mathbb{1}_{[-V,V]^d}(x) d\lambda(x)$, $\{i, j\} \subset \{1, \ldots, d\}$. However, as h varies, so could c_0^h, and thus no diagonalization of $c_0^h + \Sigma$ is possible (in general), simultaneously in all (small enough) positive h. Thus, retaining c_0^h in its totality, we should have to discretize mixed second partial derivatives, which would introduce (further) nonpositive entries in the corresponding Q-matrix Q^h of X^h. It is not clear whether these would necessarily be counterbalanced in a way that would ensure nonnegative off-diagonal entries. Retaining the diagonal terms of c_0^h, however, is of no material consequence in this respect.

It is verified just as in the univariate case, component by component, that there is some $h_\star \in (0, +\infty]$ such that for all $h \in (0, h_\star)$, L^h is indeed the infinitesimal generator of some CTMC (i.e. the off-diagonal entries of Q^h are nonnegative). Q^h is then a regular (as spatially homogeneous) Q-matrix, and X^h is a compound Poisson process (in law, see Remark 1.40), whose Lévy measure we denote λ^h.

2.2.2 Summary of results

We have (see Remark 2.21(iii) pursuant to Proposition 2.20 for proof):

Remark 2.2 (Convergence in distribution). X^h converges to X weakly in finite-dimensional distributions (hence with respect to the Skorokhod topology on the space of càdlàg paths[1] [Jacod and Shiryaev, 2003, p. 415, Corollary 3.9]) as $h \downarrow 0$.

[1] Upon the choice of such versions.

Next, in order to formulate the rates of convergence, recall that $P^h_{t,T}(x,y)$ (respectively $p_{t,T}(x,y)$) denote the transition probabilities (respectively continuous transition densities, when they exist) of X^h (respectively X) from x at time t to y at time T, $\{x,y\} \subset \mathbb{Z}^d_h$, $0 \leq t < T$. Further, for $0 \leq t < T$ define:

$$\Delta_{T-t}(h) := \sup_{\{x,y\}\subset\mathbb{Z}^d_h} D^h_{t,T}(x,y), \text{ where } D^h_{t,T}(x,y) := \left|p_{t,T}(x,y) - \frac{1}{h^d}P^h_{t,T}(x,y)\right|. \tag{2.5}$$

We now summarize the results first in the univariate, and then in the multivariate setting (Remark 2.2 holding true of both).

Univariate case

The assumption alluded to in the Introduction (Section 2.1) is the following (we state it explicitly when it is being used):

Assumption 2.3. *Either $\sigma^2 > 0$ or Orey's condition [Orey, 1968] holds:*

$$\exists\epsilon \in (0,2) \qquad \textit{such that} \qquad \liminf_{r\downarrow 0} \frac{1}{r^{2-\epsilon}} \int_{[-r,r]} u^2 d\lambda(u) > 0.$$

The usage of the two schemes and the specification of V is as summarized in Table 2.1. In short we use scheme 1 or scheme 2, according as to whether $\sigma^2 > 0$ or $\sigma^2 = 0$, and we use $V = 0$ or $V = 1$, according as to whether $\lambda(\mathbb{R}) < \infty$ or $\lambda(\mathbb{R}) = \infty$. By contrast to Assumption 2.3 we presuppose the provisions of Table 2.1 throughout this paragraph.

Lévy measure/diffusion part	$\sigma^2 > 0$	$\sigma^2 = 0$
$\lambda(\mathbb{R}) < \infty$	scheme 1, $V = 0$	scheme 2, $V = 0$
$\lambda(\mathbb{R}) = \infty$	scheme 1, $V = 1$	scheme 2, $V = 1$

Table 2.1: Usage of the two schemes and of V depending on the nature of σ^2 and λ.

Under Assumption 2.3 for every $t > 0$, $\phi_{X_t} \in L^1(m)$ where m is Lebesgue measure and (for $0 \leq t < T$, $y \in \mathbb{R}$, P_{X_t}-a.s. in $x \in \mathbb{R}$):

$$p_{t,T}(x,y) = \frac{1}{2\pi}\int_{\mathbb{R}} \exp\{ip(x-y)\}\exp\{\Psi(p)(T-t)\}\,dp \tag{2.6}$$

(see Remark 2.10). Choose $p_{t,T}(x,y)$ for which (2.6) obtains for each $x \in \mathbb{R}$.

Similarly, with Ψ^h denoting the characteristic exponent of the compound Poisson process (in law) X^h (for $0 \leq t < T$, $y \in \mathbb{Z}_h$, $\mathsf{P}_{X^h_t}$-a.s. in $x \in \mathbb{Z}_h$):

$$\frac{1}{h}P^h_{t,T}(x,y) = \frac{1}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \exp\{ip(x-y)\}\exp\{\Psi^h(p)(T-t)\}dp \tag{2.7}$$

(see Proposition 2.14). Note that the right-hand side is defined even if $\mathsf{P}(X_t^h = x) = 0$ and we let the left-hand side take this value when this is so.

The main result can now be stated.

Theorem 2.4 (Convergence of transition kernels)**.** *Under Assumption 2.3, whenever $s > 0$, the convergence of $\Delta_s(h)$ is summarized in the following table. In general convergence is no better than stipulated.*

<table>
<tr><td></td><td>$\lambda(\mathbb{R}) = 0$</td><td>$0 < \lambda(\mathbb{R}) < \infty$</td><td>$\kappa(0) < \infty = \lambda(\mathbb{R})$</td><td>$\kappa(0) = \infty$</td></tr>
<tr><td>$\sigma^2 > 0$</td><td>$\Delta_s(h) = O(h^2)$</td><td>$\Delta_s(h) = O(h)$</td><td rowspan="2">$\Delta_s(h) = O(h)$</td><td rowspan="2">$\Delta_s(h) = O(h\kappa(h/2))$</td></tr>
<tr><td>$\sigma^2 = 0$</td><td>$\times$</td><td>$\times$</td></tr>
</table>

More exhaustive statements, of which this theorem is a summary, are to be found in Propositions 2.26 and 2.27, and will be proved in Section 2.4. The proof of Theorem 2.4 itself can be found at the end of Section 2.4.

Remark 2.5. Assumption 2.3 implies that X_t, for any $t > 0$, has a smooth density [Sato, 1999, p. 190, Proposition 28.3]. It hence appears to be unlikely that this assumption constitutes a necessary condition for the convergence rates of Theorem 2.4 to hold. In particular, Assumption 2.6 with $d = 1$, stipulating a certain exponential decay of the characteristic exponents, is implied by Assumption 2.3 (see Remark 2.10 and Proposition 2.22) but sufficient for the validity of the convergence rates in Theorem 2.4 (see Theorem 2.8).

Multivariate case

The relevant technical condition here is:

Assumption 2.6. *There are $\{P, C, \epsilon\} \subset (0, \infty)$ and an $h_0 \in (0, h_\star]$, such that for all $h \in (0, h_0)$, $s > 0$ and $p \in [-\pi/h, \pi/h]^d \backslash (-P, P)^d$:*

$$|\phi_{X_s^h}(p)| \leq \exp\{-Cs|p|^\epsilon\} \tag{2.8}$$

whereas for $p \in \mathbb{R}^d \backslash (-P, P)^d$:

$$|\phi_{X_s}(p)| \leq \exp\{-Cs|p|^\epsilon\}. \tag{2.9}$$

Again we shall state it explicitly when it is being used.

Remark 2.7. It is shown, just as in the univariate case, that Assumption 2.6 holds if $l = d$, i.e. if Σ is non-degenerate. Moreover, then we may take $P = 0$, $C = \frac{1}{2}\left(\frac{2}{\pi}\right)^2 \left(\wedge_{j=1}^d \sigma_j^2\right)$, $\epsilon = 2$ and $h_0 = h_\star$.

It would be natural to expect that the same could be verified for the multivariate analogue of Orey's condition, which we suggest as being:

$$\liminf_{r\downarrow 0}\inf_{e\in S(0,1)}\int_{\overline{B}(0,r)}|\langle e,x\rangle|^2 d\lambda(x)/r^{2-\epsilon}>0$$

for some $\epsilon\in(0,2)$ (see Notation 1.2 for closed balls and spheres). Specifically, under this condition, it is easy to see that (2.9) of Assumption 2.6 still holds. However, we are unable to show the validity of (2.8).

Under Assumption 2.6, Fourier inversion yields the integral representation of the continuous transition densities for X (for $0\leq t<T$, $\{x,y\}\subset\mathbb{R}^d$):

$$p_{t,T}(x,y)=\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}e^{i\langle p,x-y\rangle}e^{(T-t)\Psi(p)}dp.$$

On the other hand, $L^2([-\pi/h,\pi/h]^d)$ Hilbert space techniques yield for the normalized transition probabilities of X^h (for $0\leq t<T$, $y\in\mathbb{Z}^d_h$ and $\mathsf{P}_{X^h_t}$-a.s. in $x\in\mathbb{Z}^d_h$):

$$\frac{1}{h^d}P_{t,T}(x,y)=\frac{1}{(2\pi)^d}\int_{[-\pi/h,\pi/h]^d}e^{i\langle p,x-y\rangle}e^{(T-t)\Psi^h(p)}dp,$$

where Ψ^h is the characteristic exponent of X^h.

Finally, we state the result with the help of the following notation:

- for $\delta\in[0,\infty)$: $\kappa(\delta):=\int_{[-1,1]^d\backslash[-\delta,\delta]^d}|x|d\lambda(x)$, $\zeta(\delta):=\delta\kappa(\delta)$ and $\chi(\delta):=\sum_{1\leq i<j\leq d}\int_{[-\delta,\delta]^d}|x_ix_j|d\lambda(x)$.
- $\hat{\sigma}^2:=\wedge_{j=1}^d\sigma_j^2$ and $\sigma^2:=\sum_{j=1}^d\sigma_j^2$.

Note that by the Dominated Convergence Theorem, $(\zeta+\chi)(\delta)\to 0$ as $\delta\downarrow 0$ (this is seen as in the univariate case, cf. Lemma 2.17).

Theorem 2.8 (Convergence — multivariate case)**.** *Let $d\in\mathbb{N}$ and suppose Assumption 2.6 holds. Then for any $s>0$, $\Delta_s(h)=O(h\vee(\zeta+\chi)(h/2))$. Moreover, if $\hat{\sigma}^2>0$, then there exists a universal constant $D_d\in(0,\infty)$, such that for any $s>0$:*

1. *If $\lambda(\mathbb{R}^d)=0$,*

$$\limsup_{h\downarrow 0}\frac{\Delta_s(h)}{h^2}\leq D_d\left[\frac{\sigma^2}{\hat{\sigma}^2}\frac{1}{\sqrt{s\hat{\sigma}^2}}+\frac{|\mu|}{\hat{\sigma}^2}\right]\frac{1}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

2. *If $0<\lambda(\mathbb{R}^d)<\infty$,*

$$\limsup_{h\downarrow 0}\frac{\Delta_s(h)}{h}\leq D_d\frac{\lambda(\mathbb{R}^d)s}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

3. If $\kappa(0) < \infty = \lambda(\mathbb{R}^d)$,

$$\limsup_{h\downarrow 0} \frac{\Delta_s(h)}{h} \leq D_d \left[\lambda(\mathbb{R}^d\backslash[-1,1]^d)s + \frac{\kappa(0)s}{\sqrt{s\hat{\sigma}^2}}\right] \frac{1}{(s\hat{\sigma}^2)^{\frac{d+1}{2}}}.$$

4. If $\kappa(0) = \infty$,

$$\limsup_{h\downarrow 0} \frac{\Delta_s(h)}{(\zeta+\chi)(h/2)} \leq D_d \frac{s}{(s\hat{\sigma}^2)^{\frac{d+2}{2}}}.$$

Remark 2.9. Notice that in the univariate case $\zeta + \chi$ reduces to ζ. The presence of χ is a consequence of the omission of non-diagonal entries of c_0^h in the multivariate approximation scheme (cf. Remark 2.1).

The proof of Theorem 2.8 is an easy extension of the arguments behind Theorem 2.4 and can be found immediately following the proof of Proposition 2.23.

2.3 Transition kernels and convergence of characteristic exponents

In the interests of space, simplicity of notation and ease of exposition, the analysis in this and in Section 2.4 is restricted to dimension $d = 1$. Proofs in the multivariate setting are, for the most part, a direct and trivial extension of those in the univariate case. However, when this is not so, necessary and explicit comments will be provided in the sequel, as appropriate.

2.3.1 Integral representations

First we note the following result (its proof is essentially by the standard inversion theorem; see also [Sato, 1999, p. 190, Proposition 28.3]).

Remark 2.10. Under Assumption 2.3, for some $\{P, C, \epsilon\} \subset (0,\infty)$ depending only on $\{\lambda, \sigma^2\}$ and then all $p \in \mathbb{R}\backslash(-P, P)$ and $t \geq 0$: $|\phi_{X_t}(p)| \leq \exp\{-Ct|p|^\epsilon\}$. Moreover, when $\sigma^2 > 0$, one may take $P = 0$, $C = \frac{1}{2}\sigma^2$ and $\epsilon = 2$, whereas otherwise ϵ may take the value from Orey's condition in Assumption 2.3. Consequently, X_t ($t > 0$) admits the continuous density $f_{X_t}(y) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-ipy}\phi_{X_t}(p)dp$ ($y \in \mathbb{R}$). In particular, the law $P_{t,T}(x, \cdot)$ is given by (2.6).

Second, to obtain (2.7), we apply some classical theory of Hilbert spaces, see e.g. [Dudley, 2004].

Definition 2.11. For $s \in \mathbb{Z}_h$ let $g_s : [-\frac{\pi}{h}, \frac{\pi}{h}] \to \mathbb{C}$ be given by $g_s(p) := \sqrt{\frac{h}{2\pi}}e^{-isp}$. The $(g_s)_{s\in\mathbb{Z}_h}$ constitute an orthonormal basis of the Hilbert space $L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$.

Let $A \in l^2(\mathbb{Z}_h)$. We define: $\mathcal{F}_h A := \sum_{s\in\mathbb{Z}_h} A(s) g_s$, so that $\mathcal{F}_h : l^2(\mathbb{Z}_h) \to L^2([-\pi/h, \pi/h])$ is a bounded linear mapping. The inverse of this transform $\mathcal{F}_h^{-1} : L^2([-\frac{\pi}{h}, \frac{\pi}{h}]) \to l^2(\mathbb{Z}_h)$ is given by:

$$(\mathcal{F}_h^{-1}\phi)(s) = \langle \phi, g_s \rangle := \int_{[-\frac{\pi}{h},\frac{\pi}{h}]} \phi(u)\overline{g_s}(u)du,$$

for $\phi \in L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$ and $s \in \mathbb{Z}_h$. It is again a bounded linear mapping.

Definition 2.12. For a bounded linear operator $A : l^2(\mathbb{Z}_h) \to l^2(\mathbb{Z}_h)$, we say $F_A : [-\pi/h, \pi/h] \to \mathbb{R}$ is its *diagonalization*, if $\mathcal{F}_h A \mathcal{F}_h^{-1}\phi = F_A\phi$ for all $\phi \in L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$.

We now diagonalize L^h, which allows us to establish (2.7). The straightforward proof is left to the reader.

Proposition 2.13. *Fix $C \in l^1(\mathbb{Z}_h)$. The following introduces a number of bounded linear operators $A : l^2(\mathbb{Z}_h) \to l^2(\mathbb{Z}_h)$ and gives their diagonalization. With $f \in l^2(\mathbb{Z}_h)$, $s \in \mathbb{Z}_h$, $p \in [-\frac{\pi}{h}, \frac{\pi}{h}]$:*

(i) $\Delta_h f(s) := \frac{f(s+h)+f(s-h)-2f(s)}{h^2}$. $F_{\Delta_h}(p) = 2\frac{\cos(hp)-1}{h^2}$.

(ii) $\nabla_h f(s) := \frac{f(s+h)-f(s-h)}{2h}$. $F_{\nabla_h}(p) = i\frac{\sin(hp)}{h}$. *Under scheme 2 we let* $\nabla_h^+ f(s) := \frac{f(s+h)-f(s)}{h}$ *(respectively* $\nabla_h^- f(s) := \frac{f(s)-f(s-h)}{h}$*) and then* $F_{\nabla_h^+}(p) = \frac{e^{ihp}-1}{h}$ *(respectively* $F_{\nabla_h^-}(p) = \frac{1-e^{-ihp}}{h}$*).*

(iii) $L_C f(s) := \sum_{s'\in\mathbb{Z}_h}(f(s+s') - f(s))C(s')$. $F_{L_C}(p) = \sum_{s\in\mathbb{Z}_h} C(s)(e^{isp} - 1)$.

As λ is finite outside any neighborhood of 0, $L^h|_{l^2(\mathbb{Z}_h)}$ (as in (2.2), respectively (2.4)) is a bounded linear mapping. We denote this restriction by L^h also. Its diagonalization is then given by $\Psi^h := F_{L^h}$, where, under scheme 1,

$$\Psi^h(p) = i(\mu - \mu^h)\frac{\sin(hp)}{h} + (\sigma^2 + c_0^h)\frac{(\cos(hp) - 1)}{h^2} + \sum_{s\in\mathbb{Z}_h\setminus\{0\}} c_s^h \left(e^{isp} - 1\right) \quad (2.10)$$

and under scheme 2,

$$\begin{aligned}\Psi^h(p) \quad &= \quad (\mu - \mu^h)\left(\frac{e^{ihp} - 1}{h}\mathbb{1}_{[0,\infty)}(\mu - \mu^h) + \frac{1 - e^{-ihp}}{h}\mathbb{1}_{(-\infty,0]}(\mu - \mu^h)\right) \\ &+ \quad (\sigma^2 + c_0^h)\frac{(\cos(hp) - 1)}{h^2} + \sum_{s\in\mathbb{Z}_h\setminus\{0\}} c_s^h \left(e^{isp} - 1\right) \qquad (2.11)\end{aligned}$$

(with $p \in [-\frac{\pi}{h}, \frac{\pi}{h}]$, but we can and will view Ψ^h as defined for all real p by the formulae above). Under either scheme, Ψ^h is bounded and continuous as the final sum converges absolutely and uniformly.

Proposition 2.14. *For scheme 1 under* (2.3)*, and always for scheme 2, for every* $0 \leq t < T$*,* $y \in \mathbb{Z}_h$ *and* $\mathsf{P}_{X_t^h}$*-a.s. in* $x \in \mathbb{Z}_h$ *(2.7) holds, i.e.:*

$$\mathsf{P}(X_T^h = y | X_t^h = x) = \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \exp\{ip(x-y)\} \exp\{\Psi^h(p)(T-t)\} dp.$$

Proof. (Condition (2.3) ensures scheme 1 is well-defined (Q^h needs to have nonnegative off-diagonal entries).) Note that: $\mathsf{P}(X_T^h = y | X_t^h = x) = (e^{(T-t)L^h} \mathbb{1}_{\{y\}})(x)$. Thus (2.7) follows directly from the relation $\mathcal{F}_h L^h \mathcal{F}_h^{-1} = \Psi^h \cdot$ (where $\Psi^h \cdot$ is the operator that multiplies functions pointwise by Ψ^h), since:

$$\begin{aligned}
(e^{(T-t)L^h} \mathbb{1}_{\{y\}})(x) &= (\mathcal{F}_h^{-1} e^{\mathcal{F}_h (T-T) L^h \mathcal{F}_h^{-1}} \mathcal{F}_h \mathbb{1}_{\{y\}})(x) \\
&= \sqrt{\frac{h}{2\pi}} \int_{[-\frac{\pi}{h}, \frac{\pi}{h}]} e^{ipx} (e^{(T-t)\mathcal{F}_h L^h \mathcal{F}_h^{-1}} g_y)(p) dp \\
&= \sqrt{\frac{h}{2\pi}} \int_{[-\frac{\pi}{h}, \frac{\pi}{h}]} e^{ipx} \left(\sum_{k=0}^{\infty} \frac{(T-t)^k F_{L^h}^k}{k!} g_y \right)(p) dp \\
&= \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \exp\{ip(x-y)\} \exp\{\Psi^h(p)(T-t)\} dp,
\end{aligned}$$

where on the right hand side of the third equality, the sum is in principle in $L^2([-\frac{\pi}{h}, \frac{\pi}{h}])$, but since the sum converges pointwise boundedly, it can be taken as such (by the very definition of convergence in L^2, and by bounded convergence). □

In what follows we study the convergence of (2.7) to (2.6) as $h \downarrow 0$. These expressions are particularly suited to such an analysis, not least because the spatial and temporal components are factorized.

Note also that, by Proposition 2.14, for every $t \geq 0$ and Lebesgue-a.e. $p \in [-\pi/h, \pi/h]$, $\phi_{X_t^h}(p) = \mathsf{E}[e^{ipX_t^h}] = \mathcal{F}_h(\mathcal{F}_h^{-1}(e^{t\Psi^h}))(p) = \exp\{t\Psi^h(p)\}$. Then, via continuity and periodicity:

$$\phi_{X_t^h}(p) = \exp\{t\Psi^h(p)\},$$

everywhere in $p \in \mathbb{R}$. Further, note that X^h is a compound Poisson processes (possibly only in law, see Remark 1.40; but we may always choose a càdlàg version if required/convenient, cf. Remark 1.36) and Ψ^h is its characteristic exponent [Sato, 1999, p. 33, Lemma 7.6].

In the multivariate scheme, by considering the Hilbert space $L^2([-\pi/h, \pi/h]^d)$ instead, X^h is again seen to be compound Poisson with characteristic exponent given by (for $p \in \mathbb{R}^d$):

$$\begin{aligned}
\Psi^h(p) &= \sum_{j=1}^{d}(\sigma_j^2 + c_{0j}^h)\frac{\cos(hp_j)-1}{h^2} + i\sum_{j=1}^{l}(\mu_j - \mu_j^h)\frac{\sin(hp_j)}{h} + \\
& \sum_{j=l+1}^{d}(\mu_j - \mu_j^h)\left(\frac{e^{ihp_j}-1}{h}\mathbb{1}_{[0,\infty)}(\mu_j - \mu_j^h) + \frac{1-e^{-ihp_j}}{h}\mathbb{1}_{(-\infty,0]}(\mu_j - \mu_j^h)\right) \\
&+ \sum_{s\in\mathbb{Z}_h^d\backslash\{0\}}\left(e^{i\langle p,s\rangle}-1\right)c_s^h. \qquad (2.12)
\end{aligned}$$

In the sequel, we shall let λ^h denote the Lévy measure of X^h.

2.3.2 Convergence of characteristic exponents

We introduce for $p \in \mathbb{R}$:

$$f_h(p) := \frac{\cos(hp)-1}{h^2} + \frac{p^2}{2}$$

and, under scheme 1:

$$\begin{aligned}
g_h(p) &:= i\left(\frac{\sin(hp)}{h} - p\right) \\
l_h(p) &:= c_0^h\frac{\cos(hp)-1}{h^2} - \mu^h i\frac{\sin(hp)}{h} + \\
&\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h\left(e^{isp}-1\right) - \int_{\mathbb{R}}\left(e^{ipu} - 1 - ipu\mathbb{1}_{[-V,V]}(u)\right)d\lambda(u),
\end{aligned}$$

respectively, under scheme 2:

$$\begin{aligned}
g_h(p) &:= \frac{e^{ihp}-1}{h}\mathbb{1}_{(0,\infty)}(\mu-\mu^h) + \frac{1-e^{-ihp}}{h}\mathbb{1}_{(-\infty,0]}(\mu-\mu^h) - ip; \\
l_h(p) &:= c_0^h\frac{\cos(hp)-1}{h^2} - \mu^h\left[\frac{e^{ihp}-1}{h}\mathbb{1}_{(0,\infty)}(\mu-\mu^h) + \frac{1-e^{-ihp}}{h}\mathbb{1}_{(-\infty,0]}(\mu-\mu^h)\right] + \\
&\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h\left(e^{isp}-1\right) - \int_{\mathbb{R}}\left(e^{ipu} - 1 - ipu\mathbb{1}_{[-V,V]}(u)\right)d\lambda(u).
\end{aligned}$$

Thus:

$$\Psi^h - \Psi = \sigma^2 f_h + \mu g_h + l_h.$$

Next, we give three elementary but key lemmas. The first concerns some elementary trigonometric inequalities as well as the Lipschitz difference for the remainder of the exponential series $f_l(x) := \sum_{k=l+1}^{\infty}\frac{(ix)^k}{k!}$ ($x \in \mathbb{R}$, $l \in \{0,1,2\}$): these estimates will be used again and again in what follows. The second is only used in the estimates pertaining to the multivariate scheme. Finally, the third lemma establishes key convergence properties relating to λ.

Lemma 2.15. *For all real x: $0 \le \cos(x) - 1 + \frac{x^2}{2} \le \frac{x^4}{4!}$, $0 \le \operatorname{sgn}(x)(x - \sin(x)) \le \operatorname{sgn}(x)\frac{x^3}{3!}$ and $0 \le x^2 + 2(1-\cos(x)) - 2x\sin(x) \le x^4/4$. Whenever $\{x, y\} \subset \mathbb{R}$ we have (with $\delta := y - x$):*

1. $|e^{ix} - 1 - (e^{iy} - 1)|^2 \le \delta^2$.
2. $|e^{ix} - 1 - ix - (e^{iy} - 1 - iy)|^2 \le \delta^4/4 + \delta^2 x^2 + |\delta|^3|x|$.
3. $|e^{ix} - 1 - ix + x^2/2 - (e^{iy} - 1 - iy + y^2/2)|^2 \le \delta^6/36 + |\delta|^5|x|/6 + (5/12)\delta^4 x^2 + |\delta|^3|x|^3/2 + \delta^2 x^4/4$.

Proof. The first set of inequalities may be proved by comparison of derivatives. Then, (1) follows from $|e^{i(x-y)} - 1|^2 = 2(1 - \cos(x - y))$ and $|e^{iy}| = 1$; (2) from

$$|e^{ix} - ix - e^{iy} + iy|^2 = \big(\delta^2 + 2(1 - \cos(\delta)) - 2\delta\sin(\delta)\big) - 2\delta(\cos(x) - 1)\sin(\delta) + 2\delta\sin(x)(1 - \cos(\delta))$$

and finally (3) from the decomposition of $|e^{ix} - ix + x^2/2 - e^{iy} + iy - y^2/2|^2$ into the following terms:

1. $2(1 - \cos(\delta)) + \delta^2 + \delta^4/4 - 2\delta\sin(\delta) - (1 - \cos(\delta))\delta^2 \le \delta^6/36$ for any real δ.
2. $\delta^3 x - \sin(x)\sin(\delta)\delta^2 = \delta^2(\delta(x - \sin(x)) + \sin(x)(\delta - \sin(\delta))) \le |\delta|^3|x|^3/6 + |\delta|^5|x|/6$.
3. $-2(1-\cos(\delta))\delta x + 2\delta x(1-\cos(x))(1-\cos(\delta)) + 2\delta(1-\cos(\delta))\sin(x) = 2\delta(1-\cos(\delta))(x(1 - \cos(x)) + \sin(x) - x) \le |\delta|^3|x|^3/3$, since for all real x one has $|\sin(x) - x\cos(x)| \le |x|^3/3$.
4. $-(\cos(x) - 1)(1 - \cos(\delta))\delta^2 \le x^2\delta^4/4$.
5. $\delta^2 x^2 - 2\delta x\sin(x)\sin(\delta) - 2\delta\sin(\delta)(\cos(x) - 1) = x^2\delta(\delta - \sin(\delta)) + 2\delta\sin(\delta)(1 - \cos(x) - x\sin(x) + x^2/2) \le \delta^4 x^2/6 + \delta^2 x^4/4$, since for all real x, one has $0 \le 1 - \cos(x) - x\sin(x) + x^2/2 \le x^4/8$.

The latter inequalities are again seen to be true by comparing derivatives. □

Lemma 2.16. *Let $\{p, x, y\} \subset \mathbb{R}^d$. Then:*

1. $|(e^{i\langle p,x\rangle} - 1) - (e^{i\langle p,y\rangle} - 1)| \le |p||x - y|$.
2. $|(e^{i\langle p,x\rangle} - i\langle p, x\rangle - 1) - (e^{i\langle p,y\rangle} - i\langle p, y\rangle - 1)| \le 2|p|^2(|x| + |y|)|x - y|$.

Proof. This is an elementary consequence of the complex Mean Value Theorem [Evard and Jafari, 1992, p. 859, Theorem 2.2] and the Cauchy-Schwartz inequality. $\square$

Lemma 2.17. *For any Lévy measure λ on $\mathbb{R}$, one has for the two functions (given for $1 \geq \delta > 0$): $M_0(\delta) := \delta^2 \int_{[-1,1]\backslash(-\delta,\delta)} d\lambda(x)$ and $M_1(\delta) := \delta \int_{[-1,1]\backslash(-\delta,\delta)} |x| d\lambda(x)$ that $M_0(\delta) \to 0$ and $M_1(\delta) \to 0$ as $\delta \downarrow 0$. If, moreover, $\int_{[-1,1]} |x| d\lambda(x) < \infty$, then $\delta \int_{[-1,1]\backslash(-\delta,\delta)} d\lambda(x) \to 0$ as $\delta \downarrow 0$.*

Proof. Indeed let μ be the finite measure on $([-1,1], \mathcal{B}([-1,1]))$ given by $\mu(A) := \int_A x^2 d\lambda(x)$ (A a Borel subset of $[-1,1]$) and let $f_\delta^0(x) := \left(\frac{\delta}{x}\right)^2 \mathbb{1}_{[-1,1]\backslash(-\delta,\delta)}(x)$ and $f_\delta^1(x) := \frac{\delta}{|x|} \mathbb{1}_{[-1,1]\backslash(-\delta,\delta)}(x)$ be functions on $[-1,1]$. Clearly $0 \leq f_\delta^0, f_\delta^1 \leq 1$ and $f_\delta^0, f_\delta^1 \to 0$ pointwise as $\delta \downarrow 0$. Hence by Lebesgue's Dominated Convergence Theorem (DCT), we have $M_0(\delta) = \int f_\delta^0 d\mu$ and $M_1(\delta) = \int f^1(\delta) d\mu$ converging to $\int 0 d\mu = 0$ as $\delta \downarrow 0$. The "finite first absolute moment" case is similar. $\square$

Proposition 2.18. *Under scheme 1, with $\sigma^2 > 0$,* (2.3) *holds for all sufficiently small h.*

Definition 2.19. Pursuant to Proposition 2.18, under either of the two schemes, we let $h_\star \in (0, +\infty]$ be such that Q^h has non-negative off-diagonal entries for all $h \in (0, h_\star)$.

Proof. If $V = 0$ this is immediate. If $V = 1$, then (via a triangle inequality):

$$\begin{aligned} h|\mu^h| &\leq h \left| \sum_{s \in \mathbb{Z}_h \backslash \{0\}} s \int_{A_s^h} \mathbb{1}_{[-1,1]}(y) d\lambda(y) \right| \leq h \sum_{s \in \mathbb{Z}_h \backslash \{0\}} \int_{A_s^h} |s - u + u| \mathbb{1}_{[-1,1]}(y) d\lambda(y) \\ &\leq h \left(\frac{h}{2} \lambda([-1,1] \backslash [-h/2, h/2]) + \int_{[-1,1] \backslash [h/2, h/2]} |u| d\lambda(u) \right) \to 0 \end{aligned}$$

as $h \downarrow 0$ by Lemma 2.17. Eventually the expression is smaller than $\sigma^2 > 0$ and the claim follows. $\square$

Furthermore, we have the following inequalities, which together imply an estimate for $|\Psi^h - \Psi|$. In the following, recall the notation ($\delta \geq 0$): $\zeta(\delta) := \delta \int_{[-1,1]\backslash[-\delta,\delta]} |x| d\lambda(x)$, $\gamma(\delta) := \delta^2 \int_{[-1,1]\backslash[-\delta,\delta]} d\lambda(x)$, $c := \lambda(\mathbb{R})$, $b := \kappa(0)$, $d := \lambda(\mathbb{R}\backslash[-1,1])$. Recall also the definition of the sets A_s^h following (2.2).

Proposition 2.20 (Convergence of characteristic exponents)**.** *For all $p \in \mathbb{R}$: $0 \leq f_h(p) \leq p^4 h^2/4!$ and $0 \leq i\mathrm{sgn}(p) g_h(p) \leq h^2 |p|^3/3!$ (respectively, under scheme 2, $|g_h(p)| \leq hp^2/2!$). Moreover:*

(i) when $c < \infty$; with $V = 0$: $|l_h(p)| \leq c|p|h/2$.

(ii) when $b < \infty = c$; with $V = 1$; for all $h \leq 2$:

$$|l_h(p)| \leq \frac{h}{2}\left(|p|d + p^2 b\right) + (p^2 + |p|^3 + p^4)o(h)$$

(respectively under scheme 2,

$$|l_h(p)| \leq \frac{h}{2}\left(|p|d + 2p^2 b\right) + (p^2 + |p|^3 + p^4)o(h)$$

) where $o(h)$ depends only on λ.[2]

(iii) when $b = \infty$; with $V = 1$; for all $h \leq 2$:

$$|l_h(p)| \leq p^2\left(\zeta(h/2) + \frac{1}{2}\gamma(h/2)\right) + (|p| + |p|^3 + p^4)O(h)$$

(respectively under scheme 2,

$$|l_h(p)| \leq p^2\left[2\zeta(h/2) + \frac{1}{2}\gamma(h/2)\right] + (|p| + p^2 + |p|^3 + p^4)O(h)$$

) where again $O(h)$ depends only on λ. Note here that we always have $\gamma \leq \zeta$ and that ζ decays strictly slower than h, as $h \downarrow 0$.

Remark 2.21.

(i) We may briefly summarize the essential findings of Proposition 2.20 in Table 2.2, by noting that the following will have been proved for $p \in \mathbb{R}$ and $h \in (0, h_\star \wedge 2)$:

$$|\Psi^h(p) - \Psi(p)| \leq f(h)R(|p|) + o(f(h))Q(|p|) \tag{2.13}$$

where R and Q are polynomials of respective degrees α and β and $f : (0, h_\star \wedge 2) \to (0, \infty)$.

(ii) An analogue of (2.13) is got in the multivariate case simply by examining directly the difference of (2.12) and (2.1). One does so either component by component (when it comes to the drift and diffusion terms), the estimates

[2]The above notation, incorporating the symbol $o(h)$, is a slight abuse. Nevertheless, it is one to which we shall gratefully adhere in the sequel: thus $o(h)$ stands for a function of h, defined on some right neighborhood of 0, which is $o(h)$ in the sense of Notation 1.5.

$(f(h),\alpha,\beta)$	$\sigma^2>0$ (scheme 1)	$\sigma^2=0$ (scheme 2)
$\lambda(\mathbb{R})=0$ $(V=0)$	$(h^2,4,-\infty)$	$(h,2,-\infty)$
$\lambda(\mathbb{R})<\infty$ $(V=0)$	$(h,1,4)$	$(h,2,-\infty)$
$\kappa(0)<\infty=\lambda(\mathbb{R})$ $(V=1)$	$(h,2,4)$	
$\kappa(0)=\infty$ $(V=1)$	$(\zeta(h/2),2,4)$	

Table 2.2: Summary of Proposition 2.20 via the triplet $(f(h),\alpha,\beta)$ introduced in (i) of Remark 2.21. We agree $\deg 0=-\infty$, where 0 is the zero polynomial.

being then the same as in the univariate case; or else one employs, in addition, Lemma 2.16 (for the part corresponding to the integral against the Lévy measure). In particular, (2.13) (with $p\in\mathbb{R}^d$) follows for suitable choices of R, Q and f, and Table 2.2 remains unaffected, apart from its last entry, wherein ζ should be replaced by $\zeta+\chi$ (one must also replace "$\sigma^2=0$" (respectively "$\sigma^2>0$") by "Σ (respectively non-) degenerate" (amalgamating scheme 1 & 2 into the multivariate one) and $\lambda(\mathbb{R})$ by $\lambda(\mathbb{R}^d)$).

(iii) The above entails, in particular, convergence of $\Psi^h(p)$ to $\Psi(p)$ as $h\downarrow 0$ pointwise in $p\in\mathbb{R}$. Lévy's continuity theorem [Dudley, 2004, p. 326] and stationarity and independence of increments yield at once Remark 2.2.

(iv) Note that we use $V=1$ rather than $V=0$ when $b<\infty=c$, because this choice yields linear convergence (locally uniformly) of $\Psi^h\to\Psi$. By contrast, retaining $V=0$, would have meant that the decay of $\Psi^h-\Psi$ would be governed, modulo terms which are $O(h)$, by the quantity $Q(h):=\sum_{s\in\mathbb{Z}_h}\int_{A_s^h\cap[-1,1]}(s-u)d\lambda(u)$ (as will become clear from the estimates in the proof of Proposition 2.20 below). But the latter can decay slower than h. In particular, consider the family of Lévy measures, indexed by $\epsilon\in[0,1)$: $\lambda_\epsilon=\sum_{n=1}^\infty w_n\delta_{-x_n}$, with $h_n=1/3^n$, $x_n=3h_n/2$, $w_n=1/x_n^\epsilon$, $n\geq 1$. For all these measures $b<\infty=c$. Furthermore, it is straightforward to verify that $\liminf_{n\to\infty}Q(h_n)/K(h_n)>0$, where $K(h)$ is $h^{1-\epsilon}$ or $h\log(1/h)$, according as to whether $\epsilon\in(0,1)$ or $\epsilon=0$.

(v) It is seen from Table 2.2 that the order of convergence goes from quadratic (at least when $\sigma^2>0$) to linear, to sublinear, according as to whether the Lévy measure is zero, $\lambda(\mathbb{R})>0$ & $\kappa(0)<\infty$, or κ becomes more and more singular at the origin. Let us attempt to offer some intuition in this respect. First, the quadratic order of convergence is due to the convergence properties of the discrete second and symmetric first derivative. Further, as soon as the Lévy measure is non-zero, the latter is aggregated over the intervals $(A_s^h)_{s\in\mathbb{Z}_h\backslash\{0\}}$, length h, which (at least in the worst case scenario) commit respective errors of order $\lambda(A_s^h)h$ or $\int_{A_s^h}(|x| \wedge 1)d\lambda(x)h$ ($s \in \mathbb{Z}_h \backslash \{0\}$) each, according as to whether $V = 0$ or $V = 1$. Hence, the more singular the κ, the bigger the overall error. Figure 2.1 depicts this progressive worsening of the convergence rate for the case of α-stable Lévy processes.

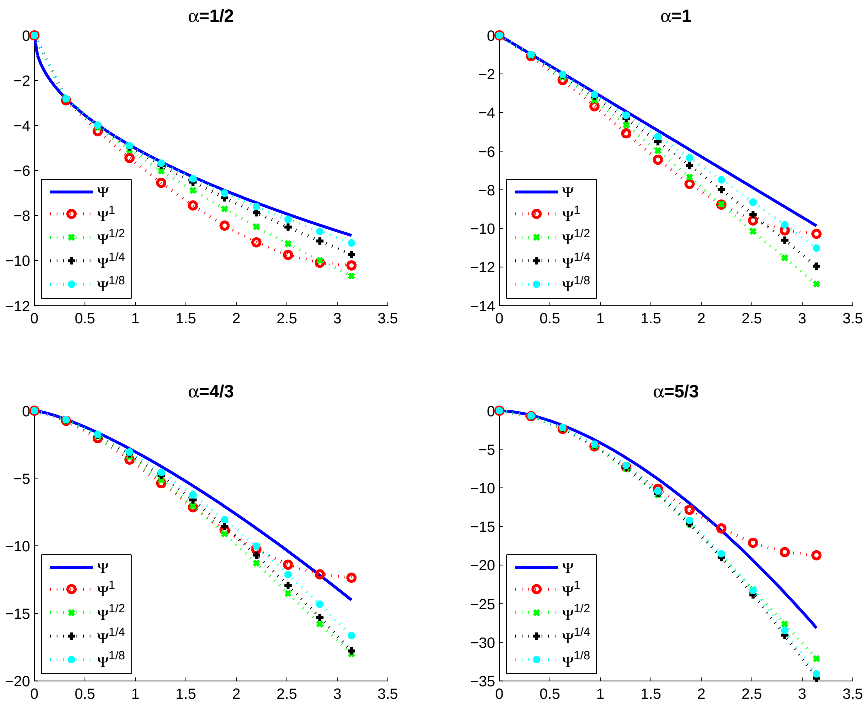

Figure 2.1: Comparison of the convergence of characteristic exponents for α-stable Lévy processes, $\alpha \in \{1/2, 1, 4/3, 5/3\}$; $\sigma^2 = 0$, $\mu = 0$ and $\lambda(dx) = dx/|x|^{1+\alpha}$ (scheme 2, $V = 1$). Each plot is of Ψ and of Ψ^h ($h \in \{1, 1/2, 1/4, 1/8\}$) on the interval $[0, \pi]$. Note that (i) $\kappa(0) = \infty$, precisely when $\alpha \geq 1$ and (ii) the characteristic exponents are real-valued for the examples shown. The plots are indeed suggestive of a progressive worsening of the rate of convergence as $\alpha \uparrow$.

Proof of Proposition 2.20. The first two assertions are transparent by Lemma 2.15 — with the exception of the estimate under scheme 2, where (with $\delta := hp$):

$$|g_h(p)| = \frac{1}{h}\sqrt{\delta^2 - 2\delta \sin(\delta) + 2(1 - \cos(\delta))} \leq \frac{1}{h}\frac{\delta^2}{2} = hp^2/2!.$$

Further, if $c < \infty$ (under $V = 0$):

$$\left|\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h(e^{isp}-1) - \int_{\mathbb{R}\backslash[-\frac{h}{2},\frac{h}{2}]}(e^{ipu}-1)d\lambda(u)\right| = \left|\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h e^{isp} - \int_{\mathbb{R}\backslash[-\frac{h}{2},-\frac{h}{2}]} e^{ipu}d\lambda(u)\right| =$$

$$= \left|\sum_{s\in\mathbb{Z}_h\backslash\{0\}}\int_{A_s^h}\left(e^{isp}-e^{ipu}\right)d\lambda(u)\right| \leq \sum_{s\in\mathbb{Z}_h\backslash\{0\}}\int_{A_s^h}\left|1-e^{ip(u-s)}\right|d\lambda(u) \leq |p|h\lambda\left(\mathbb{R}\backslash\left[-\frac{h}{2},\frac{h}{2}\right]\right)/2,$$

where in the second inequality we apply (1) of Lemma 2.15 and the first follows from the triangle inequalities. Finally, $|\int_{[-h/2,h/2]}(e^{ipu}-1)d\lambda(u)| \leq \lambda([-h/2,h/2])|p|h/2$, again by (1) of Lemma 2.15, and the claim follows.

For the remaining two claims, in addition to recalling the general results of Lemma 2.15, we prepare the following specific estimates. Whenever $\{x, y\} \subset \mathbb{R}$, with $\delta := y - x$, $0 \neq |x| \geq |\delta|$, we have:

- using the inequality $\sqrt{1+z} \leq 1 + z/2$ $(z \geq 0)$ and (2) of Lemma 2.15:

$$|e^{ix}-ix-e^{iy}+iy| \leq |\delta x|\left(1+\frac{1}{2}\left|\frac{\delta}{x}\right|+\frac{1}{8}\frac{\delta^2}{x^2}\right) = |\delta x| + \frac{1}{2}\delta^2 + \frac{1}{8}\left|\frac{\delta^3}{x}\right| \leq |\delta x| + \frac{5}{8}\delta^2. \quad (2.14)$$

- using (3) of Lemma 2.15:

$$|e^{ix}-ix-e^{iy}+iy| \leq |e^{ix}-e^{iy}-ix+iy+x^2/2-y^2/2|+\frac{1}{2}|x^2-y^2| \leq \frac{7}{6}|\delta|x^2+|\delta||x|+\frac{1}{2}\delta^2. \quad (2.15)$$

Now, when $c = \infty$ (under $V = 1$; for all $h \leq 2$), denoting $\xi(\delta) := \int_{[-\delta,\delta]} x^2 d\lambda(x)$, we have, under scheme 1, as follows:

$$\left|c_0^h\left(\frac{\cos(hp)-1}{h^2}+\frac{p^2}{2}\right)\right| \leq p^4h^2\xi(h/2)/4!. \quad (2.16)$$

$$\begin{aligned} & \left|\int_{[-\frac{h}{2},\frac{h}{2}]} u^2\left(-\frac{p^2}{2}\right)d\lambda(u) - \int_{[-\frac{h}{2},\frac{h}{2}]}\left(e^{ipu}-1-ipu\right)d\lambda(u)\right| \\ \leq & \left|\int_{[-\frac{h}{2},\frac{h}{2}]}\left(\cos(pu)-1+\frac{p^2u^2}{2!}\right)d\lambda(u)\right| + \left|\int_{[-\frac{h}{2},\frac{h}{2}]}\left(\sin(pu)-pu\right)d\lambda(u)\right| \\ \leq & \ p^4(h/2)^2\xi(h/2)/4! + |p|^3(h/2)\xi(h/2)/3!. \end{aligned} \quad (2.17)$$

$$|-\mu^h g_h(p)| = \left|-i\mu^h\left(\frac{\sin(hp)}{h}-p\right)\right| \leq \frac{1}{3!}h^2|p|^3\left(\zeta(h/2)+\kappa(h/2)\right). \quad (2.18)$$

$$\begin{aligned}
&\left|\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h(e^{isp}-1) - ip\mu^h - \int_{\mathbb{R}\backslash[-\frac{h}{2},\frac{h}{2}]}(e^{ipu}-1-ipu\mathbb{1}_{[-1,1]}(u))d\lambda(u)\right| \\
\leq\ & \sum_{s\in\mathbb{Z}_h\backslash\{0\}} \int_{A_s^h}\left|e^{ipu}-e^{ips}-ipu\mathbb{1}_{[-1,1]}(u)+ips\mathbb{1}_{[-1,1]}(u)\right|d\lambda(u) \\
\leq\ & \sum_{s\in\mathbb{Z}_h\backslash\{0\}} \left[\int_{A_s^h\cap(\mathbb{R}\backslash[-1,1])} + \int_{A_s^h\cap[-1,1]}\right]\left|e^{ipu}-e^{ips}-ipu\mathbb{1}_{[-1,1]}(u)+ips\mathbb{1}_{[-1,1]}(u)\right|d\lambda(u) \\
\leq\ & \frac{h}{2}|p|\int_{\mathbb{R}\backslash[-1,1]}d\lambda(u) + p^2\frac{h}{2}\int_{[-1,1]\backslash[-\frac{h}{2},\frac{h}{2}]}|u|d\lambda(u) + p^2\frac{5}{8}\left(\frac{h}{2}\right)^2\lambda([-1,1]\backslash[-h/2,h/2]), \quad (2.19)
\end{aligned}$$

where, in particular, we have applied (2.14) to $x = ps$, $y = pu$. If in addition $b = \infty$, we opt rather to use (2.15), again with $x = ps$ and $y = pu$, and obtain instead:

$$\begin{aligned}
&\left|\sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h(e^{isp}-1) - ip\mu^h - \int_{\mathbb{R}\backslash[-\frac{h}{2},\frac{h}{2}]}(e^{ipu}-1-ipu\mathbb{1}_{[-1,1]}(u))d\lambda(u)\right| \\
\leq\ & \frac{h}{2}|p|\int_{\mathbb{R}\backslash[-1,1]}d\lambda(u) + p^2\frac{h}{2}\int_{[-1,1]\backslash[-\frac{h}{2},\frac{h}{2}]}|u|d\lambda(u) + p^2\frac{1}{2}\left(\frac{h}{2}\right)^2\lambda([-1,1]\backslash[-h/2,h/2]) + \\
+\ & \frac{7}{6}|p|^3\frac{h}{2}\int_{[-1,1]}x^2d\lambda(x). \quad (2.20)
\end{aligned}$$

Under scheme 2, (2.16), (2.17) and (2.19)/(2.20) remain unchanged, whereas (2.18) reads:

$$\left|\mu^h g_h(p)\right| \leq \frac{h}{2}p^2\left(\zeta(h/2)+\kappa(h/2)\right). \quad (2.21)$$

Now, combining (2.16), (2.17), (2.18) and (2.19) under scheme 1 (respectively (2.16), (2.17), (2.21) and (2.19) under scheme 2), yields the desired inequalities when $b < \infty$. If $b = \infty$ use (2.20) in place of (2.19). □

2.4 Rates of convergence for transition kernels

Finally let us incorporate the estimates of Proposition 2.20 into an estimate of $D_{t,T}^h(x,y)$ (recall the notation in (2.5)). Assumption 2.3 and Table 2.1 are understood as being in effect throughout this section from this point onwards. Recall that $|\Psi^h - \Psi| \leq \sigma^2|f_h| + \mu|g_h| + |l_h|$ and that the approximation is considered for $h \in (0, h_\star)$ (cf. Definition 2.19).

First, the following observation, which is a consequence of the h-uniform growth of $-\Re\Psi^h(p)$ as $|p| \to \infty$, will be crucial to our endeavour (compare Remark 2.10).

Proposition 2.22. *For some $\{P, C, \epsilon\} \subset (0,\infty)$ and $h_0 \in (0, h_\star]$, depending only on $\{\lambda, \sigma^2\}$, and then all $h \in (0, h_0)$, $p \in [-\pi/h, \pi/h]\backslash(-P, P)$ and $t \geq 0$: $|\phi_{X_t}^h(p)| \leq$*

$\exp\{-Ct|p|^\epsilon\}$. *Moreover, when* $\sigma^2 > 0$*, we may take* $\epsilon = 2$*,* $P = 0$*,* $C = \frac{1}{2}\left(\frac{2}{\pi}\right)^2$ *and* $h_0 = h_\star$*, whereas otherwise* ϵ *may take the same value as in Orey's condition (cf. Assumption 2.3).*

Proof. Assume first $\sigma^2 > 0$, so that we are working under scheme 1. It is then clear from (2.10) that:

$$-\Re\Psi^h(p) \geq \sigma^2 \frac{1-\cos(hp)}{h^2} \geq \frac{1}{2}\left(\frac{2}{\pi}\right)^2 \sigma^2 p^2,$$

since $1 - \cos(x) = 2\sin^2(x/2) \geq 2\left(\frac{x}{\pi}\right)^2$ for all $x \in [-\pi, \pi]$. On the other hand, if $\sigma^2 = 0$, we work under scheme 2 and necessarily $V = 1$. In that case it follows from (2.11) for $h \leq 2$ and $p \in [-\pi/h, \pi/h]\backslash\{0\}$, that:

$$\begin{aligned}
-\Re\Psi^h(p) &\geq \left[c_0^h \frac{1-\cos(hp)}{h^2} + \sum_{s\in\mathbb{Z}_h\backslash\{0\}} c_s^h(1-\cos(sp))\right] \\
&\geq \frac{2}{\pi^2}p^2\left(\int_{A_0^h} u^2 d\lambda(u) + \sum_{s\in\mathbb{Z}_h\backslash\{0\},|s|\leq\frac{\pi}{|p|}} s^2 c_s^h\right) \\
&\geq \frac{2}{\pi^2}p^2\left(\int_{A_0^h} u^2 d\lambda(u) + \frac{4}{9}\sum_{s\in\mathbb{Z}_h\backslash\{0\},|s|\leq\frac{\pi}{|p|}} \int_{A_s^h} u^2 d\lambda(u)\right) \\
&\geq \frac{2}{\pi^2}p^2\left(\int_{A_0^h} u^2 d\lambda(u) + \frac{4}{9}\int_{[-\left(\frac{\pi}{|p|}-\frac{h}{2}\right),\frac{\pi}{|p|}-\frac{h}{2}]\backslash A_0^h} u^2 d\lambda(u)\right) \\
&\geq \frac{8}{9}\frac{1}{\pi^2}p^2\int_{\left[-\left(\left(\frac{\pi}{|p|}-\frac{h}{2}\right)\vee\frac{h}{2}\right),\left(\left(\frac{\pi}{|p|}-\frac{h}{2}\right)\vee\frac{h}{2}\right)\right]} u^2 d\lambda(u) \\
&\geq \frac{8}{9}\frac{1}{\pi^2}p^2\int_{[-\frac{1}{2}\frac{\pi}{|p|},\frac{1}{2}\frac{\pi}{|p|}]} u^2 d\lambda(u).
\end{aligned}$$

Now invoke Assumption 2.3. There are some $\{r_0, A_0\} \in (0, +\infty)$ such that for all $r \in (0, r_0]$: $\int_{[-r,r]} u^2 d\lambda(u) \geq A_0 r^{2-\epsilon}$. Thus for $P = \pi/(2r_0)$ and then all $p \in \mathbb{R}\backslash(-P, P)$, we obtain:

$$\int_{[-\frac{1}{2}\frac{\pi}{|p|},\frac{1}{2}\frac{\pi}{|p|}]} u^2 d\lambda(u) \geq A_0\left(\frac{1}{2}\frac{\pi}{|p|}\right)^{2-\epsilon},$$

from which the desired conclusion follows at once. Remark that, possibly, r_0 may be taken as $+\infty$, in which case P may be taken as zero. $\square$

Second, we have the following general observation which concerns the transfer of the rate of convergence from the characteristic exponents to the transition kernels.

Its validity is in fact independent of Assumption 2.3.

Proposition 2.23. *Suppose there are $\{P, C, \epsilon\} \subset (0, \infty)$, a real-valued polynomial R, an $h_0 \in (0, h_\star]$, and a function $f : (0, h_0) \to (0, \infty)$, decaying to 0 no faster than some power law, such that for all $h \in (0, h_0)$:*

1. *for all $p \in [-\pi/h, \pi/h]$: $|\Psi^h(p) - \Psi(p)| \leq f(h)R(|p|)$.*
2. *for all $s > 0$ and $p \in [-\pi/h, \pi/h] \backslash (-P, P)$: $|\phi_{X_s^h}(p)| \leq \exp\{-Cs|p|^\epsilon\}$; whereas for $p \in \mathbb{R} \backslash (-P, P)$: $|\phi_{X_s}(p)| \leq \exp\{-Cs|p|^\epsilon\}$.*

Then for any $s > 0$, $\Delta_s(h) = O(f(h))$.

Before proceeding to the proof of this proposition, we note explicitly the following elementary, but key lemma:

Lemma 2.24. *For $\{z, v\} \subset \mathbb{C}$: $|e^z - e^v| \leq (|e^z| \vee |e^v|)|z - v|$.*

Proof. This follows from the inequality $|e^z - 1| \leq |z|$ for $\Re z \leq 0$, whose validity may be seen by direct estimation. Indeed, (writing $z = \gamma_0 + is_0$, $\{\gamma_0, s_0\} \subset \mathbb{R}$), we have:

$$|e^z - 1|^2 = (e^{\gamma_0} - 1)^2 + 2e^{\gamma_0}(1 - \cos s_0).$$

Then, since $\gamma_0 \leq 0$, $e^{\gamma_0} \leq 1$ and $1 - e^{\gamma_0} \leq -\gamma_0$ (by comparing derivatives). Finally, use $1 - \cos s_0 \leq s_0^2/2$. □

Proof of Proposition 2.23. From (2.6) and (2.7) we have for the quantity $\Delta_s(h)$ from (2.5):

$$\Delta_s(h) \leq \int_{\mathbb{R}\backslash(-\pi/h,\pi/h)} |\exp\{\Psi(p)s\}|dp + \int_{[-\pi/h,\pi/h]} |\exp\{\Psi^h(p)s\} - \exp\{\Psi(p)s\}|dp.$$

Then the first term decays faster than any power law in h by (2) and L'Hôpital's rule, say, while in the second term we use the estimate of Lemma 2.24. Since $\exp\{-Ct|p|^\epsilon\}dp$ integrates every polynomial in $|p|$ absolutely, by (1) and (2) integration in the second term can then be extended to the whole of $\mathbb{R}$ and the claim follows. □

Proposition 2.23 allows us to transfer the rates of convergence directly from those of the characteristic exponents to the transition kernels. In particular, we have, immediately, the following proof of the multivariate result (which we state before the univariate case is dealt with in full detail):

Proof of Theorem 2.8. The conclusions of Theorem 2.8 follow from a straightforward extension (of the proof) of Proposition 2.23 to the multivariate setting, (ii) of Remark 2.21, Assumption 2.6 and Remark 2.7. $\square$

Returning to the univariate case, analogous conclusions could be got from Remark 2.10, Proposition 2.22 (themselves both consequences of Assumption 2.3) and Proposition 2.20. In the sequel, however, in the case when $\sigma^2 > 0$, we shall be interested in a more precise estimate of the constant in front of the leading order term (D_1 in the statement of Theorem 2.6). Moreover, we shall want to show our estimates are tight in an appropriate precise sense.

To this end we assume given a function K with the properties that:

(F) $0 \leq K(h) \to \infty$ as $h \downarrow 0$ and $K(h) \leq \frac{\pi}{h}$ for all sufficiently small h;

(E) the quantity

$$\mathcal{A}(h) := \left[\int_{-\infty}^{-K(h)} + \int_{K(h)}^{\infty}\right] |\exp\{\Psi(p)s\}|\, dp + \left[\int_{-\frac{\pi}{h}}^{-K(h)} + \int_{K(h)}^{\frac{\pi}{h}}\right] \left|\exp\{\Psi^h(p)s\}\right| dp$$

decays faster than the leading order term in the estimate of $D^h_{t,T}(x, y)$ (for which see, e.g., Table 2.2);

(C) $\sup_{[-K(h),K(h)]} |\Psi^h - \Psi| \leq 1$ for all small enough h

(suitable choices of K will be identified later, cf. Table 2.3 on p. 48). We now comment on the reasons behind these choices.

First, the constants $\{C, P, \epsilon\}$ are taken so as to satisfy simultaneously the conditions of Remark 2.10 and Proposition 2.22. In particular, if $\sigma^2 > 0$, we take $\epsilon = 2$, $P = 0$, $C = \frac{1}{2}\sigma^2$, and if $\sigma^2 = 0$, we may take ϵ from Orey's condition (cf. Assumption 2.3).

Next, we divide the integration regions in (2.6) and (2.7) into five parts (cf. property (F)): $(-\infty, -\frac{\pi}{h}]$, $(-\frac{\pi}{h}, -K(h))$, $[-K(h), K(h)]$, $(K(h), \frac{\pi}{h})$, $[\frac{\pi}{h}, \infty)$. Then we separate (via a triangle inequality) the integrals in the difference $D^h_{t,T}(x, y)$ accordingly and use the triangle inequality in the second and fourth region, thus (with $s := T - t > 0$):

$$\begin{aligned} 2\pi D^h_{t,T}(x, y) \leq\ & \left[\int_{-\infty}^{-\pi/h} + \int_{\pi/h}^{\infty}\right] |\exp\{\Psi(p)s\}|\, dp + \\ & \left[\int_{-\frac{\pi}{h}}^{-K(h)} + \int_{K(h)}^{\frac{\pi}{h}}\right] \left(\left|\exp\{\Psi^h(p)s\}\right| + |\exp\{\Psi(p)s\}|\right) dp + \\ & \int_{-K(h)}^{K(h)} \left|\exp\{\Psi(p)s\} - \exp\{\Psi^h(p)s\}\right| dp. \end{aligned}$$

Finally, we gather the terms with $|\exp\{\Psi(p)s\}|$ in the integrand and use $|e^z - 1| \leq e^{|z|} - 1$ ($z \in \mathbb{C}$) to estimate the integral over $[-K(h), K(h)]$, so as to arrive at:

$$2\pi D^h_{t,T}(x,y) \leq \mathcal{A}(h) + \int_{-K(h)}^{K(h)} |\exp\{\Psi(p)s\}| \left(\exp\left\{s\left|\Psi^h(p) - \Psi(p)\right|\right\} - 1\right) dp. \tag{2.22}$$

Now, the rate of decay of $\mathcal{A}(h)$ can be controlled by choosing $K(h)$ converging to $+\infty$ fast enough, viz. property (E). On the other hand, in order to control the second term on the right-hand side of the inequality in (2.22), we choose $K(h)$ converging to $+\infty$ slowly enough so as to guarantee (C). Table 2.3 lists suitable choices of $K(h)$. It is easily checked from Table 2.2 (respectively using L'Hôpital's rule coupled with Remark 2.10 and Proposition 2.22), that these choices of $K(h)$ do indeed satisfy (C) (respectively (E)) above. Property (F) is straightforward to verify.

	$\sigma^2 > 0$ (scheme 1)	$\sigma^2 = 0$ (scheme 2)
$\lambda(\mathbb{R}) = 0$ ($V = 0$)	$K(h) = \sqrt{\frac{2}{Cs}\log\frac{1}{h}} \to \mathcal{A}(h) = o(h^2)$	$\times$
$\lambda(\mathbb{R}) < \infty$ ($V = 0$)	$K(h) = \sqrt{\frac{1}{Cs}\log\frac{1}{h}} \to \mathcal{A}(h) = o(h)$	$\times$
$\kappa(0) < \infty = \lambda(\mathbb{R})$ ($V = 1$)	$K(h) = \sqrt{\frac{1}{Cs}\log\frac{1}{h}} \to \mathcal{A}(h) = o(h)$	$K(h) = \sqrt[\epsilon]{\frac{2}{Cs}\log\frac{1}{h}} \to \mathcal{A}(h) = o(h)$
$\kappa(0) = \infty$ ($V = 1$)	$K(h) = \left(\frac{1}{\zeta(h/2)}\right)^{1/4} \to \mathcal{A}(h) = o(\zeta(h/2))$	

Table 2.3: Suitable choices of $K(h)$. For example, the $\sigma^2 > 0$ and $\lambda(\mathbb{R}) = 0$ entry indicates that we choose $K(h) = \sqrt{\frac{2}{Cs}\log\frac{1}{h}}$ and then $\mathcal{A}(h)$ is of order $o(h^2)$.

Further, owing to (C), for all sufficiently small h, everywhere on $[-K(h), K(h)]$:

$$\begin{aligned} e^{s|\Psi^h - \Psi|} - 1 &= s|\Psi^h - \Psi| + \sum_{k=2}^{\infty} \frac{(s|\Psi^h - \Psi|)^k}{k!} \\ &\leq s|\Psi^h - \Psi| + (s|\Psi^h - \Psi|)^2 e^{s|\Psi^h - \Psi|} \leq s|\Psi^h - \Psi| + e(s|\Psi^h - \Psi|)^2. \end{aligned}$$

Manifestly the second term will always decay strictly faster than the first (so long as they are not 0). Moreover, since $\exp\{-Cs|p|^\epsilon\}dp$ integrates every polynomial in $|p|$ (cf. the findings of Proposition 2.20) absolutely, it will therefore be sufficient in the sequel to estimate (cf. (2.22)):

$$\frac{s}{2\pi}\int_{\mathbb{R}} \exp\{-Cs|p|^\epsilon\} \left|\Psi^h(p) - \Psi(p)\right| dp. \tag{2.23}$$

On the other hand, for the purposes of establishing sharpness of the rates for the quantity $D^h_{t,T}(x,y)$, we make the following:

Remark 2.25 (RD). Suppose we seek to prove that $f \geq 0$ converges to 0 no faster than $g > 0$, i.e. that $\limsup_{h\downarrow 0} f(h)/g(h) \geq C > 0$ for some C. If one can show $f(h) \geq A(h) - B(h)$ and $B = o(g)$, then to show $\limsup_{h\downarrow 0} f(h)/g(h) \geq C$, it is sufficient to establish $\limsup_{h\downarrow 0} A(h)/g(h) \geq C$. We refer to this extremely useful principle as *reduction by domination* (henceforth RD).

In particular, it follows from our discussion above, that it will be sufficient to consider (we shall always choose $x = y = 0$):

$$\frac{s}{2\pi}\int_{-K(h)}^{K(h)} e^{s\Psi(p)}\left(\Psi^h(p) - \Psi(p)\right) dp, \tag{2.24}$$

i.e. in Remark 2.25 this is A, and the difference from $D_{t,T}(0,0)$ represents B. Moreover, we can further replace $\Psi^h(p) - \Psi(p)$ in the integrand of (2.24) by any expression whose difference from $\Psi^h(p) - \Psi(p)$ decays, upon integration, faster than the leading order term. For the latter reductions we (shall) refer to the proof of Proposition 2.20.

We have now brought the general discussion as far as we can. The rest of the analysis must inevitably deal with each of the particular instances separately and we do so in the following two propositions.

Proposition 2.26 (Convergence of transition kernels — $\sigma^2 > 0$)**.** *Suppose $\sigma^2 > 0$. Then for any $s = T - t > 0$:*

1. *If $\lambda(\mathbb{R}) = 0$:*

$$\Delta_s(h) \leq h^2\left[\frac{1}{3\pi}\frac{|\mu|}{\sigma^4 s} + \frac{1}{8\sqrt{2\pi}}\frac{1}{(s\sigma^2)^{3/2}}\right] + o(h^2).$$

Moreover, with $\sigma^2 s = 1$ and $\mu = 0$ we have that $\limsup_{h\downarrow 0} D^h_{t,T}(0,0)/h^2 \geq 1/(8\sqrt{2\pi})$, thus proving that in general the convergence rate is no better than quadratic.

2. *If $0 < \lambda(\mathbb{R}) < \infty$:*

$$\Delta_s(h) \leq h\frac{1}{2\pi}\frac{c}{\sigma^2} + o(h).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}(\delta_{1/2} + \delta_{-1/2})$ one has that $\limsup_{h\downarrow 0} D^h_{t,T}(0,0)/h > 0$, showing that convergence in general is indeed no better than linear.

3. If $\kappa(0) < \infty = \lambda(\mathbb{R})$:

$$\Delta_s(h) \leq h\left[\frac{1}{2\pi}\frac{d}{\sigma^2} + \frac{1}{2\sqrt{2\pi}}\frac{bs}{(\sigma^2 s)^{3/2}}\right] + o(h).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}(\delta_{3/2} + \delta_{-3/2}) + \frac{1}{2}\sum_{k=1}^{\infty}(\delta_{1/3^k} + \delta_{-1/3^k})$, one has $\limsup_{h\downarrow 0} D^h_{t,T}(0,0)/h > 0$.

4. If $\kappa(0) = \infty$:

$$\Delta_s(h) \leq \frac{1}{\sqrt{2\pi}}\frac{s}{(\sigma^2 s)^{3/2}}\left(\zeta(h/2) + \frac{1}{2}\gamma(h/2)\right) + o(\zeta(h/2)).$$

Moreover, with $\sigma^2 = s = 1$, $\mu = 0$, and $\lambda = \sum_{k=1}^{\infty} w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/x_n$ ($n \in \mathbb{N}$), one has $\limsup_{h\downarrow 0} D^h_{t,T}(0,0)/\zeta(h/2) > 0$.

Proof. Estimates of $\Delta_s(h)$ follow at once from (2.23) and Proposition 2.20, simply by integration. As regards establishing sharpness of the estimates, however, we have as follows (recall that we always take $x = y = 0$):

1. $\lambda(\mathbb{R}) = 0$. Using (2.24) it is sufficient to consider:

$$A(h) := \frac{1}{2\pi}\left|\int_{-K(h)}^{K(h)} \exp\left\{-\frac{1}{2}p^2\right\} f_h(p)dp\right|.$$

By DCT, we have $A(h)/h^2 \to \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\{-\frac{1}{2}p^2\}p^4/4!dp$ and the claim follows.

2. $0 < \lambda(\mathbb{R}) < \infty$. Using (2.24) and further RD (recall Remark 2.25) via the estimates in the proof of Proposition 2.20, we conclude that it is sufficient to observe for the sequence $(h_n = \frac{1}{3^n})_{n\geq 1} \downarrow 0$ that:

$$\limsup_{n\to\infty}\frac{1}{2\pi h_n}\left|\int_{-K(h_n)}^{K(h_n)} \exp\left\{-\frac{1}{2}p^2 - 1 + \cos(p/2)\right\} l_{h_n}(p)dp\right| > 0.$$

It is also clear that we may express:

$$l_{h_n}(p) = 2\frac{1}{2}\Re\left(e^{ip(1/2-h_n/2)} - e^{ip/2}\right) = \cos(p/2)(\cos(ph_n/2) - 1) + \sin(p/2)\sin(ph_n/2).$$

Therefore, by further RD, it will be sufficient to consider:

$$\limsup_{n\to\infty}\frac{1}{2\pi h_n}\left|\int_{-K(h_n)}^{K(h_n)} \exp\left\{-\frac{1}{2}p^2 - 1 + \cos(p/2)\right\}\sin(ph_n/2)\sin(p/2)dp\right|.$$

By DCT it is equal to:

$$I := \frac{1}{2\pi}\int_0^\infty p\sin(p/2)\exp\{-\frac{1}{2}p^2 - 1 + \cos(p/2)\}dp.$$

The numerical value of this integral is (to one decimal place in units of $1/(2\pi)$) $0.4/(2\pi)$, but we can show that $I > 0$ analytically. Note the integrand is positive on $[0,6]$. Hence we have $2\pi e I \geq \sin(1/2)e^{\cos(3/2)}\int_1^3 pe^{-p^2/2}dp - e\int_6^\infty pe^{-p^2/2}dp = \sin(1/2)e^{\cos(3/2)}[e^{-1/2} - e^{-9/2}] - e^{-17}$. Now use $\sin(1/2) \geq (1/2)\cdot(2/\pi)$ (which follows from the concavity of sin on $[0,\pi/2]$), so that, very crudely: $2\pi e I \geq (1/\pi)e^{-1/2}(1 - e^{-4}) - e^{-17} \geq (1/\pi)e^{-1/2}(1/2) - e^{-17} \geq (1/e^2)e^{-1/2}(1/e) - e^{-17} \geq e^{-4} - e^{-17} > 0$.

3. $\kappa(0) < \infty = \lambda(\mathbb{R})$. Let $h_n = 1/3^n$, $n \geq 1$. Because the second term in λ lives on $\cup_{n\in\mathbb{N}}\mathbb{Z}_{h_n}$, it is seen quickly (via RD) that one need only consider (to within non-zero multiplicative constants):

$$\limsup_{n\to\infty}\int_{-K(h_n)}^{K(h_n)}\frac{1}{h_n}\sin(ph_n/2)\sin(3p/2)\exp\left\{-\frac{1}{2}p^2 + (\cos(3p/2)-1) + \sum_{k=1}^\infty(\cos(p/3^k)-1)\right\}dp.$$

By DCT it is sufficient to observe that:

$$\int_0^{2\pi/3}\sin(3p/2)p\exp\left\{-\frac{1}{2}p^2 + (\cos(3p/2)-1) - \frac{p^2}{2}\sum_{k=1}^\infty\frac{1}{9^k}\right\}dp - \int_{2\pi/3}^\infty p\exp\left\{-\frac{1}{2}p^2\right\}dp > 0.$$

To see the latter, note that the second integral is immediate and equal to: $e^{-(2\pi/3)^2/2}$. As for the first one, make the change of variables $u = 3p/2$. Thus we need to establish that:

$$A := (4/(9e))\int_0^\pi \sin(u)u\exp\{-u^2/4 + \cos(u)\}du - e^{-(2\pi/3)^2/2} > 0.$$

Next note that $-u^2/4 + \cos(u)$ is decreasing on $[0,\pi]$ and the integrand in A is positive. It follows that:

$$\begin{aligned}A \geq &\frac{4}{9e}\int_0^{\pi/3}u\sin(u)\exp\left\{-\frac{1}{4}\left(\frac{\pi}{3}\right)^2 + \cos\left(\frac{\pi}{3}\right)\right\}du + \\ &\frac{4}{9e}\int_{\pi/3}^{\pi/2}u\sin(u)\exp\left\{-\frac{1}{4}\left(\frac{\pi}{2}\right)^2 + \cos\left(\frac{\pi}{2}\right)\right\}du - e^{-2\pi^2/9} = \\ &\frac{4}{9e}\left[e^{-\frac{\pi^2}{36}+\frac{1}{2}}\left[\frac{\sqrt{3}}{2} - \frac{\pi}{6}\right] + e^{-\frac{\pi^2}{16}}\left[1 - \frac{\sqrt{3}}{2} + \frac{\pi}{6}\right]\right] - e^{-2\pi^2/9}\end{aligned}$$

Using series formulae/rational lower and upper bounds for π and π^2, and the series expansion of the exponential function, say, it is now an elementary exercise to verify that this explicit expression can be estimated from below by a positive quantity, so that $A > 0$.

4. $\kappa(0) = \infty$. Let again $h_n = 1/3^n$, $n \geq 1$. Notice that:

$$\sigma_1 := \int_{[-1,1]\backslash[-h_n/2,h_n/2]} u^2 d\lambda(u) = 2\sum_{k=1}^{n} x_k^2 w_k,$$

while

$$\sigma_2 := \sum_{s\in\mathbb{Z}_{h_n}\backslash\{0\}} c_s^{h_n} s^2 = 2\sum_{k=1}^{n}\left(x_k - \frac{h_n}{2}\right)^2 w_k,$$

so that $\Delta := \sigma_1 - \sigma_2 = 2\zeta(h_n/2) - \gamma(h_n/2) \geq \zeta(h_n/2)$. Using (3) of Lemma 2.15 in the estimates of Proposition 2.20, it is then not too difficult to see that it is sufficient to show $\int_{-K(h_n)}^{K(h_n)} p^2 \exp\{\Psi(p)\}dp$ converges to a strictly positive value as $n \to \infty$, which is transparent (since Ψ is real valued).

□

Proposition 2.27 (Convergence of transition kernels — $\sigma^2 = 0$)**.** *Suppose $\sigma^2 = 0$. For any $s = T - t > 0$:*

1. *If Orey's condition is satisfied and $\kappa(0) < \infty = \lambda(\mathbb{R})$, then $\Delta_s(h) = O(h)$. Moreover, with $\sigma^2 = 0$, $s = 1$, $\mu = 0$ and $\lambda = \frac{1}{2}\sum_{k=1}^{\infty} w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/\sqrt{x_n}$ ($n \in \mathbb{N}$), Orey's condition holds with $\epsilon = 1/2$ and one has* $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/h > 0$.

2. *If Orey's condition is satisfied and $\kappa(0) = \infty$, then $\Delta_s(h) = O(\zeta(h/2))$. Moreover, with $\sigma^2 = 0$, $s = 1$, $\mu = 0$, and $\lambda = \sum_{k=1}^{\infty} w_k(\delta_{x_k} + \delta_{-x_k})$, where $x_n = \frac{3}{2}\frac{1}{3^n}$ and $w_n = 1/x_n$ ($n \in \mathbb{N}$), Orey's condition holds with $\epsilon = 1$ and one has* $\limsup_{h\downarrow 0} D_{t,T}^h(0,0)/\zeta(h/2) > 0$.

Proof. Again the rates of convergence for $\Delta_s(h)$ follow at once from (2.23) and Proposition 2.20 (or, indeed, from Proposition 2.23). As regards sharpness of these rates, we have (recall that we take $x = y = 0$):

1. $\kappa(0) < \infty = \lambda(\mathbb{R})$. Let $h_n = 1/3^n$, $n \geq 1$. By checking Orey's condition on the decreasing sequence $(h_n)_{n\geq 1}$, Assumption 2.3 is satisfied with $\epsilon = 1/2$ and we have $b < \infty = c$. $\mu^h = 0$ by symmetry. Moreover by (2.24), and by further going through the estimates of Proposition 2.20 using RD, it suffices to show:

$$\limsup_{n\to\infty} \frac{1}{h_n}\left|\int_{-K(h_n)}^{K(h_n)} \exp\{\Psi(p)\}\left(\sum_{s\in\mathbb{Z}_{h_n}\backslash\{0\}}\int_{A_s^{h_n}} (\cos(ps) - \cos(pu))\, d\lambda(u)\right) dp\right| > 0.$$

Now, one can write for $s \in \mathbb{Z}_{h_n} \backslash \{0\}$ and $u \in A_s^{h_n}$,

$$\cos(sp) - \cos(pu) = \cos(pu)(\cos((s-u)p) - 1) - \sin(pu)(\sin((s-u)p) - (s-u)p) - \sin(pu)(s-u)p$$

and via RD get rid of the first two terms (i.e. they contribute to B rather than A in Remark 2.25). It follows that it is sufficient to observe:

$$\limsup_{n\to\infty} \frac{1}{h_n} \left| \int_{-K(h_n)}^{K(h_n)} \exp\left\{ \sum_{k=1}^{\infty} (\cos(px_k) - 1) w_k \right\} \left(\sum_{k=1}^{n} w_k \sin(px_k) \right) h_n p dp \right| > 0.$$

Finally, via DCT and evenness of the integrand, we need only have:

$$\int_0^\infty \left(\sum_{k=1}^{\infty} w_k \sin(px_k) \right) p \exp\left\{ \sum_{k=1}^{\infty} (\cos(px_k) - 1) w_k \right\} dp \neq 0.$$

One can in fact check that the integrand is strictly positive, as Lemma 2.28 shows, and thus the proof is complete.

2. $\kappa(0) = \infty$. The example here works for the same reasons as it did in (4) of the proof of Proposition 2.26 (but here benefiting explicitly also from $\mu^h = 0$). We only remark that Orey's condition is of course fulfilled with $\epsilon = 1$, by checking it on the decreasing sequence $(h_n)_{n\geq 1}$.

$\square$

Lemma 2.28. *Let $\psi(p) := \sum_{k=1}^{\infty} 3^{k/2} \sin(\frac{3}{2}p/3^k)$, $p \in (0, \infty)$. Then ψ is strictly positive.*

Proof. We observe that, (i) $\psi|_{(0,\frac{\pi}{2}]} > 0$ and (ii) for $p \in (\pi/2, 3\pi/2]$ we have: $\psi(p) > \sqrt{3}/(\sqrt{3} - 1) =: A_0$. Indeed, (i) is trivial since for $p \in (0, \pi/2]$, $\psi(p)$ is defined as a sum of strictly positive terms. We verify (ii) by observing that (ii.1) $\psi(\pi/2) > A_0$ and (ii.2) ψ is nondecreasing on $[\pi/2, 3\pi/2]$. Both these claims are tedious but elementary to verify by hand. Indeed, with respect to (ii.1), summing three terms of the series defining $\psi(\pi/2)$ is sufficient. Specifically we have $\psi(\pi/2) > \sqrt{3}\sin(\pi/4) + 3\sin(\pi/12) + 3\sqrt{3}\sin(\pi/36)$ and we estimate $\sin(\pi/36) \geq \frac{\pi}{36}\sin(\pi/3)/(\pi/3)$. With respect to (ii.2) we may differentiate under the summation sign, and then $\psi'(p) \geq \frac{\sqrt{3}}{2}\cos(3\pi/4) + \frac{1}{2}\cos(\pi/4) + \frac{\sqrt{3}}{6}\cos(\pi/12)$. The final details of the calculations are left to the reader.

Finally, we claim that if for some $B > 0$ we have $\psi|_{(0,B]} > 0$ and $\psi|_{(B,3B]} > A_0$, then $\psi|_{(0,3B]} > 0$ and $\psi|_{(3B,9B]} > A_0$, and hence the assertion of the lemma will follow at once (by applying the latter first to $B = \pi/2$, then $B = 3\pi/2$ and so on). So let $3p \in (3B, 9B]$, i.e. $p \in (B, 3B]$. Then $\psi(3p) = \sqrt{3}(\sin(3p/2) + \psi(p)) > \sqrt{3}(-1 + A_0) = A_0$, as required. $\square$

Finally, note that:

Proof of Theorem 2.4. The conclusions of Theorem 2.4 follow from Propositions 2.26 and 2.27. $\square$

2.5 Convergence of expectations and algorithm

2.5.1 Convergence of expectations

For the sake of generality we state the results in the multivariate setting, but only do so when this is not too burdensome on the brevity of exposition. For $d = 1$, either the multivariate or the univariate schemes may be considered.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be bounded Borel measurable and define for $t \geq 0$ and $h \in (0, h_\star)$: $p_t := p_{0,t}$ and $P_t^h := P_{0,t}^h$, whereas for $x \in \mathbb{Z}_h^d$, we let $p_t(x) := p_t(0, x)$ and $P_t^h(x) = P_t^h(0, x)$ (assuming the continuous transition densities exist). Note that for $t \geq 0$, and then for $x \in \mathbb{R}^d$ (by spatial homogeneity):

$$\mathsf{E}^x[f \circ X_t] = \int_{\mathbb{R}} f(y) p_t(x, y) dy, \tag{2.25}$$

whereas for $x \in \mathbb{Z}_h^d$ and $h \in (0, h_\star)$:

$$\mathsf{E}^x[f \circ X_t^h] = \sum_{y \in \mathbb{Z}_h^d} f(y) P_t^h(x, y) \tag{2.26}$$

(where, as usual, under P^x (which induces E^x), $X_0 = x$ and $X_0^h = x$, a.s.). Moreover, if f is continuous, we know that, as $h \downarrow 0$, $\mathsf{E}^x[f \circ X_t^h] \to \mathsf{E}^x[f \circ X_t]$, since $X_t^h \to X_t$ in distribution. Next, under additional assumptions on the function f, we are able to establish the rate of this convergence and how it relates to the convergence rate of the transition kernels, to wit:

Proposition 2.29. *Assume* (2.9) *of Assumption 2.6. Let $h_0 \in (0, \infty)$, $g : (0, h_0) \to (0, \infty)$ and $t > 0$ be such that $\Delta_t = O(g)$ (recall notation* (2.5)*). Suppose furthermore that the following two conditions on f are satisfied:*

(i) f is (piecewise[3]*, if $d = 1$) Lipschitz continuous.*

(ii) $\sup_{h \in (0, h_0)} h^d \sum_{x \in \mathbb{Z}_h^d} |f(x)| < \infty$.

[3]In the sense that there exists some natural n, and then disjoint open intervals $(I_i)_{i=1}^n$, whose union is cofinite in $\mathbb{R}$, and such that $f|_{I_i}$ is Lipschitz for each $i \in \{1, \ldots, n\}$.

Then:

$$\sup_{x\in\mathbb{Z}_h^d} |\mathsf{E}^x[f\circ X_t] - \mathsf{E}^x[f\circ X_t^h]| = O(h\vee g(h)).$$

Remark 2.30.

(1) Condition (ii) is satisfied in the univariate case $d=1$, if, e.g.: $f\in L^1(\mathbb{R})$, with respect to the Lebesgue measure, f is locally bounded and for some $K\in[0,\infty)$, $|f||_{(-\infty,-K]}$ (restriction of $|f|$ to $(-\infty,-K]$) is nondecreasing, whereas $|f||_{[K,\infty)}$ is nonincreasing.

(2) The rate of convergence of the expectations is thus got by combining the above proposition with the findings of Theorems 2.4 and 2.8.

(3) Note also that the convergence rate in Proposition 2.29 is never established as better than linear, albeit the transition kernels may converge faster, e.g. at a quadratic rate in the case of Brownian motion with drift. This is so because we are not only approximating the density with the normalized probability mass function, but also the integral is substituted by a sum (cf. (2.25) and (2.26)). One thus has to estimate $f(y)-f(z)$ for $z\in A_y^h$, $y\in\mathbb{Z}_h^d$. Excluding the trivial case of a constant f, however, this estimate can be at best linear in $|y-z|$ (Hölder continuous functions on $\mathbb{R}^d$ with Hölder exponent $\alpha>1$ being, in fact, constant). Moreover, it appears that this problem could not be avoided using Fourier inversion techniques (as opposed to the direct estimate given in the proof below). Indeed, one would then need to estimate, in particular, the difference of the Fourier transforms $\int_{\mathbb{R}} e^{-ipy}f(y)dy - h\sum_{y\in\mathbb{Z}_h^d} f(y)e^{-ipy}$, wherein again an integral is substituted by a discrete sum and a similar issue arises.

Proof. Decomposing the difference $\mathsf{E}^x[f\circ X_t]-\mathsf{E}^x[f\circ X_t^h]$ via (2.25) and (2.26), we have:

$$\begin{aligned}
\mathsf{E}^x[f\circ X_t] - \mathsf{E}^x[f\circ X_t^h] &= \sum_{y\in\mathbb{Z}_h^d}\int_{A_y^h} (f(z)-f(y))\, p_t(x,z)dz + && (2.27)\\
&+ \sum_{y\in\mathbb{Z}_h^d}\int_{A_y^h} f(y)\,(p_t(x,z)-p_t(x,y))\, dz + && (2.28)\\
&+ \sum_{y\in\mathbb{Z}_h^d} f(y)h^d\left[p_t(x,y) - \frac{1}{h^d}P_t^h(x,y)\right]. && (2.29)
\end{aligned}$$

Now, (2.29) is of order $O(g(h))$, by condition (ii) and since $\Delta_t = O(g)$. Further, (2.27) is of order $O(h)$ on account of condition (i), and since $\int p_t(x,z)dz = 1$ for any

$x \in \mathbb{R}^d$ (to see how the function f being piecewise Lipschitz continuous is sufficient in dimension one ($d = 1$), simply observe $\sup_{\{x,y\}\subset\mathbb{R}} p_t(x, y)$ is finite, as follows immediately from the integral representation of p_t). Finally, note that $p_t(x, \cdot)$ is also Lipschitz continuous (uniformly in $x \in \mathbb{R}^d$), as follows again at once from the integral representation of the transition densities. Thus, (2.28) is also of order $O(h)$, where again we benefit from condition (ii) on the function f. $\square$

In order to be able to relax condition (ii) of Proposition 2.29, we first establish the following Proposition 2.31, which concerns finiteness of moments of X_t.

In preparation thereof, recall the definition of submultiplicativity of a function $g : \mathbb{R}^d \to [0, \infty)$:

$$g \text{ is submultiplicative} \Leftrightarrow \exists a \in (0, \infty) \text{ such that } g(x + y) \leq ag(x)g(y), \text{ whenever } \{x, y\} \subset \mathbb{R}^d \tag{2.30}$$

and we refer to [Sato, 1999, p. 159, Proposition 25.4] for examples of such functions. Any submultiplicative locally bounded function g is necessarily bounded in exponential growth [Sato, 1999, p. 160, Lemma 25.5], to wit:

$$\exists\{b, c\} \subset (0, \infty) \text{ such that } g(x) \leq be^{c|x|} \text{ for } x \in \mathbb{R}^d. \tag{2.31}$$

Proposition 2.31. *Let $g : \mathbb{R}^d \to [0, \infty)$ be measurable, submultiplicative and locally bounded, and suppose $\int_{\mathbb{R}^d\setminus[-1,1]^d} gd\lambda < \infty$. Then for any $t > 0$, $\mathsf{E}[g \circ X_t] < \infty$ and, moreover, there is an $h_0 \in (0, h_\star)$ such that*

$$\sup_{h\in(0,h_0)} \mathsf{E}[g \circ X_t^h] < \infty.$$

Conversely, if $\int_{\mathbb{R}^d\setminus[-1,1]^d} gd\lambda = \infty$, then for all $t > 0$, $\mathsf{E}[g \circ X_t] = \infty$.

Proof. The argument below follows the exposition given in [Sato, 1999, pp. 159-162], modifying the latter to the extent that uniform boundedness of $\mathsf{E}[g \circ X_t^h]$ over $h \in (0, h_0)$ may be obtained. In particular, we refer to [Sato, 1999, p. 159, Theorem 25.3] for the claim that $\mathsf{E}[g \circ X_t] < \infty$, if and only if $\int_{\mathbb{R}^d\setminus[-1,1]^d} gd\lambda < \infty$. We take $\{a, b, c\} \subset (0, \infty)$ satisfying (2.30) and (2.31) above. Recall also that λ^h is the Lévy measure of the process X^h, $h \in (0, h_\star)$.

Now, decompose $X = X^1 + X^2$ and $X^h = X^{h1} + X^{h2}$, $h \in (0, h_\star)$ as independent sums, where X^1 is compound Poisson, with Lévy measure $\lambda_1 := \mathbb{1}_{\mathbb{R}^d\setminus[-1,1]^d} \cdot \lambda$, and X^{h1} are also compound Poisson, with Lévy measures $\lambda_1^h := \mathbb{1}_{\mathbb{R}^d\setminus[-1,1]^d} \cdot \lambda^h$, $h \in (0, h_\star)$. Consequently X^2 is a Lévy process with characteristic triplet $(\Sigma, \mathbb{1}_{[-1,1]^d} \cdot \lambda, \mu)_{\tilde{c}}$ and X^{h2} are compound Poisson, with Lévy measures $\mathbb{1}_{[-1,1]^d} \cdot \lambda^h$, $h \in (0, h_\star)$.

Moreover, for $h \in (0, h_\star)$, by submultiplicativity and independence:

$$\mathsf{E}[g \circ X_t^h] = \mathsf{E}[g \circ (X_t^{h1} + X_t^{h2})] \leq a\mathsf{E}[g \circ X_t^{h1}]\mathsf{E}[g \circ X_t^{h2}].$$

We first estimate $\mathsf{E}[g \circ X_t^{h1}]$. Let $(J_n)_{n\geq 1}$ (respectively N_t) be the sequence of jumps (respectively number of jumps by time t) associated to (respectively of) the compound Poisson process X^{h1}. Then $X_t^{h1} = \sum_{j=1}^{N_t} J_j$ and so by submultiplicativity:

$$\begin{aligned}
\mathsf{E}[g \circ X_t^{h1}] &\leq \mathsf{E}\left[g(0)\mathbb{1}_{\{N_t=0\}} + a^{N_t-1}\prod_{j=1}^{N_t} g(J_j)\mathbb{1}_{\{N_t>0\}}\right] \\
&= g(0)e^{-t\lambda_1^h(\mathbb{R}^d)} + \sum_{n=1}^{\infty}\frac{t^n a^{n-1}}{n!}e^{-t\lambda_1^h(\mathbb{R}^d)}\left(\int g d\lambda_1^h\right)^n.
\end{aligned}$$

We also have for all $h \in (0, 1 \wedge h_\star)$:

$$\begin{aligned}
\int g d\lambda_1^h &= \sum_{s\in\mathbb{Z}_h^d\backslash[-1,1]^d}\int_{A_s^h} g(s)d\lambda = \sum_{s\in\mathbb{Z}_h^d\backslash[-1,1]^d}\int_{A_s^h} g(u+(s-u))d\lambda(u) \\
&\leq a\left(\sup_{k\in A_h^0} g(k)\right)\sum_{s\in\mathbb{Z}_h^d\backslash[-1,1]^d}\int_{A_s^h} g d\lambda, \text{by submultiplicativity} \\
&\leq a\left(\sup_{k\in A_1^0} g(k)\right)\int_{\mathbb{R}^d\backslash[-1/2,1/2]^d} g d\lambda.
\end{aligned}$$

Now, since g is locally bounded, λ is finite outside neighborhoods of 0, and since by assumption $\int_{\mathbb{R}^d\backslash[-1,1]^d} g d\lambda < \infty$, we obtain: $\sup_{h\in(0,1\wedge h_\star)}\mathsf{E}[g \circ X_t^{h1}] < \infty$.

Second, we consider $\mathsf{E}[g \circ X_t^{h2}]$. First, by boundedness in exponential growth and the triangle inequality:

$$\mathsf{E}[g \circ X_t^{h2}] \leq b\mathsf{E}[e^{c|X_t^{h2}|}] \leq b\mathsf{E}[e^{c\sum_{j=1}^d |X_{tj}^{h2}|}] = b\mathsf{E}\left[\prod_{j=1}^d e^{c|X_{tj}^{h2}|}\right].$$

It is further seen by a repeated application of the Cauchy-Schwartz inequality that it will be sufficient to show, for each $j \in \{1, \ldots, d\}$, that for some $h_0 \in (0, h_\star]$:

$$\sup_{h\in(0,h_0)}\mathsf{E}\left[e^{2^{d-1}c|X_{tj}^{h2}|}\right] < \infty.$$

Here $X_t^{h2} = (X_{t1}^{h2}, \ldots, X_{td}^{h2})$ and likewise for X_t^2. Fix $j \in \{1, \ldots, d\}$.

The characteristic exponent of X_j^{h2}, denoted Ψ_2^h, extends to an entire function on $\mathbb{C}$. Likewise for the characteristic exponent of X_j^2, denoted Ψ_2 [Sato, 1999, p. 160, Lemma 25.6]. Moreover, since, by expansion into power series, one has, locally uniformly in $\beta \in \mathbb{C}$, as $h \downarrow 0$:

- $\frac{e^{\beta h}+e^{-\beta h}-2}{2h^2} \to \frac{1}{2}\beta^2$;
- $\frac{e^{\beta h}-e^{-\beta h}}{2h} \to \beta$;
- $\frac{e^{\beta h}-1}{h} \to \beta$ and $\frac{1-e^{-\beta h}}{h} \to \beta$;

since furthermore:

- $\left((\beta, u) \mapsto \frac{e^{\beta u}-\beta u-1}{u^2}\right) : \mathbb{R}\backslash\{0\} \times \mathbb{C} \to \mathbb{C}$ is bounded on bounded subsets of its domain;

and since finally by the complex Mean Value Theorem [Evard and Jafari, 1992, p. 859, Theorem 2.2]:

- as applied to the function $(x \mapsto e^{\beta x}) : \mathbb{C} \to \mathbb{C}$; $|e^{\beta x} - e^{\beta y}| \leq |x-y||\beta|(|e^{\beta z_1}| + |e^{\beta z_2}|)$ for some $\{z_1, z_2\} \subset \mathrm{conv}(\{x, y\})$, for all $\{x, y\} \subset \mathbb{R}$;

- as applied to the function $(x \mapsto e^{\beta x} - \beta x) : \mathbb{C} \to \mathbb{C}$; $|e^{\beta x} - \beta x - (e^{\beta y} - \beta y)| \leq |x-y||\beta| \left(|e^{\beta z_1} - 1| + |e^{\beta z_2} - 1|\right)$ for some $\{z_1, z_2\} \in \mathrm{conv}(\{x, y\})$, for all $\{x, y\} \subset \mathbb{R}$;

then the usual decomposition of the difference $\Psi_2^h - \Psi_2$ (see proof of Proposition 2.20) shows that $\Psi_2^h \to \Psi_2$ locally uniformly in $\mathbb{C}$ as $h \downarrow 0$. Next let ϕ_2^h and ϕ_2 be the characteristic functions of X_{tj}^{h2} and X_{tj}^2, respectively, $h \in (0, h_\star)$; themselves entire functions on $\mathbb{C}$. Using the estimate of Lemma 2.24, we then see, by way of corollary, that also $\phi_2^h \to \phi_2$ locally uniformly in $\mathbb{C}$ as $h \downarrow 0$.

Now, since ϕ_2^h is an entire function, for $n \in \mathbb{N} \cup \{0\}$, $i^n \mathsf{E}[(X_{tj}^{h2})^n] = (\phi_2^h)^{(n)}(0)$ and it is Cauchy's estimate [Stewart and Tall, 1983, p. 184, Lemma 10.5] that, for a fixed $r > 2^{d-1}c$, $\left|(\phi_2^h)^{(n)}(0)\right| \leq \frac{n!}{r^n} M^h$, where $M^h := \sup_{\{z \in \mathbb{C}:|z|=r\}} |\phi_2^h|$. Observe also that for some $h_0 \in (0, h_\star]$, $\sup_{h \in (0,h_0)} M^h < \infty$, since $\phi_2^h \to \phi_2$ locally uniformly as $h \downarrow 0$ and ϕ_2 is continuous (hence locally bounded).

Further to this $\mathsf{E}[|X_{tj}^{h2}|^{2k+1}] \leq 1 + \mathsf{E}[(X_{tj}^{h2})^{2k+2}]$ (for $k \in \mathbb{N} \cup \{0\}$) and $\mathsf{E}\left[e^{2^{d-1}c|X_{tj}^{h2}|}\right] = \sum_{n=0}^{\infty} \frac{1}{n!} \mathsf{E}[|X_{tj}^{h2}|^n](c2^{d-1})^n$. From this the desired conclusion finally follows. $\square$

The following result can now be established in dimension $d = 1$:

Proposition 2.32. *Let $d = 1$ and $t > 0$. Let furthermore:*

(i) $g : \mathbb{R} \to [0,\infty)$*, measurable, satisfy* $\mathsf{E}[g \circ X_t] < \infty$*, g locally bounded, submultiplicative,* $g \neq 0$*.*

(ii) $f : \mathbb{R} \to \mathbb{C}$*, measurable, be locally bounded,* $\int_{\mathbb{R}} |f| \in (0,\infty]$*,* $|f|$ *ultimately monotone (i.e.* $|f||_{[K,\infty)}$ *and* $|f||_{(-\infty,-K]}$ *monotone for some* $K \in [0,\infty)$*),* $|f|/|g|$ *ultimately nonincreasing (i.e.* $(|f|/|g|)|_{[K,\infty)}$ *and* $(-|f|/|g|)|_{(-\infty,-K]}$ *nonincreasing for some* $K \in [0,\infty)$*), and with the following Lipschitz property holding for some* $\{a, A\} \in (0,\infty)$*:* $f|_{[-A,A]}$ *is piecewise Lipschitz, whereas*

$$|f(z) - f(y)| \leq a|z-y|(g(z)+g(y)), \text{ whenever } \{z,y\} \subset \mathbb{R}\backslash(-A,A).$$

(iii) $K : (0,\infty) \to [0,\infty)$*, with* $\lim_{0+} K = +\infty$*.*

Then $|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]|$ *is of order:*

$$O\left(\left(\int_{[-K(h),K(h)]} |f(x)|dx\right)(h \vee \Delta_t(h)) + \left(\frac{|f|}{|g|} \vee \frac{|f|}{|g|} \circ (-\mathrm{id}_{\mathbb{R}})\right)(K(h) - 3h/2)\right), \quad (2.32)$$

where $\Delta_t(h)$ *is defined in* (2.5)*.*

Remark 2.33.

(1) In (2.32) there is a balance of two terms, viz. the choice of the function K. Thus, the slower (respectively faster) that K increases to $+\infty$ at $0+$, the better the convergence of the first (respectively second) term, provided $f \notin L^1(\mathbb{R})$ (respectively $|f|/|g|$ is ultimately converging to 0, rather than it just being nonincreasing). In particular, when so, then the second term can be made to decay arbitrarily fast, whereas the first term will always have a convergence which is strictly worse than $h \vee \Delta_t(h)$. But this convergence can be made arbitrarily close to $h \vee \Delta_t(h)$ by choosing K increasing more and more slowly (this since f is locally bounded). In general the choice of K would be guided by balancing the rate of decay of the two terms.

(2) Since, in the interest of relative generality, (further properties of) f and λ are not specified, g cannot be made explicit. Confronted with a specific f and Lévy process X, we should like to choose g approaching infinity (at $\pm\infty$) as fast as possible, while still ensuring $\mathsf{E}[g \circ X_t] < \infty$ (cf. Proposition 2.31). This makes, *ceteris paribus*, the second term in (2.32) decay as fast as possible.

(3) We exemplify this approach by considering two examples. Suppose for simplicity $\Delta_t(h) = O(h)$.

(a) Let first $|f|$ be bounded by $(x \mapsto A|x|^n)$ for some $A \in (0,\infty)$ and $n \in \mathbb{N}$, and assume that for some $m \in (n,\infty)$, the function $g = (x \mapsto |x|^m \vee 1)$ satisfies $\mathsf{E}[g \circ X_t] < \infty$ (so that (i) holds). Suppose furthermore condition (ii) is satisfied as well (as it is for, e.g., $f = (x \mapsto x^n)$). It is then clear that the first term of (2.32) will behave as $\sim K(h)^{n+1}h$, and the second as $\sim K(h)^{-(m-n)}$, so we choose $K(h) \sim 1/h^{1/(1+m)}$ for a rate of convergence which is of order $O(h^{\frac{m-n}{m+1}})$.

(b) Let now $|f|$ be bounded by $(x \mapsto Ae^{\alpha|x|})$ for some $\{A,\alpha\} \subset (0,\infty)$, and assume that for some $\beta \in (\alpha,\infty)$, the function $g = (x \mapsto e^{\beta|x|})$ indeed satisfies $\mathsf{E}[g \circ X_t] < \infty$ (so that (i) holds). Suppose furthermore condition (ii) is satisfied as well (as it is for, e.g., $f = (x \mapsto (e^{\alpha x} - k)^+)$, where $k \in [0,\infty)$ — use Lemma 2.24). It is then clear that the first term of (2.32) will behave as $\sim e^{\alpha K(h)}h$, and the second as $\sim e^{-(\beta-\alpha)K(h)}$, so we choose, up to a bounded additive function of h, $K(h) = \log(1/h^{1/\beta})$ for a rate of convergence which is of order $O(h^{1-\frac{\alpha}{\beta}})$.

(4) Finally, note that Proposition 2.32 can, in particular, be applied to f, which is the mapping $(x \mapsto e^{ipx})$, $p \in \mathbb{R}$, once suitable functions g and K have been identified. This, however, would give weaker results than what can be inferred regarding the rate of the convergence of the characteristic functions $\phi_{X_t^h}(p) \to \phi_{X_t}(p)$ from Remark 2.21(i) (using Lemma 2.24, say). This is so, because the characteristic exponents admit the Lévy-Khintchine representation (allowing for a very detailed analysis of the convergence), a property that is lost for a general function f (cf. Remark 2.30(3)).

Proof of Proposition 2.32. This is a simple matter of estimation; for all sufficiently small $h > 0$:

$$|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]| = \left| \int_{\mathbb{R}} f(z)p_t(z)dz - \sum_{y \in \mathbb{Z}_h} f(y)P_t^h(y) \right|$$

$$\leq \left| \sum_{y \in [-K(h),K(h)] \cap \mathbb{Z}_h} \left(\int_{A_y^h} f(z)p_t(z)dz - f(y)P_t^h(y) \right) \right| + \sum_{y \in \mathbb{Z}_h \backslash [-K(h),K(h)]} |f(y)|P_t^h(y) +$$

$$\int_{\mathbb{R}\backslash[-(K(h)-h/2),K(h)-h/2]} |f(z)|\, p_t(z)dz$$

$$\leq \underbrace{\left| \sum_{y \in \mathbb{Z}_h \cap [-K(h),K(h)]} \int_{A_y^h} (f(z) - f(y))\, p_t(z)dz \right|}_{\text{(A)}} + \underbrace{\left| \sum_{y \in \mathbb{Z}_h \cap [-K(h),K(h)]} \int_{A_y^h} f(y)\, (p_t(z) - p_t(y))\, dz \right|}_{\text{(B)}} +$$

$$\underbrace{\left|\sum_{y\in\mathbb{Z}_h\cap[-K(h),K(h)]} f(y)h\left[p_t(y)-\frac{1}{h}P^h_t(y)\right]\right|}_{\text{(C)}}+$$

$$\underbrace{\left(\frac{|f|}{|g|}\vee\frac{|f|}{|g|}\circ(-\mathrm{id}_{\mathbb{R}})\right)(K(h))\,\mathsf{E}[g\circ X^h_t]}_{\text{(D)}}+\underbrace{\left(\frac{|f|}{|g|}\vee\frac{|f|}{|g|}\circ(-\mathrm{id}_{\mathbb{R}})\right)(K(h)-h/2)\,\mathsf{E}[g\circ X_t]}_{\text{(E)}}.$$

Thanks to Proposition 2.31, and the fact that $|f|/|g|$ is ultimately nonincreasing, (D) & (E) are bounded (modulo a multiplicative constant) by $\frac{|f|}{|g|}(K(h)-h/2)\vee\frac{|f|}{|g|}(-(K(h)-h/2))$. From the Lipschitz property of f, submultiplicativity and local boundedness of g, and the fact that $\mathsf{E}[g\circ X_t]<\infty$, we obtain (A) is of order $O(h)$. By the local boundedness and eventual monotonicity of $|f|$, the Lipschitz nature of p_t and the fact that $\int|f|>0$, (B) is bounded (modulo a multiplicative constant) by $h\int_{[-(K(h)+h),K(h)+h]}|f|$. Finally, a similar remark pertains to (C), but with $\Delta_t(h)$ in place of h. Combining these, using once again $\int|f|>0$, yields the desired result, since we may finally replace $K(h)$ by $(K(h)-h)\vee 0$. □

2.5.2 Algorithm

From a numerical perspective we must ultimately consider the processes X^h on a finite state space, which we take to be $S^h_M:=\{x\in\mathbb{Z}^d_h:|x|\leq M\}$ ($M>0$, $h\in(0,h_\star)$). We let $\hat{Q}^h$ denote the sub-Markov generator got from Q^h by restriction to S^h_M, and $\hat{X}^h$ be the corresponding Markov chain got by killing X^h at the time $T^h_M:=\inf\{t\geq 0:|X^h_t|>M\}$, sending it to the coffin state ∂ thereafter [Syski, 1992].

Then the basis for the numerical evaluations is the observation that for a (finite state space) Markov chain Y with generator matrix Q, the probability $\mathsf{P}^y(Y_t=z)$ (respectively the expectation $\mathsf{E}^y[f\circ Y]$, when defined) is given by $(e^{tQ})_{yz}$ (respectively $(e^{tQ}f)(y)$). With this in mind we propose the:

Sketch algorithm

(i) Choose $\{h, M\} \subset (0, \infty)$.

(ii) Calculate, for the truncated sub-Markov generator $\hat{Q}^h$, the matrix exponential $\exp\{t\hat{Q}^h\}$ or action $\exp\{t\hat{Q}^h\}f$ thereof (where f is a suitable vector).

(iii) Adjust truncation parameter M, if needed, and discretization parameter h, until sufficient precision has been established.

Two questions now deserve attention: (1) what is the truncation error and (2) what is the expected cost of this algorithm. We address both in turn.

First, with a view to the localization/truncation error, we shall find use of the following:

Proposition 2.34. *Let $g : [0, \infty) \to [0, \infty)$ be nondecreasing, continuous and submultiplicative, with $\lim_{+\infty} g = +\infty$. Let $t > 0$ and denote by*

$$X_t^\star = \sup_{s\in[0,t]} |X_s|, \quad X_t^{h\star} = \sup_{s\in[0,t]} |X_s^h|,$$

the running suprema of $|X|$ and of $|X^h|$, $h \in (0, h_\star)$, respectively. Suppose furthermore $\mathsf{E}[g \circ |X_t|] < \infty$. Then $\mathsf{E}[g \circ X_t^\star] < \infty$ and, moreover, there is some $h_0 \in (0, h_\star]$ such that

$$\sup_{h\in(0,h_0)} \mathsf{E}[g \circ X_t^{h\star}] < \infty.$$

Remark 2.35. The function $g \circ |\cdot| : \mathbb{R}^d \to [0, \infty)$ is measurable, submultiplicative and locally bounded, so for a condition on the Lévy measure equivalent to $\mathsf{E}[g \circ |X_t|] < \infty$ see Proposition 2.31.

We prove Proposition 2.34 below, but first let us show its relation to the truncation error. For a function $f : \mathbb{Z}_h^d \to \mathbb{R}$, we extend its domain to $\mathbb{Z}_h^d \cup \{\partial\}$, by stipulating that $f(\partial) = 0$. The following (very crude) estimates may then be made:

Corollary 2.36. *Fix $t > 0$. Assume the setting of Proposition 2.34. There is some $h_0 \in (0, h_\star]$ and then $C := \sup_{h\in(0,h_0)} \mathsf{E}[g \circ X_t^{h\star}] < \infty$, such that the following two claims hold:*

(i) For all $h \in (0, h_0)$:

$$\sum_{x \in \mathbb{Z}_h^d} |\mathsf{P}(X_t^h = x) - \mathsf{P}(\hat{X}_t^h = x)| = \mathsf{P}(T_M^h < t) \leq C/g(M).$$

(ii) Let $f : \mathbb{Z}_h^d \to \mathbb{R}$ and suppose $|f| \leq \tilde{f} \circ |\cdot|$, with $\tilde{f} : [0, \infty) \to [0, \infty)$ nondecreasing and such that $\tilde{f}/g$ is (respectively ultimately) nonincreasing. Then for all (respectively sufficiently large) $M > 0$ and $h \in (0, h_0)$:

$$|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]| \leq C \left(\frac{\tilde{f}}{g}\right)(M).$$

Remark 2.37.

1. With regard to (i), note that M may be taken fixed (i.e. independent of h) and chosen so as to satisfy a prescribed level of precision. In that case such a choice may be verified explicitly at least retrospectively: the sub-Markov generator $\hat{Q}^h$ gives rise to the sub-Markov transition matrix $\hat{P}_t^h := e^{t\hat{Q}^h}$; its deficit (in the row corresponding to state 0) is precisely the probability $\mathsf{P}(T_M^h < t)$.

2. But, with respect to (ii), M may also be made to depend on h, and then made to increase to $+\infty$ as $h \downarrow 0$, in which case it is natural to balance the rate of decay of $|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]|$ against that of $|\mathsf{E}[f \circ X_t] - \mathsf{E}[f \circ X_t^h]|$ (cf. Proposition 2.32). In particular, since $\mathsf{E}[g \circ |X_t|] < \infty \Leftrightarrow \mathsf{E}[g \circ X_t^\star] \Leftrightarrow \int_{\mathbb{R}^d \backslash [-1,1]^d} g \circ |\cdot| d\lambda < \infty$ [Sato, 1999, p. 159, Theorem 25.3 & p. 166, Theorem 25.18], this problem is essentially analogous to the one in Proposition 2.32. In particular, Remark 2.33 extends in a straightforward way to account for the truncation error, with M in place of $K(h) - 3h/2$.

Proof. (i) follows from the estimate $\sum_{x \in \mathbb{Z}_h^d} |\mathsf{P}(X_t^h = x) - \mathsf{P}(\hat{X}_t^h = x)| = \mathsf{P}(T_M^h < t) = \mathsf{P}(X_t^{h\star} > M) \leq \frac{\mathsf{E}[g \circ X_t^{h\star}]}{g(M)}$, which is an application of Markov's inequality. When it comes to (ii), we have for all (respectively sufficiently large) $M > 0$:

$$\begin{aligned}
&|\mathsf{E}[f \circ X_t^h] - \mathsf{E}[f \circ \hat{X}_t^h]| \leq \mathsf{E}\left[\left(|f| \circ X_t^h\right) \mathbb{1}(T_M^h < t)\right] \leq \mathsf{E}\left[\left(\tilde{f} \circ |X_t^h|\right) \mathbb{1}(T_M^h < t)\right] \\
\leq\ & \mathsf{E}\left[\left(\tilde{f} \circ X_t^{h\star}\right) \mathbb{1}(T_M^h < t)\right] = \mathsf{E}\left[\left(\left(\frac{\tilde{f}}{g}\right) \circ X_t^{h\star}\right) \left(g \circ X_t^{h\star}\right) \mathbb{1}(X_t^{h\star} > M)\right] \\
\leq\ & \left(\frac{\tilde{f}}{g}\right)(M) \mathsf{E}[g \circ X_t^{h\star}],
\end{aligned}$$

whence the desired conclusion follows. $\square$

Proof of Proposition 2.34. We refer to [Sato, 1999, p. 166, Theorem 25.18] for the proof that $\mathsf{E}[g \circ X_t^\star] < \infty$. Next, by right continuity of the sample paths of X, we may choose $b > 0$, such that $\mathsf{P}(X_t^* \leq b/2) > 0$ and we may also insist on $b/2$ being a continuity point of the distribution function of $X_t^\star$ (there being only denumerably many points of discontinuity thereof). Now, $X^h \to X$ as $h \downarrow 0$ with respect to the Skorokhod topology on the space of càdlàg paths. Moreover, by [Jacod and Shiryaev, 2003, p. 339, Proposition 2.4], the mapping $\Phi := (\alpha \mapsto \sup_{s\in[0,t]} |\alpha(s)|) : \mathbb{D}([0,\infty), \mathbb{R}^d) \to \mathbb{R}$ is continuous at every point α in the space of càdlàg paths $\mathbb{D}([0,\infty), \mathbb{R}^d)$, which is continuous at t. In particular, Φ is continuous, a.s. with respect to the law of the process X on the Skorokhod space [Sato, 1999, p. 59, Theorem 11.1]. By the Portmanteau Theorem and since weak convergence is preserved under continuous mappings, it follows that there is some $h_0 \in (0, h_\star]$ such that $\inf_{h\in(0,h_0)} \mathsf{P}(X_t^{h\star} \leq b/2) > 0$.

Moreover, from the proof of [Sato, 1999, p. 166, Theorem 25.18], by letting $\tilde{g} : [0,\infty) \to [0,\infty)$ be nondecreasing, continuous, vanishing at zero and agreeing with g on restriction to $[1,\infty)$, we may then show for each $h \in (0, h_\star)$ that:

$$\mathsf{E}[\tilde{g} \circ (X_t^{h\star} - b); X_t^{h\star} > b] \leq \mathsf{E}[\tilde{g} \circ |X_t^h|]/\mathsf{P}(X_t^{h\star} \leq b/2).$$

Now, since $\mathsf{E}[g \circ |X_t|] < \infty$, by Proposition 2.31 (cf. Remark 2.35), there is some $h_0 \in (0, h_\star]$ such that $\sup_{h\in(0,h_0)} \mathsf{E}[g \circ |X_t^h|] < \infty$, and thus $\sup_{h\in(0,h_0)} \mathsf{E}[\tilde{g} \circ |X_t^h|] < \infty$.

Combining the above, it follows that for some $h_0 \in (0, h_\star]$, $\sup_{h\in(0,h_0)} \mathsf{E}[\tilde{g} \circ (X_t^{h\star} - b); X_t^{h_\star} > b] < \infty$ and thus $\sup_{h\in(0,h_0)} \mathsf{E}[g \circ (X_t^{h\star} - b); X_t^{h_\star} > b] < \infty$. Finally, an application of submultiplicativity of g allows us to conclude. $\square$

Having thus dealt with the truncation error, let us briefly discuss the cost of our algorithm.

The latter is clearly governed by the calculation of the matrix exponential, or, respectively, of its action on some vector. Indeed, if we consider as fixed the generator matrix $\hat{Q}^h$, and, in particular, its dimension $n \sim (M/h)^d$, then this may typically require $O(n^3)$ [Moler and Loan, 2003; Higham, 2005], respectively $O(n^2)$ [Al-Mohy and Higham, 2011], floating point operations. Note, however, that this is a notional complexity analysis of the algorithm. A more detailed argument would ultimately have to specify precisely the particular method used to determine the (respectively action of a) matrix exponential, and, moreover, take into account how $\hat{Q}^h$ (and, possibly, the truncation parameter M, cf. Remark 2.37) behave as $h \downarrow 0$.

Further analysis in this respect goes beyond the desired scope of this thesis.

We finish off by giving some numerical experiments in the univariate case. To compute the action of $\hat{Q}^h$ on a vector we use the MATLAB function `expmv.m` [Al-Mohy and Higham, 2011], unless $\hat{Q}^h$ is sparse, in which case we use the MATLAB function `expv.m` from [Sidje, 1998].

We begin with transition densities. To shorten notation, fix the time $t = 1$ and allow $p := p_1(0, \cdot)$ and $p^h := \frac{1}{h}\hat{P}_1^h(0, \cdot)$ ($\hat{P}^h$ being the analogue of P^h for the process $\hat{X}^h$). Note that to evaluate the latter, it is sufficient to compute $(e^{\hat{Q}^h t})_{0\cdot} = (e^{(\hat{Q}^h)' t}\mathbb{1}_{\{0\}})'$, where $(\hat{Q}^h)'$ denotes transposition.

Example 2.38. Consider first Brownian motion with drift, $\sigma^2 = 1$, $\mu = 1$, $\lambda = 0$ (scheme 1, $V = 0$). We compare the density p with the approximation p^h ($h \in \{1/2^n : n \in \{0, 1, 2, 3\}\}$) on the interval $[0, 2]$ (see Figure 2.2 on p. 65), choosing $M = 5$. The vector of deficit probabilities $(\mathsf{P}(T_M^{1/2^n} < t))_{n=0}^3$ corresponding to using this truncation was $(5.9 \cdot 10^{-4}, 1.5 \cdot 10^{-4}, 5.8 \cdot 10^{-5}, 4.4 \cdot 10^{-5})$. In this case the matrix $\hat{Q}^h$ is sparse.

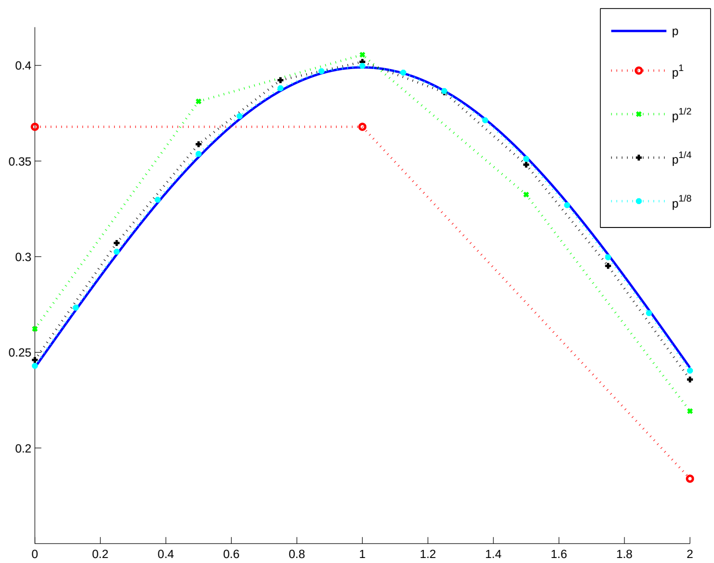

Figure 2.2: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0, 2]$ for Brownian motion with drift ($\sigma^2 = \mu = 1$, $\lambda = 0$, scheme 1, $V = 0$). See Example 2.38 for details.

Example 2.39. Consider now α-stable Lévy processes, $\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = dx/|x|^{1+\alpha}$ (scheme 2, $V = 1$). We compare the density p with p^h on the interval

$[0,1]$ (see Figure 2.3 on p. 67). Computations are made for the vector of alphas given by $(\alpha_k)_{k=1}^4 := (1/2, 1, 4/3, 5/3)$ with corresponding truncation parameters $(M_k)_{k=1}^4 = (500, 100, 30, 20)$ resulting in the deficit probabilities (uniformly over the h considered) of $(\mathsf{P}(T_{M_k}^h < t))_{k=1}^4 = (1.7 \cdot 10^{-1}, 2.0 \cdot 10^{-2}, (\text{from } 1.7 \text{ to } 1.8) \cdot 10^{-2}, (\text{from } 0.94 \text{ to } 1.01) \cdot 10^{-2})$. The heavy tails of the Lévy density necessitate a relatively high value of M. Nevertheless, excluding the case $\alpha = 5/3$, a reduction of M by a factor of 5 resulted in an absolute change of the approximating densities, which was at most of the order of magnitude of the discretization error itself. Conversely, for $\alpha = 1/2$, when the deficit probability is highest and appreciable, increasing M by a factor of 2, resulted in an absolute change of the calculated densities of the order 10^{-6} (uniformly over $h \in \{1, 1/2, 1/4\}$). Finally, note that $\alpha = 1$ gives rise to the Cauchy distribution, whereas otherwise we use the MATLAB function `stblpdf.m` to get a benchmark density against which a comparison can be made.

Example 2.40. A particular VG model [Carr et al., 2002; Madan et al., 1998] has $\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = \frac{e^{-|x|}}{|x|}\mathbb{1}_{\mathbb{R}\backslash\{0\}}(x)dx$ (scheme 2, $V = 1$). Again we compare p with p^h ($h \in \{1/2^n : n \in \{0, 1, 2, 3\}\}$) on the interval $[0, 1]$ (see Figure 2.4 on p. 68), choosing $M = 5$. The vector of deficit probabilities $(\mathsf{P}(T_M^{1/2^n} < t))_{n=0}^3$ corresponding to using this truncation was $(5.2 \cdot 10^{-3}, 6.4 \cdot 10^{-3}, 7.2 \cdot 10^{-3}, 7.6 \cdot 10^{-3})$. The density p is given explicitly by $(x \mapsto e^{-|x|}/2)$.

Finally, to illustrate convergence of expectations, we consider a particular option pricing problem.

Example 2.41. Suppose that, under the pricing measure, the stock price process $S = (S_t)_{t\geq 0}$ is given by $S_t = S_0 e^{rt+X_t}$, $t \geq 0$, where S_0 is the initial price, r is the interest rate, and X is a tempered stable process with Lévy measure given by:

$$\lambda(dx) = c\left(\frac{e^{-\lambda_+ x}}{x^{1+\alpha}}\mathbb{1}_{(0,\infty)}(x) + \frac{e^{-\lambda_-|x|}}{|x|^{1+\alpha}}\mathbb{1}_{(-\infty,0)}(x)\right)dx.$$

To satisfy the martingale condition, we must have $\mathsf{E}[e^{X_t}] \equiv 1$, which in turn uniquely determines the drift μ (we have, of course, $\sigma^2 = 0$). The price of the European put option with maturity T and strike K at time zero is then given by:

$$P(T, K) = e^{-rT}\mathsf{E}[(K - S_T)^+].$$

We choose the same value for the parameters as [Poirot and Tankov, 2006], namely $S_0 = 100$, $r = 4\%$, $\alpha = 1/2$, $c = 1/2$, $\lambda_+ = 3.5$, $\lambda_- = 2$ and $T = 0.25$, so that we may quote the reference values $P(T, K)$ from there.

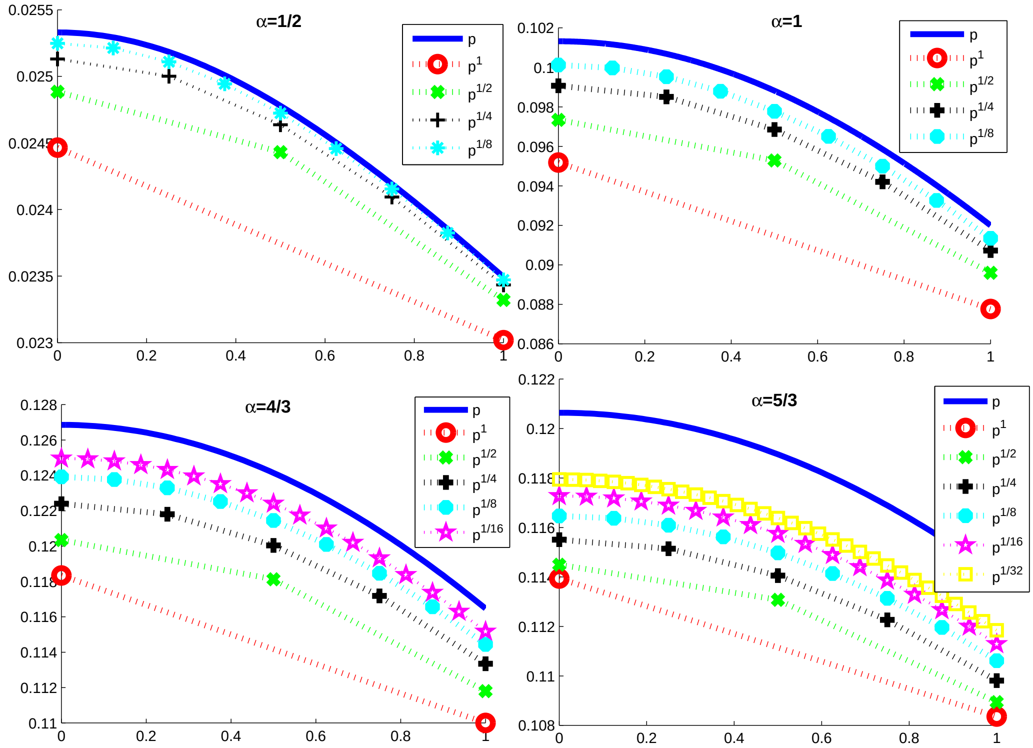

Figure 2.3: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0,1]$ for α-stable Lévy processes ($\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = dx/|x|^{1+\alpha}$, scheme 2, $V = 1$), $\alpha \in \{1/2, 1, 4/3, 5/3\}$. See Example 2.39 for details. Note that convergence becomes progressively worse as $\alpha \uparrow$, which is *precisely consistent* with Figure 2.1 and the theoretical order of convergence, this being $O(h^{(2-\alpha)\wedge 1})$ (up to a slowly varying factor $\log(1/h)$, when $\alpha = 1$; and noting that Orey's condition is satisfied with $\epsilon = \alpha$). For example, when $\alpha = 5/3$ each successive approximation should be closer to the limit by a factor of $\left(\frac{1}{2}\right)^{1/3} \doteq 0.8$, as it is.

Now, in the present case, X is a process of finite variation, i.e. $\kappa(0) < \infty$, hence convergence of densities is of order $O(h)$, since Orey's condition holds with $\epsilon = 1/2$ (scheme 2, $V = 1$). Moreover, $\mathbb{1}_{\mathbb{R}\backslash[-1,1]} \cdot \lambda$ integrates $(x \mapsto e^{2|x|})$, whereas the function $(x \mapsto (K - e^{rt+x})^+)$ is bounded. Pursuant to (2) of Remark 2.37 we thus choose $M = M(h) := \left(\frac{1}{2}\log(1/h)\right) \vee 1$, which by Corollary 2.36 and Proposition 2.32 (with $K(h) = M(h)$) (cf. also ((3)b) of Remark 2.33) ensures that:

$$|\hat{P}^h(T,K) - P(T,K)| = O(h \log(1/h)),$$

where $\hat{P}^h(T,K) := e^{-rT}\mathsf{E}[(K - S_0 e^{rT+\hat{X}^h_T})^+]$. Table 2.4 on p. 69 summarizes this

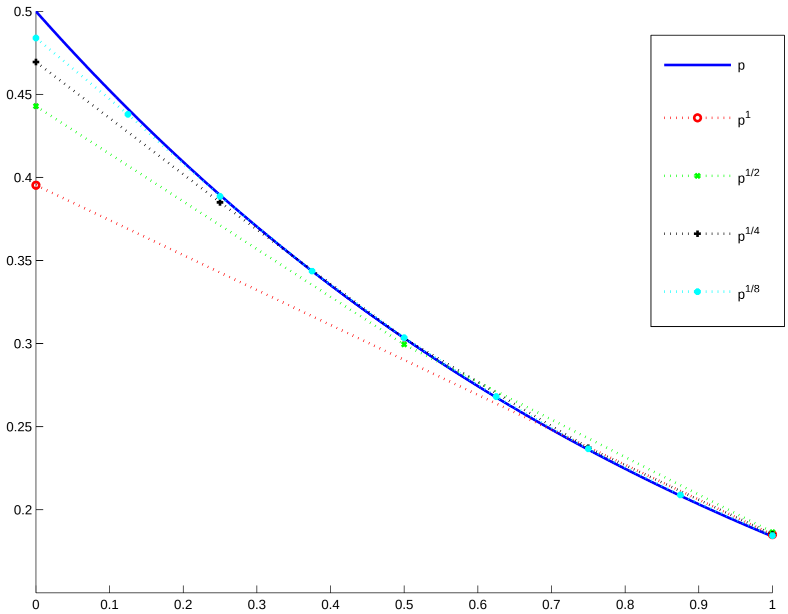

Figure 2.4: Convergence of p^h to p (as $h \downarrow 0$) on the interval $[0,1]$ for the VG model ($\sigma^2 = 0$, $\mu = 0$, $\lambda(dx) = \frac{e^{-|x|}}{|x|}\mathbb{1}_{\mathbb{R}\backslash\{0\}}(x)dx$, scheme 2, $V = 1$). Note that in this case Orey's condition fails, but (at least as evidenced numerically) linear convergence does not. See Example 2.40 for details.

convergence on the decreasing sequence $h_n := 1/2^n$, $n \geq 1$.

In particular, we wish to emphasize that the computations were all (reasonably) fast. For example, to compute the vector $(\hat{P}^{h_n}(T,K))_{n=1}^{9}$ with $K = 80$, the times (in seconds; entry-by-entry) (0.0106, 0.0038, 0.0044, 0.0078, 0.0457, 0.0367, 0.0925, 0.4504, 2.4219) were required on an Intel 2.53 GHz processor (times obtained using MATLAB's `tic-toc` facility). This is much better than, e.g., the Monte Carlo method of [Poirot and Tankov, 2006] and comparable with the finite difference method of [Cont and Voltchkova, 2005] (VG2 model in [Cont and Voltchkova, 2005, p. 1617, Section 7]).

In conclusion, the above numerical experiments serve to indicate that our method behaves robustly when the Blumenthal-Getoor index of the Lévy measure is not too close to 2 (in particular, if the pure-jump part has finite variation). It does less well if this is not the case, since then the discretisation parameter h must be chosen small, which is expensive in terms of numerics (viz. the size of $\hat{Q}^h$).

$K \rightarrow$	80	85	90	95	100	105	110	115	120
$P(T,K) \rightarrow$	1.7444	2.3926	3.2835	4.5366	6.3711	9.1430	12.7631	16.8429	21.1855
n	$\hat{P}^{h_n}(T,K) - P(T,K)$								
1	0.6411	0.5422	0.2006	-0.5033	-1.7885	-0.8227	0.0970	0.5570	0.7542
2	-0.1089	0.2816	0.4295	0.2151	-0.5806	0.0975	0.5341	0.5109	0.2250
3	-0.2271	-0.1596	-0.1928	0.0920	-0.2046	0.1405	0.0348	-0.4356	-0.3937
4	-0.0904	-0.0753	-0.0517	-0.0442	0.0652	0.1487	0.0057	-0.1511	-0.1838
5	-0.0411	-0.0338	-0.0193	-0.0053	0.0679	0.0569	-0.0073	-0.0616	-0.0833
6	-0.0184	-0.0163	-0.0081	0.0022	0.0347	0.0314	-0.0033	-0.0244	-0.0384
7	-0.0079	-0.0069	-0.0040	0.0019	0.0152	0.0109	-0.0034	-0.0108	-0.0164
8	-0.0034	-0.0029	-0.0016	0.0011	0.0072	0.0053	-0.0012	-0.0048	-0.0070
9	-0.0014	-0.0012	-0.0007	0.0006	0.0033	0.0026	-0.0004	-0.0020	-0.0030

Table 2.4: Convergence of the put option price for a CGMY model (scheme 2, $V = 1$). See Example 2.41 for details.

Chapter 3

Some fluctuation results in the theory of Lévy processes

The class of Lévy processes for which overshoots are almost surely constant quantities is precisely characterized. A fluctuation theory and, in particular, a theory of scale functions is developed for upwards skip-free Lévy chains, i.e. for right-continuous random walks embedded into continuous time as compound Poisson processes. This is done by analogy to the spectrally negative class of Lévy processes.

Throughout this chapter we work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$, which satisfies the standard assumptions (see Definition 1.16). We let $X = (X_t)_{t\geq 0}$ be a Lévy process on this space (X is assumed to be $\mathbb{F}$-adapted and to have independent increments relative to $\mathbb{F}$) with characteristic triplet $(\sigma^2, \lambda, \mu)_{\tilde{c}}$ relative to some cut-off function $\tilde{c}$. Recall the notation regarding the supremum and infimum processes $\overline{X}$ and $\underline{X}$, as well as the first passage times T_x, $\hat{T}_x$ and T_x^-, $x \in \mathbb{R}$ for the process X (see Definitions 1.20 and 1.22).

3.1 Non-random overshoots

3.1.1 Introduction

Fluctuation theory represents one of the most important areas within the study of Lévy processes, with applications in finance, insurance, dam theory etc. [Kyprianou, 2006] A key result, then, is the Wiener-Hopf factorization, particularly explicit in the spectrally negative case, when there are no positive jumps, a.s. [Sato, 1999, Section 9.46] [Bertoin, 1996, Chapter VII].

What makes the analysis so much easier in the latter instance, is the fact that the overshoots $(R_x)_{x\geq 0}$ [Sato, 1999, p. 369] over a given level are known *a priori* to be constant and equal to zero. As we shall see, this is also the only class of Lévy processes for which this is true (see Lemma 3.8). But it is not so much the exact values of the overshoots that matter, as does the fact that these values are non-random (and known). It is therefore natural to ask if there are any other Lévy processes having constant overshoots (a.s.) and, moreover, what *precisely* is the class having this property.

Of course, in the existing literature one finds expressions regarding the distribution of the overshoots. For example, [Sato, 1999, p. 369, Theorem 49.1] gives the double Laplace transform $\int_{(0,\infty)} e^{-ux}\mathsf{E}[e^{-qR_x}]dx$ ($\{u,q\} \subset (0,\infty)$) in terms of the Wiener-Hopf factors. Similarly, in [Doney and Kyprianou, 2006] we find an expression for the law of the overshoot in terms of the Lévy measure, but only after it has been integrated against the bivariate renewal functions. Unfortunately, neither of these seem immediately useful in answering the question posed above.

Further to this, the asymptotic study of quantities at first passage above a given level has been undertaken in [Doney and Kyprianou, 2006; Kyprianou et al., 2010] and behaviour just prior to first passage has also been investigated, see, e.g. [Sato, 1999, p. 378, Remark 49.9] and [Kyprianou, 2006, Chapter 7]. On the other hand it appears that the (natural) question, outlined above, has not yet received due attention.

The answer to it, presented in this section, is as follows: for the overshoots of a Lévy process to be almost surely constant (conditionally on the process going above the level in question), it is both necessary and sufficient that *either* the process has no positive jumps (a.s.) *or* for some $h > 0$, it is compound Poisson, living on the lattice $\mathbb{Z}_h = h\mathbb{Z}$, and can only jump up by h.

A more exhaustive statement of this result, which derives the same conclusion from substantially weakened hypotheses, is contained in Theorem 3.3 of Subsection 3.1.2, which also introduces the relevant notions. Subsection 3.1.3 supplies the proof. Finally, Appendix A contains a result concerning conditional expectation, Proposition A.2, which is used in the proof, but is also interesting in its own right.

3.1.2 Statement of result

First we introduce the continuous-time analogue (modulo a spatial scaling) of a right-continuous integer-valued random walk (for which see, e.g., [Brown et al., 2010]):

Definition 3.1 (Upwards skip-free Lévy chain)**.** X is said to be an *upwards skip-free Lévy chain*, if it is a compound Poisson process, and for some $h > 0$, $\mathrm{supp}(\lambda) \subset \mathbb{Z}_h$ and $\mathrm{supp}(\lambda|_{\mathcal{B}((0,\infty))}) = \{h\}$.

Second, the following notion, which is a rephrasing of "being almost surely constant conditionally on a given event", will prove useful:

Definition 3.2 (P-triviality)**.** Let $S \neq \emptyset$ be any measurable space, whose σ-algebra $\mathcal{S}$ contains the singletons. An S-valued random element R is said to be P-*trivial on* an event $A \in \mathcal{F}$ if there exists $r \in S$ such that $R = r$ P-a.s. on A (i.e. $\mathsf{P}(\{R = r\} \cap A) = \mathsf{P}(A)$; equivalently, the push-forward measure $(B \mapsto \mathsf{P}(A \cap R^{-1}(B)))$, defined on $\mathcal{S}$, is carried by $\{r\}$, not excluding the case when $\mathsf{P}(A) = 0$). The random element R may only be defined on some set $B \supset A$ (in which case R should be measurable with respect to the trace σ-algebra $\{B \cap G : G \in \mathcal{F}\}$ on B).

Thanks to Definitions 3.1 and 3.2, we can now state succinctly the main result of this section:

Theorem 3.3 (Non-random position at first passage time)**.** *The following are equivalent:*

(a) For some $x > 0$, $X(T_x)$ is P-trivial on $\{T_x < \infty\}$.

(b) For all $x \in \mathbb{R}$, $X(T_x)$ is P-trivial on $\{T_x < \infty\}$.

(c) For some $x \geq 0$, $X(\hat{T}_x)$ is P-trivial on $\{\hat{T}_x < \infty\}$ and P-a.s. strictly positive thereon.

(d) For all $x \in \mathbb{R}$, $X(\hat{T}_x)$ is P-trivial on $\{\hat{T}_x < \infty\}$.

(e) Either $\lambda((0,\infty)) = 0$ or X is an upwards skip-free Lévy chain.

If so, then the exceptional sets in (b) and (d) can actually be chosen not to depend on *x; i.e. outside a P-negligible set,* for each *$x \in \mathbb{R}$, $X(T_x)$ (respectively $X(\hat{T}_x)$) is constant on $\{T_x < \infty\}$ (respectively $\{\hat{T}_x < \infty\}$).*

Remark 3.4. In (c), if $x > 0$, then $X(\hat{T}_x)$ is automatically P-a.s. strictly positive on $\{\hat{T}_x < \infty\}$.

Finally, it will at times be convenient to work with the canonical space $\mathbb{D} := \{\omega \in \mathbb{R}^{[0,\infty)} : \omega \text{ is càdlàg}\}$ of càdlàg paths, mapping $[0,\infty)$ into $\mathbb{R}$. Then $\mathcal{H}$ will denote the σ-field generated by all the evaluation maps, whereas for $\omega \in \mathbb{D}$, $\overline{\omega}$ will be the supremum process of ω (i.e. $\overline{\omega}(t) := \sup\{\omega(s) : s \in [0,t]\}$, $t \geq 0$), and further for $a \in \mathbb{R}$, $T_a(\omega) := \inf\{t \geq 0 : \omega(t) \geq a\}$ will be the first entrance time of ω into the set $[a,\infty)$. Context shall make it clear when T_a will be seen as the latter mapping, $T_a : \mathbb{D} \to [0,\infty]$, and when as the first entrance time of X into $[a,\infty)$, as per above.

3.1.3 Proof of theorem

Remark 3.5. We note that $\mathsf{P}(T_x = 0 \text{ for all } x \in \mathbb{R}_-) = 1$. Moreover, $\mathsf{P}(T_x < \infty \text{ for all } x \in \mathbb{R}) = 1$, whenever X either drifts to $+\infty$ or oscillates. If not, then either X is the zero process, or else X drifts to $-\infty$ [Sato, 1999, p. 255, Proposition 37.10] and on the event $\{T_x = \infty\}$ one has $\lim_{t\to T_x} X(t) = -\infty$ for each $x \in \mathbb{R}$, P-a.s.

For the most part we find it more convenient to work with the collection $(T_x)_{x\in\mathbb{R}}$, rather than $(\hat{T}_x)_{x\in\mathbb{R}}$, even though this makes certain measurability issues more involved.

Remark 3.6. Note that whenever 0 is regular for $(0,\infty)$ (i.e. $\mathsf{P}(\hat{T}_0 = 0) = 1$), then for each $x \in \mathbb{R}$, $T_x = \hat{T}_x$ P-a.s. (apply the strong Markov property at the time T_x). For conditions equivalent to this, see [Kyprianou, 2006, p. 142, Theorem 6.5]. Conversely, if 0 is irregular for $(0,\infty)$, then by Blumenthal's $0-1$ law [Sato, 1999, p. 275, Proposition 40.4], P-a.s., $\hat{T}_0 > 0 = T_0$.

We now give two lemmas. The second concerns continuity of the supremum process $\overline{X}$. Since its formulation requires the relevant subsets of the sample space to be measurable, the first lemma establishes this.

Notation-wise, in the following lemma, for a process $Y = (Y_t)_{t\geq 0}$, we agree $Y_{0-} := Y_0$ and $Y_{t-} = \lim_{s\uparrow t} Y_s$ $(t > 0)$, whenever these limits exist.

Lemma 3.7. *Let $(\Omega', \mathcal{G}, \mathbb{G} = (\mathcal{G}_t)_{t\geq 0}, \mathsf{Q})$ be a filtered probability space. Suppose Y is a $\mathbb{G}$-adapted process. Then, with Ω_0 being the set on which Y is càdlàg, for each $\epsilon > 0$ and $t \geq 0$, $A_\epsilon := \cup_{s\in[0,t]}\{Y_s - Y_{s-} > \epsilon\} \cap \Omega_0 \in \mathcal{G}_t|_{\Omega_0}$, the trace σ-field. As a consequence of this, if Y is either (i) càdlàg or (ii), with $(\mathcal{G}, \mathsf{Q})$ complete, Q-a.s. càdlàg, then the sets $\{Y \text{ is continuous}\} = \{Y_{t-} = Y_t \text{ for all } t \geq 0\}$ and $\{Y \text{ has no positive jumps}\} = \{Y_{t-} \geq Y_t \text{ for all } t \geq 0\}$ belong to $\mathcal{G}$.*

Proof. Define in addition $B_\epsilon := \cup_{s\in[0,t]}\{Y_s - Y_{s-} \geq \epsilon\} \cap \Omega_0$ $(\epsilon > 0)$. Then, on the one hand, by the càdlàg property:

$$A_\epsilon \subset \cup_{n\in\mathbb{N}} F_n, \tag{3.1}$$

where $F_n := \Omega_0 \cap \left(\cap_{N\in\mathbb{N}} \cup_{\{s,r\}\subset(\mathbb{Q}\cap[0,t])\cup\{t\}, s<r, r-s<1/N} \{Y_r - Y_s > \epsilon + 1/n\}\right)$. On the other hand, again by the càdlàg property, for each $n \in \mathbb{N}$:

$$F_n \subset B_{\epsilon+1/n}. \tag{3.2}$$

Indeed, if $\omega \in F_n$, then for each $N \in \mathbb{N}$ we may choose a pair of real numbers (s_N, r_N), $0 \leq s_N < r_N \leq t$, $r_N - s_N < 1/N$, with $Y_{r_N}(\omega) - Y_{s_N}(\omega) > \epsilon + 1/n$. Since

$[0,t]$ is compact, there is some accumulation point $s^\star$ for the sequence $(s_N)_{N\geq 1}$, and, by passing to a subsequence, we may assume without loss of generality that $s_N \to s^\star$ as $N \to \infty$. Moreover, by right-continuity, it is necessary that there is some natural M, with $s_N < s^\star$ for all $N \geq M$; whereas by the existence of left-hand limits, it will also be necessary that there is some natural M, with $s^\star < r_N$ for all $N \geq M$. Then, by passing to the limit, it follows that $Y_{s^\star}(\omega) - Y_{s^\star -}(\omega) \geq \epsilon + 1/n$. From (3.2), we conclude that:

$$\cup_{n\in\mathbb{N}} F_n \subset \cup_{n\in\mathbb{N}} B_{\epsilon+1/n} = A_\epsilon. \tag{3.3}$$

Combining (3.1) and (3.3) we obtain $A_\epsilon \in \mathcal{G}_t|_{\Omega_0}$.

The final assertion of the lemma follows at once. □

Lemma 3.8 (Continuity of the running supremum)**.** *The supremum process $\overline{X}$ is continuous (*P*-a.s.), if and only if X has no positive jumps (*P*-a.s). In particular, if $X(T_x) = x$ *P*-a.s. on $\{T_x < \infty\}$ for each $x > 0$, then $\overline{X}$ is continuous and hence X has no positive jumps,* P*-a.s.*

Proof. We first show the validity of the equivalence. Indeed, sufficiency of the "no positive jumps" condition is immediate. We prove necessity by contradiction: suppose then, that X had positive jumps with a positive probability and (*per absurdum*) its supremum process was P-a.s. continuous. Then, for some $a > 0$, X would have a jump exceeding a with a positive probability and necessarily we would have $\lambda((a,\infty)) > 0$. Moreover, by the Lévy-Itô decomposition, one may write, P-a.s., $X = X^1 + X^2$ as an independent sum, where X^2 is a compound Poisson process of the positive jumps of X exceeding (i.e. of height $>$) a and $X^1 := X - X^2$ is whatever remains (see e.g. [Applebaum, 2009, p. 126, Theorem 2.4.16] and the results leading thereto, in particular [Applebaum, 2009, p. 116, Theorem 2.4.6]).

Next, let S be the supremum process of $|X^1|$ and T be the first jump time of X^2. By right-continuity of the sample paths, for some $t > 0$, $\mathsf{P}(\{S_t < a/2\}) > 0$. Further, by independence, and the fact that $T \sim \mathrm{Exp}(\lambda((a,\infty)))$ [Applebaum, 2009, p. 101, Theorem 2.3.5(1)], one has $\mathsf{P}(\{S_t < a/2\} \cap \{T < t\}) > 0$. Hence, with a positive probability, X will attain a new supremum (on $[0,t]$) by a jump in $\overline{X}$, which is a contradiction.

Finally, suppose $X(T_x) = x$ P-a.s. on $\{T_x < \infty\}$ for each $x > 0$. Then the supremum process $\overline{X}$ is a.s. continuous. Indeed, suppose not. Then with a positive probability $\overline{X}$ would have a jump, and therefore, for some pair of rationals r_1, r_2 with $0 < r_1 < r_2$, there would be a jump of $\overline{X}$ over (r_1, r_2) with a positive probability. Then, on this event $X(T_{(r_1+r_2)/2}) \geq r_2 > (r_1 + r_2)/2$, a contradiction. □

Having established this lemma, the first main step towards the proof of Theorem 3.3 is the following:

Proposition 3.9 (P-triviality of $X(T_x)$)**.** *The random variable* $X(T_x)$ *(defined on* $\{T_x < \infty\}$*) is* P*-trivial on* $\{T_x < \infty\}$ *for each* $x > 0$*, if and only if:*
either

(a) X *has no positive jumps (*P*-a.s.) (equivalently:* $\lambda((0, \infty)) = 0$*)*

or

(b) X *is compound Poisson and for some* $h > 0$*, we have* $\mathrm{supp}(\lambda) \subset \mathbb{Z}_h$*, while* $\mathrm{supp}(\lambda|_{\mathcal{B}((0,\infty))}) = \{h\}$

(conditions (a) and (b) being mutually exclusive). If so, then $X(T_x) = x$ *on* $\{T_x < \infty\}$ *for each* $x \geq 0$ *(P-a.s.) under (a) and* $X(T_x) = h\lceil x/h \rceil$ *on* $\{T_x < \infty\}$ *for each* $x \geq 0$ *(*P*-a.s.) under (b).*

Remark 3.10. Note that, under (b), $\mathsf{P}(\{X_t \in \mathbb{Z}_h \text{ for all } t \geq 0\}) = 1$. This follows by [Sato, 1999, p. 149, Corollary 24.6] and sample path right-continuity.

The main idea behind the proof of Proposition 3.9 is to appeal first to Lemma 3.8 for the case when, for all $x > 0$, $X(T_x) = x$ P-a.s. on $\{T_x < \infty\}$. This gives (a). Then we treat separately the compound Poisson case; in all other instances the Lévy-Itô decomposition and the well-established path properties of Lévy processes yield the claim. Intuitively, for a Lévy process to cross over every level in a non-random fashion, either it does so necessarily continuously when there are no positive jumps (cf. also [Kolokoltsov, 2011, p. 274, Proposition 6.1.2]), or, if there are, then it must be forced to live on the lattice $\mathbb{Z}_h$ for some $h > 0$ and only jump up by h. Formally:

Proof. Assume, without loss of generality, that X is càdlàg with certainty (rather than just P-a.s.). Clearly conditions (a) and (b) are mutually exclusive, sufficiency of the conditions and the final remark of Proposition 3.9 obtain by sample path right-continuity. With regard to the equivalence noted parenthetically in (a) see [Sato, 1999, p. 346, Remark 46.1].

Necessity of the conditions from Proposition 3.9 is shown as follows. Let $X(T_x)$ be P-trivial on $\{T_x < \infty\}$ for each $x > 0$.

Suppose first that for each $x > 0$, $X(T_x) = x$ (P-a.s.) on $\{T_x < \infty\}$. Then by Lemma 3.8, (a) must hold.

There remains the case when, for some $x > 0$, $\mathsf{P}(T_x < \infty) > 0$ and there is a non-random $f(x)$ with $f(x) = X(T_x) > x$ P-a.s. on $\{T_x < \infty\}$. In particular, X

must have positive jumps, and for some $a > 0$, $\beta := \lambda((a, \infty)) > 0$. Use again the Lévy-Itô decomposition as in the proof of Lemma 3.8 with S denoting the supremum process of $|X^1|$ and T the first jump time of X^2 (note that $T \sim \mathrm{Exp}(\beta)$). We will consider the following two cases separately:

(Case 1) X is *not* compound Poisson, i.e. either λ has infinite mass or $\sigma^2 > 0$, or if this fails (with $\tilde{c} = 0$ as the cut-off function) $\mu \neq 0$.

(Case 2) X is compound Poisson, i.e. the diffusion coefficient vanishes, $\sigma^2 = 0$, λ is finite and (with $\tilde{c} = 0$ as the cut-off function) the drift $\mu = 0$.

Consider first Case 1. By right-continuity of the sample paths, there is a $t > 0$ with $\mathsf{P}(\{S_t < a/4\}) > 0$.

We next argue that, on the event:

$$C := \{T < t\} \cap \{S_t < a/4\},$$

which has positive probability, $X^1(T)$ is not P-trivial. We prove this by contradiction. More precisely, we shall find that assuming the converse will contradict the following observation regarding the sample paths of X^1: the set of jump times of X^1 is dense, a.s., by [Sato, 1999, p. 136, Theorem 21.3] when λ has infinite mass; the sample paths of X^1 have locally infinite variation, a.s., by [Sato, 1999, p. 140, Theorem 21.9(ii)] when $\sigma^2 > 0$; finally, X^1 has no non-degenerate intervals of constancy, a.s., when $\sigma^2 = 0$, $\lambda(\mathbb{R}) < \infty$ but the drift is non-zero.

Indeed, suppose that $X^1(T)$ were to be P-trivial on the event C, so that there would be a (necessarily unique) $b \in (-a/4, a/4)$ with $X^1(T) = b$ P-a.s. on C, i.e. $\mathsf{P}(\{X^1(T) = b\} \cap C) = \mathsf{P}(C)$. We next condition on $\mathcal{G} := \sigma(T)$ by applying Proposition A.1 from Appendix A. Specifically, we take, discarding, without loss of generality, the P-negligible set $\{T = \infty\}$, $Y := T$ (so that $Y : (\Omega, \mathcal{F}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ and, of course, $\sigma(Y) \subset \mathcal{G}$) and $Z := X^1$ (so that $\sigma(Z) \perp \mathcal{G}$ and $Z : (\Omega, \mathcal{F}) \to (\mathbb{D}, \mathcal{H})$ — recall from the end of Subsection 3.1.2 notation pertaining to the space $(\mathbb{D}, \mathcal{H})$). Finally, $f : \mathbb{R}_+ \times \mathbb{D} \to \mathbb{R}$ is given by:

$$f(s, \omega) := \mathbb{1}_{\{b\}}(\omega(s)) \mathbb{1}_{[0,t)}(s) \mathbb{1}_{[0,a/4)}(\max\{\overline{\omega}(t), \overline{-\omega}(t)\}), \quad (s, \omega) \in \mathbb{R}_+ \times \mathbb{D}.$$

Note that the latter is bounded and $\mathcal{B}(\mathbb{R}) \otimes \mathcal{H}/\mathcal{B}(\mathbb{R})$-measurable by [Karatzas and Shreve, 1988, p. 5, Remark 1.14] and since, owing to sample path right-continuity, $(\omega \mapsto \overline{\omega}(t))$ is $\mathcal{H}/\mathcal{B}(\mathbb{R}_+)$-measurable. Proposition A.1 thus yields:

$$\mathsf{E}[f \circ (Y, Z) | \mathcal{G}] = g \circ Y,$$

where $g := (y \mapsto \mathsf{E}[f \circ (y, Z)])$, $g : \mathbb{R}_+ \to \mathbb{R}$, is Borel measurable. Now, on the one hand:

$$\mathsf{E}[g \circ Y] = \int g d\mathsf{P}_T = \int_0^\infty ds\beta e^{-\beta s}\mathsf{E}[f \circ (s, Z)] = \int_0^t ds\beta e^{-\beta s}\mathsf{P}(\{X^1(s) = b\} \cap \{S_t < a/4\}).$$

On the other hand:

$$\begin{aligned} \mathsf{E}[\mathsf{E}[f \circ (Y, Z)|\mathcal{G}]] &= \mathsf{E}[f \circ (Y, Z)] = \mathsf{P}(\{X^1(T) = b\} \cap C) = \mathsf{P}(C) \\ &= \mathsf{P}(T < t)\mathsf{P}(S_t < a/4) = \int_0^t ds\beta e^{-\beta s}\mathsf{P}(S_t < a/4). \end{aligned}$$

In summary, it follows that:

$$\int_0^t ds\beta e^{-\beta s}\mathsf{P}(\{X^1(s) = b\} \cap \{S_t < a/4\}) = \int_0^t ds\beta e^{-\beta s}\mathsf{P}(\{S_t < a/4\}).$$

Hence, Lebesgue-a.e. in $s \in (0, t)$, a.s. on $\{S_t < a/4\}$, $X^1(s) = b$. Now we can find for each rational $r \in (0, t)$ and $n \in \mathbb{N}$ an $x_n^r \in B(r, 1/n)$ for which a.s. on $\{S_t < a/4\}$, $X^1(x_n^r) = b$. So a.s. on $\{S_t < a/4\}$, on a dense countable subset of $(0, t)$, $X^1 = b$. Thus by sample path right-continuity a.s. on $\{S_t < a/4\}$, $X^1 = b$ everywhere on $[0, t)$. Hence, on an event of positive probability, there are *no* jump times on the whole of the non-degenerate interval $[0, t)$, the path has *zero* variation over $[0, t)$ and is, moreover, *constant* thereon, a contradiction.

We have thus established that $X^1(T)$ is *not* P-trivial on the event C.

Observe now that $X^2(T)$ is independent of T, both being jointly independent of X^1. Then $X^2(T) \perp \sigma(\mathbb{1}_C, X^1(T))$, so that (for Borel subsets A and B of $\mathbb{R}$):

$$\mathsf{P}(C \cap \{X^1(T) \in A\} \cap \{X^2(T) \in B\})\mathsf{P}(C) = \mathsf{P}(C \cap \{X^1(T) \in A\})\mathsf{P}(C \cap \{X^2(T) \in B\}).$$

We conclude that the first jump of X^2, $X^2(T)$, is independent of $X^1(T)$, conditionally on C. The support of their sum $X(T) = X^1(T) + X^2(T)$ on C, is therefore the closure of the sum of their respective supports [Sato, 1999, p. 148, Lemma 24.1] and as such contains at least two points. It follows that, on the stipulated event of positive probability, which is contained in $\{T_{a/2} < \infty\}$ and on which $T_{a/2} = T$, $X(T_{a/2}) = X(T)$ is not P-trivial, a contradiction.

Consider now Case 2. Suppose furthermore that the support of $\lambda|_{\mathcal{B}((0,\infty))}$ were to contain at least two points $b < c$, say. Choose $\delta < b/2$ small enough such that $B(b, \delta) \cap B(c, \delta) = \emptyset$. The measure λ must charge both these open balls, and hence the first jump can be in either one, each with a positive probability. Thus $X(T_{b/2})$ would not be P-trivial on the event $\{T_{b/2} < \infty\}$, a contradiction. Plainly, then, the support of $\lambda|_{\mathcal{B}((0,\infty))}$ is $\{h\}$ for some $h > 0$.

It only remains to show that λ is supported by $\mathbb{Z}_h$. To see this, suppose it were not. Then there would be an $x < 0$ and a $\delta > 0$, with $B(x, \delta)$ having a non-empty intersection with the support of λ and an empty intersection with $\mathbb{Z}_h$. With a positive probability X would jump for the first time into $B(x, \delta)$ and then have a sequence of jumps of size h upwards going above h for the first time at a level distinct from h. With a positive probability, X also goes above h by making its first jump to h, a contradiction.

The proof is complete. $\square$

The second (and last) main step towards the proof of Theorem 3.3 consists in taking advantage of the temporal and spatial homogeneity of Lévy processes. Thus the condition in Proposition 3.9 is relaxed to one in which the P-triviality of the position at first passage is required for one $x > 0$, rather than all. To shorten notation let us introduce:

Definition 3.11. For $x \in \mathbb{R}$, let $\mathsf{Q}^x : \mathcal{B}(\mathbb{R}) \to [0, \mathsf{P}(T_x < \infty)]$,

$$\mathsf{Q}^x(B) := \mathsf{P}(\{X(T_x) \in B\} \cap \{T_x < \infty\}), \quad B \in \mathcal{B}(\mathbb{R}),$$

be the (possibly subprobability) law of $X(T_x)$ on $\{T_x < \infty\}$ under P on the space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. We also introduce the set:

$$\mathcal{A} := \{x \in \mathbb{R} : \mathsf{Q}^x, \text{ which may have zero mass, is carried by a singleton}\}.$$

Remark 3.12. Clearly $(-\infty, 0] \subset \mathcal{A}$ and for each $a \in \mathcal{A}$, there exists an (unique, if $\mathsf{P}(T_a < \infty) > 0$) $f(a)$ such that:

$$\mathsf{Q}^a = \mathsf{P}(T_a < \infty)\delta_{f(a)}.$$

With this at our disposal, we can formulate our claim as:

Proposition 3.13. *Suppose $\mathcal{A} \cap \mathbb{R}^+ \neq \emptyset$. Then $\mathcal{A} = \mathbb{R}$.*

The proof of Proposition 3.13 will proceed in several steps, but the essence of it consists in establishing the intuitively appealing identity

$$\mathsf{Q}^b(A) = \int d\overline{\mathsf{Q}^c}(x_c)\mathsf{Q}^{b-x_c}(A - x_c)$$

for Borel sets A and $c \in (0, b)$, see Lemma 3.14 below. This identity puts a constraint on the family of measures $(\mathsf{Q}^a)_{a \in \mathbb{R}}$. In particular, it allows to demonstrate (under the hypothesis $\mathcal{A} \cap \mathbb{R}^+ \neq \emptyset$) that $\mathcal{A}$ is dense in the reals. Then we can appeal to quasi left-continuity to conclude the proof. The main argument is thus fairly short, and a substantial amount of time is spent on measurability issues.

Lemma 3.14. *Let $b \in \mathbb{R}^+$, $c \in (0, b)$ and $A \in \mathcal{B}(\mathbb{R})$. Then:*

$$\mathsf{Q}^b(A) = \int d\overline{\mathsf{Q}^c}(x_c)\mathsf{Q}^{b-x_c}(A - x_c). \tag{3.4}$$

Proof. If $\mathsf{P}(T_c < \infty) = 0$, then $\mathsf{P}(T_b < \infty) = 0$, $\mathsf{Q}^b = \mathsf{Q}^c = 0$, and the claim is trivial. So assume, without loss of generality, that $\mathsf{P}(T_c < \infty) > 0$ and that X is càdlàg with certainty (rather than just P-a.s.).

Let (on $\{T_c < \infty\}$): $\overset{\triangle}{X} := (X(T_c+t) - X(T_c))_{t \geq 0}$ and $\overset{\triangle}{T}_y := \inf\{t \geq 0 : \overset{\triangle}{X}_t \geq y\}$ ($y \in \mathbb{R}$), while $\mathcal{F}'_{T_c} := \{B \cap \{T_c < \infty\} : B \in \mathcal{F}_{T_c}\}$ is $\mathcal{F}_{T_c}$ lowered onto $\{T_c < \infty\}$. By the strong Markov property, $\overset{\triangle}{X}$ is independent of $\mathcal{F}'_{T_c}$ under $\mathsf{P}(\cdot|\{T_c < \infty\})$. Then:

$$\begin{aligned}
\mathsf{Q}^b(A) &= \mathsf{E}[\mathbb{1}_A \circ X(T_b)\mathbb{1}_{\{T_b<\infty\}}], \text{ by the definition of } \mathsf{Q}^b, \\
&= \mathsf{E}[\mathbb{1}_{\{\overset{\triangle}{X}(\overset{\triangle}{T}_{b-X(T_c)})+X(T_c)\in A\}}\mathbb{1}_{\{\overset{\triangle}{T}_{b-X(T_c)}<\infty\}}\mathbb{1}_{\{T_c<\infty\}}], \text{ since } T_b = T_c + \overset{\triangle}{T}_{b-X(T_c)}, \\
&= \mathsf{P}(T_c < \infty) \times \mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}\left[\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}\left[\mathbb{1}_{\{\overset{\triangle}{X}(\overset{\triangle}{T}_{b-X(T_c)})+X(T_c)\in A\}}\mathbb{1}_{\{\overset{\triangle}{T}_{b-X(T_c)}<\infty\}}|\mathcal{F}'_{T_c}\right]\right], \\
&\quad \text{by the tower property and the definition of the conditional measure } \mathsf{P}(\cdot|\{T_c < \infty\}), \\
&= \int d\overline{\mathsf{Q}^c}(x_c)\mathsf{Q}^{b-x_c}(A - x_c), \\
&\quad \text{by the strong Markov property \& Proposition A.2 (see below).}
\end{aligned}$$

We now specify precisely how the strong Markov property and Proposition A.2 from Appendix A are applied here, this not being completely trivial. Recall again from the end of Subsection 3.1.2 the notation pertaining to the space $(\mathbb{D}, \mathcal{H})$.

The probability space we will be working on is $(\{T_c < \infty\}, \mathcal{F}_{\{T_c<\infty\}}, \mathsf{P}(\cdot|\{T_c < \infty\}))$, where $\mathcal{F}_{\{T_c<\infty\}} := \{B \cap \{T_c < \infty\} : B \in \mathcal{F}\}$, and it is complete, since $(\Omega, \mathcal{F}, \mathsf{P})$ is. Further, define (on $\{T_c < \infty\}$) $Y := X(T_c)$; $Z := \overset{\triangle}{X}$ and $f : \mathbb{R} \times \mathbb{D} \to \mathbb{R}$ by:

$$f(x, \omega) = \mathbb{1}_A(x + \omega(T_{b-x}(\omega)))\mathbb{1}_{[0,\infty)}(T_{b-x}(\omega)), \quad (x, \omega) \in \mathbb{R} \times \mathbb{D},$$

where we let $\omega(\infty) = \omega(0)$ for definiteness.[1]

Now, the random element $Z : (\{T_c < \infty\}, \mathcal{F}_{\{T_c<\infty\}}) \to (\mathbb{D}, \mathcal{H})$ is independent of $\mathcal{G} := \mathcal{F}'_{T_c}$, whereas the random element $Y : (\{T_c < \infty\}, \mathcal{F}_{\{T_c<\infty\}}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is

[1]The reader is cautioned not to confuse the mapping f, which is introduced here solely for the purposes of establishing how Proposition A.2 is applied in obtaining (3.4), with the notation from Remark 3.12. Indeed, the context will always make it clear which f we are referring to.

measurable with respect to $\mathcal{F}'_{T_c}$. Measurability of Y is a consequence of [Karatzas and Shreve, 1988, p. 5, Proposition 1.13 & p. 9, Proposition 2.18] and the Début Theorem [Kallenberg, 1997, p. 101, Theorem 6.7] and measurability of Z follows similarly.

We next show that f is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{H})^\star/\mathcal{B}(\mathbb{R})$-measurable. First note that:

1. $(x, \omega) \mapsto \omega + x$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{H}/\mathcal{H}$-measurable (in fact continuous, see [Jacod and Shiryaev, 2003, p. 328, Proposition 1.17 & p. 329, Proposition 1.23]), hence $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{H})^\star/\mathcal{H}^\star$-measurable, by [Meyer, 1966, (2) on p. 23].

2. By the Début Theorem, for every $b \in \mathbb{R}$, T_b is a stopping time of the augmented (with respect to *any* probability measure) right-continuous modification of the canonical filtration $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ on $\mathbb{D}$ /where $\mathcal{H}_t$ is generated by the evaluation maps up to, and including, time t, $t \geq 0$/. Hence $(\omega \mapsto T_b(\omega))$ is $\mathcal{H}^\star/\mathcal{B}([0, \infty])$-measurable.

It follows that $(x, \omega) \mapsto T_b(\omega + x) = T_{b-x}(\omega)$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{H})^\star/\mathcal{B}([0, \infty])$-measurable (as a composition). Next:

1. $(x, \omega) \mapsto (\omega, \mathbb{1}_{[0,\infty)}(T_{b-x}(\omega))T_{b-x}(\omega))$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{H})^\star/\mathcal{H} \otimes \mathcal{B}(\mathbb{R}_+)$-measurable.

2. $(\omega, t) \mapsto \omega(t)$ is $\mathcal{H} \otimes \mathcal{B}(\mathbb{R}_+)/\mathcal{B}(\mathbb{R})$-measurable (indeed, if X is the coordinate process on $\mathbb{D}$, then this is the mapping $(\omega, t) \mapsto X(\omega, t)$, which is measurable by [Karatzas and Shreve, 1988, p. 5, Remark 1.14]).

Therefore $(x, \omega) \mapsto \omega(T_{b-x}(\omega))$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{H})^\star/\mathcal{B}(\mathbb{R})$-measurable (as a composition, with the above convention for $\omega(\infty)$). The required measurability of f now follows from measurability of addition and multiplication.

We are now in a position to apply Proposition A.2. We have:

$$
\begin{aligned}
& \mathsf{P}(T_c < \infty)\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[f \circ (Y, Z)|\mathcal{F}'_{T_c}]] = \\
= \; & \mathsf{P}(T_c < \infty)\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[(y \mapsto \mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[f \circ (y, Z)]) \circ X(T_c)], \text{ by Proposition A.2,} \\
= \; & \int d\overline{\mathsf{Q}^c}(y)\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[f \circ (y, Z)], \text{ by the Image Measure Theorem} \\
& \text{[Dudley, 2004, p. 121, Theorem 4.1.11], since } \overline{\mathsf{Q}^c} \text{ coincides with the (subprobability)} \\
& \text{law of } X(T_c) \text{ on } \left(\mathbb{R}, \overline{\mathcal{B}(\mathbb{R})}^{\mathsf{Q}^c}\right).
\end{aligned}
$$

Note here that we need to work with the (subprobability) law of $X(T_c)$ on the space $(\mathbb{R}, \overline{\mathcal{B}(\mathbb{R})}^{\mathsf{Q}^c})$ /rather than $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$/, since we only know the integrand to be measurable with respect to $\overline{\mathcal{B}(\mathbb{R})}^{\mathsf{Q}^c}$.

Now, by the strong Markov property, Z is also identical in law under the measure $\mathsf{P}(\cdot|\{T_c < \infty\})$ to X under the measure P on the space $(\mathbb{D}, \mathcal{H})$ and hence on the space $(\mathbb{D}, \mathcal{H}^\star)$ /the extension of a law to the universal completion being unique [Meyer, 1966, (1) on p. 23]/. Moreover, for any real d and Borel set $D \subset \mathbb{R}$, the mapping $g_{d,D} : \mathbb{D} \to \mathbb{R}$ given by $(\omega \mapsto \mathbb{1}_D(\omega(T_d(\omega)))\mathbb{1}_{[0,\infty)}(T_d(\omega)))$ is $\mathcal{H}^\star/\mathcal{B}(\mathbb{R})$-measurable, by the same reasoning as above. Hence:

$$\begin{aligned}
\mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[f \circ (y, Z)] &= \mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[\mathbb{1}_{A-y} \circ \overset{\triangle}{X}(\overset{\triangle}{T}_{b-y})\mathbb{1}_{[0,\infty)} \circ \overset{\triangle}{T}_{b-y}] \\
&= \mathsf{E}^{\mathsf{P}(\cdot|\{T_c<\infty\})}[g_{b-y,A-y} \circ Z] \\
&= \mathsf{E}^{\mathsf{P}}[g_{b-y,A-y} \circ X] = \mathsf{Q}^{b-y}(A - y),
\end{aligned}$$

as required. □

Proof of Proposition 3.13. Given $\mathcal{A} \cap \mathbb{R}^+ \neq \emptyset$, we wish to show the inclusion $\mathbb{R}^+ \subset \mathcal{A}$. Assume, again without loss of generality, that X is càdlàg with certainty (rather than just P-a.s.).

(i) First observe that $\mathsf{P}(T_x = \infty) = 1$ for some $x > 0$, precisely when $\mathsf{P}(T_x = \infty) = 1$ for all $x > 0$. This follows either by the strong Markov property of Lévy processes and mathematical induction or, alternatively, one can appeal directly to [Sato, 1999, p. 155, Proposition 24.14(i)]. Therefore it is sufficient to consider the case when $\mathsf{P}(T_x < \infty) > 0$ for all $x \in \mathbb{R}$.

(ii) Claim:

($\natural$) If $b \in \mathcal{A}$, then for every $c \in (0, b)$: either $c \in \mathcal{A}$ or $(0, b - c] \cap \mathcal{A} \neq \emptyset$.

To show this, let $b \in \mathcal{A}$, $c \in (0, b)$ and take any $A \in \mathcal{B}(\mathbb{R})$. By Lemma 3.14:

$$\mathsf{Q}^b(A) = \int d\overline{\mathsf{Q}^c}(x_c)\mathsf{Q}^{b-x_c}(A - x_c). \tag{3.5}$$

On the other hand, since $b \in \mathcal{A}$:

$$\mathsf{Q}^b(A) = \mathsf{P}(T_b < \infty)\delta_{f(b)}(A). \tag{3.6}$$

Combining (3.5) and (3.6), we have:

$$\int d\overline{\mathsf{Q}^c}(x_c)\mathsf{Q}^{b-x_c}(A - x_c) = \mathsf{P}(T_b < \infty)\delta_{f(b)}(A),$$

from which we conclude that $\overline{\mathsf{Q}^c}$-a.e. in $x_c \in \mathbb{R}$, Q^{b-x_c} assigns all its mass to $\{f(b) - x_c\}$. (Suppose not, then with $\overline{\mathsf{Q}^c}$-positive measure in $x_c \in \mathbb{R}$,

$\mathsf{Q}^{b-x_c}(\mathbb{R}\backslash\{f(b)-x_c\})>0$, and hence $Q^b(\mathbb{R}\backslash\{f(b)\})>0$, a contradiction.)

Next, if $b'\in\mathcal{A}$ and $c'\in(0,b']$:

(*) $\mathsf{Q}^{c'}$ assigns all its mass to $[c',b')\cup\{f(b')\}$.

Therefore $c\in\mathcal{A}$, or Q^c cannot ascribe all its mass to $\{f(b)\}$ and hence $\mathsf{Q}^c([c,b))>0$. In the latter case, for some $x_c\in[c,b)$, Q^{b-x_c} is carried by $\{f(b)-x_c\}$, whence $b-x_c\in\mathcal{A}\cap(0,b-c]$.

(iii) Let $x_0:=\inf\mathcal{A}\cap\mathbb{R}^+$. Then $x_0=0$. Indeed, if not, then ($\natural$) of (ii), applied to some $[x_0,\infty)\cap\mathcal{A}\ni b<3x_0/2$ and $c=3x_0/4$ (say), yields a contradiction. Therefore there exists a decreasing sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathcal{A}\cap\mathbb{R}^+$ converging to 0.

(iv) Claim: $\mathcal{A}$ is dense in $\mathbb{R}$. If $f(x_n)\to 0$ as $n\to\infty$, this is obvious, since,

(**) with any $x\in\mathcal{A}$, $\cup_{n\in\mathbb{N}_0}[x+nf(x),(n+1)f(x)]\subset\mathcal{A}$,

by the strong Markov property and mathematical induction. Suppose the nonincreasing sequence $(f(x_n))_{n\in\mathbb{N}_0}$ does not converge to 0. Then there is an $\epsilon>0$ and a natural N, such that $f(x_n)\geq\epsilon$ and $x_n<\epsilon$ for all $n\geq N$. In particular, by (*), $f(x_n)=f(x_N)$ for all $n\geq N$. Therefore $[x_n,f(x_N)]\subset\mathcal{A}$ for all $n\geq N$ by (**). Therefore $[0,f(x_N)]\subset\mathcal{A}$ and upon exceeding any positive level less than or equal to $f(x_N)$ we land at $f(x_N)$ a.s. Hence, by the strong Markov property and mathematical induction, $\mathcal{A}=\mathbb{R}$.

(v) So we may assume $\mathcal{A}$ is dense. Now we use quasi left-continuity of Lévy processes [Bertoin, 1996, p. 21, Proposition 7] as follows. Take any $x\in\mathbb{R}^+$ and a sequence $\mathcal{A}\cap(0,x)\supset(x_n)_{n\geq 1}\uparrow x$. Introduce the $\mathbb{F}$-stopping time $S:=\inf\{t\geq 0:\overline{X}_t\geq x\}$. We then have $T_{x_n}\uparrow S$ (as $n\to\infty$). By quasi left-continuity, it follows that $\lim_{n\to\infty}X(T_{x_n})=X(S)$ P-a.s. on $\{S<\infty\}$. Therefore, in fact, $S=T_x$ P-a.s. on $\{S<\infty\}$ (and hence on $\{T_x<\infty\}$), and, moreover, $X(T_x)=\lim_{n\to\infty}f(x_n)$ P-a.s. on $\{T_x<\infty\}$. But this means, precisely, that $x\in\mathcal{A}$.

The proof is complete. $\square$

Finally we can combine the above into a proof of Theorem 3.3.

Proof of Theorem 3.3. The statement is essentially contained in Propositions 3.9 and 3.13. We only have to worry about (c) and (d), since so far we have only considered the stopping times T_x.

Now, (c) implies for some $f(x) > 0$, $X(\hat{T}_x) = f(x)$ P-a.s. on $\{\hat{T}_x < \infty\}$, therefore $X(T_{f(x)}) = f(x)$ P-a.s. on $\{T_{f(x)} < \infty\}$ and hence (a). Conversely, (e) implies (d) by sample path right-continuity. $\square$

Remark 3.15. Theorem 3.3 characterizes the class of Lévy processes for which overshoots are known *a priori* and are non-random. Moreover, the original motivation for this investigation is validated by the fact that upwards skip-free Lévy chains admit a fluctuation theory, which is just as explicit, almost (but not entirely) analogous to the spectrally negative case and which embeds (existing) results for right-continuous random walks into continuous time. We expound on this in the next section.

3.2 Fluctuation theory for upwards skip-free Lévy chains

3.2.1 Introduction

We have seen in Section 3.1 that precisely two types of Lévy processes exhibit the property of non-random overshoots: those with no positive jumps a.s., and upwards skip-free Lévy chains (see Definition 3.1). We have also remarked that this common property which the two classes share results in a more explicit fluctuation theory (including the Wiener-Hopf factorization) than for a general Lévy process, this being rarely the case (cf. [Kyprianou, 2006, p. 172, Subsection 6.5.4]).

Now, with reference to existing literature on fluctuation theory, the spectrally negative case (when there are no positive jumps, a.s.) is dealt with in detail in [Bertoin, 1996, Chapter VII] [Sato, 1999, Section 9.46] and especially [Kyprianou, 2006, Chapter 8]. On the other hand no equally exhaustive treatment of the right-continuous random walk seems to have been presented thus far, but see [Quine, 2004; Brown et al., 2010; Marchal, 2001] [Doney and Picard, 2007, p. 99, Section 9.3] [Spitzer, 2001, *passim*]. In particular, no such exposition appears present for the continuous-time analogue of such random walks, wherein the connection and analogy to the spectrally negative class of Lévy processes becomes most transparent and direct.

In the present section we proceed to do just that, i.e. we develop, by analogy to the spectrally negative case, a complete fluctuation theory (including theory of scale functions) for upwards skip-free Lévy chains. Indeed, the transposition of the results from the spectrally negative to the skip-free setting is essentially straightforward. Over and above this, however, and beyond what is purely analogous to the exposition of the spectrally negative case, further specifics of the reflected process (see Theorem 3.16(i)) and of the excursions from the supremum (see Theorem 3.16(iii)) are identified, and a linear recursion is presented which allows us to directly compute the families of scale functions (see (3.25), (3.26) and Proposition 3.35).

The organisation of the rest of this section is as follows. Subsection 3.2.2 specifies the setting. Then Subsection 3.2.3 develops the relevant fluctuation theory, in particular details of the Wiener-Hopf factorization. Finally, Subsection 3.2.4 deals with the two-sided exit problem and the accompanying families of scale functions.

3.2.2 Setting

For the remainder of this section, X will be assumed throughout an upwards skip-free Lévy chain, with $\lambda(\{h\}) > 0$ ($h > 0$) and characteristic exponent $\Psi(p) = \int(e^{ipx} - 1)\lambda(dx)$ ($p \in \mathbb{R}$). In general, we insist on every sample path of X being càdlàg (i.e. right-continuous, admitting left limits). We shall, however, sometimes and then only provisionally, relax the assumption on the filtered probability space satisfying the standard assumptions, by transferring X as the coordinate process onto the canonical space $\mathbb{D}_h := \{\omega \in \mathbb{Z}_h^{[0,\infty)} : \omega \text{ is càdlàg}\}$ of càdlàg paths, mapping $[0, \infty) \to \mathbb{Z}_h$, equipping $\mathbb{D}_h$ with the σ-algebra and natural filtration of evaluation maps; this, however, will always be made explicit. We allow $X \perp e_1 \sim \mathrm{Exp}(1)$; then define $e_p := e_1/p$ ($p \in (0, \infty)\backslash\{1\}$).

3.2.3 Fluctuation theory

In the following, to fully appreciate the similarity (and eventual differences) with the spectrally negative case, the reader is invited to directly compare the exposition of this subsection with that of [Bertoin, 1996, Section VII.1] and [Kyprianou, 2006, Section 8.1].

Laplace exponent, the reflected process, local times and excursions from the supremum, supremum process and long-term behaviour, exponential change of measure

Since the Poisson process admits exponential moments of all orders, it follows that $\mathsf{E}[e^{\beta\overline{X}_t}] < \infty$ and, in particular, $\mathsf{E}[e^{\beta X_t}] < \infty$ for all $\{\beta, t\} \subset [0, \infty)$. Indeed, it may be seen by a direct computation that for $\beta \in \overline{\mathbb{C}^{\rightarrow}}$, $t \geq 0$, $\mathsf{E}[e^{\beta X_t}] = \exp\{t\psi(\beta)\}$, where $\psi(\beta) := \int_{\mathbb{R}}(e^{\beta x} - 1)\lambda(dx)$ is the Laplace exponent of X. Moreover, ψ is continuous (by the DCT) on $\overline{\mathbb{C}^{\rightarrow}}$ and analytic in $\mathbb{C}^{\rightarrow}$ (use the theorems of Cauchy [Rudin, 1970, p. 206, 10.13 Cauchy's theorem for triangle], Morera [Rudin, 1970, p. 209, 10.17 Morera's theorem] and Fubini).

Next, note that $\psi(\beta)$ tends to $+\infty$ as $\beta \to \infty$ over the reals, due to the presence of the atom of λ at h. Upon restriction to $[0, \infty)$, ψ is strictly convex, as follows first on $(0, \infty)$ by using differentiation under the integral sign and noting that the second derivative is strictly positive, and then extends to $[0, \infty)$ by continuity.

Denote then by $\Phi(0)$ the largest root of $\psi|_{[0,\infty)}$. Indeed, 0 is always a root, and due to strict convexity, if $\Phi(0) > 0$, then 0 and $\Phi(0)$ are the only two roots. The two cases occur, according as to whether $\psi'(0+) \geq 0$ or $\psi'(0+) < 0$, which is clear. It is less obvious, but nevertheless true, that this right derivative at 0 actually exists, indeed $\psi'(0+) = \int_{\mathbb{R}} x\lambda(dx) \in [-\infty, \infty)$. This follows from the fact that $(e^{\beta x} - 1)/\beta$ is nonincreasing as $\beta \downarrow 0$ for $x \in \mathbb{R}_-$ and hence monotone convergence applies. Continuing from this, and with a similar justification, one also gets the equality $\psi''(0+) = \int x^2\lambda(dx) \in (0, +\infty]$ (where we agree $\psi''(0+) = +\infty$ if $\psi'(0+) = -\infty$). In any case, $\psi : [\Phi(0), \infty) \to [0, \infty)$ is continuous and increasing, it is a bijection and we let $\Phi : [0, \infty) \to [\Phi(0), \infty)$ be the inverse bijection, so that $\psi \circ \Phi = \mathrm{id}_{\mathbb{R}_+}$.

With these preliminaries having been established, our first theorem identifies characteristics of the reflected process, the local time of X at the maximum, as well as the expected length of excursions and the probability of an infinite excursion therefrom (for definitions of these terms see Subsection 1.2.2 and e.g. [Kyprianou, 2006, pp. 140-147]; we agree that an excursion (from the maximum) starts immediately X leaves its running maximum and ends immediately it returns to it; by its length we mean the amount of time between these two time points).

Theorem 3.16 (Reflected process; (inverse) local time; excursions)**.**

(i) *The generator matrix $\tilde{Q}$ of the Markov process $Y := \overline{X} - X$ on $\mathbb{Z}_h^+$ is given by (with $\{s, s'\} \subset \mathbb{Z}_h^+$): $\tilde{Q}_{ss'} = \lambda(\{s - s'\}) - \delta_{ss'}\lambda(\mathbb{R})$, unless $s = s' = 0$, in which case we have $\tilde{Q}_{ss'} = -\lambda((-\infty, 0))$.*

(ii) *For the reflected process Y, 0 is a holding point. The actual time spent at 0 by Y is a local time at the maximum. Its inverse S is then a (possibly killed) compound Poisson subordinator with unit positive drift.*

(iii) *Assuming that $\lambda((-\infty, 0)) > 0$ to avoid the trivial case, the expected length of an excursion away from the supremum is equal to $\frac{\lambda(\{h\})h - \psi'(0+)}{(\psi'(0+) \vee 0)\lambda((-\infty,0))}$; whereas the probability of such an excursion being infinite is $\frac{\lambda(\{h\})}{\lambda((-\infty,0))}(e^{\Phi(0)h} - 1)$.*

Proof. (i) is clear, since, e.g. Y transitions away from 0 at the rate at which X makes a negative jump, and from $s \in \mathbb{Z}_h^+ \backslash \{0\}$ to 0 at the rate at which X jumps up by s or more etc.; however see Appendix B for the technical details.

(ii) is standard [Kyprianou, 2006, p. 141, Example 6.3 & p. 149, Theorem 6.10].

Finally, we establish (iii). Denote $q_n := \lambda(\{-nh\})/\lambda(\mathbb{R})$ for $n \in \mathbb{N}$ and $p := \lambda(\{h\})/\lambda(\mathbb{R})$; β provisionally denoting the expected excursion length. Further, let the discrete-time Markov chain W (on the state space $\mathbb{N}_0$) be endowed with the initial distribution $w_j := \frac{q_j}{1-p}$ for $j \in \mathbb{N}$, $w_0 := 0$; and transition matrix P, given by $P_{0i} = \delta_{0i}$, whereas for $i \geq 1$: $P_{ij} = p$, if $j = i-1$; $P_{ij} = q_{j-i}$, if $j > i$; and $P_{ij} = 0$ otherwise (W jumps down with probability p, up i steps with probability q_i, $i \geq 1$, until it reaches 0, where it gets stuck). Let further N be the first hitting time for W of $\{0\}$, so that a typical excursion length of X is equal in distribution to an independent sum of N (possibly infinite) $\mathrm{Exp}(\lambda(\mathbb{R}))$-random variables. It is Wald's identity that $\beta = (1/\lambda(\mathbb{R}))\mathsf{E}[N]$. Then (in the obvious notation, where $\underline{\infty}$ indicates the sum is inclusive of ∞), by Fubini: $\mathsf{E}[N] = \sum_{n=1}^{\underline{\infty}} n \sum_{l=1}^{\infty} w_l \mathsf{P}_l(N = n) = \sum_{l=1}^{\infty} w_l k_l$, where k_l is the mean hitting time of $\{0\}$ for W, if it starts from $l \in \mathbb{N}_0$, as in [Norris, 1997, p. 12]. From the skip-free property of the chain W it is moreover transparent that $k_i = \alpha i$, $i \in \mathbb{N}_0$, for some $0 < \alpha \leq \infty$ (with the usual convention $0 \cdot \infty = 0$). Moreover we know [Norris, 1997, p. 17, Theorem 1.3.5] that $(k_i : i \in \mathbb{N}_0)$ is the minimal solution to $k_0 = 0$ and $k_i = 1 + \sum_{j=1}^{\infty} P_{ij} k_j$ $(i \in \mathbb{N})$. Plugging in $k_i = \alpha i$, the last system of linear equations is equivalent to (provided $\alpha < \infty$) $0 = 1 - p\alpha + \alpha\zeta$, where $\zeta := \sum_{j=1}^{\infty} j q_j$. Thus, if $\zeta < p$, the minimal solution to the system is $k_i = i/(p-\zeta)$, $i \in \mathbb{N}_0$, from which $\beta = \zeta/(\lambda((-\infty,0))(p-\zeta))$ follows at once. If $\zeta \geq p$, clearly we must have $\alpha = +\infty$, hence $\mathsf{E}[N] = +\infty$ and thus $\beta = +\infty$.

To establish the probability of an excursion being infinite, i.e. $\sum_{i=1}^{\infty} q_i(1-\alpha_i)/\sum_{i=1}^{\infty} q_i$, where $\alpha_i := \mathsf{P}_i(N < \infty) > 0$, we see that (by the skip-free property) $\alpha_i = \alpha_1^i$, $i \in \mathbb{N}_0$, and by the strong Markov property, for $i \in \mathbb{N}$, $\alpha_i = p\alpha_{i-1} + \sum_{j=1}^{\infty} q_j \alpha_{i+j}$. It follows that $1 = p\alpha_1^{-1} + \sum_{j=1}^{\infty} q_j \alpha_1^j$, i.e. $0 = \psi(\log(\alpha_1^{-1})/h)$. Hence, by Theorem 3.17(ii), whose proof will be independent of this one, $\alpha_1 = e^{-\Phi(0)h}$ (since $\alpha_1 < 1$, if and only if X drifts to $-\infty$). $\square$

We turn our attention now to the supremum process $\overline{X}$. First, using the lack of memory property of the exponential law and the skip-free nature of X, we deduce from the strong Markov property applied at the time T_a, that for every $a, b \in \mathbb{Z}_h^+$, $p > 0$: $\mathsf{P}(T_{a+b} < e_p) = \mathsf{P}(T_a < e_p)\mathsf{P}(T_b < e_p)$. In particular, for any $n \in \mathbb{N}_0$: $\mathsf{P}(T_{nh} < e_p) = \mathsf{P}(T_h < e_p)^n$. And since for $s \in \mathbb{Z}_h^+$, $\{T_s < e_p\} = \{\overline{X}_{e_p} \geq s\}$ (P-a.s.), one has (for $n \in \mathbb{N}_0$): $\mathsf{P}(\overline{X}_{e_p} \geq nh) = \mathsf{P}(\overline{X}_{e_p} \geq h)^n$. Therefore $\overline{X}_{e_p}/h \sim \mathrm{geom}(1 - \mathsf{P}(\overline{X}_{e_p} \geq h))$.

Next, to identify $\mathsf{P}(\overline{X}_{e_p} \geq h)$, $p > 0$, observe that (for $\beta \geq 0$, $t \geq 0$): $\mathsf{E}[\exp\{\Phi(\beta)X_t\}] = e^{t\beta}$ and hence $(\exp\{\Phi(\beta)X_t - \beta t\})_{t\geq 0}$ is an $(\mathbb{F}, \mathsf{P})$-martingale by stationary independent increments of X, for each $\beta \geq 0$. Then apply the Optional Sampling Theorem at the bounded stopping time $T_x \wedge t$ $(t, x \geq 0)$ to get:

$$\mathsf{E}[\exp\{\Phi(\beta)X(T_x \wedge t) - \beta(T_x \wedge t)\}] = 1.$$

Note that $X(T_x \wedge t) \leq h\lceil x/h \rceil$ and $\Phi(\beta)X(T_x \wedge t) - \beta(T_x \wedge t)$ converges to $\Phi(\beta)h\lceil x/h\rceil - \beta T_x$ (P-a.s.) as $t \to \infty$ on $\{T_x < \infty\}$. It converges to $-\infty$ on the complement of this event, P-a.s., provided $\beta + \Phi(\beta) > 0$. Therefore we deduce by dominated convergence, first for $\beta > 0$ and then also for $\beta = 0$, by taking limits:

$$\mathsf{E}[\exp\{-\beta T_x\}\mathbb{1}_{\{T_x<\infty\}}] = \exp\{-\Phi(\beta)h\lceil x/h\rceil\}. \tag{3.7}$$

Before we formulate out next theorem, recall also that any non-zero Lévy process either drifts to $+\infty$, oscillates or drifts to $-\infty$ [Sato, 1999, pp. 255-256, Proposition 37.10 and Definition 37.11].

Theorem 3.17 (Supremum process and long-term behaviour)**.**

(i) The failure probability for the geometrically distributed $\overline{X}_{e_p}/h$ *is* $\exp\{-\Phi(p)h\}$ *(*$p > 0$*).*

(ii) X *drifts to* $+\infty$*, oscillates or drifts to* $-\infty$ *according as to whether* $\psi'(0+)$ *is positive, zero, or negative. In the latter case* $\overline{X}_\infty/h$ *has a geometric distribution with failure probability* $\exp\{-\Phi(0)h\}$*.*

(iii) $(T_{nh})_{n\in\mathbb{N}_0}$ *is a discrete-time increasing stochastic process, vanishing at* 0 *and having stationary independent increments up to the explosion time, which is an independent geometric random variable; it is a killed random walk.*

Remark 3.18. Unlike in the spectrally negative case [Bertoin, 1996, p. 189], the supremum process cannot be obtained from the reflected process, since the latter does not discern a point of increase in X when the latter is at its running maximum.

Proof. We have for every $s \in \mathbb{Z}_h^+$:

$$\mathsf{P}(\overline{X}_{e_p} \geq s) = \mathsf{P}(T_s < e_p) = \mathsf{E}[\exp\{-pT_s\}\mathbb{1}_{\{T_s<\infty\}}] = \exp\{-\Phi(p)s\}. \tag{3.8}$$

Thus (i) obtains.

For (ii) note that letting $p \downarrow 0$ in (3.8), we obtain $\overline{X}_\infty < \infty$ (P-a.s.), if and only if $\Phi(0) > 0$, which is equivalent to $\psi'(0+) < 0$. If so, $\overline{X}_\infty/h$ is geometrically distributed with failure probability $\exp\{-\Phi(0)h\}$ and then (and only then) does X drift to $-\infty$.

It remains to consider drifting to $+\infty$ (the cases being mutually exclusive and exhaustive). Indeed, X drifts to $+\infty$, if and only if $\mathsf{E}[T_s]$ is finite for each $s \in \mathbb{Z}_h^+$ [Bertoin, 1996, p. 172, Proposition VI.17]. Using again the nondecreasingness of $(e^{-\beta T_s} - 1)/\beta$ in $\beta \in [0, \infty)$, we deduce from (3.7), by monotone convergence, that one may differentiate under the integral sign, to get $\mathsf{E}[T_s \mathbb{1}_{\{T_s<\infty\}}] = (\beta \mapsto -\exp\{-\Phi(\beta)s\})'(0+)$. So the $\mathsf{E}[T_s]$ are finite, if and only if $\Phi(0) = 0$ (so that $T_s < \infty$ P-a.s.) and $\Phi'(0+) < \infty$. Since Φ is the inverse of $\psi|_{[\Phi(0),\infty)}$, this is equivalent to saying $\psi'(0+) > 0$.

Finally, (iii) is clear. $\square$

$\psi'(0+)$	$\Phi(0)$	$\Phi'(0+)$	Long-term behaviour	Excursion length
$\in (0, \infty)$	0	$\in (0, \infty)$	drifts to $+\infty$	finite expectation
0	0	$+\infty$	oscillates	a.s. finite with infinite expectation
$\in [-\infty, 0)$	$\in (0, \infty)$	$\in (0, \infty)$	drifts to $-\infty$	infinite with a positive probability

Table 3.1: Connections between the quantities $\psi'(0+)$, $\Phi(0)$, $\Phi'(0+)$. Behaviour of X at large times and of its excursions away from the running supremum (the latter if $\lambda((-\infty, 0)) > 0$).

We conclude this paragraph by offering a way to reduce the general case of an upwards skip-free Lévy chain to one which necessarily drifts to $+\infty$. This will prove useful in the sequel (more specifically, in the proof of Theorem 3.24). First, there is a pathwise approximation of an oscillating X, by (what is again) an upwards skip-free Lévy chain, but drifting to infinity.

Remark 3.19. Suppose X oscillates. Let (possibly by enlarging the probability space to accommodate for it) N be an independent Poisson process with intensity 1 and $N_t^\epsilon := N_{t\epsilon}$ $(t \geq 0)$ so that N^ϵ is a Poisson process with intensity ϵ, independent of X. Define $X^\epsilon := X + hN^\epsilon$. Then, as $\epsilon \downarrow 0$, X^ϵ converges to X, uniformly on bounded time sets, almost surely, and is clearly an upwards skip-free Lévy chain drifting to $+\infty$.

The reduction of the case when X drifts to $-\infty$ is somewhat more involved and is done by a change of measure. For this purpose assume until the end of this paragraph (i.e. up to and inclusive of Proposition 3.21), that X is already the coordinate process on the canonical space $\Omega = \mathbb{D}_h$, equipped with the σ-algebra $\mathcal{F}$ and filtration $\mathbb{F}$ of evaluation maps (so that P *coincides* with the law of X on $\mathbb{D}_h$ and $\mathcal{F} = \sigma(\mathrm{pr}_s : s \in [0, +\infty))$, whilst for $t \geq 0$, $\mathcal{F}_t = \sigma(\mathrm{pr}_s : s \in [0, t])$, where $\mathrm{pr}_s(\omega) = \omega(s)$, for $(s, \omega) \in [0, +\infty) \times \mathbb{D}_h$). We make this transition in order to be able to apply the Kolmogorov extension theorem in the proposition, which follows. Note, however, that we are no longer able to assume standard conditions on $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$. Notwithstanding this, $(T_x)_{x\in\mathbb{R}}$ remain $\mathbb{F}$-stopping times, since by the nature of the space $\mathbb{D}_h$, for $x \in \mathbb{R}$, $t \geq 0$, $\{T_x \leq t\} = \{\overline{X}_t \geq x\} \in \mathcal{F}_t$.

Proposition 3.20 (Exponential change of measure)**.** *Let $c \geq 0$. Then, demanding:*

$$\mathsf{P}_c(\Lambda) = \mathsf{E}[\exp\{cX_t - \psi(c)t\}\mathbb{1}_\Lambda] \quad (\Lambda \in \mathcal{F}_t, t \geq 0) \tag{3.9}$$

this introduces a unique measure P_c on $\mathcal{F}$. Under the new measure, X remains an upwards skip-free Lévy chain with Laplace exponent $\psi_c = \psi(\cdot + c) - \psi(c)$, drifting to $+\infty$, if $c \geq \Phi(0)$, unless $c = \psi'(0+) = 0$. Moreover, if λ_c is the new Lévy measure of X under P_c, then $\lambda_c \ll \lambda$ and $\frac{d\lambda_c}{d\lambda}(x) = e^{cx}$ λ-a.e. in $x \in \mathbb{R}$. Finally, for every $\mathbb{F}$-stopping time T, $\mathsf{P}_c \ll \mathsf{P}$ on restriction to $\mathcal{F}'_T := \{A \cap \{T < \infty\} : A \in \mathcal{F}_T\}$, and:

$$\frac{d\mathsf{P}_c|_{\mathcal{F}'_T}}{d\mathsf{P}|_{\mathcal{F}'_T}} = \exp\{cX_T - \psi(c)T\}.$$

Proof. That P_c is introduced consistently as a probability measure on $\mathcal{F}$ follows from the Kolmogorov extension theorem [Parthasarathy, 1967, p. 143, Theorem 4.2] (see Appendix C for details). Indeed, $M := (\exp\{cX_t - \psi(c)t\})_{t\geq 0}$ is a nonnegative martingale (use independence and stationarity of increments of X and the definition of the Laplace exponent), equal identically to 1 at time 0.[2]

Further, for all $\beta \in \overline{\mathbb{C}^{\rightarrow}}$, $\{t, s\} \subset \mathbb{R}_+$ and $\Lambda \in \mathcal{F}_t$:

$$\begin{aligned}
\mathsf{E}_c[\exp\{\beta(X_{t+s} - X_t)\}\mathbb{1}_\Lambda] &= \mathsf{E}[\exp\{cX_{t+s} - \psi(c)(t+s)\}\exp\{\beta(X_{t+s} - X_t)\}\mathbb{1}_\Lambda] \\
&= \mathsf{E}[\exp\{(c+\beta)(X_{t+s} - X_t) - \psi(c)s\}\mathsf{E}[\exp\{cX_t - \psi(c)t\}\mathbb{1}_\Lambda] \\
&= \exp\{s(\psi(c+\beta) - \psi(c))\}\mathsf{P}_c(\Lambda).
\end{aligned}$$

An application of the Functional Monotone Class Theorem then shows that X is indeed a Lévy process on $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P}_c)$ and its Laplace exponent under P_c is as stipulated (that $X_0 = 0$ P_c-a.s. follows from the absolute continuity of P_c with respect to P on restriction to $\mathcal{F}_0$).

Next, from the expression for ψ_c, the claim regarding λ_c follows at once. Then clearly X remains an upwards skip-free Lévy chain under P_c, drifting to $+\infty$, if $\psi'(c+) > 0$.

Finally, let $A \in \mathcal{F}_T$ and $t \geq 0$. Then $A \cap \{T \leq t\} \in \mathcal{F}_{T\wedge t}$, and by the Optional Sampling Theorem:

$$\mathsf{P}_c(A \cap \{T \leq t\}) = \mathsf{E}[M_t \mathbb{1}_{A\cap\{T\leq t\}}] = \mathsf{E}[\mathsf{E}[M_t \mathbb{1}_{A\cap\{T\leq t\}}|\mathcal{F}_{T\wedge t}]] = \mathsf{E}[M_{T\wedge t}\mathbb{1}_{A\cap\{T\leq t\}}] = \mathsf{E}[M_T \mathbb{1}_{A\cap\{T\leq t\}}].$$

[2]Remark that M is, in general, *not* uniformly integrable. For example, if X drifts to $-\infty$ and $c > 0$, then $M_\infty := \lim_{t\to\infty} M_t = 0$ exists a.s. (see, e.g., [Karatzas and Shreve, 1988, Subsection 1.3.B]). In particular, then, a martingale change of measure cannot be applied directly (cf. Remark C.7).

Using the MCT, letting $t \to \infty$, we obtain the equality $\mathsf{P}_c(A \cap \{T < \infty\}) = \mathsf{E}[M_T \mathbb{1}_{A\cap\{T<\infty\}}]$. $\square$

Proposition 3.21 (Conditioning to drift to $+\infty$)**.** *Assume $\Phi(0) > 0$ and denote $\mathsf{P}^\natural := \mathsf{P}_{\Phi(0)}$ (see (3.9)). We then have as follows.*

(1) For every $\Lambda \in \mathcal{A} := \cup_{t\geq 0}\mathcal{F}_t$, $\lim_{n\to\infty} \mathsf{P}(\Lambda|\overline{X}_\infty \geq nh) = \mathsf{P}^\natural(\Lambda)$.

(2) For every $x \geq 0$, the stopped process $X^{T_x} = (X_{t\wedge T_x})_{t\geq 0}$ is identical in law under the measures $\mathsf{P}^\natural$ and $\mathsf{P}(\cdot|T_x < \infty)$ on the canonical space $\mathbb{D}_h$.

Proof. With regard to (1), we have as follows. Let $t \geq 0$. By the Markov property of X at time t, the process $\overset{\triangle}{X} := (X_{t+s} - X_t)_{s\geq 0}$ is identical in law with X on $\mathbb{D}_h$ and independent of $\mathcal{F}_t$ under P. Thus, letting $\overset{\triangle}{T}_y := \inf\{t \geq 0 : \overset{\triangle}{X}_t \geq y\}$ ($y \in \mathbb{R}$), one has for $\Lambda \in \mathcal{F}_t$ and $n \in \mathbb{N}_0$, by conditioning:

$$\mathsf{P}(\Lambda \cap \{t < T_{nh} < \infty\}) = \mathsf{E}[\mathsf{E}[\mathbb{1}_\Lambda \mathbb{1}_{\{t<T_{nh}\}} \mathbb{1}_{\{\overset{\triangle}{T}_{nh-X_t}<\infty\}} |\mathcal{F}_t]] = \mathsf{E}[e^{\Phi(0)(X_t-nh)} \mathbb{1}_{\Lambda\cap\{t<T_{nh}\}}],$$

since $\{\Lambda, \{t < T_{nh}\}\} \cup \sigma(X_t) \subset \mathcal{F}_t$. Next, noting that $\{\overline{X}_\infty \geq nh\} = \{T_{nh} < \infty\}$:

$$\begin{aligned}
\mathsf{P}(\Lambda|\overline{X}_\infty > nh) &= e^{\Phi(0)nh} \left(\mathsf{P}(\Lambda \cap \{T_{nh} \leq t\}) + \mathsf{P}(\Lambda \cap \{t < T_{nh} < \infty\})\right) \\
&= e^{\Phi(0)nh} \left(\mathsf{P}(\Lambda \cap \{T_{nh} \leq t\}) + \mathsf{E}[e^{\Phi(0)(X_t-nh)} \mathbb{1}_{\Lambda\cap\{t<T_{nh}\}}]\right) \\
&= e^{\Phi(0)nh}\mathsf{P}(\Lambda \cap \{T_{nh} \leq t\}) + \mathsf{P}^\natural(\Lambda \cap \{t < T_{nh}\}).
\end{aligned}$$

The second term clearly converges to $\mathsf{P}^\natural(\Lambda)$ as $n \to \infty$. The first converges to 0, because by (3.8) $\mathsf{P}(\overline{X}_{e_1} \geq nh) = e^{-nh\Phi(1)} = o(e^{-nh\Phi(0)})$, as $n \to \infty$, and we have the estimate $\mathsf{P}(T_{nh} \leq t) = \mathsf{P}(\overline{X}_t \geq nh) = \mathsf{P}(\overline{X}_t \geq nh|e_1 \geq t) \leq \mathsf{P}(\overline{X}_{e_1} \geq nh|e_1 \geq t) \leq e^t\mathsf{P}(\overline{X}_{e_1} \geq nh)$.

We next show (2). Note first that X is $\mathbb{F}$-progressively measurable (in particular, measurable), hence the stopped process X^{T_x} is measurable as a mapping into $\mathbb{D}_h$ [Karatzas and Shreve, 1988, p. 5, Problem 1.16].

Further, by the strong Markov property, conditionally on $\{T_x < \infty\}$, $\mathcal{F}_{T_x}$ is independent of the future increments of X after T_x, hence also of $\{T_{x'} < \infty\}$ for any $x' > x$. We deduce that the law of X^{T_x} is the same under $\mathsf{P}(\cdot|T_x < \infty)$ as it is under $\mathsf{P}(\cdot|T_{x'} < \infty)$ for any $x' > x$. (2) then follows from (1) by letting x' tend to $+\infty$, the algebra $\mathcal{A}$ being sufficient to determine equality in law by a π/λ-argument. $\square$

Wiener-Hopf factorization

Recall the notation and terminology of Subsection 1.2.2. Thanks to the skip-free nature of the compound Poisson process X, we can expand on the contents of

Proposition 1.26, by offering further details of its Wiener-Hopf factorization (the first of its two versions, at least). Indeed, if we let $N_t := \overline{X}_t/h$ and $T_k := T_{kh}$ ($t \geq 0$, $k \in \mathbb{N}_0$) then clearly $T := (T_k)_{k\geq 0}$ are the arrival times of a renewal process (with a possibly defective inter-arrival time distribution) and $N := (N_t)_{t\geq 0}$ is the 'number of arrivals' process. One also has the relation: $\overline{G}_t^* = T_{N_t}$, $t \geq 0$ (P-a.s.). Thus the random variables entering the Wiener-Hopf factorization are determined in terms of the renewal process (T, N).

Moreover, we can proceed to calculate explicitly the Wiener-Hopf factors as well as $\hat{\kappa}$ and κ^*. Let $p > 0$. First, since $\overline{X}_{e_p}/h$ is a geometrically distributed random variable we have, for any $\beta \in \overline{\mathbb{C}^{\rightarrow}}$:

$$\mathsf{E}[e^{-\beta \overline{X}_{e_p}}] = \sum_{k=0}^{\infty} e^{-\beta hk}(1 - e^{-\Phi(p)h})e^{-\Phi(p)hk} = \frac{1 - e^{-\Phi(p)h}}{1 - e^{-\beta h - \Phi(p)h}}. \tag{3.10}$$

Note here that $\Phi(p) > 0$ for all $p > 0$. On the other hand, using conditioning (Lemma A.1), for any $\alpha \geq 0$:

$$\begin{aligned}
\mathsf{E}\left[e^{-\alpha \overline{G}^*_{e_p}}\right] &= \mathsf{E}\left[\left((u,t) \mapsto \sum_{k=0}^{\infty} \mathbb{1}_{[0,\infty)}(t_k) e^{-\alpha t_k} \mathbb{1}_{[t_k, t_{k+1})}(u)\right) \circ (e_p, T)\right] \\
&= \mathsf{E}\left[\left(t \mapsto \sum_{k=0}^{\infty} \mathbb{1}_{[0,\infty)}(t_k) e^{-\alpha t_k}(e^{-pt_k} - e^{-pt_{k+1}})\right) \circ T\right], \text{ since } e_p \perp T \\
&= \mathsf{E}\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{T_k<\infty\}} \left(e^{-(p+\alpha)T_k} - e^{-(p+\alpha)T_k} e^{-p(T_{k+1}-T_k)}\right)\right] \\
&= \mathsf{E}\left[\sum_{k=0}^{\infty} e^{-(p+\alpha)T_k} \mathbb{1}_{\{T_k<\infty\}} \left(1 - e^{-p(T_{k+1}-T_k)}\right)\right].
\end{aligned}$$

Now, conditionally on $T_k < \infty$, $T_{k+1} - T_k$ is independent of T_k and has the same distribution as T_1. Therefore, by (3.7) and the theorem of Fubini:

$$\mathsf{E}[e^{-\alpha \overline{G}^*_{e_p}}] = \sum_{k=0}^{\infty} e^{-\Phi(p+\alpha)hk}(1 - e^{-\Phi(p)h}) = \frac{1 - e^{-\Phi(p)h}}{1 - e^{-\Phi(p+\alpha)h}}. \tag{3.11}$$

We identify from (3.10) for any $\beta \in \overline{\mathbb{C}^{\rightarrow}}$: $\frac{\kappa^*(p,0)}{\kappa^*(p,\beta)} = \frac{1-e^{-\Phi(p)h}}{1-e^{-\beta h-\Phi(p)h}}$ and therefore for any $\alpha \geq 0$: $\frac{\kappa^*(p+\alpha,0)}{\kappa^*(p+\alpha,\beta)} = \frac{1-e^{-\Phi(p+\alpha)h}}{1-e^{-\beta h-\Phi(p+\alpha)h}}$. We identify from (3.11) for any $\alpha \geq 0$: $\frac{\kappa^*(p,0)}{\kappa^*(p+\alpha,0)} = \frac{1-e^{-h\Phi(p)}}{1-e^{-\Phi(p+\alpha)h}}$. Therefore, multiplying the last two equalities, for $\alpha \geq 0$

and $\beta \in \overline{\mathbb{C}^{\rightarrow}}$, the equality:

$$\frac{\kappa^*(p,0)}{\kappa^*(p+\alpha,\beta)} = \frac{1-e^{-\Phi(p)h}}{1-e^{-\beta h-\Phi(p+\alpha)h}} \tag{3.12}$$

obtains. In particular, for $\alpha > 0$ and $\beta \in \overline{\mathbb{C}^{\rightarrow}}$, we recognize for some constant $k^* \in (0,\infty)$: $\kappa^*(\alpha,\beta) = k^*(1-e^{-(\beta+\Phi(\alpha))h})$. Next, observe that by independence and duality (for $\alpha \geq 0$ and $\theta \in \mathbb{R}$):

$$\mathsf{E}[\exp\{-\alpha\overline{G}^*_{e_p} + i\theta\overline{X}_{e_p}\}]\mathsf{E}[\exp\{-\alpha\underline{G}_{e_p} + i\theta\underline{X}_{e_p}\}] = \int_0^\infty dtpe^{-pt}\mathsf{E}[\exp\{-\alpha t + i\theta X_t\}] =$$
$$\int_0^\infty dtpe^{-pt-\alpha t+\Psi(\theta)t} = \frac{p}{p+\alpha-\Psi(\theta)}.$$

Therefore:

$$(p+\alpha-\psi(i\theta))\frac{\hat{\kappa}(p,0)}{\hat{\kappa}(p+\alpha,i\theta)} = p\frac{1-e^{i\theta h-\Phi(p+\alpha)h}}{1-e^{-\Phi(p)h}}.$$

Both sides of this equality are continuous in $\theta \in \overline{\mathbb{C}^{\downarrow}}$ and analytic in $\theta \in \mathbb{C}^{\downarrow}$. They agree on $\mathbb{R}$, hence agree on $\overline{\mathbb{C}^{\downarrow}}$ by analytic continuation. Therefore (for all $\alpha \geq 0$, $\beta \in \overline{\mathbb{C}^{\rightarrow}}$):

$$(p+\alpha-\psi(\beta))\frac{\hat{\kappa}(p,0)}{\hat{\kappa}(p+\alpha,\beta)} = p\frac{1-e^{\beta h-\Phi(p+\alpha)h}}{1-e^{-\Phi(p)h}}, \tag{3.13}$$

i.e. for all $\beta \in \overline{\mathbb{C}^{\rightarrow}}$ and $\alpha \geq 0$ for which $p+\alpha \neq \psi(\beta)$ one has:

$$\mathsf{E}[\exp\{-\alpha\underline{G}_{e_p} + \beta\underline{X}_{e_p}\}] = \frac{p}{p+\alpha-\psi(\beta)}\frac{1-e^{(\beta-\Phi(p+\alpha))h}}{1-e^{-\Phi(p)h}}.$$

Moreover, for the unique $\beta_0 > 0$, for which $\psi(\beta_0) = p+\alpha$, one can take the limit $\beta \to \beta_0$ in the above to obtain: $\mathsf{E}[\exp\{-\alpha\underline{G}_{e_p} + \beta_0\underline{X}_{e_p}\}] = \frac{ph}{\psi'(\beta_0)(1-e^{-\Phi(p)h})} = \frac{ph\Phi'(p+\alpha)}{1-e^{-\Phi(p)h}}$. We also recognize from (3.13) for $\alpha > 0$ and $\beta \in \overline{\mathbb{C}^{\rightarrow}}$ with $\alpha \neq \psi(\beta)$, and some constant $\hat{k} \in (0,\infty)$: $\hat{\kappa}(\alpha,\beta) = \hat{k}\frac{\alpha-\psi(\beta)}{1-e^{(\beta-\Phi(\alpha))h}}$. With $\beta_0 = \Phi(\alpha)$ one can take the limit in the latter as $\beta \to \beta_0$ to obtain: $\hat{\kappa}(\alpha,\beta_0) = \hat{k}\psi'(\beta_0)/h = \frac{\hat{k}}{h\Phi'(\alpha)}$.

In summary:

Theorem 3.22 (Wiener-Hopf factorization for upwards skip-free Lévy chains)**.** *We have the following identities in terms of ψ and Φ:*

(i) For every $\alpha \geq 0$ and $\beta \in \overline{\mathbb{C}^{\rightarrow}}$:

$$\mathsf{E}[\exp\{-\alpha\overline{G}^*_{e_p} - \beta\overline{X}_{e_p}\}] = \frac{1-e^{-\Phi(p)h}}{1-e^{-(\beta+\Phi(p+\alpha))h}}$$

and

$$\mathsf{E}[\exp\{-\alpha\underline{G}_{e_p} + \beta\underline{X}_{e_p}\}] = \frac{p}{p+\alpha-\psi(\beta)}\frac{1-e^{(\beta-\Phi(p+\alpha))h}}{1-e^{-\Phi(p)h}}$$

(the latter whenever $p+\alpha \neq \psi(\beta)$; for the unique $\beta_0 > 0$ such that $\psi(\beta_0) = p+\alpha$, i.e. for $\beta_0 = \Phi(p+\alpha)$, the right-hand side is given by $\frac{ph}{\psi'(\beta_0)(1-e^{-\Phi(p)h})} = \frac{ph\Phi'(p+\alpha)}{1-e^{-\Phi(p)h}}$).

(ii) For some $\{k^*, \hat{k}\} \subset \mathbb{R}^+$ and then for every $\alpha > 0$ and $\beta \in \overline{\mathbb{C}^{\rightarrow}}$:

$$\kappa^*(\alpha, \beta) = k^*(1 - e^{-(\beta+\Phi(\alpha))h})$$

and

$$\hat{\kappa}(\alpha, \beta) = \hat{k}\frac{\alpha - \psi(\beta)}{1 - e^{(\beta-\Phi(\alpha))h}}$$

(the latter whenever $\alpha \neq \psi(\beta)$; for the unique $\beta_0 > 0$ such that $\psi(\beta_0) = \alpha$, i.e. for $\beta_0 = \Phi(\alpha)$, one has the right-hand side given by $\hat{k}\psi'(\beta_0)/h = \frac{\hat{k}}{h\Phi'(\alpha)}$).

As a consequence of Theorem 3.22(i), we obtain the formula for the Laplace transform of the running infimum evaluated at an independent exponentially distributed random time e_p:

$$\mathsf{E}[e^{\beta \underline{X}_{e_p}}] = \frac{p}{p-\psi(\beta)} \frac{1-e^{(\beta-\Phi(p))h}}{1-e^{-\Phi(p)h}} \quad (\beta \in \mathbb{R}_+ \backslash \{\Phi(p)\}) \tag{3.14}$$

(and $\mathsf{E}[e^{\Phi(p)\underline{X}_{e_p}}] = \frac{p\Phi'(p)h}{1-e^{-\Phi(p)h}}$). In particular, if $\psi'(0+) > 0$, then letting $p \downarrow 0$ in (3.14), one obtains by the DCT:

$$\mathsf{E}[e^{\beta \underline{X}_\infty}] = \frac{e^{\beta h}-1}{\Phi'(0+)h\psi(\beta)} \quad (\beta > 0). \tag{3.15}$$

3.2.4 Theory of scale functions

Again the reader is invited to compare the exposition of the following subsection with that of [Bertoin, 1996, Section VII.2] and [Kyprianou, 2006, Section 8.2], which deal with the spectrally negative case.

The scale function W

It will be convenient to consider in this paragraph the times at which X attains a new maximum. We let D_1, D_2 and so on, denote the depths (possibly zero, or infinity) of the excursions below these new maxima. For $k \in \mathbb{N}$, it is agreed that $D_k = +\infty$ if the process X never reaches the level $(k-1)h$. Then it is clear that for $y \in \mathbb{Z}_h^+$, $x \geq 0$ (cf. [Bühlmann, 1970, p. 137, Paragraph 6.2.4(a)] [Doney and Picard, 2007, p. 99, Section 9.3]):

$$\mathsf{P}(\underline{X}_{T_y} \geq -x) = \mathsf{P}(D_1 \leq x, D_2 \leq x+h, \ldots, D_{y/h} \leq x+y-h) =$$

$$\mathsf{P}(D_1 \leq x) \cdot \mathsf{P}(D_1 \leq x+h) \cdots \mathsf{P}(D_1 \leq x+y-h) = \frac{\prod_{r=1}^{\lfloor (y+x)/h \rfloor} \mathsf{P}(D_1 \leq (r-1)h)}{\prod_{r=1}^{\lfloor x/h \rfloor h} \mathsf{P}(D_1 \leq (r-1)h)} = \frac{W(x)}{W(x+y)},$$

where we have introduced (up to a multiplicative constant) the *scale function*:

$$W(x) := 1/\prod_{r=1}^{\lfloor x/h \rfloor} \mathsf{P}(D_1 \leq (r-1)h) \quad (x \geq 0). \tag{3.16}$$

(When convenient, we extend W by 0 on $(-\infty, 0)$.)

Remark 3.23. If needed, we can of course express $\mathsf{P}(D_1 \leq hk)$, $k \in \mathbb{N}_0$, in terms of the usual excursions away from the maximum. Thus, let $\tilde{D}_1$ be the depth of the first excursion away from the current maximum. By the time the process attains a new maximum (that is to say h), conditionally on this event, it will make a total of N departures away from the maximum, where (with J_1 the first jump time of X, $p := \lambda(\{h\})/\lambda(\mathbb{R})$, $\tilde{p} := \mathsf{P}(X_{J_1} = h | T_h < \infty) = p/\mathsf{P}(T_h < \infty)$) $N \sim \text{geom}(\tilde{p})$. So, denoting $\tilde{\theta}_k := \mathsf{P}(\tilde{D}_1 \leq hk)$, one has $\mathsf{P}(D_1 \leq hk) = \mathsf{P}(T_h < \infty) \sum_{l=0}^{\infty} \tilde{p}(1-\tilde{p})^l \tilde{\theta}_k^l = \frac{p}{1-(1-e^{\Phi(0)h}p)\tilde{\theta}_k}$, $k \in \mathbb{N}_0$.

The following theorem characterizes the scale function in terms of its Laplace transform.

Theorem 3.24 (The scale function)**.** *For every $y \in \mathbb{Z}_h^+$ and $x \geq 0$ one has:*

$$\mathsf{P}(\underline{X}_{T_y} \geq -x) = \frac{W(x)}{W(x+y)} \tag{3.17}$$

and $W : [0, \infty) \to [0, \infty)$ is (up to a multiplicative constant) the unique right-continuous and piecewise continuous function of exponential order with Laplace transform:

$$\hat{W}(\beta) = \int_0^{\infty} e^{-\beta x} W(x) dx = \frac{e^{\beta h} - 1}{\beta h \psi(\beta)} \quad (\beta > \Phi(0)). \tag{3.18}$$

Proof. (For uniqueness see e.g. [Engelberg, 2005, p. 14, Theorem 10]. It is clear that W is of exponential order, simply from the definition (3.16).)

Suppose first X tends to $+\infty$. Then, letting $y \to \infty$ in (3.17) above, we obtain $\mathsf{P}(-\underline{X}_\infty \leq x) = W(x)/W(+\infty)$. Here, since the left-hand side limit exists by the DCT, is finite and non-zero at least for all large enough x, so does the right-hand side, and $W(+\infty) \in (0, \infty)$.

Therefore $W(x) = W(+\infty)\mathsf{P}(-\underline{X}_\infty \leq x)$ and hence the Laplace-Stieltjes transform of W is given by (3.15) — here we consider W as being extended by 0 on $(-\infty, 0)$:

$$\int_{[0,\infty)} e^{-\beta x} dW(x) = W(+\infty)\frac{e^{\beta h}-1}{\Phi'(0+)h\psi(\beta)} \quad (\beta > 0).$$

Since (integration by parts [Revuz and Yor, 1999, Chapter 0, Proposition 4.5]) $\int_{[0,\infty)} e^{-\beta x} dW(x) = \beta \int_{(0,\infty)} e^{-\beta x} W(x)dx$,

$$\int_0^\infty e^{-\beta x} W(x)dx = \frac{W(+\infty)}{\Phi'(0+)}\frac{e^{\beta h}-1}{\beta h\psi(\beta)} \quad (\beta > 0). \tag{3.19}$$

Suppose now that X oscillates. Via Remark 3.19, approximate X by the processes X^ϵ, $\epsilon > 0$. In (3.19), fix β, carry over everything except for $\frac{W(+\infty)}{\Phi'(0+)}$, divide both sides by $W(0)$, and then apply this equality to X^ϵ. Then on the left-hand side, the quantities pertaining to X^ϵ will converge to the ones for the process X as $\epsilon \downarrow 0$ by the MCT. Indeed, for $y \in \mathbb{Z}_h^+$, $\mathsf{P}(\underline{X}_{T_y} = 0) = W(0)/W(y)$ and (in the obvious notation): $1/\mathsf{P}(\underline{X^\epsilon}_{T_y^\epsilon} = 0) \uparrow 1/\mathsf{P}(\underline{X}_{T_y} = 0) = W(y)/W(0)$, since $X^\epsilon \downarrow X$, uniformly on bounded time sets, almost surely, as $\epsilon \downarrow 0$. (It is enough to have convergence for $y \in \mathbb{Z}_h^+$, as this implies convergence for all $y \geq 0$, W being the right-continuous piecewise constant extension of $W|_{\mathbb{Z}_h^+}$.) Thus we obtain in the oscillating case, for some $\alpha \in (0,\infty)$ which is the limit of the right-hand side as $\epsilon \downarrow 0$:

$$\int_0^\infty e^{-\beta x} W(x)dx = \alpha\frac{e^{\beta h}-1}{\beta h\psi(\beta)} \quad (\beta > 0). \tag{3.20}$$

Finally, we are left with the case when X drifts to $-\infty$. We treat this case by a change of measure (see Proposition 3.20 and the paragraph immediately preceding it). To this end assume, provisionally, that X is already the coordinate process on the canonical filtered space $\mathbb{D}_h$. Then we calculate by Proposition 3.21(2) (for $y \in \mathbb{Z}_h^+$, $x \geq 0$):

$$\mathsf{P}(\underline{X}_{T_y} \geq -x) = \mathsf{P}(T_y < \infty)\mathsf{P}(\underline{X}_{T_y} \geq -x | T_y < \infty) = e^{-\Phi(0)y}\mathsf{P}(\underline{X^{T_y}}_\infty \geq -x | T_y < \infty) = \\ e^{-\Phi(0)y}\mathsf{P}^\natural(\underline{X^{T_y}}_\infty \geq -x) = e^{-\Phi(0)y}\mathsf{P}^\natural(\underline{X}_{T(y)} \geq -x) = e^{-\Phi(0)y}W^\natural(x)/W^\natural(x+y),$$

where the third equality uses the fact that $(\omega \mapsto \inf\{\omega(s) : s \in [0,\infty)\}) : (\mathbb{D}_h, \mathcal{F}) \to ([-\infty,\infty), \mathcal{B}([-\infty,\infty))$ is a measurable transformation. Here $W^\natural$ is the scale function corresponding to X under the measure $\mathsf{P}^\natural$, with Laplace transform:

$$\int_0^\infty e^{-\beta x} W^\natural(x)dx = \frac{e^{\beta h}-1}{\beta h\psi(\Phi(0)+\beta)} \quad (\beta > 0).$$

Note that the equality $\mathsf{P}(\underline{X}_{T_y} \geq -x) = e^{-\Phi(0)y}W^\natural(x)/W^\natural(x+y)$ remains true if we revert back to our original X (no longer assumed to be in its canonical guise). This is so because we can always go from X to its canonical counter-part by taking an image measure. Then the law of the process, hence the Laplace exponent and the probability $\mathsf{P}(\underline{X}_{T_y} \geq -x)$ do not change in this transformation.

Now define $\tilde{W}(x) := e^{\Phi(0)\lfloor 1+x/h \rfloor h} W^\natural(x)$ $(x \geq 0)$. Then $\tilde{W}$ is the right-continuous piecewise-constant extension of $\tilde{W}|_{\mathbb{Z}_h^+}$. Moreover, for all $y \in \mathbb{Z}_h^+$ and $x \geq 0$, (3.17) obtains with W replaced by $\tilde{W}$. Plugging in $x = 0$ into (3.17), $\tilde{W}|_{\mathbb{Z}_h}$ and $W|_{\mathbb{Z}_h}$ coincide up to a multiplicative constant, hence $\tilde{W}$ and W do as well. Moreover, for all $\beta > \Phi(0)$, by the MCT:

$$\begin{aligned}
\int_0^\infty e^{-\beta x} \tilde{W}(x) dx &= e^{\Phi(0)h} \sum_{k=0}^\infty \int_{kh}^{(k+1)h} e^{-\beta x} e^{\Phi(0)kh} W^\natural(kh) dx \\
&= e^{\Phi(0)h} \sum_{k=0}^\infty \frac{1}{\beta} e^{-\beta kh} (1 - e^{-\beta h}) e^{\Phi(0)kh} W^\natural(kh) \\
&= e^{\Phi(0)h} \frac{\beta - \Phi(0)}{\beta} \frac{1 - e^{-\beta h}}{1 - e^{-(\beta - \Phi(0))h}} \int_0^\infty e^{-(\beta - \Phi(0))x} W^\natural(x) dx \\
&= e^{\Phi(0)h} \frac{\beta - \Phi(0)}{\beta} \frac{1 - e^{-\beta h}}{1 - e^{-(\beta - \Phi(0))h}} \frac{e^{(\beta - \Phi(0))h} - 1}{(\beta - \Phi(0))h\psi(\beta)} = \frac{(e^{\beta h} - 1)}{\beta h \psi(\beta)}.
\end{aligned}$$

□

Remark 3.25. Henceforth the normalization of the scale function W will be understood so as to enforce the validity of (3.18).

Proposition 3.26. *$W(0) = 1/(h\lambda(\{h\}))$, and $W(+\infty) = 1/\psi'(0+)$ if $\Phi(0) = 0$. If $\Phi(0) > 0$, then $W(+\infty) = +\infty$.*

Proof. Integration by parts and the DCT yield $W(0) = \lim_{\beta \to \infty} \beta \hat{W}(\beta)$. (3.18) and another application of the DCT then show that $W(0) = 1/(h\lambda(\{h\}))$. Similarly, integration by parts and the MCT give the identity $W(+\infty) = \lim_{\beta \downarrow 0} \beta \hat{W}(\beta)$. The conclusion $W(+\infty) = 1/\psi'(0+)$ is then immediate from (3.18) when $\Phi(0) = 0$. If $\Phi(0) > 0$, then the right-hand side of (3.18) tends to infinity as $\beta \downarrow \Phi(0)$ and thus, by the MCT, necessarily $W(+\infty) = +\infty$. □

The scale functions $W^{(q)}$, $q \geq 0$

Definition 3.27. For $q \geq 0$, let $W^{(q)}(x) := e^{\Phi(q)\lfloor 1+x/h \rfloor h} W_{\Phi(q)}(x)$ $(x \geq 0)$, where W_c plays the role of W but for the process (X, P_c) ($c \geq 0$; see Proposition 3.20). Note that $W^{(0)} = W$. When convenient we extend $W^{(q)}$ by 0 on $(-\infty, 0)$.

Theorem 3.28. *For each $q \geq 0$, $W^{(q)} : [0, \infty) \to [0, \infty)$ is the unique right-continuous and piecewise continuous function of exponential order with Laplace* *transform:*

$$\widehat{W^{(q)}}(\beta) = \int_0^\infty e^{-\beta x} W^{(q)}(x)dx = \frac{e^{\beta h} - 1}{\beta h(\psi(\beta) - q)} \quad (\beta > \Phi(q)). \tag{3.21}$$

Moreover, for all $y \in \mathbb{Z}_h^+$ and $x \geq 0$:

$$\mathsf{E}[e^{-qT_y} \mathbb{1}_{\{\underline{X}_{T_y} \geq -x\}}] = \frac{W^{(q)}(x)}{W^{(q)}(x+y)}. \tag{3.22}$$

Proof. The claim regarding the Laplace transform follows from Proposition 3.20, Theorem 3.24 and Definition 3.27 as it did in the case of the scale function W (cf. final paragraph of the proof of Theorem 3.24). For the second assertion, let us calculate (moving onto the canonical space $\mathbb{D}_h$ as usual, using Proposition 3.20 and noting that $X_{T_y} = y$ on $\{T_y < \infty\}$):

$$\mathsf{E}[e^{-qT_y} \mathbb{1}_{\{\underline{X}_{T_y} \geq -x\}}] = \mathsf{E}[e^{\Phi(q)X_{T_y} - qT_y} \mathbb{1}_{\{\underline{X}_{T_y} \geq -x\}}] e^{-\Phi(q)y} =$$

$$e^{-\Phi(q)y} \mathsf{P}_{\Phi(q)}(\underline{X}_{T_y} \geq -x) = e^{-\Phi(q)y} \frac{W_{\Phi(q)}(x)}{W_{\Phi(q)}(x+y)} = \frac{W^{(q)}(x)}{W^{(q)}(x+y)}.$$

□

Proposition 3.29. *For all $q > 0$: $W^{(q)}(0) = 1/(h\lambda(\{h\}))$ and $W^{(q)}(+\infty) = +\infty$.*

Proof. As in Proposition 3.26, $W^{(q)}(0) = \lim_{\beta\to\infty} \beta \widehat{W^{(q)}}(\beta) = 1/(h\lambda(\{h\}))$. Since $\Phi(q) > 0$, $W^{(q)}(+\infty) = +\infty$ also follows at once from the expression for $\widehat{W^{(q)}}$. □

Moreover:

Proposition 3.30. *For $q \geq 0$:*

(i) If $\Phi(q) > 0$ or $\psi'(0+) > 0$, then $\lim_{x\to\infty} W^{(q)}(x) e^{-\Phi(q)\lfloor 1+x/h \rfloor h} = 1/\psi'(\Phi(q))$.

(ii) If $\Phi(q) = \psi'(0+) = 0$ (hence $q = 0$), then $W^{(q)}(+\infty) = +\infty$, however $\limsup_{x\to\infty} W^{(q)}(x)/x < \infty$. Indeed, $\lim_{x\to\infty} W^{(q)}(x)/x = 2/m_2$, if $m_2 := \int y^2 \lambda(dy) < \infty$ and $\lim_{x\to\infty} W^{(q)}(x)/x = 0$, if $m_2 = \infty$.

Proof. The first claim is immediate from Proposition 3.26, Definition 3.27 and Proposition 3.20. To handle the second claim, let us calculate, for the Laplace transform $\widehat{dW}$ of the measure dW, the quantity (using integration by parts, Theorem 3.24 and the fact that (since $\psi'(0+) = 0$) $\int y\lambda(dy) = 0$):

$$\lim_{\beta\downarrow 0} \beta \widehat{dW}(\beta) = \lim_{\beta\downarrow 0} \frac{\beta^2}{\psi(\beta)} = \frac{2}{m_2} \in [0, +\infty).$$

For:
$$\lim_{\beta\downarrow 0}\int (e^{\beta y}-1)\lambda(dy)/\beta^2 = \lim_{\beta\downarrow 0}\int \frac{e^{\beta y}-\beta y-1}{\beta^2 y^2}y^2\lambda(dy) = \frac{m_2}{2},$$
by the MCT, since $(u \mapsto \frac{e^{-u}+u-1}{u^2})$ is nonincreasing on $(0,\infty)$ (the latter can be checked by comparing derivatives). The claim then follows by the Karamata Tauberian Theorem [Bingham et al., 1987, p. 37, Theorem 1.7.1 with $\rho = 1$]. □

The functions $Z^{(q)}$, $q \geq 0$

Definition 3.31. For each $q \geq 0$, let $Z^{(q)}(x) := 1 + q\int_0^{\lfloor x/h\rfloor h} W^{(q)}(z)dz$ $(x \geq 0)$. When convenient we extend these functions by 1 on $(-\infty, 0)$.

Proposition 3.32. *In the sense of measures on the real line, for every $q > 0$:*
$$\mathsf{P}_{-\underline{X}_{e_q}} = \frac{qh}{e^{\Phi(q)h}-1}dW^{(q)} - qW^{(q)}(\cdot - h)\cdot\Delta,$$
where $\Delta := h\sum_{k=1}^\infty \delta_{kh}$ is the normalized counting measure on $\mathbb{Z}_h^{++} \subset \mathbb{R}$, $\mathsf{P}_{-\underline{X}_{e_q}}$ is the law of $-\underline{X}_{e_q}$ under P, and $(W^{(q)}(\cdot - h)\cdot\Delta)(A) = \int_A W^{(q)}(y-h)\Delta(dy)$ for Borel subsets A of $\mathbb{R}$.

Theorem 3.33. *For each $x \geq 0$,*
$$\mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<\infty\}}] = Z^{(q)}(x) - \frac{qh}{e^{\Phi(q)h}-1}W^{(q)}(x) \tag{3.23}$$
when $q > 0$, and $\mathsf{P}(T_x^- < \infty) = 1 - W(x)/W(+\infty)$. The Laplace transform of $Z^{(q)}$, $q \geq 0$, is given by:
$$\widehat{Z^{(q)}}(\beta) = \int_0^\infty Z^{(q)}(x)e^{-\beta x}dx = \frac{1}{\beta}\left(1+\frac{q}{\psi(\beta)-q}\right), \quad (\beta > \Phi(q)). \tag{3.24}$$

Proof of Proposition 3.32 and Theorem 3.33. First, with regard to the Laplace transform of $Z^{(q)}$, we have the following derivation (using integration by parts, for every $\beta > \Phi(q)$):
$$\begin{aligned}\int_0^\infty Z^{(q)}(x)e^{-\beta x}dx &= \int_0^\infty \frac{e^{-\beta x}}{\beta}dZ^{(q)}(x) = \frac{1}{\beta}\left(1+q\sum_{k=1}^\infty e^{-\beta kh}W^{(q)}((k-1)h)h\right)\\ &= \frac{1}{\beta}\left(1+\frac{qe^{-\beta h}\beta h}{1-e^{-\beta h}}\sum_{k=1}^\infty\frac{(1-e^{-\beta h})}{\beta}e^{-\beta(k-1)h}W^{(q)}((k-1)h)\right)\\ &= \frac{1}{\beta}\left(1+q\frac{\beta h}{e^{\beta h}-1}\widehat{W^{(q)}}(\beta)\right) = \frac{1}{\beta}\left(1+\frac{q}{\psi(\beta)-q}\right).\end{aligned}$$

Next, to prove Proposition 3.32, note that it will be sufficient to check the equality of the Laplace transforms [Bhattacharya and Waymire, 2007, p. 109, Theorem 8.4]. By what we have just shown, (3.14), integration by parts, and Theorem 3.28, we need then only establish, for $\beta > \Phi(q)$:

$$\frac{q}{\psi(\beta)-q}\frac{e^{(\beta-\Phi(q))h}-1}{1-e^{-\Phi(q)h}} = \frac{qh}{e^{\Phi(q)h}-1}\frac{\beta(e^{\beta h}-1)}{(\psi(\beta)-q)\beta h} - \frac{q}{\psi(\beta)-q},$$

which is clear.

Finally, let $x \in \mathbb{Z}_h^+$. For $q > 0$, evaluate the measures in Proposition 3.32 at $[0, x]$, to obtain:

$$\begin{aligned} \mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<\infty\}}] &= \mathsf{P}(e_q \geq T_x^-) = \mathsf{P}(\underline{X}_{e_q} < -x) = 1 - \mathsf{P}(\underline{X}_{e_q} \geq -x) \\ &= 1 + q\int_0^x W^{(q)}(z)dz - \frac{qh}{e^{\Phi(q)h}-1}W^{(q)}(x), \end{aligned}$$

whence the claim follows. On the other hand, when $q = 0$, the following calculation is straightforward: $\mathsf{P}(T_x^- < \infty) = \mathsf{P}(\underline{X}_\infty < -x) = 1 - \mathsf{P}(\underline{X}_\infty \geq -x) = 1 - W(x)/W(+\infty)$ (we have passed to the limit $y \to \infty$ in (3.17) and used the DCT on the left-hand side of this equality). □

Proposition 3.34. *Let $q \geq 0$, $x \geq 0$, $y \in \mathbb{Z}_h^+$. Then:*

$$\mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<T_y\}}] = Z^{(q)}(x) - Z^{(q)}(x+y)\frac{W^{(q)}(x)}{W^{(q)}(x+y)}.$$

Proof. Observe that $\{T_x^- = T_y\} = \emptyset$, P-a.s. The case when $q = 0$ is immediate and indeed contained in Theorem 3.24, since, P-a.s., $\Omega\backslash\{T_x^- < T_y\} = \{T_x^- \geq T_y\} = \{\underline{X}_{T_y} \geq -x\}$. For $q > 0$ we observe that by the strong Markov property, Theorem 3.28 and Theorem 3.33:

$$\begin{aligned} & \mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<T_y\}}] = \mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_x^-<\infty\}}] - \mathsf{E}[e^{-qT_x^-}\mathbb{1}_{\{T_y<T_x^-<\infty\}}] \\ = \ & Z^{(q)}(x) - \frac{qh}{e^{\Phi(q)h}-1}W^{(q)}(x) - \mathsf{E}[e^{-qT_y}\mathbb{1}_{\{T_y<T_x^-\}}]\mathsf{E}[e^{-qT_{x+y}^-}\mathbb{1}_{\{T_{x+y}^-<\infty\}}] \\ = \ & Z^{(q)}(x) - \frac{qh}{e^{\Phi(q)h}-1}W^{(q)}(x) - \frac{W^{(q)}(x)}{W^{(q)}(x+y)}\left(Z^{(q)}(x+y) - \frac{qh}{e^{\Phi(q)h}-1}W^{(q)}(x+y)\right) \\ = \ & Z^{(q)}(x) - Z^{(q)}(x+y)\frac{W^{(q)}(x)}{W^{(q)}(x+y)}. \end{aligned}$$

□

Calculating scale functions

In this paragraph it will be assumed for notational convenience, but without loss of generality, that $h = 1$ and that X is the canonical process on $\Omega = \mathbb{D}_h$ equipped with the usual σ-algebra and filtration. We define:

$$\gamma := \lambda(\mathbb{R}), \quad p := \lambda(\{1\})/\gamma, \quad q_k := \lambda(\{-k\})/\gamma, \; k \geq 1.$$

Fix $q \geq 0$. Then denote, provisionally, $e_{m,k} := \mathsf{E}[e^{-qT_k}\mathbb{1}_{\{\underline{X}_{T_k} \geq -m\}}]$, and $e_k := e_{0,k}$, where $\{m, k\} \subset \mathbb{N}_0$ and note that, thanks to Theorem 3.28, $e_{m,k} = \frac{e_{m+k}}{e_m}$ for all $\{m, k\} \subset \mathbb{N}_0$. Now, $e_0 = 1$. Moreover, by the strong Markov property and using Lemma A.1, for each $k \in \mathbb{N}_0$, by conditioning on $\mathcal{F}_{T_k}$ and then on $\mathcal{F}_J$, where J is the time of the first jump after T_k (so that, conditionally on $T_k < \infty$, $J - T_k \sim \mathrm{Exp}(\gamma)$):

$$\begin{aligned} e_{k+1} &= \mathsf{E}\Big[e^{-qT_k}\mathbb{1}_{\{\underline{X}_{T_k} \geq 0\}}e^{-q(J-T_k)}\big(\mathbb{1}(\text{next jump after } T_k \text{ up}) + \\ &\quad \mathbb{1}(\text{next jump after } T_k \text{ 1 down, then up 2 before down more than } k-1) + \cdots + \\ &\quad \mathbb{1}(\text{next jump after } T_k\ k \text{ down \& then up } k+1 \text{ before down more than } 0)\big)e^{-q(T_{k+1}-J)}\Big] \\ &= e_k\frac{\gamma}{\gamma+q}[p + q_1e_{k-1,2} + \cdots + q_ke_{0,k+1}] = e_k\frac{\gamma}{\gamma+q}[p + q_1\frac{e_{k+1}}{e_{k-1}} + \cdots + q_k\frac{e_{k+1}}{e_0}]. \end{aligned}$$

Upon division by e_ke_{k+1}, we obtain:

$$W^{(q)}(k) = \frac{\gamma}{\gamma+q}[pW^{(q)}(k+1) + q_1W^{(q)}(k-1) + \cdots + q_kW^{(q)}(0)].$$

Put another way, for all $k \in \mathbb{Z}_+$:

$$pW^{(q)}(k+1) = \left(1 + \frac{q}{\gamma}\right)W^{(q)}(k) - \sum_{l=1}^{k} q_lW^{(q)}(k-l). \tag{3.25}$$

Coupled with the initial condition $W^{(q)}(0) = 1/(\gamma p)$ (from Proposition 3.29 and Proposition 3.26), this is an explicit recursion scheme by which the values of $W^{(q)}$ can be obtained (cf. [Vylder and Goovaerts, 1988, Section 4, Equations (6) & (7)] [Dickson and Waters, 1991, Section 7, Equations (7.1) & (7.5)] [Marchal, 2001, p. 255, Proposition 3.1]). We can also see the vector $W^{(q)} = (W^{(q)}(k))_{k\in\mathbb{Z}}$ as a suitable eigenvector of the transition matrix P associated to the jump chain of X. Namely, we have for all $k \in \mathbb{Z}_+$: $\left(1 + \frac{q}{\gamma}\right)W^{(q)}(k) = \sum_{l\in\mathbb{Z}} P_{kl}W^{(q)}(l)$.

Now, with regard to the function $Z^{(q)}$, its values can be computed directly from the values of $W^{(q)}$ by a straightforward summation, indeed: $Z^{(q)}(n) = 1 + q\sum_{k=0}^{n-1}W^{(q)}(k)$ ($n \in \mathbb{N}_0$). Alternatively, (3.25) yields immediately its analogue, valid for each $n \in \mathbb{Z}^+$ (make a summation $\sum_{k=0}^{n-1}$ and multiply by q, using Fubini's theorem for the last sum):

$$pZ^{(q)}(n+1) - p - pqW^{(q)}(0) = \left(1+\frac{q}{\gamma}\right)(Z^{(q)}(n)-1) - \sum_{l=1}^{n-1} q_l (Z^{(q)}(n-l)-1),$$

i.e. for all $k \in \mathbb{Z}_+$:

$$pZ^{(q)}(k+1) + \left(1 - p - \sum_{l=1}^{k-1} q_l\right) = \left(1+\frac{q}{\gamma}\right) Z^{(q)}(k) - \sum_{l=1}^{k-1} q_l Z^{(q)}(k-l). \tag{3.26}$$

Again this can be seen as an eigenvalue problem. Namely, for all $k \in \mathbb{Z}_+$ we have: $\left(1+\frac{q}{\gamma}\right) Z^{(q)}(k) = \sum_{l\in\mathbb{Z}} P_{kl} Z^{(q)}(l)$. In summary:

Proposition 3.35 (Calculation of $W^{(q)}$ and $Z^{(q)}$)**.** *Let $h = 1$ and $q \geq 0$. Seen as vectors, $W^{(q)} := (W^{(q)}(k))_{k\in\mathbb{Z}}$ and $Z^{(q)} := (Z^{(q)}(k))_{k\in\mathbb{Z}}$ satisfy, entry-by-entry (P being the transition matrix associated to the jump chain of X; $\lambda_q := 1 + q/\lambda(\mathbb{R})$):*

$$(PW^{(q)})|_{\mathbb{Z}_+} = \lambda_q W^{(q)}|_{\mathbb{Z}_+} \text{ and } (PZ^{(q)})|_{\mathbb{Z}_+} = \lambda_q Z^{(q)}|_{\mathbb{Z}_+}, \tag{3.27}$$

i.e. (3.25) and (3.26) hold true for $k \in \mathbb{Z}_+$. Additionally, $W^{(q)}|_{\mathbb{Z}^-} = 0$ with $W^{(q)}(0) = 1/\lambda(\{1\})$, whereas $Z^{(q)}|_{\mathbb{Z}_-} = 1$.

For the purposes of the following remark and corollary it is no longer assumed that $h = 1$ or, indeed, that the underlying filtered probability space is the canonical one, i.e. we revert back to our original setting.

Remark 3.36. Let L be the infinitesimal generator [Sato, 1999, p. 208, Theorem 31.5] of X. It is seen from (3.27), that for each $q \geq 0$, $((L-q)W^{(q)})|_{\mathbb{R}_+} = ((L-q)Z^{(q)})|_{\mathbb{R}_+} = 0$.

Corollary 3.37. *For each $q \geq 0$, the stopped processes Y and Z, defined by $Y_t := e^{-q(t\wedge T_0^-)} W^{(q)} \circ X_{t\wedge T_0^-}$ and $Z_t := e^{-q(t\wedge T_0^-)} W^{(q)} \circ X_{t\wedge T_0^-}$, $t \geq 0$, are nonnegative P-martingales with respect to the natural filtration $\mathbb{F}^X = (\mathcal{F}_s^X)_{s\geq 0}$ of X.*

Proof. We argue for the case of the process Y, the justification for Z being similar. Let $(H_k)_{k\geq 1}$, $H_0 := 0$, be the sequence of jump times of X (where, possibly by discarding a P-negligible set, we may insist on all of the H_k, $k \in \mathbb{N}_0$, being finite and increasing to $+\infty$ as $k \to \infty$). Let $0 \leq s < t$, $A \in \mathcal{F}_s^X$. By the MCT it will be sufficient to establish for $\{l, k\} \subset \mathbb{N}_0$, $l \leq k$, that:

$$\mathsf{E}[\mathbb{1}(H_l \leq s < H_{l+1})\mathbb{1}_A Y_t \mathbb{1}(H_k \leq t < H_{k+1})] = \mathsf{E}[\mathbb{1}(H_l \leq s < H_{l+1})\mathbb{1}_A Y_s \mathbb{1}(H_k \leq t < H_{k+1})]. \tag{3.28}$$

On the left-hand (respectively right-hand) side of (3.28) we may now replace Y_t (respectively Y_s) by Y_{H_k} (respectively Y_{H_l}) and then harmlessly insist on $l < k$.

Moreover, up to a completion, $\mathcal{F}^X_s \subset \sigma((H_m \wedge s, X(H_m \wedge s))_{m\geq 0})$. Therefore, by a π/λ-argument, we need only verify (3.28) for sets A of the form: $A = \bigcap_{m=1}^M \{H_m \wedge s \in A_m\} \cap \{X(H_m \wedge s) \in B_m\}$, A_m, B_m Borel subsets of $\mathbb{R}$, $1 \leq m \leq M$, $M \in \mathbb{N}$. Due to the presence of the indicator $\mathbb{1}(H_l \leq s < H_{l+1})$, we may also take, without loss of generality, $M = l$ and hence $A \in \mathcal{F}^X_{H_l}$. Further, $\mathcal{H} := \sigma(H_{l+1} - H_l, H_k - H_l, H_{k+1} - H_l)$ is independent of $\mathcal{F}^X_{H_l} \vee \sigma(Y_{H_k})$ and then $\mathsf{E}[Y_{H_k}|\mathcal{F}^X_{H_l} \vee \mathcal{H}] = \mathsf{E}[Y_{H_k}|\mathcal{F}^X_{H_l}] = Y_{H_l}$, P-a.s. (as follows at once from (3.27) of Proposition 3.35), whence (3.28) obtains. $\square$

Chapter 4

Application to the numerical evaluation of scale functions for spectrally negative Lévy processes

> We introduce a general algorithm for the computation of the scale functions of a spectrally negative Lévy process X, based on the weak approximation of X via upwards skip-free continuous-time Markov chains with stationary independent increments from Chapter 2. The algorithm consists of evaluating a finite linear recursion with coefficients given explicitly in terms of the Lévy triplet of X, thus providing an explicit link between the semimartingale characteristics of X and its scale functions. In the interest of space we forgo making the analysis of the algorithm explicit in the present thesis; the interested reader is referred instead to the preprint [Mijatović, Vidmar, and Jacka, 2013b].

Throughout this chapter we let X be a spectrally negative Lévy process (see Definition 1.27). The Laplace exponent ψ of X can then be expressed as (see e.g. [Bertoin, 1996, p. 188]):

$$\psi(\beta) = \frac{1}{2}\sigma^2\beta^2 + \mu\beta + \int_{(-\infty,0)} \left(e^{\beta y} - \beta y\tilde{c}(y) - 1\right)\lambda(dy), \quad \beta \in \overline{\mathbb{C}^{\rightarrow}}.$$

The Lévy triplet of X is thus given by $(\sigma^2, \lambda, \mu)_{\tilde{c}}$, $\tilde{c} := \mathbb{1}_{[-V,0)}$ with V equal to either 0 or 1, the former only if $\int_{[-1,0)} |x|\lambda(dx) < \infty$. Further, when the Lévy measure satisfies $\int_{[-1,0)} |x|\lambda(dx) < \infty$, we may always express ψ in the form $\psi(\beta) = \frac{1}{2}\sigma^2\beta^2 + \mu_0\beta + \int_{(-\infty,0)} \left(e^{\beta y} - 1\right) \lambda(dy)$ for $\beta \in \overline{\mathbb{C}^{\rightarrow}}$. If in addition $\sigma^2 = 0$, then necessarily the drift μ_0 must be positive, $\mu_0 > 0$ [Kyprianou, 2006, p. 212].

4.1 Introduction

For a spectrally negative Lévy process X, fluctuation theory in terms of the two families of scale functions, $(W^{(q)})_{q\in[0,\infty)}$ and $(Z^{(q)})_{q\in[0,\infty)}$, has been developed (see Subsection 1.2.2 and the references therein). Of particular importance is the function $W := W^{(0)}$, in terms of which the others may be defined, and which features in the solution of many important problems of applied probability (see Section 4.2 below). It is central to these applications to be able to evaluate scale functions numerically for any spectrally negative Lévy process X.

Analytically, W is characterized by its Laplace transform. Typically, however, it is not possible to perform the inversion explicitly and the user is faced with a Laplace inversion algorithm, involving, usually, complex numerical integration of a function of the Laplace exponent of X, and certainly the evaluation of the Laplace exponent (for complex arguments). While such a procedure provides a way of numerically evaluating W, it says little about the dependence of the scale function on the Lévy triplet of X: recall that the characteristic exponent of X depends on a parametric complex integral of the Lévy measure, a function of which the algorithm integrates numerically in the parameter along a curve in the complex plane (in the case of the Bromwich integral), making it hard to discern how a perturbation in the Lévy measure influences the values taken by the scale function. Moreover, a Laplace inversion algorithm fails to ensure that the computed values of the scale function are probabilistically meaningful. Put differently, given an output of a numerical Laplace inversion it is not necessary that the formulae involving W in Section 4.2 below yield probabilities of events, i.e. values in the interval $[0, 1]$.

The goal of the present chapter is to define a very simple novel algorithm for computing W, based on a purely probabilistic idea of the weak approximation from Chapter 2 and the findings of Section 3.2 (Chapter 3), which avoids all the issues mentioned in the paragraph above. Indeed, this weak approximation of X (as a Markov process) by a CTMC, which (as it emerges) is skip-free to the right, provides a natural way of encoding the underlying probabilistic structure of the problem in the design of the algorithm. In particular, to compute $W(x)$ for some $x > 0$, choose small $h > 0$ such that x/h is an integer and define the approximation

$W_h(x)$ by the formula:

$$pW_h(y+h) = W_h(y) - \sum_{k=1}^{y/h} q_k W_h(y-kh), \qquad W_h(0) = (p\gamma h)^{-1}, \tag{4.1}$$

for $y = 0, h, 2h, \ldots, x-h$, where the coefficients p, $(q_k)_{k\geq 1}$ and γ are given *explicitly* in terms of the Lévy measure λ, (possibly vanishing) Gaussian component σ^2 and drift μ of the spectrally negative Lévy process X (see (4.4)–(4.5) in Section 4.3 below). Furthermore, these coefficients have a natural probabilistic interpretation in terms of the upwards skip-free Lévy chain (see Definition 3.1), which is used to approximate X: p (resp. q_k, $k \in \mathbb{N}$) is the probability that the jump of the chain is of size h (resp. $-kh$, $k \in \mathbb{N}$) and γ is the total mass of the Lévy measure of the chain. An algorithm, completely analogous to (4.1), for the computation of the scale functions $W^{(q)}$ and $Z^{(q)}$, $q \geq 0$, also follows from our results.

Now, it is clear from (4.1) that the values of W_h may be computed by a simple finite *linear recursion* with coefficients given explicitly in terms of the characteristics of X. Algorithm (4.1) yields, as a by-product of the evaluation of $W_h(x)$, values $W_h(y)$ for all $y = 0, h, 2h, \cdots, x-h, x$ (see MATLAB code for the algorithm in [Mijatović, Vidmar, and Jacka, 2013a]).

Furthermore, Algorithm (4.1) provides an explicit link between the (deterministic) semimartingale characteristics of X, and in particular its Lévy measure, and the scale function W (see (4.4)–(4.5)). This is analogous in spirit to the one-dimensional Itô diffusion setting, where the computation of the scale function requires numerical evaluation of certain integrals of the coefficients of the SDE driving the diffusion, thus linking the deterministic characteristics of the process with its scale function (for the explicit formulae of the integrals see e.g. [Borodin and Salminen, 2002, Chapters 2 and 3]). Moreover, Algorithm (4.1) gives a precise evaluation (modulo computer arithmetic) of a scale function, introduced in Section 3.2 of Chapter 3, of the approximating upwards skip-free Lévy chain and hence yields probabilistically consistent outputs whenever it is well-defined.

Finally, one may show, albeit we avoid making this explicit in the present thesis (but refer the reader to the preprint [Mijatović, Vidmar, and Jacka, 2013b]), that, as $h \downarrow 0$, the pointwise convergence of the approximating scale functions to those of the original Lévy process holds; further, under mild additional assumptions, it is possible to establish the sharp rate at which this convergence transpires on $\mathbb{Z}_h$ (again see [Mijatović, Vidmar, and Jacka, 2013b]) .

4.2 Literature overview and applications

An excellent overview of available numerical methods for computing the scale functions can be found in [Kuznetsov et al., 2013, Chapter 5]. Except possibly for special subclasses of the spectrally negative family, these are one or another of the many Laplace inversion methods, which have stood the test of time. They require, thus, the evaluation of the Laplace exponent at complex, rarely only real, values of its argument. This makes our proposed approach qualitatively different from the techniques in the literature.

Nevertheless, in the special case when X is a positive drift minus a compound Poisson subordinator, numerical schemes for (finite time) ruin/survival probabilities that very much parallel our approach have been proposed (see e.g. [Vylder and Goovaerts, 1988] and [Dickson and Waters, 1991]; note that ruin probabilities are intimately related to scale functions, see (i)-(ii) below). Indeed, discrete-time Markov chain approximations of one sort or another for this, modulo the starting value, classical insurance surplus process in the collective model, are quite ubiquitous in literature (see further e.g. [Cardoso and Waters, 2003; Dickson and Gray, 1984] and the references therein).

Further, for an overview of (the few, but important) examples when the scale functions *can* be given analytically, see e.g. [Hubalek and Kyprianou, 2011]. Indeed, in the case of meromorphic Lévy processes [Kuznetsov et al., 2012], a computational method for the (finite-time) Gerber-Shiu measure (which is related to scale functions [Kyprianou, 2013, Theorem 5.5]) can be found in [Kuznetsov and Morales, 2014]. We note that it is also possible to construct various scale functions indirectly, see e.g. [Kuznetsov et al., 2013, Chapter 4], i.e. not starting from the basic datum, which we consider to be the characteristic triplet of X.

Finally, in terms of applications there are numerous identities concerning boundary crossing problems and related path decompositions in which scale functions feature [Kuznetsov et al., 2013, p. 100]. They do so either indirectly (usually as Laplace transforms of quantities which are ultimately of interest), or even directly. We list a few important problems in applied probability, the solutions of which are given explicitly in terms of the scale function W (see Subsection 1.2.2 for notation regarding first passage times, the supremum process etc.):

(i) **Two-sided exit problem**. For $a \geq 0$, recall T_a (respectively T_a^-) is the first entrance time of X to $[a, \infty)$ (respectively $(-\infty, -a)$), see Definition 1.20. Then:

$$\mathsf{P}(T_x^- > T_a) = \frac{W(x)}{W(a+x)},$$

whenever $\{a, x\} \subset \mathbb{R}^+$, see e.g. [Bertoin, 1996, Chapter VII, Theorem 8]. These quantities are of interest for the insurance industry, where capital may be modeled, e.g., by a positive drift (representing accrued premiums) minus a subordinator (representing claims), in which case x corresponds to the initial capital.

(ii) **Ruin probabilities**. Of particular relevance in the insurance context is the probability of eventual bankruptcy. In the case that the modeling Lévy process X drifts to $+\infty$, we have for $x \in \mathbb{R}^+$, the generalised Cramér-Lundberg identity:

$$\mathsf{P}(T_x^- = \infty) = W(x)\psi'(0+),$$

where ψ is the Laplace exponent of X, see e.g. [Kyprianou, 2006, p. 217, Equation (8.15)].

(iii) **Continuous-state branching processes**. Under mild conditions, the law of the supremum of a continuous-state branching process Y is given by the identity (for $x \in \mathbb{R}^+$, $y \in \mathbb{R}$):

$$\mathsf{P}_y(\sup_{s\geq 0} Y_s \leq x) = \frac{W(x-y)}{W(x)},$$

where W is the scale function of the associated Lévy process, see [Bingham, 1976].

(iv) **Population biology**. The typical branch length H between two consecutive individuals alive at time $t \in \mathbb{R}^+$, conditionally on there being at least two extant individuals at said time, satisfies the identity:

$$\mathsf{P}(H < s) = \frac{1 - W(s)^{-1}}{1 - W(t)^{-1}},$$

whenever $s \in (0, t]$, and with W the scale function associated to the jumping chronological contour process. We refer to [Lambert, 2011] for details.

Miscellaneous other areas featuring scale functions (together with their derivatives and the integrals $Z^{(q)}$) include queuing theory, optimal stopping and control problems, fragmentation processes etc. For example in:

(a) **Optimal stopping**. Consider the Shepp-Shiryaev optimal stopping problem $v(x) = \sup_\tau \mathsf{E}[e^{-q\tau + (\overline{X}_\tau \vee x)}]$ (which was solved for a spectrally negative Lévy process X in [Avram et al., 2104]). Here $\overline{X}$ denotes the running supremum process of X, and the supremum is taken over all a.s. finite stopping times τ relative to the natural filtration of X. Then, under relatively mild additional conditions:

$$v(x) = e^x Z^{(q)}(x^* - x),$$

with $x^* = \inf\{x \geq 0 : Z^{(q)}(x) - qW^{(q)}(x) \leq 0\}$. Note that our algorithm is particularly suited to the calculation of an approximation for x^*, since the values of the scale functions are computed recursively, and one can simply stop the first time a nonpositive value of the difference $Z_h^{(q)}(x) - qW_h^{(q)}(x)$ has been found.

(b) **Optimal control**. In an optimal dividend problem involving a spectrally negative Lévy process X and a discounting rate q, the value function u under a barrier strategy at level $a > 0$, is given, under suitable conditions, by (for details see [Loeffen, 2008]):

$$u(x) = \begin{cases} W^{(q)}(x)/W^{(q)\prime}(a), & \text{for } 0 \leq x \leq a, \\ x - a + \frac{W^{(q)}(a)}{W^{(q)\prime}(a)}, & \text{for } x > a. \end{cases}$$

For a comprehensive overview of these and further applications we refer to [Kuznetsov et al., 2013, Section 1.2] and the references therein. A suite of identities involving Laplace transforms of quantities pertaining to the reflected process of X can be found in [Mijatović and Pistorius, 2012].

4.3 Genesis of the algorithm

The key idea leading to the algorithm in (4.1) is best described by the following three steps: (i) approximate the spectrally negative Lévy process X by a family of continuous-time Markov chains X^h with state space $\mathbb{Z}_h$ ($h \in (0, h_\star)$ for some $h_\star > 0$, cf. Definition 2.19), as described in Chapter 2; (ii) recognize that the X^h are (upon the choice of càdlàg versions) upwards skip-free Lévy chains and (iii) apply the results of Proposition 3.35 to the processes X^h/h, $h \in (0, h_\star)$. One thus obtains for each $q \geq 0$, a family of scale functions $(W_h^{(q)})_{h\in(0,h_\star)}$, which approximate $W^{(q)}$ (likewise for $Z^{(q)}$) and converge thereto as $h \downarrow 0$ (see [Mijatović, Vidmar, and Jacka, 2013b] for a precise analysis of this convergence). Moreover, a finite recursion for computing these approximating scale functions is readily available.

We now explicate the three steps in some detail.

Consider first step (i). As we have seen in Chapter 2, we shall use two approximating schemes, scheme 1 and 2, according as to whether $\sigma^2 > 0$ or $\sigma^2 = 0$. To construct the processes $(X^h)_{h\in(0,h_\star)}$, we further let $V = 0$, if λ is finite and $V = 1$, if λ is infinite (see Table 2.1). Note $h_\star \in (0, +\infty]$ needs to be chosen small enough to make the approximations well-defined (recall statement of Proposition 2.18). For $h > 0$, define also $c_y^h := \lambda(A_y^h)$ with $A_y^h := [y - h/2, y + h/2)$ $(y \in \mathbb{Z}_h^{--})$; $A_0^h := [-h/2, 0)$;

$$c_0^h := \int_{A_0^h} y^2 \mathbb{1}_{[-V,0)}(y)\lambda(dy) \quad \text{and} \quad \mu^h := \sum_{y\in\mathbb{Z}_h^{--}} y \int_{A_y^h} \mathbb{1}_{[-V,0)}(z)\lambda(dz).$$

It is clear from Chapter 2 that, for each $h \in (0, h_\star)$, X^h will be a CP process (we shall insist on càdlàg versions, as we may, cf. Remark 1.36), with $X_0^h = 0$, a.s., and whose positive jumps do not exceed h. Thus each X^h admits its Laplace exponent $\psi^h(\beta) := \log \mathsf{E}[e^{\beta X_1^h}]$ $(\beta \in \overline{\mathbb{C}^{\rightarrow}})$, which in turn uniquely determines its law. Moreover, ψ^h may be obtained from the characteristic exponent by analytic continuation, and we have from Equations (2.10) and (2.11) (for $\beta \in \overline{\mathbb{C}^{\rightarrow}}$) under scheme 1,

$$\psi^h(\beta) = (\mu - \mu^h)\frac{e^{\beta h} - e^{-\beta h}}{2h} + (\sigma^2 + c_0^h)\frac{e^{\beta h} + e^{-\beta h} - 2}{2h^2} + \sum_{y\in\mathbb{Z}_h^{--}} c_y^h \left(e^{\beta y} - 1\right), \quad (4.2)$$

and under scheme 2,

$$\psi^h(\beta) = (\mu - \mu^h)\frac{e^{\beta h} - 1}{h} + c_0^h \frac{e^{\beta h} + e^{-\beta h} - 2}{2h^2} + \sum_{y\in\mathbb{Z}_h^{--}} c_y^h \left(e^{\beta y} - 1\right). \quad (4.3)$$

Note that, starting directly from (2.11), the term $(\mu - \mu^h)\frac{e^{\beta h}-1}{h}$ in (4.3) should actually read as:

$$(\mu - \mu^h)\left(\frac{e^{\beta h} - 1}{h}\mathbb{1}_{[0,\infty)}(\mu - \mu^h) + \frac{1 - e^{-\beta h}}{h}\mathbb{1}_{(-\infty,0]}(\mu - \mu^h)\right).$$

However, since X is a spectrally negative Lévy process, we have $\mu - \mu^h \geq 0$, at least for all sufficiently small h. For, if $\int_{[-1,0)} |y|\lambda(dy) < \infty$, then $\mu_0 > 0$ and by dominated convergence $\mu - \mu^h \to \mu_0$ as $h \downarrow 0$. On the other hand, if $\int_{[-1,0)} |y|\lambda(dy) = \infty$, then we deduce by monotone convergence $-\mu^h \geq \frac{1}{2}\int_{[-1,-h/2)} |y|\lambda(dy) \to \infty$ as $h \downarrow 0$. We may therefore assume that $h_\star$ is already chosen small enough, so that $\mu - \mu^h \geq 0$ holds for all $h \in (0, h_\star)$.

In summary, then, $h_\star$ is chosen so small as to guarantee that, for all $h \in$

$(0, h_\star)$: (I) $\mu - \mu^h \geq 0$ and (II) ψ^h is the Laplace exponent of some CP process X^h, which is also a CTMC with state space $\mathbb{Z}_h$. Equations (4.2) and (4.3) then determine the *weak* approximation $(X^h)_{h \in (0,h_\star)}$ precisely. This concludes step (i).

Next, it is easily seen that, for each $h \in (0, h_\star)$, X^h is in fact an upwards skip-free Lévy chain (establishing step (ii)): one need only check that $\lambda^h(\{h\}) > 0$, where λ^h is the Lévy measure of X^h, and this can be seen, e.g., from the explicit expression for ψ^h.

Finally, step (iii) is nothing other than a direct application of Proposition 3.35. In particular, we can express explicitly the coefficients of the linear recursion in (4.1) in terms of the Lévy triplet of X. Define:

$$\tilde{\sigma}_h^2 := \frac{1}{2h^2}\left(\sigma^2 + c_0^h\right), \qquad \tilde{\mu}^h := \frac{1}{2h}\left(\mu - \mu^h\right),$$

and note that $\tilde{\mu}^h$ is non-negative for $h \in (0, h_\star)$. Recall that V equals 0 or 1 according as to whether λ is finite or infinite and note that, if $V = 0$, we have $\tilde{\sigma}_h^2 = \sigma^2/2h^2$ and $\tilde{\mu}^h = \mu/2h$. We can now define the coefficients in (4.1) by:

$$\gamma := \lambda(-\infty, -h/2) + 2\tilde{\sigma}_h^2 + \mathbb{1}_{\{0\}}(\sigma^2)2\tilde{\mu}^h, \quad p := \left(\tilde{\sigma}_h^2 + \mathbb{1}_{(0,\infty)}(\sigma^2)\tilde{\mu}^h + \mathbb{1}_{\{0\}}(\sigma^2)2\tilde{\mu}^h\right)/\gamma, \tag{4.4}$$

$$q_1 := \left(\tilde{\sigma}_h^2 - \mathbb{1}_{(0,\infty)}(\sigma^2)\tilde{\mu}^h + c_{-h}^h\right)/\gamma, \quad q_k := c_y^h/\gamma, \quad \text{where} \quad y = -kh, \quad k \geq 2. \tag{4.5}$$

(observe that necessarily $\gamma > 0$).

4.4 Key attractions of algorithm

We conclude by summarising the key attractions of our algorithm (also in relation to the numerical procedures for Laplace inversion currently available for the evaluation of scale functions, see e.g. [Cohen, 2007] and [Kuznetsov et al., 2013, Chapter 5]):

(a) *consistency*: for each fixed $h > 0$, our algorithm calculates precisely (i.e. without numerical error, only modulo rounding) the values of a scale function for the process X^h, which weakly approximates X; i.e. no approximation is required to evaluate a scale function of X^h;

(b) *conceptual simplicity*: a weak approximation of X (as a Markov process) by a CTMC, which is skip-free to the right, provides a natural way of encoding the underlying probabilistic structure of the problem in the design of the algorithm;

(c) *robustness*: the method in (4.1) is valid for all spectrally negative Lévy processes;

(d) *straightforwardness of the algorithm*: the implementation requires only to solve a lower triangular system of linear equations (Matlab code [Mijatović, Vidmar, and Jacka, 2013a]) avoiding e.g. numerical complex integration;

(e) *convergence*: the rates of convergence of this algorithm have been established under mild assumptions, together with their optimality [Mijatović, Vidmar, and Jacka, 2013b] — these rates depend on the behaviour of the tail of the Lévy measure at the origin; by contrast behaviour of Laplace inversion algorithms tends to be susceptible to the degree of smoothness of the scale function itself (for which see [Chan et al., 2011]) [Abate and Whitt, 2006]).

Finally, it is worth (in part re-) emphasizing the key difference between our algorithm on the one hand, and *any* of the Laplace inversion techniques on the other. Indeed, the latter start with the Laplace transform of the scale functions (thus the Laplace exponent) as their basic datum, whilst the former derives the coefficients needed for its computation directly from the characteristic triplet (modulo the computation of the jump intensities of the approximating chain, of course). When the Laplace exponent is not given explicitly in terms of elementary/special functions, then with Laplace inversion techniques one would have necessarily to resort to an evaluation of the Laplace exponent (at complex values of its argument) via numerical integration — which appears disadvantageous as compared to the computation of the jump intensities of the approximating chain in our algorithm. In a sense, then, the scale functions go from being a Laplace exponent and then an inversion, to being just a limit process, away from the characteristic triplet — a result also of purely theoretical significance.

Bibliography

J. Abate and W. Whitt. A Unified Framework for Numerically Inverting Laplace Transforms. *INFORMS Journal on Computing*, 18(4):408–421, 2006.

A. H. Al-Mohy and N. J. Higham. Computing the Action of the Matrix Exponential, with an Application to Exponential Integrators. *SIAM J. Sci. Comput.*, 33 (2):488–511, 2011. MATLAB codes at `http://www.maths.manchester.ac.uk/~higham/papers/matrix-functions.php`.

D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2009.

F. Avram, A. E. Kyprianou, and M. R. Pistorius. Exit Problems for Spectrally Negative Lévy Processes and Applications to (Canadized) Russian Options. *Annals of Applied Probability*, 14(1):215–238, 2104.

V. Bally and D. Talay. The Law of the Euler Scheme for Stochastic Differential Equations: II. Convergence Rate of the Density. *Monte Carlo Methods and Applications*, 2(2):93–128, 2009.

J. Bertoin. *Lévy Processes.* Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1996.

R. N. Bhattacharya and E. C. Waymire. *A Basic Course in Probability Theory.* Universitext - Springer-Verlag. Springer, 2007.

N. H. Bingham. Continuous branching processes and spectral positivity. *Stochastic Processes and their Applications*, 4(3):217 – 242, 1976.

N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular Variation.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1987.

A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae.* Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002.

B. Böttcher and R. L. Schilling. Approximation of Feller processes by Markov chains with Lévy increments. *Stochastics and Dynamics*, 9(1):71–80, 2009.

M. Brown, E. A. Peköz, and S. M. Ross. Some results for skip-free random walk. *Probability in the Engineering and Informational Sciences*, 24(4):491–507, 2010.

H. Bühlmann. *Mathematical Methods in Risk Theory*. Grundlehren der mathematischen Wissenschaft: A series of comprehensive studies in mathematics. Springer, 1970.

R. M. R. Cardoso and H. R. Waters. Recursive calculation of finite time ruin probabilities under interest force. *Insurance: Mathematics and Economics* , 33 (3):659 – 676, 2003.

P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

T. Chan, A. E. Kyprianou, and M. Savov. Smoothness of scale functions for spectrally negative Lévy processes. *Probability Theory and Related Fields*, 150(3-4): 691–708, 2011.

K. L. Chung. *Markov chains with stationary transition probabilities*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin-Gottingen-Heidelberg, 1960.

E. Çinlar. *Probability and stochastics*. Graduate texts in mathematics. Springer New York, 2011.

A. M. Cohen. *Numerical methods for Laplace transform inversion*, volume 5 of *Numerical Methods and Algorithms*. Springer, New York, 2007.

S. Cohen and J. Rosiński. Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered and operator stable processes. *Bernoulli*, 13(1):195–210, 2007.

R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, 2004.

R. Cont and E. Voltchkova. A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models. *SIAM J. Numerical Analysis*, 43(4): 1596–1626, 2005.

J. Crosby, N. Le Saux, and A. Mijatović. Approximating Lévy processes with a view to option pricing. *International Journal of Theoretical and Applied Finance*, 13 (1):63–91, 2010.

D. C. M. Dickson and J. R. Gray. Approximations to ruin probability in the presence of an upper absorbing barrier. *Scandinavian Actuarial Journal*, 1984(2):105–115, 1984.

D. C. M. Dickson and H. R. Waters. Recursive calculation of survival probabilities. *ASTIN Bulletin*, 21(2):199–221, 1991.

R. A. Doney and A. E. Kyprianou. Overshoots and undershoots of Lévy processes. *Annals of Applied Probabilty*, 16(1):91–106, 2006.

R. A. Doney and J. Picard. *Fluctuation theory for Lévy processes: Ecole d'Eté de Probabilités de Saint-Flour XXXV-2005*. Number 1897 in Ecole d'Eté de Probabilités de Saint-Flour. Springer-Verlag, Berlin Heidelberg, 2007.

R. M. Dudley. *Real Analysis and Probability*. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, 2004.

S. Engelberg. *A Mathematical Introduction to Control Theory, Volume 2*. Series in Electrical And Computer Engineering. Imperial College Press, 2005.

J.-Cl. Evard and F. Jafari. A Complex Rolle's Theorem. *The American Mathematical Monthly*, 99(9):858–861, 1992.

J. E. Figueroa-López. Approximations for the distributions of bounded variation Lévy processes. *Statistics & Probability Letters*, 80(23-24):1744–1757, 2010.

D. Filipovic, M. Eberhard, and P. Schneider. Density Approximations for Multivariate Affine Jump-Diffusion Processes. *Journal of Econometrics*, 176:93–111, 2013.

M. G. Garroni and J. L. Menaldi. *Green functions for second order parabolic integro-differential problems*. Pitman research notes in mathematics series. Longman Scientific & Technical, 1992.

P. Glasserman. *Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability) (v. 53)*. Springer, 1st edition, 2003.

G. R. Grimmett and D. R. Stirzaker. *Probability and random processes (3rd edition)*. Oxford scientific publications, Oxford, 2001.

N. J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM J. Matrix Anal. Appl.*, 26(4):1179–1193, 2005.

F. Hubalek and A. E. Kyprianou. Old and New Examples of Scale Functions for Spectrally Negative Lévy Processes. In R. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, Progress in Probability, pages 119–145. Springer, 2011. [Online version; accessed 10 July 2013 at `http://arxiv.org/pdf/0801.0393v2.pdf`.].

J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes (2nd edition)*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, 2003.

O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York Berlin Heidelberg, 1997.

I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer, 1988.

J. Kiessling and R. Tempone. Diffusion approximation of Lévy processes with a view towards finance. *Monte Carlo Methods and Applications*, 17(1):11–45, 2011.

A. Klenke. *Probability theory: a comprehensive course*. Springer-Verlag London, 2008.

P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer-Verlag, Berlin; New York, 1992.

V. Knopova and R. L. Schilling. Transition density estimates for a class of Lévy and Lévy-type processes. *J. Theor. Probab.*, 25(1):144–170, 2012.

A. Kohatsu-Higa, S. Ortiz-Latorre, and P. Tankov. Optimal simulation schemes for Lévy driven stochastic differential equations. *Math. Comp.*, to appear. arXiv:1204.4877 [math.PR].

V. N. Kolokoltsov. *Markov Processes, Semigroups, and Generators*. De Gruyter Studies in Mathematics. De Gruyter, 2011.

A. Kuznetsov and M. Morales. Computing the finite-time expected discounted penalty function for a family of Lévy risk processes. *Scandinavian Actuarial Journal*, 2014(1):1–31, 2014.

A. Kuznetsov, A. E. Kyprianou, and J. C. Pardo. Meromorphic Lévy processes and their fluctuation identities. *The Annals of Applied Probability*, 22(3):1101–1135, 2012.

A. Kuznetsov, A. E. Kyprianou, and V. Rivero. The theory of scale functions for spectrally negative Lévy processes. In *Lévy Matters II*, Lecture Notes in Mathematics, pages 97–186. Springer Berlin Heidelberg, 2013.

A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications*. Springer-Verlag, Berlin Heidelberg, 2006.

A. E. Kyprianou. *Gerber-Shiu Risk Theory*. EEA Series. Springer, 2013.

A. E. Kyprianou, J. C. Pardo, and V. Rivero. Exact and asymptotic n-tuple laws at first and last passage. *Annals of Applied Probabilty*, 20(2):522–564, 2010.

A. Lambert. Species abundance distributions in neutral models with immigration or mutation and general lifetimes. *Journal of Mathematical Biology*, 63(1):57–72, 2011.

R. L. Loeffen. On optimality of the barrier strategy in de Finetti's dividend problem for spectrally negative Lévy processes. *Annals of Applied Probability*, 18(5):1669–1680, 2008.

D. B. Madan, P. Carr, and E. C. Chang. The Variance Gamma Process and Option Pricing. *European Finance Review*, 2:79–105, 1998.

P. Marchal. A Combinatorial Approach to the Two-Sided Exit Problem for Left-Continuous Random Walks. *Combinatorics, Probability and Computing*, 10:251–266, 2001.

P. A. Meyer. *Probability and potentials*. Blaisdell book in pure and applied mathematics. Blaisdell Pub. Co., 1966.

A. Mijatović. Spectral properties of trinomial trees. *Proceedings of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, 463(2083): 1681–1696, 2007.

A. Mijatović and M. R. Pistorius. On the drawdown of completely asymmetric Lévy processes. *Stochastic Processes and their Applications*, 122(11):3812–3836, 2012.

A. Mijatović, M. Vidmar, and S. Jacka. Matlab code for the scale function algorithm in Equation (4.1) above, 2013a. Available at `http://www.ma.ic.ac.uk/~amijatov/Abstracts/ScaleFun_Levy.html`.

A. Mijatović, M. Vidmar, and S. Jacka. Markov chain approximations to scale functions of Lévy processes. arXiv:1310.1737 [math.PR], 2013b.

C. Moler and C. V. Loan. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. *SIAM Review*, 45(1):3–49, 2003.

J. R. Norris. *Markov chains*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, 1997.

S. Orey. On continuity properties of infinitely divisible distribution functions. *Annals of Mathematical Statistics*, 39(3):936–937, 1968.

K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York and London, 1967.

J. Poirot and P. Tankov. Monte Carlo Option Pricing for Tempered Stable (CGMY) Processes. *Asia-Pacific Financial Markets*, 13(4):327–344, 2006.

M. P. Quine. On the escape probability for a left or right continuous random walk. *Annals of Combinatorics*, 8:221–223, 2004.

M. Reed and B. Simon. *Methods of modern mathematical physics. I.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin Heidelberg, 1999.

J. Rosiński. Simulation of Lévy processes. In *Encyclopedia of Statistics in Quality and Reliability: Computationally Intensive Methods and Simulation*. Wiley, 2008.

W. Rudin. *Real and complex analysis*. International student edition. McGraw-Hill, 1970.

K. I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, 1999.

R. B. Sidje. Expokit: a software package for computing matrix exponentials. *ACM Trans. Math. Softw.*, 24(1):130–156, 1998. MATLAB codes at `http://www.maths.uq.edu.au/expokit/`.

F. Spitzer. *Principles of Random Walk*. Graduate texts in mathematics. Springer, 2001.

I. Stewart and D. Tall. *Complex analysis*. Cambridge University Press, Cambridge, 1983.

R. Syski. *Passage Times for Markov Chains*. Studien Zur Osterreichischen Philosophie. IOS Press, 1992.

A. Szimayer and R. A. Maller. Finite approximation schemes for Lévy processes, and their application to optimal stopping problems. *Stochastic Processes and their Applications*, 117(10):1422–1447, 2007.

P. Sztonyk. Transition density estimates for jump Lévy processes. *Stochastic Processes and their Applications*, 121(6):1245–1265, 2011.

H. Tanaka and A. Kohatsu-Higa. An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs. *The Annals of Applied Probability*, 19(3):1026–1062, 2009.

F. De Vylder and M. J. Goovaerts. Recursive calculation of finite-time ruin probabilities. *Insurance: Mathematics and Economics*, 7(1):1–7, 1988.

J. Yeh. *Real Analysis: Theory of Measure And Integration*. World Scientific, second edition, 2006.

Appendix A

Two lemmas on conditioning

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space. Recall that the symbol $\perp$ is used to indicate stochastic independence relative to the probability measure P, whereas the completion of a σ-field $\mathcal{S}$ relative to the measure μ is denoted $\overline{\mathcal{S}}^{\mu}$, $\overline{\mu}$ being the unique extension of μ to $\overline{\mathcal{S}}^{\mu}$.

Proposition A.1 (Basic lemma on conditioning)**.** *Let $Y : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ and $Z : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$ be two random elements, and $\mathcal{G}$ any sub-σ-algebra of $\mathcal{F}$, such that $\sigma(Y) \subset \mathcal{G}$ and $\sigma(Z) \perp \mathcal{G}$. Let f be any bounded (or nonnegative, or nonpositive) $\mathcal{S} \otimes \mathcal{T}/\mathcal{B}([-\infty, +\infty])$-measurable mapping. Then for any $y \in S$, $f \circ (y, Z)$ is $\mathcal{F}/\mathcal{B}([-\infty, +\infty])$-measurable, $g := (y \mapsto \mathsf{E}[f \circ (y, Z)])$ is $\mathcal{S}/\mathcal{B}([-\infty, +\infty])$-measurable and, P-a.s.,*

$$\mathsf{E}[f \circ (Y, Z)|\mathcal{G}] = g \circ Y. \tag{A.1}$$

Proof. Linearity and monotonicity of conditional expectation [Çinlar, 2011, p. 143] show that the class of functions f for which the conclusion of the lemma holds true is a monotone class. By the Functional Monotone Class Theorem [Çinlar, 2011, p. 10, Theorem 2.19], it is then sufficient to check its validity for $f = \mathbb{1}_\Lambda$ with Λ belonging to the π-system $\{A \times B : (A, B) \in \mathcal{S} \times \mathcal{T}\}$ generating $\mathcal{S} \otimes \mathcal{T}$. In that case (A.1) (measurability being clear) follows at once by independence of Y and Z [Klenke, 2008, p. 174, Theorem 8.14(vi)] and the "taking out what is known" property (conditional determinism [Çinlar, 2011, p. 144, Theorem 1.10(a)]) of conditional expectation. $\square$

There is a modification of this proposition, which allows for completions, to wit:

Proposition A.2 (Lemma on conditioning with completions)**.** *Assume now* $(\mathcal{F}, \mathsf{P})$ *is complete. Let* $Y : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ *and* $Z : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$ *again be two random elements, and* $\mathcal{G}$ *any sub-*σ*-algebra of* $\mathcal{F}$*, such that* $\sigma(Y) \subset \mathcal{G}$ *and* $\sigma(Z) \perp \mathcal{G}$*. Let* f *be any bounded (or nonnegative, or nonpositive)* $\overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}/\mathcal{B}([-\infty, +\infty])$*-measurable mapping. Then:*

(i) (Y, Z) *is* $\mathcal{F}/\overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}$*-measurable,*

(ii) Y *(respectively* Z*) is* $\mathcal{F}/\overline{\mathcal{S}}^{\mathsf{P}_Y}$*-measurable (respectively* $\mathcal{F}/\overline{\mathcal{T}}^{\mathsf{P}_Z}$*-measurable),*

(iii) $\overline{\mathsf{P}_Y}$*-a.s. in* $y \in S$*,* $f \circ (y, Z)$ *is* $\mathcal{F}/\mathcal{B}([-\infty, +\infty])$*-measurable,*

(iv) $(y \mapsto \mathsf{E}[f \circ (y, Z)])$ *is* $\overline{\mathcal{S}}^{\mathsf{P}_Y}/\mathcal{B}([-\infty, +\infty])$*-measurable (defining* $\mathsf{E}[f \circ (y, Z)]$ *to be, say,* 0*, on the* $\overline{\mathsf{P}_Y}$*-negligible set in* $y \in S$*, on which* $f \circ (y, Z)$ *is not* $\mathcal{F}/\mathcal{B}([-\infty, +\infty])$*-measurable)*

and, P*-a.s.,*

$$\mathsf{E}[f \circ (Y, Z)|\mathcal{G}] = (y \mapsto \mathsf{E}[f \circ (y, Z)]) \circ Y. \tag{A.2}$$

Proof. Throughout we use the Image Measure Theorem [Dudley, 2004, p. 121, Theorem 4.1.11].

First note that (Y, Z) is $\mathcal{F}/\mathcal{S} \otimes \mathcal{T}$-measurable, hence it is $\mathcal{F}/\overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}$-measurable, since $\mathcal{F}$ is P-complete. Similarly for Y and Z. (In both cases apply a generating class argument combining [Dudley, 2004, pp. 101-102, Theorem 3.3.1 and Propositions 3.3.2 & 3.3.3], cf. also [Kallenberg, 1997, p. 21, Exercise 8].) Thus we have (i) and (ii).

Next, the measure spaces $(S, \overline{\mathcal{S}}^{\mathsf{P}_Y}, \overline{\mathsf{P}_Y})$ and $(T, \overline{\mathcal{T}}^{\mathsf{P}_Z}, \overline{\mathsf{P}_Z})$ are complete and, by [Yeh, 2006, p. 543, Theorem 23.23], $\overline{\overline{\mathcal{S}}^{\mathsf{P}_Y} \otimes \overline{\mathcal{T}}^{\mathsf{P}_Z}}^{\overline{\mathsf{P}_Y} \times \overline{\mathsf{P}_Z}} = \overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}$, since $\mathsf{P}_Y \times \mathsf{P}_Z = \mathsf{P}_{(Y,Z)}$, owing to independence of Y and Z. It follows that f is $\overline{\overline{\mathcal{S}}^{\mathsf{P}_Y} \otimes \overline{\mathcal{T}}^{\mathsf{P}_Z}}^{\overline{\mathsf{P}_Y} \times \overline{\mathsf{P}_Z}}/\mathcal{B}([-\infty, +\infty])$-measurable. The latter allows us to conclude (iii) and (iv), as follows.

First, by [Yeh, 2006, p. 545, Theorem 23.25(b)], $f(y, \cdot)$ is $\overline{\mathcal{T}}^{\mathsf{P}_Z}/\mathcal{B}([-\infty, +\infty])$-measurable, $\overline{\mathsf{P}_Y}$-a.s. in $y \in S$. Coupled with (ii), this yields (iii). Second, note that for any $y \in S$ for which $f(y, \cdot)$ is $\overline{\mathcal{T}}^{\mathsf{P}_Z}/\mathcal{B}([-\infty, +\infty])$-measurable, $\mathsf{E}[f \circ (y, Z)] = \int f(y, \cdot) d\overline{\mathsf{P}_Z}$. Thus (iv) follows by Tonelli's Theorem [Yeh, 2006, p. 546, Theorem 23.26(a)].

Finally we wish to establish (A.2). As in Lemma A.1, linearity and monotonicity of conditional expectation show that the class of $\overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}/\mathcal{B}([-\infty, +\infty])$-measurable functions f for which (A.2) holds is a monotone class. By the Functional

Monotone Class Theorem it will thus be sufficient to consider $f = \mathbb{1}_\Lambda$ with Λ belonging to the π-system $\{A \times B : (A, B) \in \mathcal{S} \times \mathcal{T}\} \cup \mathcal{N}$, where $\mathcal{N}$ is the set of all $\overline{\mathsf{P}_{(Y,Z)}}$-null sets, generating $\overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}}$ [Dudley, 2004, p. 102, Proposition 3.3.2].

Now, for Λ belonging to $\{A \times B : (A, B) \in \mathcal{S} \times \mathcal{T}\}$, (A.2) is the contents of Proposition A.1. On the other hand suppose Λ is $\overline{\mathsf{P}_{(Y,Z)}}$-null. Then, P-a.s., the left-hand side of (A.2) is equal to 0, since $\overline{\mathsf{P}_{(Y,Z)}}$ coincides with the law of (Y, Z) on $(S \times T, \overline{\mathcal{S} \otimes \mathcal{T}}^{\mathsf{P}_{(Y,Z)}})$ (the extension of a law to its completed σ-field being unique), and hence $\mathsf{E}[f \circ (Y, Z)] = \int f d\overline{\mathsf{P}_{(Y,Z)}} = 0$. The right-hand side of (A.2) is nonnegative. To show that it too is 0, P-a.s., compute again its expectation using Tonelli's Theorem [Yeh, 2006, p. 546, Theorem 23.26] and the fact that by [Yeh, 2006, p. 543, Theorem 23.23] $\overline{\mathsf{P}_{(Y,Z)}} = \overline{\overline{\mathsf{P}_Y} \times \overline{\mathsf{P}_Z}}$ (where $\overline{\mathsf{P}_Y}$ and $\overline{\mathsf{P}_Z}$ are also the laws of Y and Z on the completed spaces $(S, \overline{\mathcal{S}}^{\mathsf{P}_Y})$ and $(T, \overline{\mathcal{T}}^{\mathsf{P}_Z})$, respectively):

$$\int d\overline{\mathsf{P}_Y}(y) \int d\overline{\mathsf{P}_Z}(z) f(y, z) = \int f d\overline{\overline{\mathsf{P}_Y} \times \overline{\mathsf{P}_Z}} = \int f d\overline{\mathsf{P}_{(Y,Z)}} = 0. \tag{A.3}$$

Thus indeed also the right-hand side of (A.2) equals 0, P-a.s., and the proof is complete. $\square$

Appendix B

Continuous-time random walk reflected at its maximum

Given a compound Poisson process on $\mathbb{Z}_h := \{hk : k \in \mathbb{Z}\}$ (i.e. (modulo a spatial scaling) given a random walk embedded into continuous time as a CP process), we rigorously establish the infinitesimal generator of its reflected process at the maximum (for the definition of the latter, see Definition 1.24). In the sequel $l^\infty(S)$ will denote the Banach space of bounded complex-valued functions on a denumerable set S, endowed with the supremum norm; a Q-matrix is said to be regular when its entries are uniformly bounded.

To begin with, let $Y = (Y_t)_{t\geq 0}$ be a Markov process [Çinlar, 2011, Chapter IX] on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ with state space $(S, 2^S)$, where S is denumerable, and define for $0 \leq s \leq t$, $f \in l^\infty(S)$ and $u \in S$:

$$(P(s,t)f)(u) := \sum_{u' \in S} f(u')\mathsf{P}(Y_t = u'|Y_s = u),$$

if $\mathsf{P}(Y_s = u) \neq 0$ and let $(P(s,t)f)(u) := f(u)$ otherwise. Clearly $P(s,t) : l^\infty(S) \to l^\infty(S)$, $\|P(s,t)\| = 1$. Moreover, $\mathsf{E}[f \circ Y_t|\mathcal{F}_s] = \mathsf{E}[f \circ Y_t|Y_s] = (P(s,t)f) \circ Y_s$ (P-a.s.) by the Markov property.

If we now let $l^\infty_t(S)$ denote the set of equivalence classes of $l^\infty(S)$ with respect to the measure $\mathsf{P} \circ Y_t^{-1}$, then clearly we have a canonical way of forcing $P(s,t) : l^\infty_t(S) \to l^\infty_s(S)$ and we shall use the same symbol for this enforcement. Moreover, for $f \in l^\infty(S)$ and $s \geq 0$, we shall let $\|f\|_s$ stand for the essential supremum of f with respect to the measure P_{Y_s} and we observe that this is (in a canonical way) a norm on $l^\infty_s(S)$ and a seminorm on $l^\infty(S)$; $P(s,t)$ is a bounded linear operator from $l^\infty_t(S)$ to $l^\infty_s(S)$ and its norm is 1. When viewed as such, clearly for $0 \leq s \leq t \leq u$ one has

$P_{s,t}P_{t,u} = P_{s,u}$, since for any $f \in l^\infty_u(S)$, P-a.s., $(P_{s,t}P_{t,u}f)\circ Y_s = \mathsf{E}[(P_{t,u}f)\circ Y_t|\mathcal{F}_s] = \mathsf{E}[\mathsf{E}[f \circ Y_u|\mathcal{F}_t]|\mathcal{F}_s] = \mathsf{E}[f\circ Y_u|\mathcal{F}_s] = (P_{s,u}f)\circ Y_s$ and hence $P_{s,t}P_{t,u}f = P_{s,u}f$ (in $l^\infty_s(S)$). In the sequel, however, we shall view $P(s,t)$ as mapping on quotient spaces only if we explicitly say so.

With these preliminaries having been established, we have the following key:

Proposition B.1 (Extracting the generator)**.** *Let $L : l^\infty(S) \to l^\infty(S)$ be a bounded linear operator. Suppose:*

$$\lim_{u\downarrow 0}\sup_{t\geq 0}\sup_{f\in l^\infty(S),\|f\|\leq 1}\left\|\left(\frac{P(t,t+u)-I}{u}-L\right)f\right\|_t = 0 \tag{B.1}$$

(i.e. $P(t,t+u)$ admits L as a P_{Y_t}-essential right derivative at $u=0$, uniformly in t). Then for every $0 \leq s \leq t$, $P(s,t)f = e^{L(t-s)}f$ (P_{Y_s}-a.s.) for every $f \in l^\infty(S)$, i.e. L is the infinitesimal generator of Y and Y is a time-homogeneous Markov process with transition semigroup $(P_t)_{t\geq 0} = (e^{Lt})_{t\geq 0}$.

Proof. Fix $0 \leq s \leq t$. Condition (B.1) says that for each $n \in \mathbb{N}$ and each $0 \leq k \leq n-1$, $P(s+\frac{k}{n}(t-s), s+\frac{k+1}{n}(t-s)) = I + L(t-s)/n + ((t-s)/n)B_n(k)$ where $\alpha_n := \sup_{0\leq k\leq n-1}\sup_{f\in l^\infty(S),\|f\|\leq 1}\|B_n(k)f\|_{s+\frac{k}{n}(t-s)} \to 0$ as $n\to\infty$. Viewed as operators on appropriate quotient spaces, we have $P(s,t) = \prod_{k=0}^{n-1}P(s+\frac{k}{n}(t-s), s+\frac{k+1}{n}(t-s))$; consequently for every $f \in l^\infty(S)$, $\|P(s,t)f - \prod_{k=0}^{n-1}P(s+\frac{k}{n}(t-s), s+\frac{k+1}{n}(t-s))f\|_s = 0$. Next, proceeding via a telescopic sum:

$$\begin{aligned}
&\prod_{k=0}^{n-1}P\left(s+\frac{k}{n}(t-s), s+\frac{k+1}{n}(t-s)\right) - \left(I+\frac{t-s}{n}L\right)^n = \\
= \;& P\left(s,s+\frac{1}{n}(t-s)\right)\cdots P\left(s+\frac{n-2}{n}(t-s), s+\frac{n-1}{n}(t-s)\right)P\left(s+\frac{n-1}{n}(t-s),t\right)\\
- \;& P\left(s,s+\frac{1}{n}(t-s)\right)\cdots P\left(s+\frac{n-2}{n}(t-s), s+\frac{n-1}{n}(t-s)\right)\left(I+\frac{t-s}{n}L\right)+\cdots+\\
+ \;& P\left(s,s+\frac{1}{n}(t-s)\right)\left(I+\frac{t-s}{n}L\right)^{n-1} - \left(I+\frac{t-s}{n}L\right)^n
\end{aligned}$$

so that for every $f \in l^\infty(S)$,

$$\begin{aligned}
&\left\|\left(\prod_{k=0}^{n-1}P\left(s+\frac{k}{n}(t-s), s+\frac{k+1}{n}(t-s)\right)\right)f - \left(I+\frac{t-s}{n}L\right)^n f\right\|_s\\
\leq\;& \frac{t-s}{n}\|f\|\alpha_n\left(1+(1+(t-s)\|L\|/n)+\cdots+(1+(t-s)\|L\|/n)^{n-1}\right)\\
\leq\;& \alpha_n(t-s)\left(1+(t-s)\|L\|/n\right)^n\|f\| \to 0
\end{aligned}$$

as $n\to\infty$ (where we have used the seminorm triangle inequality, and, in addition, the fact that $\|P(u_1,u_2)g\|_{u_1} \leq \|g\|_{u_2}$, for all $0 \leq u_1 \leq u_2$ and $g \in l^\infty(S)$). Finally

$\|(I + \frac{t-s}{n}L)^n f - e^{(t-s)L} f\|_s \to 0$ as $n \to \infty$, since in fact $(I + \frac{t-s}{n}L)^n \to e^{(t-s)L}$ in $l^\infty(S)$.[1] $\square$

Now take X to be a compound Poisson process with values in $\mathbb{Z}_h = h\mathbb{Z}$, càdlàg with certainty, Lévy measure λ, living on $(\Omega, \mathcal{F}, \mathsf{P})$. Denote the running supremum process by $\overline{X}$ and the reflected process by $Y := \overline{X} - X$. On account of the stationarity and independence of increments of X, the latter is a Markov process in turn [Bertoin, 1996, p. 156, Proposition 1], with values in $\mathbb{Z}_h^+$. Define also the Q-matrix $\tilde{Q} : \mathbb{Z}_h^+ \times \mathbb{Z}_h^+ \to \mathbb{R}$ by demanding:

$$\begin{aligned} \tilde{Q}_{uu'} &:= \lambda([u, \infty)) - \delta_{uu'}\lambda(\mathbb{R}), \text{ if } u' = 0 \\ \tilde{Q}_{uu'} &:= \lambda(\{u - u'\}) - \delta_{uu'}\lambda(\mathbb{R}), \text{ if } u' > 0 \end{aligned}$$

($\{u, u'\} \subset \mathbb{Z}_h^+$). Manifestly $\tilde{Q}$ is regular, albeit it is not spatially homogeneous (but it does verify the Feller condition, cf. Proposition 1.35).

Proposition B.2 (Generator of the reflected process)**.** *Let* $0 \leq v$. *Then:*

$$\lim_{v\downarrow 0} \sup_{t\geq 0} \mathsf{P}_{Y_t}-\text{ess}\sup_{u\in\mathbb{Z}_h^+} \sum_{u'\in\mathbb{Z}_h^+} \left| \frac{\mathsf{P}(Y_{t+v} = u' | Y_t = u) - \delta_{uu'}}{v} - \tilde{Q}_{uu'} \right| = 0.$$

Consequently, if we consider Y *as living on* $\mathbb{Z}_h^+$ *in Proposition B.1 and associate to* $\tilde{Q}$ *the mapping* $\tilde{L} : l^\infty(\mathbb{Z}_h^+) \to l^\infty(\mathbb{Z}_h^+)$ *via* $\tilde{L}f(u) = \sum_{u'\in\mathbb{Z}_h^+} \tilde{Q}_{uu'} f(u')$ *(*$f \in l^\infty(\mathbb{Z}_h^+)$*,* $u \in \mathbb{Z}_h^+$*), we have also:*

$$\lim_{v\downarrow 0} \sup_{t\geq 0} \sup_{f\in l^\infty(\mathbb{Z}_h^+), \|f\|\leq 1} \left\| \left(\frac{P(t, t+v) - I}{v} - \tilde{L} \right) f \right\|_t = 0.$$

Thus $\tilde{L}$ *is the infinitesimal generator of* Y*.*

Proof. Let $\{u, u'\} \subset \mathbb{Z}_h^+$, suppose $\mathsf{P}(Y_t = u) > 0$, denote $q := \lambda(\mathbb{R})$. Le T be the time *to* the second jump of X strictly after t (that is to say, T is the second jump time of the incremental process $\overset{\triangle}{X} := (X_{t+s} - X_t)_{s\geq 0}$, which is independent of $X|_{[0,t]}$ and thus of $Y|_{[0,t]}$). Then $\sum_{u'\in\mathbb{Z}_h^+} \mathsf{P}(\{Y_{t+v} = u'\} \cap \{T \leq v\} | Y_t = u) = \mathsf{P}(T \leq v) = \int_0^v qse^{-qs} d(qs) \leq (qv)^2/2$ (since T is independent of Y_t and has law $\text{Exp}(q) \star \text{Exp}(q)$). Next note that $\{T > v\}$ is the disjoint union of the two events

[1]Indeed, if A is a bounded linear operator on a Banach space, one has $(I + A/n)^n \to e^A$ as $n \to \infty$. This follows from the same relation for real numbers: $\|e^A - (I + A/n)^n\| = \|\sum_{k=0}^\infty \frac{A^k}{k!} - \sum_{k=0}^n \binom{n}{k} \frac{A^k}{n^k}\| \leq \sum_{k=0}^n \frac{\|A\|^k}{k!}(1 - \frac{n\cdots(n-k+1)}{n^k}) + \sum_{k=n+1}^\infty \frac{\|A\|^k}{k!} = e^{\|A\|} - (1 + \frac{\|A\|}{n})^n \to 0$ as $n \to \infty$.

corresponding to the incremental process $\overset{\triangle}{X}$ having no jumps and having precisely one jump in the interval $[0, t]$, respectively. Thus if $u' = 0$, then:

$$\mathsf{P}(\{Y_{t+v} = u'\} \cap \{T > v\} | Y_t = u) = \delta_{uu'} e^{-qv} + ve^{-qv} \lambda([u, \infty)),$$

whilst if $u' > 0$, then:

$$\mathsf{P}(\{Y_{t+v} = u'\} \cap \{T > t\} | Y_t = u) = \delta_{uu'} e^{-qv} + \lambda(\{u - u'\}) ve^{-qv}.$$

The first claim follows, with it the second, and the final one obtains from Proposition B.1. $\square$

Appendix C

The Kolmogorov consistency theorem and martingale change of measure

We prepare first the notational apparatus.

Notation C.1 (The space $\mathbb{D}$). Let $\mathbb{D}$ denote the Skorohod space [Jacod and Shiryaev, 2003, Chapter VI, Section 1] of càdlàg paths in $\mathbb{R}^{[0,\infty)}$ endowed with the Skorohod metric (making it into a Polish (so separable and metrizable for a complete metric) topological space), and the corresponding Borel σ-algebra $\mathcal{F}$, which coincides with the σ-algebra $\sigma(\mathrm{pr}_t : t \geq 0)$ generated by all the evaluation maps. We equip $\mathbb{D}$ also with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, where $\mathcal{F}_t := \sigma(\mathrm{pr}_s : s \in [0,t])$ is generated by evaluations up to time t. Clearly $\mathcal{F}_\infty = \mathcal{F}$ is the terminal σ-field in this setting.

In the same vein, for each $t \geq 0$, let $\mathbb{D}[0,t]$ be the space of càdlàg paths in $\mathbb{R}^{[0,t]}$ [Parthasarathy, 1967, Chapter VII, Section 6] endowed with the Skorohod metric (making it into a Polish topological space) and the corresponding Borel σ-algebra $\mathcal{F}[0,t]$, which coincides with the σ-algebra $\sigma(\mathrm{pr}_s : s \in [0,t])$ generated by all the evaluation maps.

Definition C.2 (Standard space and atoms)**.** A measurable space $(X, \mathcal{B})$ is standard, if $\mathcal{B}$ is σ-isomorphic to the Borel σ-algebra $\mathcal{B}_0$ of some Polish space, i.e. there exists a mapping $\tau : \mathcal{B} \to \mathcal{B}_0$, one-to-one and onto, and preserving countable set operations [Parthasarathy, 1967, p. 133, Definition 2.2]. An atom of $(X, \mathcal{B})$ is a set $A_0 \in \mathcal{B}\backslash\{\emptyset\}$, such that $\mathcal{B} \ni A \subset A_0$ implies $A \in \{A_0, \emptyset\}$.

Remark C.3. Every standard measurable space is automatically countably generated (i.e. there is a countable subset of its σ-algebra $\mathcal{B}$, generating $\mathcal{B}$). Further, the σ-algebras $\mathcal{F}_t$ and $\mathcal{F}[0,t]$ are not the same, but they are σ-isomorphic (and hence $(\mathbb{D}, \mathcal{F}_t)$ is standard in the sense of Definition C.2 for every $t \geq 0$). Indeed, the mapping $\tau_t := (A \mapsto \mathrm{pr}_{[0,t]}^{-1}(A) = \{\omega \in \mathbb{D} : \omega|_{[0,t]} \in A\})$ preserves countable set operations, hence is from $\mathcal{F}[0,t]$ *into* $\mathcal{F}_t$ (apply a generating class argument), it is *one-to-one* (note here that one can always extend $\omega \in \mathbb{D}[0,t]$ by $\omega(t)$ to get an element of $\mathbb{D}$), and it is *onto* (again apply a generating class argument). Consequently, the atoms of $(\mathbb{D}, \mathcal{F}_t)$ are precisely sets of the form $\tau_t(\{\omega\})$ for $\omega \in \mathbb{D}[0,t]$. We have used the fact that singletons are measurable in $(\mathbb{D}[0,t], \mathcal{F}[0,t])$ which is clear, since we are dealing with the Borel σ-algebra of a T_1 topological space.

Next we cite from [Parthasarathy, 1967, p. 143, Theorem 4.2] the following theorem, allowing one to extend a consistent family of probability measures defined on some collection of sub-σ-algebras, to the one generated by them:

Theorem C.4 (Kolmogorov extension theorem)**.** *Let $(X, \mathcal{B})$ be a measurable space, Δ a directed set*[1] *under an ordering $<$ and let there be given a family $(\mathcal{B}_\alpha)_{\alpha \in \Delta}$ of sub-σ-algebras of $\mathcal{B}$. Suppose:*

(a) the family is a filtration, i.e. $\mathcal{B}_\alpha \subset \mathcal{B}_{\alpha'}$ whenever $\alpha < \alpha'$ are from Δ;

(b) each $(X, \mathcal{B}_\alpha)$ is standard ($\alpha \in \Delta$);

(c) $\sigma(\mathcal{B}_\alpha : \alpha \in \Delta) = \mathcal{B}$;

(d) for any sequence $A_1, A_2, \ldots$ with $A_1 \supset A_2 \supset \ldots$ and with A_n an atom of $\mathcal{B}_{\alpha_n}$, $\alpha_1 < \alpha_2 < \ldots$, one has $\cap_{n \in \mathbb{N}} A_n \neq \emptyset$.

Then, given any consistent family of probability measures $(\mu_\alpha : \alpha \in \Delta)$ (i.e. whenever $\alpha < \alpha'$ are from Δ, $\mu_{\alpha'}|_{\mathcal{B}_\alpha} = \mu_\alpha$), there exists a unique probability measure μ on $\mathcal{B}$ extending this family (i.e. $\mu|_{\mathcal{B}_\alpha} = \mu_\alpha$ for every $\alpha \in \Delta$).

This allows us to establish:

Proposition C.5 (Local martingale change of measure)**.** *Assume the process $M = (M_t)_{t \geq 0}$ is a martingale on the filtered space $(\mathbb{D}, \mathcal{F}, \mathbb{F})$ under some measure P, with $\mathsf{E}[M_0] = 1$ and $M_t \geq 0$ P-a.s. for every $t \geq 0$. Then the family $(\mathsf{P}_t^\natural)_{t \geq 0}$ given by*

$$\mathsf{P}_t^\natural(A) = \mathsf{E}[M_t \mathbb{1}_A] \quad (A \in \mathcal{F}_t, t \geq 0)$$

extends uniquely to a probability measure $\mathsf{P}^\natural$ on $\mathcal{F}$.

[1] Meaning that Δ is non-empty, and $<$ is an irreflexive, transitive binary operation with any two elements having an upper bound.

Proof. Since $\mathsf{E}[M_t] = 1$ for all $t \geq 0$, we are indeed dealing with a family of probability measures (apply DCT). The family is consistent by the martingale property of M. Hence, applying Remark C.3, the conditions of the Kolmogorov extension theorem are fulfilled for the family $(\mathsf{P}^\natural_n)_{n\in\mathbb{N}_0}$ and the result obtains at once.[2] $\square$

Remark C.6. In Proposition C.5, $\mathsf{P}^\natural \ll \mathsf{P}$ on restriction to the algebra $\mathcal{A} := \cup_{t\geq 0}\mathcal{F}_t$. If $M_t > 0$, P-a.s., for each $t \geq 0$, then also $\mathsf{P} \ll \mathsf{P}^\natural$ on restriction to $\mathcal{A}$.

Remark C.7 (Martingale change of measure). Provided M is uniformly integrable (UI), then in Proposition C.5 the filtered probability space need not be the canonical one, and the conclusion still holds in the sense that the unique extension exists on the terminal σ-field $\mathcal{F}_\infty := \sigma(\mathcal{A})$. Indeed, note that by the martingale property, the family of measures $(\mathsf{P}^\natural_t)_{t\geq 0}$ clearly extends to a unique finitely-additive set function $\mathsf{P}^\natural_\infty$ on the algebra $\mathcal{A}$. By the theorem of Carathéodory, it is sufficient to show that (*) for every sequence $(A_i)_{i\geq 1}$ of disjoint sets in $\mathcal{A}$, with $A := \cup_{i\geq 1}A_i \in \mathcal{A}$, $\mathsf{P}^\natural_\infty(A) = \sum_{i\geq 1}\mathsf{P}^\natural_\infty(A_i)$. We then get existence of $\mathsf{P}^\natural$, uniqueness being clear by a π/λ-argument. To show (*), let $A_i \in \mathcal{F}_{t_i}$ in nondecreasing order, $i \geq 1$, $A := \cup_{i\geq 1}A_i \in \mathcal{F}_t$. Then for each $n \in \mathbb{N}$, by finite additivity $\mathsf{P}^\natural_\infty(A) = \mathsf{P}^\natural_\infty((\cup_{1\leq i\leq n}A_i) \cup (\cup_{i\geq n+1}A_i)) = \left(\sum_{i=1}^n \mathsf{P}^\natural_\infty(A_i)\right) + \mathsf{E}[M_{t_n\vee t}\mathbb{1}_{\cup_{i\geq n+1}A_i}]$. Now let $n \to \infty$. By the UI property, the second term converges to 0, since $\mathsf{P}(\cup_{i\geq n+1}A_i) \to 0$ as $n \to \infty$ [Klenke, 2008, p. 137, Theorem 6.24(iii)]. In the general case, the same argument shows that $\mathsf{P}^\natural_\infty$ is a countably super-additive (and finitely additive) set-function on the algebra $\mathcal{A}$, in the sense that for every sequence $(A_i)_{i\geq 1}$ of disjoint sets in $\mathcal{A}$, with $A := \cup_{i\geq 1}A_i \in \mathcal{A}$, $\mathsf{P}^\natural_\infty(A) \geq \sum_{i\geq 1}\mathsf{P}^\natural_\infty(A_i)$ (with equality, if all but finitely many A_i, $i \geq 1$, are empty).

Moreover, when this is so (i.e. M is UI), by taking a sequence $0 \leq t_n \uparrow \infty$, $(M_{t_n})_{n\in\mathbb{N}}$ is a UI discrete-time martingale which converges P-a.s. to some random variable M_∞, necessarily nonnegative P-a.s., and with $\mathsf{E}[M_\infty] = 1$. We conclude that $\mathsf{P}^\natural(A) = \mathsf{E}[M_\infty\mathbb{1}_A]$ for each $A \in \mathcal{A}$ and hence each $A \in \mathcal{F}_\infty$ by a π/λ-argument, i.e. $\mathsf{P}^\natural \ll \mathsf{P}$ and M_∞ is the Radon-Nikodym derivative. If $M_\infty > 0$ P-a.s., then also $\mathsf{P} \ll \mathsf{P}^\natural$.

See also e.g. [Revuz and Yor, 1999, pp. 325-326].

Remark C.8 (Smaller spaces). Finally note that $C[0,\infty) := \{\omega \in \mathbb{D} : \omega \text{ continuous}\}$ and $\mathbb{D}_h := \{\omega \in \mathbb{D} : \omega \text{ has values in } \mathbb{Z}_h\}$ are measurable subsets of $(\mathbb{D}, \mathcal{F})$. The second is so by right-continuity of the sample paths, and the first was shown to be as such in Lemma 3.7 of Subsection 3.1.3. Let $\mathbb{H}$ be generic for $\mathbb{D}_h$ or $C[0,\infty)$. Then

[2] Note that, in applying Theorem C.4, it is necessary to restrict oneself to a sequence of times $t_n \uparrow \infty$ (say, $t_n = n$, as was the case), as $n \to \infty$, because only then is it the case that (d) thereof obtains. There are issues with having $t_n \uparrow t^* < \infty$, as $n \to \infty$!

it is easy to see that for every $0 \le t \le \infty$, the trace σ-algebra of $\mathcal{F}_t$ on $\mathbb{H}$ is the same as the σ-algebra of evaluations on $\mathbb{H}$ up to time t. Indeed one need only check that $\{\mathbb{H} \cap F : F \in \mathcal{F}_t\} = \sigma(\{\mathbb{H} \cap \mathrm{pr}_s^{-1}(B) : 0 \le s \le t, B \in \mathcal{B}(\mathbb{R})\}) =: \mathcal{H}_t$ (where on the right-hand side one excludes $s = \infty$ if $t = \infty$). Thus all the spaces $(\mathbb{H}, \mathcal{H}_t)$ are standard by [Parthasarathy, 1967, p. 135, Theorem 2.3]. Moreover, every atom H of $\mathcal{H}_t$ clearly fixes the entire path up to time t, in the sense that $|\mathrm{pr}_{[0,t]}(H)| = 1$. It follows that in the above we could just as easily have worked with the space $(\mathbb{H}, \mathcal{H}_\infty, (\mathcal{H}_t)_{t\ge 0})$ in place of $(\mathbb{D}, \mathcal{F}_\infty, \mathbb{F})$ and none of the results would have been affected.