AN INTRODUCTION TO LÉVY PROCESSES

ERIK BAURDOUX AND ANTONIS PAPAPANTOLEON

Abstract. Lecture notes from courses at TU Berlin, LSE and CIMAT

Contents

Date: February 6, 2015.

1. Introduction

These lecture notes aim at providing a (mostly) self-contained introduction to Lévy processes. We start by defining Lévy processes in Section 2 and study the simple but very important example of a Lévy jump-diffusion in Section 3. In Section 4 we discuss infinitely divisible distributions, as the distribution of a Lévy process at any fixed time has this property. In the subsequent section we prove the Lévy–Khintchine formula, which characterises all infinitely divisible distributions. We then prove the Lévy–Itô decomposition of a Lévy process, which is an explicit existence result explaining how to construct a Lévy process based on a given infinitely divisible distribution. In Section 7, we make use of the Lévy–Itô decomposition to derive necesarry and sufficient conditions for a Lévy process to be of (in)finite variation and a subordinator, respectively. In Section 8 we study elementary operations such as linear transformations, projections and subordination. We then move on to moments and martingales followed by a brief section on simulation followed by a number of popular models used mostly in mathematical finance. Then there is a section on simulation and we finish with a brief treatment of stochastic integration.

2. Definition of Lévy processes

2.1. **Notation and auxiliary definitions.** Let $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ denote a stochastic basis, or filtered probability space, i.e. a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $\mathbf{F} = (\mathcal{F}_t)_{t\geq 0}$.

A filtration is an increasing family of sub-σ-algebras of $\mathcal{F}$, i.e. $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$. By convention $\mathcal{F}_\infty = \mathcal{F}$ and $\mathcal{F}_{\infty-} = \bigvee_{s\geq 0} \mathcal{F}_s$.

A stochastic basis satisfies the *usual conditions* if it is *right-continuous*, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$, where $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$, and is *complete*, i.e. the σ-algebra $\mathcal{F}$ is $\mathbb{P}$-complete and every $\mathcal{F}_t$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$.

Definition 2.1. A stochastic process $X = (X_t)_{t\geq 0}$ has *independent increments* if, for any $n \geq 1$ and $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

Alternatively, we say that X has independent increments if, for any $0 \leq s < t$, $X_t - X_s$ is independent of $\mathcal{F}_s$.

Definition 2.2. A stochastic process $X = (X_t)_{t\geq 0}$ has *stationary increments* if, for any $s, t \geq 0$, the distribution of $X_{t+s} - X_s$ does not depend on s. Alternatively, we say that X has stationary increments if, for any $0 \leq s \leq t$, $X_t - X_s$ is equal in distribution to X_{t-s}.

Definition 2.3. A stochastic process $X = (X_t)_{t\geq 0}$ is *stochastically continuous* if, for every $t \geq 0$ and $\epsilon > 0$

$$\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) = 0.$$

2.2. **Definition of Lévy processes.** We will now define Lévy processes and then present some well-known examples, like the Brownian motion and the Poisson process.

Definition 2.4 (Lévy process)**.** An adapted, $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{t\geq 0}$ with $X_0 = 0$ a.s. is called a *Lévy process* if:

(L1) X has independent increments,
(L2) X has stationary increments,
(L3) X is stochastically continuous.

In the sequel, we will always assume that X has *càdlàg paths*. The next two results provide the justification.

Lemma 2.5. *If X is a Lévy process and Y is a modification of X (i.e. $\mathbb{P}(X_t \neq Y_t) = 0$ a.s. for each $t \geq 0$), then Y is a Lévy process and has the same characteristics as X.*

Proof. [App09, Lemma 1.4.8]. □

Theorem 2.6. *Every Lévy process has a unique càdlàg modification that is itself a Lévy process.*

Proof. [App09, Theorem 2.1.8] or [Pro04, Theorem I.30]. □

2.3. **Examples.** The following are some well-known examples of Lévy processes:

- The *linear drift* is the simplest Lévy process, a deterministic process; see Figure 5.3 for a sample path.
- The *Brownian motion* is the only non-deterministic Lévy process with continuous sample paths; see Figure 5.3 for a sample path.
- The *Poisson*, the *compound Poisson* and the *compensated (compound) Poisson* processes are also examples of Lévy processes; see Figure 5.3 for a sample path of a compound Poisson process.

The sum of a linear drift, a Brownian motion and a (compound or compensated) Poisson process is again a Lévy process. It is often called a "jump-diffusion" process. We shall call it a *Lévy jump-diffusion* process, since there exist jump-diffusion processes which are not Lévy processes. See Figure 5.3 for a sample path of a Lévy jump-diffusion process.

3. Toy example: a Lévy jump-diffusion

Let us study the Lévy jump-diffusion process more closely; it is the simplest Lévy process we have encountered so far that contains both a diffusive part and a jump part. We will calculate the characteristic function of the Lévy jump-diffusion, since it offers significant insight into the structure of the characteristic function of general Lévy processes.

Assume that the process $X = (X_t)_{t\geq 0}$ is a Lévy jump-diffusion, i.e. a linear deterministic process, plus a Brownian motion, plus a compensated compound Poisson process. The paths of this process are described by

$$X_t = bt + \sigma W_t + \Big(\sum_{k=1}^{N_t} J_k - t\lambda\beta \Big), \tag{3.1}$$

where $b \in \mathbb{R}$, $\sigma \in \mathbb{R}_{\geqslant 0}$, $W = (W_t)_{t\geq 0}$ is a standard Brownian motion, $N = (N_t)_{t\geq 0}$ is a Poisson process with intensity $\lambda \in \mathbb{R}_{\geqslant 0}$ (i.e. $\mathbb{E}[N_t] = \lambda t$), and $J = (J_k)_{k\geq 1}$ is an i.i.d. sequence of random variables with probability distribution F and $\mathbb{E}[J_k] = \beta < \infty$. Here F describes the distribution of the jumps, which arrive according to the Poisson process N. All sources of randomness are assumed *mutually independent*.

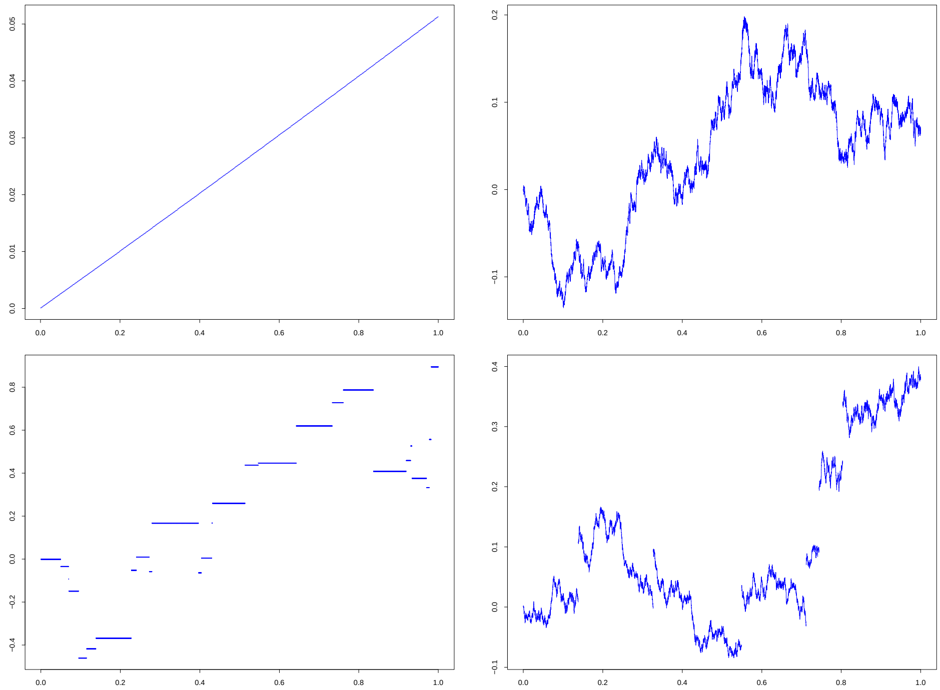

FIGURE 2.1. Sample paths of a linear drift processs (top-left), a Brownian motion (top-right), a compound Poisson process (bottom-left) and a Lévy jump-diffusion.

The characteristic function of X_t, taking into account that all sources of randomness are independent, is

$$\begin{aligned}\mathbb{E}\big[\mathrm{e}^{iuX_t}\big] &= \mathbb{E}\Big[\exp\big(iu\big(bt+\sigma W_t+\sum_{k=1}^{N_t}J_k-t\lambda\beta\big)\big)\Big]\\ &= \exp\big[iubt\big]\,\mathbb{E}\Big[\exp\big(iu\sigma W_t\big)\Big]\,\mathbb{E}\Big[\exp\big(iu\sum_{k=1}^{N_t}J_k-iut\lambda\beta\big)\Big];\end{aligned}$$

recalling that the characteristic functions of the normal and the compound Poisson distributions are

$$\mathbb{E}[\mathrm{e}^{iu\sigma W_t}]=\mathrm{e}^{-\frac{1}{2}\sigma^2u^2t},\quad W_t\sim\mathcal{N}(0,t)$$

$$\mathbb{E}[\mathrm{e}^{iu\sum_{k=1}^{N_t}J_k}]=\mathrm{e}^{\lambda t(\mathbb{E}[\mathrm{e}^{iuJ_k}-1])},\quad N_t\sim\mathrm{Poi}(\lambda t)$$

(cf. Example 4.14 and Exercise 1), we get

$$\begin{aligned}&=\exp\big[iubt\big]\exp\Big[-\frac{1}{2}u^2\sigma^2t\Big]\exp\Big[\lambda t\big(\mathbb{E}[\mathrm{e}^{iuJ_k}-1]-iu\mathbb{E}[J_k]\big)\Big]\\ &=\exp\big[iubt\big]\exp\Big[-\frac{1}{2}u^2\sigma^2t\Big]\exp\Big[\lambda t\big(\mathbb{E}[\mathrm{e}^{iuJ_k}-1-iuJ_k]\big)\Big];\end{aligned}$$

and since the distribution of J_k is F we have

$$= \exp\big[iubt\big] \exp\Big[-\frac{1}{2}u^2\sigma^2 t\Big] \exp\Big[\lambda t \int_{\mathbb{R}} \big(\mathrm{e}^{iux} - 1 - iux\big) F(\mathrm{d}x)\Big].$$

Finally, since t is a common factor, we can rewrite the above equation as

$$\mathbb{E}\big[\mathrm{e}^{iuX_t}\big] = \exp\Big[t\Big(iub - \frac{u^2\sigma^2}{2} + \int_{\mathbb{R}} (\mathrm{e}^{iux} - 1 - iux)\lambda F(\mathrm{d}x)\Big)\Big]. \tag{3.2}$$

We can make the following observations based on the structure of the characteristic function of the random variable X_t from the Lévy jump-diffusion:

(O1) time and space *factorize*;
(O2) the drift, the diffusion and the jump parts are *separated*;
(O3) the jump part *decomposes* to $\lambda \times F$, where λ is the expected number of jumps and F is the distribution of jump size.

One would naturally ask if these observations are true for *any* Lévy process. The answer for (O1) and (O2) is *yes*, because Lévy processes have stationary and independent increments. The answer for (O3) is *no*, because there exist Lévy processes with *infinitely* many jumps (on any compact time interval), thus their expected number of jumps is also infinite.

Since the characteristic function of a random variable determines its distribution, (3.2) provides a characterization of the distribution of the random variables X_t from the Lévy jump-diffusion X. We will soon see that this distribution belongs to the class of *infinitely divisible distributions* and that equation (3.2) is a special case of the celebrated *Lévy-Khintchine formula.*

3.1. **The basic connections.** The next sections will be devoted to establishing the connection between the following mathematical objects:

- Lévy processes $X = (X_t)_{t\geq 0}$
- infinitely divisible distributions $\rho = \mathcal{L}(X_1)$
- Lévy triplets (b, c, ν).

The following commutative diagram displays how these connections can be proved, where LK stands for the Lévy–Khintchine formula, LI for the Lévy–Itô decomposition, CFE for the Cauchy functional equation and SII for stationary and independent increments.

$$\begin{tikzcd}
(X_t)_{t\geq 0} \arrow[rr, Rightarrow, shift left, "SII"] \arrow[dr, Rightarrow, shift left, "CFE"] & & \mathcal{L}(X_1) = \rho \arrow[ll, Rightarrow, shift left, "LI"] \arrow[dl, Rightarrow, shift left] \\
& (b, c, \nu) \arrow[ul, Rightarrow, shift left, "LI"] \arrow[ur, Rightarrow, shift left, "LK"'] &
\end{tikzcd}$$

FIGURE 3.2. The basic connections between Lévy processes, infinitely divisible distributions and Lévy triplets.

Exercise 1. Let $X = (X_t)_{t\geq 0}$ be a compound Poisson process with intensity $\lambda > 0$ and jump distribution F, i.e.

$$X_t = \sum_{k=1}^{N_t} J_k,$$

where $N = (N_t)_{t\geq 0}$ is a Poisson process with $\mathbb{E}[N_t] = \lambda t$ and $J = (J_k)_{k\geq 0}$ is an i.i.d. sequence of random variables with distribution F. Show that

$$\mathbb{E}\left[\mathrm{e}^{iuX_t}\right] = \exp\left(\lambda t \int_{\mathbb{R}} (\mathrm{e}^{iux} - 1)F(\mathrm{d}x)\right).$$

Exercise 2. Consider the setting of the previous exercise and assume that $\mathbb{E}[J_k] = \beta < \infty$. Show that the compensated compound Poisson process $\overline{X} = (\overline{X}_t)_{t\geq 0}$ is a martingale, where

$$\overline{X}_t = X_t - \lambda\beta t.$$

4. Infinitely Divisible distributions

4.1. **Notation and auxiliary results.** Let X be a random variable and denote by $\mathbb{P}_X$ its law, by φ_X its characteristic function, and by M_X its moment generating function. They are related as follows:

$$\varphi_X(u) = \mathbb{E}\big[\mathrm{e}^{i\langle u,X\rangle}\big] = \int_\Omega \mathrm{e}^{i\langle u,X\rangle}\mathrm{d}\mathbb{P} = \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x\rangle}\mathbb{P}_X(\mathrm{d}x) = M_X(iu), \tag{4.1}$$

for all $u \in \mathbb{R}^d$.

Let ρ be a probability measure; we will denote by $\widehat{\rho}$ its characteristic function (or Fourier transform), i.e.

$$\widehat{\rho}(u) = \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x\rangle}\rho(\mathrm{d}x). \tag{4.2}$$

Let $S \subseteq \mathbb{R}^d$, we will denote by $\mathcal{B}(S)$ the Borel σ-algebra of S and by $B_b(S)$ the space of bounded, Borel measurable functions from S to $\mathbb{R}$. We will also denote convergence in law by $\xrightarrow{d}$, weak convergence by $\xrightarrow{w}$ and uniform convergence on compact sets by $\xrightarrow{\text{uc}}$.

We also recall some results from probability theory and complex analysis.

Proposition 4.1. *Let ρ, ρ_n, $n \in \mathbb{N}$, be probability measures on $\mathbb{R}^d$.*

(1) If $\rho_n \xrightarrow{w} \rho$ then $\widehat{\rho}_n(u) \xrightarrow{uc} \widehat{\rho}(u)$.
(2) If $\widehat{\rho}_n(u) \longrightarrow \widehat{\rho}(u)$ for every u, then $\rho_n \xrightarrow{w} \rho$.
(3) Let $f, f_n : \mathbb{R}^d \to \mathbb{C}$, $n = 1, 2, \ldots$, be continuous functions such that $f(0) = f_n(0) = 1$ and $f(u) \neq 0$ and $f_n(u) \neq 0$, for any u and any n. If $f_n(u) \xrightarrow{uc} f(u)$, then also $\log f_n(u) \xrightarrow{uc} \log f(u)$.

Proof. For (1) and (2) see [Shi96, p. 325], for (3) see [Sat99, Lemma 7.7]. □

Theorem 4.2 (Lévy continuity theorem)**.** *Let $(\rho_n)_{n\in\mathbb{N}}$ be probability measures on $\mathbb{R}^d$ whose characteristic functions $\widehat{\rho}_n(u)$ converge to some function $\widehat{\rho}(u)$, for all u, where $\widehat{\rho}$ is continuous at 0. Then, $\widehat{\rho}$ is the characteristic function of a probability distribution ρ and $\rho_n \xrightarrow{d} \rho$.*

Proof. [Dud02, Theorem 9.8.2] □

4.2. **Convolution.** Let μ, ρ be two probability measures on $\mathbb{R}^d$. We define the *convolution* of μ and ρ as

$$(\mu * \rho)(A) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} 1_A(x+y)\mu(\mathrm{d}x)\rho(\mathrm{d}y), \tag{4.3}$$

for each $A \in \mathcal{B}(\mathbb{R}^d)$.

Denote by $A - x := \{y - x : y \in A\}$, then we have that $1_A(x+y) = 1_{A-x}(y) = 1_{A-y}(x)$, and Fubini's theorem yields

$$(\mu * \rho)(A) = \int_{\mathbb{R}^d} \mu(A-x)\rho(\mathrm{d}x) = \int_{\mathbb{R}^d} \rho(A-y)\mu(\mathrm{d}y). \tag{4.4}$$

Proposition 4.3. *The convolution of two probability measures is again a probability measure.*

Proof. [App09, Proposition 1.2.1]. □

Definition 4.4. We define the n-fold convolution of a measure ρ as

$$\rho^{*n} = \underbrace{\rho * \cdots * \rho}_{n \text{ times}}. \tag{4.5}$$

We say that the measure ρ has a *convolution n-th root* if there exists a measure ρ_n such that

$$\rho = (\rho_n)^{*n}. \tag{4.6}$$

In the sequel we will make use of the following results.

Proposition 4.5. *Let ρ_1, ρ_2, be Borel probability measures on $\mathbb{R}^d$ and let $f \in B_b(\mathbb{R}^d)$, then*

$$\int_{\mathbb{R}^d} f(y)(\rho_1 * \rho_2)(\mathrm{d}y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x+y)\rho_1(\mathrm{d}x)\rho_2(\mathrm{d}y). \tag{4.7}$$

Proof. [App09, Proposition 1.2.2(1)]. □

Corollary 4.6. *Let X_1, X_2 be independent random variables with marginals ρ_1, ρ_2. Then, for any $f \in B_b(\mathbb{R}^d)$*

$$\mathbb{E}[f(X_1 + X_2)] = \int_{\mathbb{R}^d} f(x)(\rho_1 * \rho_2)(\mathrm{d}x). \tag{4.8}$$

In particular, for the indicator function we get

$$\mathbb{P}(X_1 + X_2 \in A) = \mathbb{E}[1_A(X_1 + X_2)] = (\rho_1 * \rho_2)(A), \tag{4.9}$$

where $A \in \mathcal{B}(\mathbb{R}^d)$.

Proof. Direct consequences of independence and Proposition 4.5. □

4.3. **Infinite divisibility.** We start by defining infinitely divisible random variables and then discuss some properties of their probability distributions and characteristic functions.

Definition 4.7. A random variable X is *infinitely divisible* if, for all $n \in \mathbb{N}$, there exist i.i.d. random variables $X_1^{(n)}, \dots, X_n^{(n)}$ such that

$$X \overset{\mathrm{d}}{=} X_1^{(n)} + \cdots + X_n^{(n)}. \tag{4.10}$$

The next result provides some insight into the structure of infinitely divisible distributions.

Proposition 4.8. *The following are equivalent:*

(1) X is infinitely divisible;
(2) $\mathbb{P}_X$ has a convolution n-th root that is itself the law of a random variable, for all $n \in \mathbb{N}$;
(3) φ_X has an n-th root that is itself the characteristic function of a random variable, for all $n \in \mathbb{N}$.

Proof. (1) ⇒ (2) Since X is infinitely divisible, we have for any $A \in \mathcal{B}(\mathbb{R}^d)$

$$\begin{aligned}
\mathbb{P}_X(A) = \mathbb{P}(X \in A) &= \mathbb{P}(X_1^{(n)} + \cdots + X_n^{(n)} \in A) \\
&= (\mathbb{P}_{X_1^{(n)}} * \cdots * \mathbb{P}_{X_n^{(n)}})(A) \qquad \text{(by independence and (4.9))} \\
&= (\mathbb{P}_{X^{(n)}} * \cdots * \mathbb{P}_{X^{(n)}})(A) \qquad \text{(by identical laws)} \\
&= (\mathbb{P}_{X^{(n)}})^{*n}(A).
\end{aligned} \tag{4.11}$$

(2) ⇒ (3) Since $\mathbb{P}_X$ has a convolution n-th root, we have

$$\begin{aligned}
\varphi_X(u) &= \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x\rangle} \mathbb{P}_X(\mathrm{d}x) = \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x\rangle} (\mathbb{P}_{X^{(n)}} * \cdots * \mathbb{P}_{X^{(n)}})(\mathrm{d}x) \\
&= \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x_1+\cdots+x_n\rangle} \mathbb{P}_{X^{(n)}}(\mathrm{d}x_1) \dots \mathbb{P}_{X^{(n)}}(\mathrm{d}x_n) \qquad \text{(Prop. 4.5)} \\
&= \prod_{i=1}^n \int_{\mathbb{R}^d} \mathrm{e}^{i\langle u,x_i\rangle} \mathbb{P}_{X^{(n)}}(\mathrm{d}x_i) \qquad \text{(by independence)} \\
&= \prod_{i=1}^n \varphi_{X^{(n)}}(u) = \Big(\varphi_{X^{(n)}}(u)\Big)^n.
\end{aligned} \tag{4.12}$$

(3) ⇒ (1) Choose $X_1^{(n)}, \dots, X_n^{(n)}$ to be independent copies of a given r.v. $X^{(n)}$. Since the characteristic function has an n-th root, we get

$$\begin{aligned}
\mathbb{E}\big[\mathrm{e}^{i\langle u,X\rangle}\big] &= \varphi_X(u) \\
&= \Big(\varphi_{X^{(n)}}(u)\Big)^n = \prod_{i=1}^n \varphi_{X_i^{(n)}}(u) \\
&= \mathbb{E}\Big[\mathrm{e}^{i\big\langle u,X_1^{(n)}+\cdots+X_n^{(n)}\big\rangle}\Big] \qquad \text{(by independence)},
\end{aligned} \tag{4.13}$$

and the result follows, since the characteristic function deteremines the law of a random variable. □

These results motivate us to give the following more general definition of infinite divisibility.

Definition 4.9. A probability measure ρ is *infinitely divisible* if, for all $n \in \mathbb{N}$, there exists another probability measure ρ_n such that

$$\rho = \underbrace{\rho_n * \cdots * \rho_n}_{n \text{ times}}. \tag{4.14}$$

Proposition 4.10. *A probability measure ρ is infinitely divisible if and only if, for all $n \in \mathbb{N}$, there exists another probability measure ρ_n such that*

$$\widehat{\rho}(u) = \Big(\widehat{\rho_n}(u)\Big)^n. \tag{4.15}$$

Proof. Similar to the proof of Proposition 4.8, thus left as an exercise. □

Next, we will discuss some properties of infinitely divisible distributions, in particular that they are closed under convolutions and weak limits.

Lemma 4.11. *If μ, ρ are infinitely divisible probability measures then $\mu * \rho$ is also infinitely divisible.*

Proof. Since μ and ρ are infinitely divisible, we know that for any $n \in \mathbb{N}$ it holds

$$\mu = (\mu_n)^{*n} \quad \text{and} \quad \rho = (\rho_n)^{*n}. \tag{4.16}$$

Hence, from the commutativity of the convolution we get that

$$\mu * \rho = (\mu_n)^{*n} * (\rho_n)^{*n} = (\mu_n * \rho_n)^{*n}.$$ □

Lemma 4.12. *If ρ is infinitely divisible then $\widehat{\rho}(u) \neq 0$ for any $u \in \mathbb{R}^d$.*

Proof. Since ρ is infinitely divisible, we know that for every $n \in \mathbb{N}$ there exists a measure ρ_n such that $\widehat{\rho} = (\widehat{\rho}_n)^n$. Using [Sat99, Prop. 2.5(v)] we have that $|\widehat{\rho}_n(u)|^2 = |\widehat{\rho}(u)|^{2/n}$ is a characteristic function. Define the function

$$\varphi(u) = \lim_{n\to\infty} |\widehat{\rho}_n(u)|^2 = \lim_{n\to\infty} |\widehat{\rho}(u)|^{2/n} = \begin{cases} 1, & \text{if } \widehat{\rho}(u) \neq 0 \\ 0, & \text{if } \widehat{\rho}(u) = 0. \end{cases}$$

Since $\widehat{\rho}(0) = 1$ and $\widehat{\rho}$ is a continuous function, we get that $\varphi(u) = 1$ in a neighborhood of 0. Now, using Lévy's continuity theorem we get that $\varphi(u)$ is a continuous function, thus $\varphi(u) = 1$ for all $u \in \mathbb{R}^d$. Hence $\widehat{\rho}(u) \neq 0$ for any $u \in \mathbb{R}^d$. □

Lemma 4.13. *If $(\rho_k)_{k\geq 0}$ is a sequence of infinitely divisible distributions and $\rho_k \xrightarrow{w} \rho$, then ρ is also infinitely divisible.*

Proof. Since $\rho_k \xrightarrow{w} \rho$ as $k \to \infty$ we get from Proposition 4.1(1) that $\widehat{\rho}_k(z) \xrightarrow{uc} \widehat{\rho}(z)$ and $\widehat{\rho}$ is the characteristic function of the probability measure ρ. In order to prove the claim, it suffices to show that $\widehat{\rho}^{1/n}$ is well-defined and the characteristic function of a probability measure. Then, the trivial equality $\widehat{\rho}(z) = \big(\widehat{\rho}(z)^{1/n}\big)^n$ yields that ρ is infinitely divisible.

Since $\widehat{\rho}_k$ and $\widehat{\rho}$ are characteristic functions, we know that they are continuous and $\widehat{\rho}_k(0) = \widehat{\rho}(0) = 1$ for every k. Moreover, $\widehat{\rho}_k$ is the characteristic function of an infinitely divisible distribution, thus from Lemma 4.12 we get that $\widehat{\rho}_k(u) \neq 0$ for any k, u. One can also show that $\widehat{\rho}(u) \neq 0$ for every u, see

[Sat99, Lemma 7.8]. Therefore, we can apply Proposition 4.1(3) and we get that $\log \widehat{\rho}_k(u) \xrightarrow{uc} \log \widehat{\rho}(u)$, hence also $\widehat{\rho}_k(u)^{1/n} \xrightarrow{uc} \widehat{\rho}(u)^{1/n}$, for every n, as $k \to \infty$. We have that $\widehat{\rho}_k^{1/n}$ is a continuous function, and using the uniform convergence to $\widehat{\rho}^{1/n}$, we can conclude that this is also continuous (around zero). Now, an application of Lévy's continuity theorem yields that $\widehat{\rho}^{1/n}$ is the characteristic function of a probability distribution. □

4.4. **Examples.** Below we present some examples of infinitely divisible distributions. In particular, using Proposition 4.8 we can easily show that the normal, the Poisson and the exponential distributions are infinitely divisible.

Example 4.14 (Normal distribution)**.** Let $X \sim \mathcal{N}(\mu, \sigma^2)$, then we have

$$\begin{aligned}\varphi_X(u) &= \exp\left(iu\mu - \frac{1}{2}u^2\sigma^2\right) = \exp\left(n\Big[iu\frac{\mu}{n} - \frac{1}{2}u^2\frac{\sigma^2}{n}\Big]\right) \\ &= \left(\exp\left[iu\frac{\mu}{n} - \frac{1}{2}u^2\frac{\sigma^2}{n}\right]\right)^n \\ &= \Big(\varphi_{X^{(n)}}(u)\Big)^n,\end{aligned} \tag{4.17}$$

where $X^{(n)} \sim \mathcal{N}(\frac{\mu}{n}, \frac{\sigma^2}{n})$.

Example 4.15 (Poisson distribution)**.** Let $X \sim \mathrm{Poi}(\lambda)$, then we have

$$\begin{aligned}\varphi_X(u) &= \exp\Big(\lambda(\mathrm{e}^{iu} - 1)\Big) = \left(\exp\left[\frac{\lambda}{n}(\mathrm{e}^{iu} - 1)\right]\right)^n \\ &= \Big(\varphi_{X^{(n)}}(u)\Big)^n,\end{aligned} \tag{4.18}$$

where $X^{(n)} \sim \mathrm{Poi}(\frac{\lambda}{n})$.

Example 4.16 (Exponential distribution)**.** Let $X \sim \mathrm{Exp}(\lambda)$, then we have

$$\begin{aligned}\varphi_X(u) &= \left(1 - \frac{iu}{\lambda}\right)^{-1} = \left[\left(1 - \frac{iu}{\lambda}\right)^{-\frac{1}{n}}\right]^n \\ &= \Big(\varphi_{X^{(n)}}(u)\Big)^n,\end{aligned} \tag{4.19}$$

where $X^{(n)} \sim \Gamma(\frac{1}{n}, \lambda)$.

Remark 4.17. Other examples of infinitely divisible distributions are the geometric, the negative binomial, the Cauchy and the strictly stable distributions. Counter-examples are the uniform and the binomial distributions.

Exercise 3. Show that the law of the random variable

$$X_t = bt + \sigma W_t + \sum_{k=1}^{N_t} J_k, \quad (t \geq 0, \quad \text{fixed}) \tag{4.20}$$

is infinitely divisible, *without* using Proposition 4.8.

4.5. **Lévy processes have infinitely divisible laws.** We close this section by taking a glimpse of the deep connections between infinitely divisible distributions and Lévy processes. In particular, we will show that if $X = (X_t)_{t\geq 0}$ is a Lévy process then X_t is an infinitely divisible random variable (for all $t \geq 0$).

Lemma 4.18. *Let $X = (X_t)_{t\geq 0}$ be a Lévy process. Then the random variables X_t, $t \geq 0$, are infinitely divisible.*

Proof. Let $X = (X_t)_{t\geq 0}$ be a Lévy process; for any $n \in \mathbb{N}$ and any $t > 0$ we trivially have that

$$X_t = X_{\frac{t}{n}} + (X_{\frac{2t}{n}} - X_{\frac{t}{n}}) + \ldots + (X_t - X_{\frac{(n-1)t}{n}}). \tag{4.21}$$

The stationarity of the increments of the Lévy process yields that

$$X_{\frac{tk}{n}} - X_{\frac{t(k-1)}{n}} \stackrel{\mathrm{d}}{=} X_{\frac{t}{n}},$$

for any $k \geq 1$, while the independence of the increments yields that the random variables $X_{\frac{tk}{n}} - X_{\frac{t(k-1)}{n}}$, $k \geq 1$, are indepedent of each other. Thus, $(X_{\frac{tk}{n}} - X_{\frac{t(k-1)}{n}})_{k\geq 1}$ is an i.i.d. sequence of random variables and, from Definition 4.7, we conclude that the random variable X_t is infinitely divisible. □

5. The Lévy–Khintchine representation

The next result provides a complete characterization of infinitely divisible distributions in terms of their characteristic functions. This is the celebrated *Lévy-Khintchine formula.* B. de Finetti and A. Kolmogorov were the first to prove versions of this representation under certain assumptions. P. Lévy and A. Khintchine indepedently proved it in the general case, the former by analyzing the sample paths of the process and the latter by a direct analytic method.

5.1. **Statement, "if" part.** We first define a Lévy measure and then state the Lévy–Khintchine representation and prove the "if part" of the theorem.

Definition 5.1 (Lévy measure)**.** Let ν be a Borel measure on $\mathbb{R}^d$. We say that ν is a *Lévy measure* if it satisfies

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu(\mathrm{d}x) < \infty. \tag{5.1}$$

Remark 5.2. Since $|x|^2 \wedge \varepsilon \leq |x|^2 \wedge 1$ for all $0 < \varepsilon \leq 1$, it follows from (5.1) that $\nu\big((-\varepsilon, \varepsilon)^c\big) < \infty$ for all $\varepsilon > 0$. In other words, any Lévy measure becomes a probability measure once restricted to the complement of an ε-neighborhood of the origin (after an appropriate normalization).

Theorem 5.3 (Lévy–Khintchine)**.** *A measure ρ is* infinitely divisible *if and only if there exists a triplet (b, c, ν) with $b \in \mathbb{R}^d$, c a symmetric, non-negative definite, $d \times d$ matrix, and ν a Lévy measure, such that*

$$\widehat{\rho}(u) = \exp\left(i\,\langle u, b\rangle - \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}^d} \big(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u, x\rangle\, 1_D\big)\nu(\mathrm{d}x) \right). \tag{5.2}$$

Here D denotes the closed unit ball in $\mathbb{R}^d$, i.e. $D := \{|x| \leq 1\}$.

Definition 5.4. We will call (b, c, ν) the *Lévy* or *characteristic triplet* of the infinitely divisible measure ρ. We call b the *drift* characteristic and c the *Gaussian* or *diffusion* characteristic.

Example 5.5. An immediate consequence of Definitions 5.1 and 5.4 and Theorem 5.3 is that the distribution of the r.v. X_1 from the Lévy jump-diffusion is infinitely divisible with Lévy triplet

$$\left(b - \int_{D^c} x\lambda F(\mathrm{d}x), \sigma^2, \lambda \times F\right).$$

Proof of Theorem 5.3, "If" part. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$, monotonic and decreasing to zero (e.g. $\varepsilon_n = \frac{1}{n}$). Define for all $u \in \mathbb{R}^d$ and $n \in \mathbb{N}$

$$\widehat{\rho}_n(u) = \exp\left(i\Big\langle u, b - \int\limits_{\varepsilon_n < |x| \leq 1} x\nu(\mathrm{d}x)\Big\rangle - \frac{\langle u, cu\rangle}{2} + \int\limits_{|x| > \varepsilon_n} (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x)\right).$$

Each $\widehat{\rho}_n$ is the characteristic function of the convolution of a normal and a compound Poisson distribution, hence $\widehat{\rho}_n$ is the characteristic function of an infinitely divisible probability measure ρ_n. We clearly have that

$$\lim_{n\to\infty} \widehat{\rho}_n(u) = \widehat{\rho}(u).$$

Then, Lévy's continuity theorem and Lemma 4.13 yield that $\widehat{\rho}$ is the characteristic function of an infinitely divisible probability measure ρ, provided that $\widehat{\rho}$ is continuous at 0.

Now, continuity of $\widehat{\rho}$ at 0 boils down to the continuity of the integral term in (5.2), i.e.

$$\begin{aligned}
\psi_\nu(u) &:= \int\limits_{\mathbb{R}^d} (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u, x\rangle\, 1_D(x))\nu(\mathrm{d}x) \\
&= \int\limits_{D} (\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u, x\rangle)\nu(\mathrm{d}x) + \int\limits_{D^c} (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x).
\end{aligned}$$

Using Taylor's expansion, the Cauchy–Schwarz inequality, the definition of the Lévy measure and dominated convergence, we get

$$\begin{aligned}
|\psi_\nu(u)| &\leq \frac{1}{2}\int\limits_{D} |\,\langle u, x\rangle\,|^2\nu(\mathrm{d}x) + \int\limits_{D^c} |\mathrm{e}^{i\langle u,x\rangle} - 1|\nu(\mathrm{d}x) \\
&\leq \frac{|u|^2}{2}\int\limits_{D} |x^2|\nu(\mathrm{d}x) + \int\limits_{D^c} |\mathrm{e}^{i\langle u,x\rangle} - 1|\nu(\mathrm{d}x) \\
&\longrightarrow 0 \quad \text{as} \quad u \to 0.
\end{aligned}$$

□

Exercise 4 (Frullani integral)**.** (i) Consider a function f such that f' exists and is continuous, and $f(0), f(\infty)$ are finite. Show that

$$\int\limits_0^\infty \frac{f(ax) - f(bx)}{x}\mathrm{d}x = (f(0) - f(\infty))\log\left(\frac{b}{a}\right),$$

for $b > a > 0$.

(*ii*) Consider the function $f(x) = \mathrm{e}^{-x}$ and set $a = \alpha > 0$ and $b = \beta = \alpha - z$ with $z < 0$. Show that

$$\exp\Big(\int_0^\infty (\mathrm{e}^{zx} - 1)\frac{\beta}{x}\mathrm{e}^{-\alpha x}\mathrm{d}x\Big) = \frac{1}{(1 - z/\alpha)^\beta}.$$

Explain why this equality remains true for $z \in \mathbb{C}$ with $\Re z \le 0$.

Exercise 5. Consider the $\Gamma(\alpha, \beta)$ distribution, described by the density

$$f_{\alpha,\beta}(x) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1}\mathrm{e}^{-\alpha x},$$

concentrated on $(0, \infty)$.

(*i*) Compute the characteristic function of the $\Gamma(\alpha, \beta)$ distribution and show it is infinitely divisible.

(*ii*) Show that the Lévy triplet of the $\Gamma(\alpha, \beta)$ distribution is

$$b = \int_0^1 x\nu(\mathrm{d}x), \quad c = 0, \quad \nu(\mathrm{d}x) = \beta x^{-1}\mathrm{e}^{-\alpha x}\mathrm{d}x.$$

5.2. **Truncation functions and uniqueness.** We will now introduce truncation functions and discuss about the uniqueness of the representation (5.2).

The integrand in (5.2) is integrable with respect to the Lévy measure ν because it is bounded outside any neighborhoud of zero and

$$\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u, x\rangle\, 1_D(x) = O(|x|^2) \quad \text{as } |x| \to 0, \tag{5.3}$$

for any fixed u. There are many other ways to construct an integrable integrand and we will be particularly interested in *continuous* ones, because they are suitable for limit arguments. This leads to the notion of a *truncation function*. The following definitions are taken from [JS03] and [Sat99] respectively.

Definition 5.6. A truncation function is a bounded function $h : \mathbb{R}^d \to \mathbb{R}^d$ that satisfies $h(x) = x$ in a neighborhood of zero.

Definition 5.7. A truncation function $h' : \mathbb{R}^d \to \mathbb{R}$ is a bounded and measurable function, satisfying

$$\begin{aligned} h'(x) &= 1 + o(|x|), && \text{as } |x| \to 0, \\ h'(x) &= O(1/|x|), && \text{as } |x| \to \infty. \end{aligned} \tag{5.4}$$

Remark 5.8. The two definitions are related via $h(x) = x \cdot h'(x)$.

Example 5.9. The following are some well-known examples of truncation functions:

(i) $h(x) = x1_D(x)$, typically called the *canonical* truncation function;
(ii) $h(x) \equiv 0$ and $h(x) \equiv x$, are also commonly used truncation functions; note that contrary to the other two examples, these are not always permissible choices;
(iii) $h(x) = \frac{x}{1+|x|^2}$, a continuous truncation function.

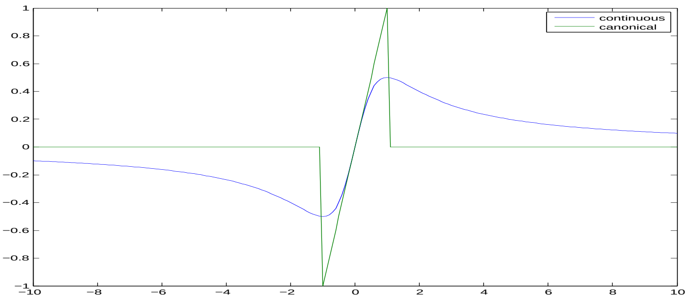

FIGURE 5.3. Illustration of the canonical and the continuous truncation functions from Example 5.9.

The Lévy–Khintchine representation of $\widehat{\rho}$ in (5.2) depends on the choice of the truncation function. Indeed, if we use another truncation function h instead of the canonical one, then (5.2) can be rewritten as

$$\widehat{\rho}(u) = \exp\left(i\,\langle u, b_h\rangle - \frac{\langle u, cu\rangle}{2} + \int\limits_{\mathbb{R}^d} \left(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u, h(x)\rangle\,\right)\nu(\mathrm{d}x)\right), \tag{5.5}$$

with b_h defined as follows:

$$b_h = b + \int\limits_{\mathbb{R}^d} \big(h(x) - x1_D(x)\big)\nu(\mathrm{d}x). \tag{5.6}$$

If we want to stress the dependence of the Lévy triplet on the truncation function, we will denote it by

$$(b_h, c, \nu)_h \qquad \text{or} \qquad (b, c, \nu)_h.$$

Note that the diffusion characteristic c and the Lévy measure ν are invariant with respect to the choice of the truncation function.

Remark 5.10. There is no rule about which truncation function to use, among the permissible choices. One simply has to be *consistent* with ones choice of a truncation function. That is, the same choice should be made for the Lévy–Khintchine representation of the characteristic function, the Lévy triplet and the path decomposition of the Lévy process.

Example 5.11. Let us revisit the Lévy jump-diffusion process (3.1). In this example, since the Lévy measure is finite and we have assumed that $\mathbb{E}[J_k] < \infty$, all the truncation functions of Example 5.9 are permissible. The Lévy triplet of this process with respect to the canonical truncation function was presented in Example 5.5. The triplets with respect to the zero and the linear truncation functions are

$$\Big(b - \int_{\mathbb{R}} x\lambda F(\mathrm{d}x), \sigma^2, \lambda \times F\Big)_0 \qquad \text{and} \qquad \big(b, \sigma^2, \lambda \times F\big)_{\mathrm{id}}\,.$$

Although the Lévy–Khintchine representation depends on the choice of the truncation function, the Lévy triplet determines the law of the distribution uniquely (once the truncation function has been fixed).

Proposition 5.12. *The representation of $\widehat{\rho}$ by (b, c, ν) in* (5.2) *is unique.*

Sketch of Proof. We will outline the argument for the diffusion coefficient c; the complete proof can be found in [Sat99, Theorem 8.1(ii)].

Let $\widehat{\rho}$ be expressed by (b, c, ν) according to (5.2). By Taylor's theorem we get that

$$|e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle 1_D(x)| \leq \frac{1}{2}|u|^2|x|^2 1_D(x) + 2 1_{D^c}(x). \tag{5.7}$$

Since the exponent in (5.2) is continuous in u, we have

$$\log \widehat{\rho}(su) = -s^2 \frac{\langle u, cu\rangle}{2} + is\langle u, b\rangle + \int_{\mathbb{R}^d} \big(e^{is\langle u,x\rangle} - 1 - is\langle u, x\rangle 1_D(x)\big)\nu(dx),$$

for $s \in \mathbb{R}$. Now, by (5.7) and dominated convergence we get

$$s^{-2}\log \widehat{\rho}(su) \longrightarrow -\frac{\langle u, cu\rangle}{2}, \quad \text{as} \quad s \to \infty. \tag{5.8}$$

Therefore, the diffusion coefficient c is uniquely identified by ρ. The proof for ν is similar in spirit, while once c, ν are uniquely determined then b is identified as well. □

5.3. **Proof, "only if" part.** The next theorem contains an essential step in the proof of the "only if" part of the Lévy–Khintchine representation (Theorem 5.3). We denote by $C_\sharp$ the space of bounded continuous functions $f : \mathbb{R}^d \to \mathbb{R}$, vanishing in a neighborhood of 0.

Theorem 5.13. *Let $h' : \mathbb{R}^d \to \mathbb{R}$ be a continuous truncation function, i.e. satisfying* (5.4)*. Suppose that ρ_n, $n \in \mathbb{N}$, are infinitely divisible distributions on $\mathbb{R}^d$ and that each $\widehat{\rho}_n$ has the Lévy–Khintchine representation with triplet $(\beta_n, c_n, \nu_n)_h$. Let ρ be a probability distribution on $\mathbb{R}^d$. Then $\rho_n \xrightarrow{w} \rho$* if and only if *(i) ρ is infinitely divisible and (ii) $\widehat{\rho}$ has the Lévy–Khintchine representation with triplet $(\beta, c, \nu)_h$, where β, c and ν satisfy the following conditions:*

(1) If $f \in C_\sharp$ then

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\nu_n(dx) = \int_{\mathbb{R}^d} f(x)\nu(dx). \tag{5.9}$$

(2) Define the symmetric, non-negative definite matrices $c_{n,\varepsilon}$ via

$$\langle u, c_{n,\varepsilon}u\rangle = \langle u, c_n u\rangle + \int_{|x|\leq\varepsilon} \langle u, x\rangle^2 \nu_n(dx). \tag{5.10}$$

Then

$$\lim_{\varepsilon\downarrow 0} \limsup_{n\to\infty} \big| \langle u, c_{n,\varepsilon}u\rangle - \langle u, cu\rangle \big| = 0 \quad \textit{for } u \in \mathbb{R}^d. \tag{5.11}$$

(3) $\beta_n \to \beta$.

Proof. "Only If" part. Assume that $\rho_n \longrightarrow \rho$. Then, using Lemmata 4.13 and 4.12 we get that ρ is infinitely divisible and $\widehat{\rho}(u) \neq 0$ for all u. Moreover, it follows from Proposition 4.1 that

$$\log \widehat{\rho}_n(u) \longrightarrow \log \widehat{\rho}(u) \tag{5.12}$$

uniformly on compact sets.

Define the measure $\phi_n(\mathrm{d}x) = (|x|^2 \wedge 1)\nu_n(\mathrm{d}x)$, and note that $\phi_n(\mathbb{R}^d) = \int_{\mathbb{R}^d} \phi_n(\mathrm{d}x) < \infty$ by the definition of the Lévy measure. We claim that (ϕ_n) is *tight*, i.e. that

$$\sup_n \phi_n(\mathbb{R}^d) < \infty \quad \text{and} \quad \lim_{l\to\infty} \sup_n \int\limits_{|x|>l} \phi_n(\mathrm{d}x) = 0. \tag{5.13}$$

A proof of tightness can be found in [Sat99, pp. 43-44]. Then, by Prokhorov's selection theorem there exists a subsequence (ϕ_{n_k}) that converges to some finite measure ϕ; cf. [Bil99, §1.6].

Next, define the measure ν via: $\nu(\{0\}) = 0$ and $\nu(\mathrm{d}x) = (|x|^2 \wedge 1)^{-1}\phi(\mathrm{d}x)$ on the set $\{|x| > 0\}$. The measure ϕ might have a point mass at 0, but this is ignored when defining ν. Let us denote

$$g(u, x) = \mathrm{e}^{i\langle u,x\rangle} - 1 - i \langle u, h(x)\rangle , \tag{5.14}$$

which is bounded and continuous in x, for fixed u, due to the choice of a continuous truncation function h. We have that

$$\begin{aligned} \log \widehat{\rho}_n(u) &= i \langle u, \beta_n\rangle - \frac{1}{2} \langle u, c_n u\rangle + \int\limits_{\mathbb{R}^d} g(u, x)\nu_n(\mathrm{d}x) \\ &= i \langle u, \beta_n\rangle - \frac{1}{2} \langle u, c_{n,\epsilon} u\rangle + I_{n,\epsilon} + J_{n,\epsilon}, \end{aligned} \tag{5.15}$$

where

$$I_{n,\epsilon} := \int\limits_{|x|\le\epsilon} \left(g(u, x) + \frac{1}{2} \langle u, x\rangle^2 \right) (|x|^2 \wedge 1)^{-1}\rho_n(\mathrm{d}x) \tag{5.16}$$

and

$$J_{n,\epsilon} := \int\limits_{|x|>\epsilon} g(u, x)(|x|^2 \wedge 1)^{-1}\rho_n(\mathrm{d}x). \tag{5.17}$$

Consider the set $E := \{\epsilon > 0 : \int_{|x|=\epsilon} \rho(\mathrm{d}x) = 0\}$, then

$$\lim_{k\to\infty} J_{n_k,\epsilon} = \int\limits_{|x|>\epsilon} g(u, x)(|x|^2 \wedge 1)^{-1}\rho(\mathrm{d}x) \tag{5.18}$$

hence

$$\lim_{E\ni\epsilon\downarrow 0} \lim_{k\to\infty} J_{n_k,\epsilon} = \int\limits_{\mathbb{R}^d} g(u, x)\nu(\mathrm{d}x), \tag{5.19}$$

because $g \in C_\sharp$. Furthermore, we have that

$$\lim_{\epsilon\downarrow 0} \sup_n |I_{n,\epsilon}| = 0, \tag{5.20}$$

since

$$\left(g(u, x) + \frac{1}{2} \langle u, x\rangle^2 \right) (|x|^2 \wedge 1)^{-1} \underset{x\to 0}{\longrightarrow} 0, \tag{5.21}$$

by the definition of the truncation function h.

Now, we consider the real and imaginary parts of (5.15) separately; then, using (5.13), (5.19) and (5.20), we get that

$$\lim_{E\ni\epsilon\downarrow 0}\limsup_{k\to\infty}\langle u, c_{n_k,\varepsilon}u\rangle = \lim_{E\ni\epsilon\downarrow 0}\liminf_{k\to\infty}\langle u, c_{n_k,\varepsilon}u\rangle \tag{5.22}$$

$$\limsup_{k\to\infty}\langle u, \beta_{n_k}\rangle = \liminf_{k\to\infty}\langle u, \beta_{n_k}\rangle \tag{5.23}$$

and both sides of these equations are finite. By (5.23) we conclude that there exists some β such that $\beta_n \longrightarrow \beta$. Moreover, since each side of (5.22) is a nonnegative quadratic form of u, it is equal to $\langle u, cu\rangle$ for some symmetric, nonnegative definite matrix. In addition, we can drop the requirement $^\varepsilon \in E$ in (5.22) using monotonicity. Hence, it follows that $\widehat{\rho}(u)$ has the representation (5.2) for these β, c and ν and that (1)-(3) hold via the subsequence (ϕ_{n_k}). However, (β, c, ν) in the Lévy–Khintchine representation are unique, see Proposition 5.12, thus (1) and (3) hold for the whole sequence (ϕ_n). In addition, we can show by revisiting the arguments above that

$$\lim_{\epsilon\downarrow 0}\limsup_{n\to\infty}\langle u, c_{n,\varepsilon}u\rangle = \lim_{\epsilon\downarrow 0}\liminf_{n\to\infty}\langle u, c_{n,\varepsilon}u\rangle = \langle u, cu\rangle \tag{5.24}$$

which is equivalent to (2).

"If" part. Define the measures $\phi_n(\mathrm{d}x) = (|x|^2\wedge 1)\nu_n(\mathrm{d}x)$ and $\phi(\mathrm{d}x) = (|x|^2\wedge 1)\nu(\mathrm{d}x)$ and consider the set E defined above. Condition (1) yields (5.19) immediately. Moreover, conditions (1) and (2) imply that (ϕ_n) is uniformly bounded, thus we get (5.19) as well. Now, using (2) and (3) we get that

$$\lim_{n\to\infty}\log\widehat{\rho}_n(u) = i\langle u, \beta\rangle - \frac{1}{2}\langle u, cu\rangle + \int_{\mathbb{R}^d} g(u,x)\nu(\mathrm{d}x).$$

Since the right hand side equals $\log\widehat{\rho}(u)$, we conclude that $\rho_n \longrightarrow \rho$. □

Finally, using the "only if" part of Theorem 5.13 we are ready to complete the proof of Theorem 5.3.

Proof of Theorem 5.3, "Only If" part. Let ρ be an infinitely divisible distribution. Choose a sequence $t_n \downarrow 0$ arbitrarily, and define ρ_n via

$$\widehat{\rho}_n(u) = \exp\Big(t_n^{-1}\left(\widehat{\rho}(u)^{t_n} - 1\right)\Big) = \exp\left(t_n^{-1}\int_{\mathbb{R}^d\setminus\{0\}}\left(\mathrm{e}^{i\langle u,x\rangle} - 1\right)\rho^{t_n}(\mathrm{d}x)\right). \tag{5.25}$$

Clearly, the distribution of ρ_n is compound Poisson and thus also infinitely divisible. Moreover, Taylor's expansion yields

$$\begin{aligned}\widehat{\rho}_n(u) &= \exp\Big(t_n^{-1}\Big(\mathrm{e}^{t_n\log\widehat{\rho}(u)} - 1\Big)\Big) = \exp\Big(t_n^{-1}\big(t_n\log\widehat{\rho}(u) + O(t_n^2)\big)\Big)\\ &= \exp\Big(\log\widehat{\rho}(u) + O(t_n)\Big),\end{aligned} \tag{5.26}$$

for fixed u, as $n\to\infty$. Hence $\widehat{\rho}_n(u)\longrightarrow\widehat{\rho}(u)$ as $n\to\infty$.

Since ρ_n is infinitely divisible it has the Lévy–Khintchine representation (5.2) for some triplet $(b_n, c_n, \nu_n)_h$ (in this case with $h\equiv 0$). However, $\widehat{\rho}_n(u) \longrightarrow \widehat{\rho}(u)$ implies that $\rho_n \xrightarrow{w} \rho$, by Proposition 4.1. Hence, using Theorem 5.13 yields that $\widehat{\rho}$ has the Lévy–Khintchine representation with some triplet $(b, c, \nu)_h$. Now, we can rewrite this as (5.1) and the result is proved. $\square$

Corollary 5.14. *Every infinitely divisible distribution is the limit of a sequence of compound Poisson distributions.*

5.4. **The Lévy–Khintchine formula for Lévy processes.** In section 4.5 we showed that for any Lévy process $X = (X_t)_{t\geq 0}$, the random variables X_t are infinitely divisible. Next, we would like to compute the characteristic function of X_t. Since X_t is infinitely divisible for any $t \geq 0$, we know that X_1 is also infinitely divisible and has the Lévy–Khintchine representation in terms of some triplet (b, c, ν).

Definition 5.15. We define the *Lévy exponent* ψ of X by

$$\psi(u) = i \langle u, b\rangle - \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}} \big(\mathrm{e}^{i\langle u,x\rangle} - 1 - i \langle u, x\rangle 1_D(x)\big)\nu(\mathrm{d}x), \qquad (5.27)$$

where

$$\mathbb{E}\big[\mathrm{e}^{i\langle u,X_1\rangle}\big] = \mathrm{e}^{\psi(u)}. \qquad (5.28)$$

Theorem 5.16. *Let $X = (X_t)_{t\geq 0}$ be a Lévy process, then*

$$\mathbb{E}\big[\mathrm{e}^{i\langle u,X_t\rangle}\big] = \mathrm{e}^{t\psi(u)}, \qquad (5.29)$$

where ψ is the Lévy exponent of X.

Proof. Define the function $\phi_u(t) = \mathbb{E}[\mathrm{e}^{i\langle u,X_t\rangle}]$. Using the independence and stationarity of the increments we have that

$$\begin{aligned}\phi_u(t+s) &= \mathbb{E}[\mathrm{e}^{i\langle u,X_{t+s}\rangle}] = \mathbb{E}[\mathrm{e}^{i\langle u,X_{t+s}-X_s\rangle}\mathrm{e}^{\langle iu,X_s\rangle}] \\ &= \mathbb{E}[\mathrm{e}^{i\langle u,X_{t+s}-X_s\rangle}]\mathbb{E}[\mathrm{e}^{\langle iu,X_s\rangle}] = \phi_u(t)\phi_u(s). \end{aligned} \qquad (5.30)$$

Moreover, $\phi_u(0) = \mathbb{E}[\mathrm{e}^{i\langle u,X_0\rangle}] = 1$ by definition. Since X is stochastically continuous we can show that $t \mapsto \phi_u(t)$ is continuous (cf. Exercise 6).

Note that (5.30) is Cauchy's second functional equation, and the unique continuous solution to this equation has the form

$$\phi_u(t) = \mathrm{e}^{t\vartheta(u)}, \qquad \text{where} \qquad \vartheta : \mathbb{R}^d \to \mathbb{C}.$$

Now the result follows since X_1 is infinitely divisible, which yields

$$\phi_u(1) = \mathbb{E}[\mathrm{e}^{i\langle u,X_1\rangle}] = \mathrm{e}^{\psi(u)}.$$

$\square$

Corollary 5.17. *The infinitely divisible random variable X_t has the Lévy triplet $(bt, ct, \nu t)$.*

Exercise 6. Let $X = (X_t)_{t\geq 0}$ be a stochastically continuous process. Show that the map $t \mapsto \varphi_{X_t}(u)$ is continuous for every $u \in \mathbb{R}^d$.

Exercise 7. Let X be a Lévy process with triplet (b, c, ν). Show that $-X$ is also a Lévy process and determine its triplet.

6. The Lévy–Itô decomposition

In the previous sections, we showed that for any Lévy process $X = (X_t)_{t\geq 0}$ the random variables X_t, $t \geq 0$, have an infinitely divisible distribution and determined this distribution using the Lévy–Khintchine representation. The aim of this section is to prove an "inverse" result: starting from an infinitely divisible distribution ρ, or equivalently from a Lévy triplet (b, c, ν), we want to construct a Lévy process $X = (X_t)_{t\geq 0}$ such that $\mathbb{P}_{X_1} = \rho$.

Theorem 6.1. *Let ρ be an infinitely divisible distribution with Lévy triplet (b, c, ν), where $b \in \mathbb{R}^d$, $c \in \mathbb{S}^d_{\geqslant 0}$ and ν is a Lévy measure. Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which four independent Lévy processes exist, $X^{(1)}, \dots, X^{(4)}$, where: $X^{(1)}$ is a constant drift, $X^{(2)}$ is a Brownian motion, $X^{(3)}$ is a compound Poisson process and $X^{(4)}$ is a square integrable, pure jump martingale with an a.s. countable number of jumps of magnitude less than 1 on each finite time interval. Setting $X = X^{(1)} + \cdots + X^{(4)}$, we have that there exists a probability space on which a Lévy process $X = (X_t)_{t\geq 0}$ is defined, with Lévy exponent*

$$\psi(u) = i\langle u, b\rangle - \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}^d} \big(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle 1_D(x)\big)\nu(\mathrm{d}x) \tag{6.1}$$

for all $u \in \mathbb{R}^d$, and path, or Lévy–Itô, decomposition

$$X_t = bt + \sqrt{c}W_t + \int_0^t \int_{D^c} x\mu^X(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_D x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x), \tag{6.2}$$

where $\nu^X = \mathrm{Leb} \otimes \nu$.

6.1. **Roadmap of the Proof.** We first provide an informal description of the proof, in order to motivate the mathematical tools required. Consider the exponent in the Lévy–Khintchine formula and rewrite it as follows:

$$\begin{aligned}
\psi(u) &= \psi^{(1)}(u) + \psi^{(2)}(u) + \psi^{(3)}(u) + \psi^{(4)}(u) \\
&= i\langle u, b\rangle - \frac{\langle u, cu\rangle}{2} + \nu(D^c)\int_{D^c} \big(\mathrm{e}^{i\langle u,x\rangle} - 1\big)\frac{\nu(\mathrm{d}x)}{\nu(D^c)} \\
&\qquad + \int_D \big(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\big)\nu(\mathrm{d}x).
\end{aligned} \tag{6.3}$$

Clearly $\psi^{(1)}$ corresponds to the characteristic exponent of a linear drift process with rate b, $\psi^{(2)}$ to a Brownian motion with covariance matrix c, and $\psi^{(3)}$ to a compound Poisson process with intensity $\lambda = \nu(D^c)$ and jump distribution $F(\mathrm{d}x) = \frac{\nu(\mathrm{d}x)}{\nu(D^c)} 1_{D^c}(\mathrm{d}x)$.

The most difficult part is to handle the process with characteristic exponent $\psi^{(4)}$. We can express this as follows:

$$\begin{aligned}\psi^{(4)}(u) &= \int\limits_D \left(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u,x\rangle\,\right)\nu(\mathrm{d}x)\\ &= \sum_{n\geq 0}\left(\lambda_n \int\limits_{D_n} \left(\mathrm{e}^{i\langle u,x\rangle} - 1\right)\nu_n(\mathrm{d}x) - i\Big\langle u, \lambda_n \int\limits_{D_n} x\nu_n(\mathrm{d}x)\Big\rangle\right),\end{aligned}$$

where we define the discs $D_n = \{2^{-(n+1)} \leq |x| < 2^{-n}\}$, the intensities $\lambda_n = \nu(D_n)$ and the probability measures $\nu_n(\mathrm{d}x) = \frac{\nu(\mathrm{d}x)}{\lambda_n} 1_{D_n}(\mathrm{d}x)$ (see again Remark 5.2). We can intuitively understand this as the Lévy exponent of a superposition of compound Poisson processes with arrival rates λ_n and jump distributions ν_n, and an additional drift term that turns these processes into martingales. In order to convert this intuition into precise mathematical statements, we will need results on Poisson random measures and square integrable martingales.

6.2. **Poisson random measures.** Let us first consider a compound Poisson process with drift $X = (X_t)_{t\geq 0}$, with

$$X_t = bt + \sum_{k=1}^{N_t} J_k,$$

where $b \in \mathbb{R}$, N is a Poisson process with intensity λ and $J = (J_k)_{k\geq 0}$ is an i.i.d. sequence of random variables with distribution F. This process has a finite number of jumps in any finite time interval, and the time between consecutive jumps is exponentially distributed with parameter λ, the rate of the Poisson process. Denote the jump times of X by $(T_k)_{k\geq 1}$ and, for a set $A \in \mathcal{B}(\mathbb{R}_+) \times \mathcal{B}(\mathbb{R}\backslash\{0\})$, define the random variable $\mu(A)$ via

$$\mu(A) := \#\{k \geq 1 : (T_k, J_k) \in A\} = \sum_{k\geq 1} 1_{\{(T_k,J_k)\in A\}}.$$

The random variable $\mu(A)$ takes values in $\mathbb{N}$ and counts the total number of jumps that belong to the time-space set A. The following lemma provides some important properties of μ.

Lemma 6.2. *Suppose that $A_1, \dots, A_k$, $k \geq 1$ are disjoint subsets of $\mathcal{B}(\mathbb{R}_+) \times \mathcal{B}(\mathbb{R}\backslash\{0\})$. Then $\mu(A_1), \dots, \mu(A_k)$ are mutually independent random variables, and for each $i \in \{1, \dots, k\}$ the random variable $\mu(A_i)$ has a Poisson distribution with intensity*

$$\lambda_i = \lambda \int_{A_i} \mathrm{d}t \times F(\mathrm{d}x).$$

Moreover, for $\mathbb{P}$-a.e. realization of X, $\mu : \mathcal{B}(\mathbb{R}_+) \times \mathcal{B}(\mathbb{R}\backslash\{0\}) \to \mathbb{N} \cup \{\infty\}$ is a measure.

Exercise 8. Prove Lemma 6.2. Steps and Hints:

(i) Recall that the law of $\{T_1, \dots, T_n\}$ conditioned on the event $\{N_t = n\}$ has the same law as the ordered independent sample from n uniformly distributed r.v. on $[0, t]$.

(ii) Use (i) and the independence of J_k to show that the law of $\{(T_k, J_k), k = 1, \dots, n\}$ conditioned on $\{N_t = n\}$ equals the law of n independent bivariate r.v. with common distribution $t^{-1}\mathrm{d}s \times F(\mathrm{d}x)$ on $[0,t] \times \mathbb{R}$, ordered in time.

(iii) Show that, for $A \in \mathcal{B}([0,t]) \times \mathcal{B}(\mathbb{R})$, $\mu(A)$ conditioned on $\{N_t = n\}$ is a Binomial r.v. with probability of success $\int_A t^{-1}\mathrm{d}s \times F(\mathrm{d}x)$.

(iv) Show that

$$\mathbb{P}(\mu(A_1) = n_1, \dots, \mu(A_k) = n_k | N_t = n) = \frac{n!}{n_0! n_1! \dots n_k!} \prod_{i=0}^{k} \left(\frac{\lambda_i}{\lambda t}\right)^{n_i},$$

where $n_0 = n - \sum_{i=1}^k n_i$ and $\lambda_0 = \lambda t - \sum_{i=1}^k \lambda_i$.

(v) Finally, integrate out the conditioning to show that

$$\mathbb{P}(\mu(A_1) = n_1, \dots, \mu(A_k) = n_k) = \prod_{i=1}^{k} \mathrm{e}^{-\lambda_i} \frac{(\lambda_i)^{n_i}}{n_i!}.$$

The random measure introduced above is a special case of the more general notion of a Poisson random measure, defined as follows.

Definition 6.3 (Poisson random measure)**.** Let $(E, \mathcal{E}, \nu)$ be a σ-finite measure space. Consider a mapping $\mu : \mathcal{E} \to \mathbb{N} \cup \{\infty\}$ such that $\{\mu(A) : A \in \mathcal{E}\}$ is a family of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then μ is called a *Poisson random measure with intensity* ν if

(1) μ is $\mathbb{P}$-a.s. a measure on $(E, \mathcal{E})$;
(2) for each $A \in \mathcal{E}$, $\mu(A)$ is Poisson distributed with parameter $\nu(A)$, where $\nu(A) \in [0, \infty]$;
(3) for mutually disjoint sets $A_1, \dots, A_n$ in $\mathcal{E}$, the random variables $\mu(A_1), \dots, \mu(A_n)$ are independent.

Remark 6.4. Note that if $\nu(A) = 0$ then we get that $\mathbb{P}(\mu(A) = 0) = 1$, while if $\nu(A) = \infty$ then we have that $\mathbb{P}(\mu(A) = \infty) = 1$.

Exercise 9. Show that every Lévy measure is a σ-finite measure on $\mathbb{R}^d \backslash \{0\}$, i.e. there exist sets $(A_i)_{i \in \mathbb{N}}$ such that $\cup_i A_i = \mathbb{R}^d \backslash \{0\}$ and $\nu(A_i) < \infty$, for all $i \in \mathbb{N}$.

Theorem 6.5. *Let $(E, \mathcal{E}, \nu)$ be a σ-finite measure space. Then, a Poisson random measure μ as defined above always exists.*

Proof. Step 1. Assume that $\nu(E) < \infty$. There exists a standard construction of an infinite product space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the following independent random variables are defined:

$$N \quad \text{and} \quad \{v_1, v_2, \dots\},$$

such that N is Poisson distributed with intensity $\nu(E)$ and each v_i has the probability distribution $\frac{\nu(\mathrm{d}x)}{\nu(E)}$. Define, for every $A \in \mathcal{E}$

$$\mu(A) = \sum_{i=1}^{N} 1_{\{v_i \in A\}}, \tag{6.4}$$

such that $N = \mu(E)$. For each $A \in \mathcal{E}$ and $i \geq 1$, the random variables $1_{\{v_i \in A\}}$ are $\mathcal{F}$-measurable, hence $\mu(A)$ is also $\mathcal{F}$-measurable. Let $A_1, \dots, A_k$ be mutually disjoint sets, then we can show that

$$\mathbb{P}(\mu(A_1) = n_1, \dots, \mu(A_k) = n_k) = \prod_{i=1}^{k} \mathrm{e}^{-\nu(A_i)} \frac{(\nu(A_i))^{n_i}}{n_i!}; \tag{6.5}$$

the derivation is similar to the proof of Lemma 6.2. Now, we can directly deduce that conditions (1)–(3) in the definition of a Poisson random measure are satisfied.

Step 2. Let ν be a σ-finite measure on $(E, \mathcal{E})$. Then, there exist subsets $(A_i)_{i \geq 1}$ of E such that $\cup_i A_i = E$ and $\nu(A_i) < \infty$. Define the measures

$$\nu_i(\cdot) := \nu(\cdot \cap A_i), \quad i \geq 1.$$

The first step yields that for each $i \geq 1$ there exists a probability space $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ such that a Poisson random measure μ_i can be defined on $(A_i, \mathcal{E}_i, \nu_i)$, where $\mathcal{E}_i := \{B \cap A_i, B \in \mathcal{E}\}$. Now, we just have to show that

$$\mu(\cdot) := \sum_{i \geq 1} \mu(\cdot \cap A_i),$$

is a Poisson random measure on E with intensity ν, defined on the product space

$$(\Omega, \mathcal{F}, \mathbb{P}) := \bigotimes_{i \geq 1} (\Omega_i, \mathcal{F}_i, \mathbb{P}_i).$$

... □

The construction of the Poisson random measure leads immediately to the following corollaries.

Corollary 6.6. *Let μ be a Poisson random measure on $(E, \mathcal{E}, \nu)$. Then, for every $A \in \mathcal{E}$, we have that $\mu(\cdot \cap A)$ is a Poisson random measure on $(E \cap A, \mathcal{E} \cap A, \nu(\cdot \cap A))$. Moreover, if $A, B \in \mathcal{E}$ are disjoint, then the random variables $\mu(\cdot \cap A)$ and $\mu(\cdot \cap B)$ are independent.*

Corollary 6.7. *Let μ be a Poisson random measure on $(E, \mathcal{E}, \nu)$. Then, the support of μ is $\mathbb{P}$-a.s. countable. If, in addition, ν is a finite measure, then the support of μ is $\mathbb{P}$-a.s. finite.*

Corollary 6.8. *Assume that the measure ν has an atom, say at the point $\varepsilon \in E$. Then, it follows from the construction of the Poisson random measure μ that $\mathbb{P}(\mu(\{\varepsilon\}) \geq 1) > 0$. Conversely, if ν has no atoms then $\mathbb{P}(\mu(\{\varepsilon\}) = 0) = 1$ for all singletons $\varepsilon \in E$.*

6.3. **Integrals wrt Poisson random measures.** Let μ be a Poisson random measure defined on the space $(E, \mathcal{E}, \nu)$. The fact that μ is $\mathbb{P}$-a.s. a measure allows us to use Lebesgue's theory of integration and consider, for a measurable function $f : E \to [0, \infty)$,

$$\int_E f(x)\mu(\mathrm{d}x),$$

which is then a well-defined, $[0, \infty]$-valued random variable. The same holds true for a signed function f, which yields a $[-\infty, \infty]$-valued random variable, provided that either f^+ or f^- are finite. This integral can be understood as follows:

$$\int_E f(x)\mu(\mathrm{d}x) = \sum_{v\in\operatorname{supp}(\mu)} f(v)\cdot m_v,$$

where m_v denotes the multiplicity of points at v (e.g., if μ has no atoms then $m_v = 1$ for every $v \in \operatorname{supp}(\mu)$). Convergence of integrals with respect to Poisson random measures and related properties are provided by the following result.

Theorem 6.9. *Let μ be a Poisson random measure on $(E, \mathcal{E}, \nu)$ and $f : E \to \mathbb{R}^d$ be a measurable function. Then:*

(i) $X = \int_E f(x)\mu(\mathrm{d}x)$ is almost surely absolutely convergent if and only if

$$\int_E (1 \wedge |f(x)|)\nu(\mathrm{d}x) < \infty. \tag{6.6}$$

(ii) If (6.6) holds then

$$\mathbb{E}\left[\mathrm{e}^{i\langle u,X\rangle}\right] = \exp\left(\int_E \left(\mathrm{e}^{i\langle u,f(x)\rangle} - 1\right)\nu(\mathrm{d}x)\right). \tag{6.7}$$

(iii) Moreover, if $f \in L^1(\nu)$ then

$$\mathbb{E}[X] = \int_E f(x)\nu(\mathrm{d}x), \tag{6.8}$$

while if $f \in L^2(\nu)$ then

$$\operatorname{Var}[X] = \int_E f(x)^2\nu(\mathrm{d}x). \tag{6.9}$$

Proof. Define simple functions

$$f(x) = \sum_{i=1}^n f_i 1_{A_i}(x) \tag{6.10}$$

where, for $i = 1, \ldots, n$, f_i are constants and $A_i \subset E$ are disjoint subsets of E such that $\nu(A_1 \cup \cdots \cup A_n) < \infty$. Then

$$\begin{aligned} X = \int_E f(x)\mu(\mathrm{d}x) &= \int_E \sum_{i=1}^n f_i 1_{A_i}(x)\mu(\mathrm{d}x) = \sum_{i=1}^n f_i \int_E 1_{A_i}(x)\mu(\mathrm{d}x) \\ &= \sum_{i=1}^n f_i\mu(A_i) \end{aligned} \tag{6.11}$$

and $\mathbb{P}(X<\infty)=1$ since each $\mu(A_i)\sim \mathrm{Poi}(\nu(A_i))$. Now

$$\begin{aligned}
\mathbb{E}\big[\mathrm{e}^{i\langle u,X\rangle}\big] &= \mathbb{E}\left[\mathrm{e}^{i\left\langle u,\sum_{i=1}^n f_i\mu(A_i)\right\rangle}\right] = \prod_{i=1}^n \mathbb{E}\left[\mathrm{e}^{i\langle u,f_i\mu(A_i)\rangle}\right]\\
&= \prod_{i=1}^n \exp\left[\left(\mathrm{e}^{i\langle u,f_i\rangle}-1\right)\nu(A_i)\right]\\
&= \exp\left[\sum_{i=1}^n \left(\mathrm{e}^{i\langle u,f_i\rangle}-1\right)\nu(A_i)\right]
\end{aligned}$$

and since $f\equiv 0$ on $E\setminus(A_1\cup\dots\cup A_n)$

$$= \exp\left(\int_E \left(\mathrm{e}^{i\langle u,f(x)\rangle}-1\right)\nu(\mathrm{d}x)\right). \tag{6.12}$$

The remainder of the proof follows the "usual" measure-theoretic recipe of first approximating positive functions by simple, positive and increasing ones and using monotone convergence, and then, for a general function, by writing it as the difference of two positive functions and using the measures $\nu(\cdot\cap\{f\ge 0\})$ and $\nu(\cdot\cap\{f\le 0\})$; see [Kyp06, Thm. 2.7]. This shows (ii), while (iii) follows from (ii) by differentiation and using the classical formula

$$\mathbb{E}[X^k] = (-i)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k}\varphi_X(u)\big|_{u=0}. \tag{6.13}$$

In order to show (i) note that the integral on the RHS of (6.12) is infinite for every u if $\mathbb{P}(X=\infty)=1$, while it is finite for every u if $\mathbb{P}(X=\infty)=0$. Conversely, assume that

$$\int_E \left(\mathrm{e}^{i\langle u,f(x)\rangle}-1\right)\nu(\mathrm{d}x)<\infty \tag{6.14}$$

for every u. Since $\mathrm{e}^{i\langle u,f(x)\rangle}-1\le \mathrm{e}^{i\langle 1,f(x)\rangle}-1$ for all $0\le u\le 1$ applying dominated convergence twice yields that

$$\lim_{u\downarrow 0}\int_E \left(\mathrm{e}^{i\langle u,f(x)\rangle}-1\right)\nu(\mathrm{d}x)=0 \tag{6.15}$$

and thus

$$\mathbb{P}(X=\infty)=0. \tag{6.16}$$

... $\square$

Exercise 10. Let ν be a measure on the space $(E,\mathcal{E})$ and $f:E\to[0,\infty)$ be a measurable function. Then, for all $u>0$, show that

$$\int_E (\mathrm{e}^{uf(x)}-1)\nu(\mathrm{d}x)<\infty \iff \int_E (1\wedge f(x))\nu(\mathrm{d}x)<\infty. \tag{6.17}$$

6.4. **Poisson random measures and stochastic processes.** In the sequel, we want to make the connection between Poisson random measures and stochastic processes. We will work in the following σ-finite space:

$$\big(E, \mathcal{E}, \nu^X\big) = \big(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_{\geqslant 0}) \times \mathcal{B}(\mathbb{R}^d), \mathrm{Leb} \otimes \nu\big)$$

where ν is a Lévy measure; see again Definition 5.1. We will denote the Poisson random measure on this space by μ^X. If we consider a time-space interval of the form $[s,t] \times A$, $s \leq t$, where $A \subset \mathbb{R}^d$ such that $0 \notin \overline{A}$, then the integral with respect to μ^X, denoted by

$$\int\limits_{[s,t]} \int\limits_{A} x \mu^X(\mathrm{d}s, \mathrm{d}x) =: X, \tag{6.18}$$

is a compound Poisson random variable with intensity $(t-s) \cdot \nu(A)$. This follows directly from Theorem 6.9, while we also get that

$$\mathbb{E}\big[\mathrm{e}^{i\langle u, X\rangle}\big] = \exp\left((t-s)\int_A (\mathrm{e}^{\langle u,x\rangle} - 1)\nu(\mathrm{d}x)\right). \tag{6.19}$$

Let us consider the collection of random variables

$$\left(\int\limits_0^t \int\limits_A x \mu^X(\mathrm{d}s, \mathrm{d}x)\right)_{t \geq 0}. \tag{6.20}$$

Then, one would naturally expect that this is a compound Poisson *stochastic process.*

Lemma 6.10. *Let μ^X be a Poisson random measure with intensity* $\mathrm{Leb} \otimes \nu$ *and assume that $A \subset \mathcal{B}(\mathbb{R}^d)$ such that $\nu(A) < \infty$. Then*

$$X_t = \int\limits_0^t \int\limits_A x \mu^X(\mathrm{d}s, \mathrm{d}x), \quad t \geq 0$$

is a compound Poisson process with arrival rate $\nu(A)$ and jump distribution $\frac{\nu(\mathrm{d}x)}{\nu(A)}|_A$.

Proof. Since $\nu(A) < \infty$, we have from Corollary 6.7 that the support of μ^X is finite. Hence, we can write X_t as follows

$$X_t = \sum_{0 \leq s \leq t} x \mu^X(\{s\} \times A) = \sum_{0 \leq s \leq t} \Delta X_s 1_{\{\Delta X_s \in A\}},$$

which shows that $t \mapsto X_t$ is a càdlàg function. Let $0 \leq s \leq t$, then the random variable

$$X_t - X_s = \int\limits_{(s,t]} \int\limits_A x \mu^X(\mathrm{d}s, \mathrm{d}x)$$

is independent from $\{X_u : u \leq s\}$, since Poisson random measures over disjoint sets are independent; cf. Corollary 6.6. From Theorem 6.9 we know that

$$\mathbb{E}\big[\mathrm{e}^{i\langle u, X_t\rangle}\big] = \exp\left(t \int\limits_A (\mathrm{e}^{i\langle u,x\rangle} - 1)\nu(\mathrm{d}x)\right). \tag{6.21}$$

The independence of increments allows us to deduce that

$$\begin{aligned}\mathbb{E}\big[\mathrm{e}^{i\langle u,X_t-X_s\rangle}\big] &= \frac{\mathbb{E}\big[\mathrm{e}^{i\langle u,X_t\rangle}\big]}{\mathbb{E}\big[\mathrm{e}^{i\langle u,X_s\rangle}\big]} \\ &= \exp\left((t-s)\int_A(\mathrm{e}^{i\langle u,x\rangle}-1)\nu(\mathrm{d}x)\right) \\ &= \mathbb{E}\big[\mathrm{e}^{i\langle u,X_{t-s}\rangle}\big],\end{aligned}$$

which yields that the increments are also stationary. Moreover, from (6.21) we have that X_t is compound Poisson distributed with arrival rate $t\cdot\nu(A)$ and jump distribution $\frac{\nu(\mathrm{d}x)}{\nu(A)}|_A$. Finally, we have that $X=(X_t)_{t\geq0}$ is a compound Poisson process since it is a process with stationary and independent increments, whose increment distributions are compound Poisson. $\square$

Lemma 6.11. *Consider the setting of the previous lemma and assume that $\int_A|x|\nu(\mathrm{d}x)<\infty$. Then*

$$M_t=\int_0^t\int_A x\mu^X(\mathrm{d}s,\mathrm{d}x)-t\int_A x\nu(\mathrm{d}x),\quad t\geq0 \tag{6.22}$$

is a $\mathbb{P}$-martingale relative to the filtration generated by the Poisson random measure μ^X

$$\mathcal{F}_t:=\sigma\left(\mu^X(G):G\in\mathcal{B}([0,t])\times\mathcal{B}(\mathbb{R}^d)\right),\quad t\geq0. \tag{6.23}$$

If, in addition, $\int_A|x|^2\nu(\mathrm{d}x)<\infty$ then M is a square-integrable martingale.

Proof. The process $M=(M_t)_{t\geq0}$ is clearly adapted to the filtration $(\mathcal{F}_t)_{t\geq0}$ generated by μ^X. Moreover, Theorem 6.9 together with the assumption $\int_A|x|\nu(\mathrm{d}x)<\infty$ immediately yield that

$$\mathbb{E}|M_t|\leq\mathbb{E}\left(\int_0^t\int_A|x|\mu^X(\mathrm{d}s,\mathrm{d}x)\right)-t\int_A|x|\nu(\mathrm{d}x)<\infty.$$

Using that M has stationary and independent increments, which follows directly from Lemma 6.10, we get that, for $0\leq s<t$,

$$\begin{aligned}\mathbb{E}[M_t-M_s|\mathcal{F}_s] &= \mathbb{E}[M_{t-s}] \\ &= \mathbb{E}\left(\int_s^t\int_A x\mu^X(\mathrm{d}s,\mathrm{d}x)\right)-(t-s)\int_A x\nu(\mathrm{d}x)=0,\end{aligned}$$

again using Theorem 6.9. This shows the martingale property.

Next, we just have to show that M is square integrable. We have, using the martingale property of M, the properties of the variance and Theorem

6.9 once more, that

$$\mathbb{E}\big[M_t^2\big] = \mathrm{Var}\big[M_t\big] = \mathrm{Var}\left(\int_s^t \int_A x\mu^X(\mathrm{d}s,\mathrm{d}x)\right)$$
$$= t\int_A |x|^2\nu(\mathrm{d}x) < \infty,$$

which concludes the proof. □

The results of this section allow us to construct compound Poisson processes with jumps taking values in discs of the form $D_\varepsilon := \{\varepsilon < |x| \leq 1\}$, for any $\varepsilon \in (0,1)$. However, we cannot consider the ball $D = \{|x| \leq 1\}$, i.e. set $\varepsilon = 0$, since there exist Lévy measures such that $\int_D |x|\nu(\mathrm{d}x) = \infty$. We will thus study the limit of the martingale M in Lemma 6.11 when the jumps belong to D_ε for $\varepsilon \downarrow 0$.

Exercise 11. Consider the measure on $\mathbb{R}^d \setminus \{0\}$ provided by

$$\nu(\mathrm{d}x) = |x|^{-(1+\alpha)}1_{\{x<0\}}\mathrm{d}x + x^{-(1+\alpha)}1_{\{x>0\}}\mathrm{d}x,$$

for $\alpha \in (1,2)$. Show that it is a Lévy measure, such that $\int_D |x|\nu(\mathrm{d}x) = \infty$.

6.5. **Square integrable martingales.** Denote by $\mathcal{M}_T^2$ the space of right-continuous, zero mean, square integrable martingales. This is a Hilbert space with inner product defined by

$$\langle M, N\rangle := \mathbb{E}[M_T N_T].$$

Therefore, for any Cauchy sequence M^n in $\mathcal{M}_T^2$ there exists an element $M \in \mathcal{M}_T^2$ such that $\|M^n - M\| \longrightarrow 0$ as $n \to \infty$, where $\|\cdot\| = \langle\cdot,\cdot\rangle$. A proof of this result can be found in Section 2.4 of [Kyp06]. In the sequel, we will make use of *Doob's martingale inequality* which states that for any $M \in \mathcal{M}_T^2$ it holds that

$$\mathbb{E}\big[\sup_{0\leq s\leq T} M_s^2\big] \leq 4\mathbb{E}\big[M_T^2\big].$$

The following result is crucial for the proof of the Lévy–Itô decomposition.

Theorem 6.12. *Consider the setting of Lemma 6.10 and recall that for any Lévy measure* $\int_{|x|\leq 1}|x|^2\nu(\mathrm{d}x) < \infty$. *For each* $\varepsilon \in (0,1)$ *define the martingale*

$$M_t^\varepsilon = \int_0^t \int_{D_\varepsilon} x\mu^X(\mathrm{d}s,\mathrm{d}x) - t\int_{D_\varepsilon} x\nu(\mathrm{d}x), \tag{6.24}$$

where $D_\varepsilon = \{\varepsilon < |x| \leq 1\}$. *Let* $\overline{\mathcal{F}}_t$ *denote the completion of* $\bigcap_{s>t}\mathcal{F}_s$ *by all the* $\mathbb{P}$*-null sets. Then, there exists a square integrable martingale* $M = (M_t)_{t\geq 0}$ *that satisfies:*

(i) *for each* $T > 0$, *there exists a deterministic subsequence* $(\varepsilon_n^T)_{n\in\mathbb{N}}$ *with* $\varepsilon_n^T \downarrow 0$, *along which*

$$\mathbb{P}\left(\lim_{n\to\infty}\sup_{0\leq s\leq T}\left(M_s^{\varepsilon_n^T} - M_s\right)^2 = 0\right) = 1,$$

(ii) it is adapted to the filtration $(\overline{\mathcal{F}}_t)_{t\geq 0}$,
(iii) it has a.s. càdlàg paths,
(iv) it has stationary and independent increments,
(v) it has an a.s. countable number of jumps on each compact time interval.

Henceforth, there exists a Lévy process $M = (M_t)_{t\geq 0}$, which is a square integrable martingale with an a.s. countable number of jumps such that, for each fixed $T > 0$, the sequence of martingales $(M_t^\varepsilon)_{0\leq t\leq T}$ converges uniformly to M on $[0, T]$ a.s. along a subsequence in ε.

Proof. (i) Consider a fixed $T > 0$ and set $0 < \eta < \varepsilon < 1$, then

$$\begin{aligned}\|M^\eta - M^\varepsilon\| &= \mathbb{E}\big[(M_T^\eta - M_T^\varepsilon)^2\big] \\ &= \mathbb{E}\left(\int_0^T \int_{\eta<|x|\leq\varepsilon} x\mu^X(\mathrm{d}s, \mathrm{d}x) - T \int_{\eta<|x|\leq\varepsilon} x\nu(\mathrm{d}x)\right)^2 \\ &= T \int_{\eta<|x|\leq\varepsilon} x^2\nu(\mathrm{d}x);\end{aligned} \tag{6.25}$$

see also Exercise 13. Since $\int_D |x|^2\nu(\mathrm{d}x) < \infty$, we have that

$$\|M^\eta - M^\varepsilon\| \longrightarrow 0, \text{ as } \varepsilon \downarrow 0, \tag{6.26}$$

hence (M^ε) is a Cauchy sequence on $\mathcal{M}_T^2$. Moreover, since $\mathcal{M}_T^2$ is a Hilbert space, there exists a martingale $M = (M_t)_{0\leq t\leq T}$ in $\mathcal{M}_T^2$ such that

$$\lim_{\varepsilon\downarrow 0} \|M - M^\varepsilon\| = 0. \tag{6.27}$$

Using Doob's maximal inequality, we get that

$$\lim_{\varepsilon\downarrow 0} \mathbb{E}\left[\sup_{0\leq s\leq T} (M_s^\varepsilon - M_s)^2\right] \leq 4 \lim_{\varepsilon\downarrow 0} \|M - M^\varepsilon\| = 0. \tag{6.28}$$

This allows us to conclude that the limit does not depend on T, thus we have a well-defined martingale limit $M = (M_t)_{t\geq 0}$. In addition, (6.28) yields that there exists a deterministic subsequence $(\varepsilon_n^T)_{n\geq 0}$, possibly depending on T, such that

$$\lim_{\varepsilon_n^T\downarrow 0} \sup_{0\leq s\leq T} \left(M_s^{\varepsilon_n^T} - M_s\right)^2 = 0, \qquad \mathbb{P}\text{-a.s.} \tag{6.29}$$

(ii) Follows directly from the definition of the filtration.
(iii) We can use the following facts:

- M^ε has càdlàg paths and converges uniformly to M, $\mathbb{P}$-a.s.;
- the space of càdlàg functions is closed under the supremum metric.

These yield immediately that M has càdlàg paths.
(iv) We have that a.s. uniform convergence along a subsequence implies also convergence in distribution along the same subsequence. Let $0 \leq q < r <$

$s < t \leq\ T$ and $u, v \in \mathbb{R}^d$, then using dominated convergence we get

$$\begin{aligned}
&\mathbb{E}\big[\exp\left(i\left\langle u, M_t - M_s\right\rangle + i\left\langle v, M_r - M_q\right\rangle\right)\big]\\
&\quad = \lim_{n\to\infty}\mathbb{E}\left[\exp\left(i\left\langle u, M_t^{\varepsilon_n^T} - M_s^{\varepsilon_n^T}\right\rangle + i\left\langle v, M_r^{\varepsilon_n^T} - M_q^{\varepsilon_n^T}\right\rangle\right)\right]\\
&\quad = \lim_{n\to\infty}\mathbb{E}\left[\exp\left(i\left\langle u, M_{t-s}^{\varepsilon_n^T}\right\rangle\right)\right]\mathbb{E}\left[\exp\left(i\left\langle v, M_{r-q}^{\varepsilon_n^T}\right\rangle\right)\right]\\
&\quad = \mathbb{E}\big[\exp\left(i\left\langle u, M_{t-s}\right\rangle\right)\big]\mathbb{E}\big[\exp\left(i\left\langle v, M_{r-q}\right\rangle\right)\big],
\end{aligned}$$

which yields that M has stationary and independent increments.
(v) According to Corollary 6.7, there exist, at most, an a.s. countable number of points in the support of the Poisson random measure μ^X. Moreover, since $\mathrm{Leb}\otimes\nu$ has no atoms, we get that μ^X takes values in $\{0,1\}$ at singletons. Hence, every discontinuity of $M = (M_t)_{t\geq 0}$ corresponds to a single point in the support of μ^X, which yields that M has an a.s. countable number of jumps in every compact time interval. □

6.6. **Proof of the Lévy–Itô decomposition.** Now, we are ready to complete the proof of the Lévy–Itô decomposition.

Proof of Theorem 6.1. Step 1. We first consider the processes $X^{(1)}$ and $X^{(2)}$ with characteristic exponents

$$\psi^{(1)}(u) = i\left\langle u, b\right\rangle \quad \text{and} \quad \psi^{(2)}(u) = \frac{\left\langle u, cu\right\rangle}{2}, \tag{6.30}$$

which correspond to a linear drift and a Brownian motion, i.e.

$$X_t^{(1)} = bt \quad \text{and} \quad X_t^{(2)} = \sqrt{c}W_t, \tag{6.31}$$

defined on some probability space $(\Omega^\natural, \mathcal{F}^\natural, \mathbb{P}^\natural)$.

Step 2. Given a Lévy measure ν, we know from Theorem 6.5 that there exists a probability space, denoted by $(\Omega^\sharp, \mathcal{F}^\sharp, \mathbb{P}^\sharp)$, such that we can construct a Poisson random measure μ^X on $(\mathbb{R}_{\geqslant 0}\times\mathbb{R}^d, \mathcal{B}(\mathbb{R}_{\geqslant 0})\times\mathcal{B}(\mathbb{R}^d), \mathrm{Leb}\otimes\nu)$. Let us define the process $X^{(3)} = (X_t^{(3)})_{t\geq 0}$ with

$$X^{(3)} = \int_0^t\int_{D^c} x\mu^X(\mathrm{d}s, \mathrm{d}x). \tag{6.32}$$

Using Lemma 6.10 we can deduce that $X^{(3)}$ is a compound Poisson process with intensity $\lambda := \nu(D^c)$ and jump distribution $F(\mathrm{d}x) := \frac{\nu(\mathrm{d}x)}{\nu(D^c)}1_{D^c}(\mathrm{d}x)$.

Step 3. Next, from the Lévy measure ν we construct a process having only jumps less than 1. For each $0 < \varepsilon \leq 1$, define the compensated compound Poisson process $X^{(4,\varepsilon)} = (X_t^{(4,\varepsilon)})_{t\geq 0}$ with

$$X^{(4,\varepsilon)} = \int_0^t\int_{\varepsilon<|x|\leq 1} x\mu^X(\mathrm{d}s, \mathrm{d}x) - t\int_{\varepsilon<|x|\leq 1} x\nu(\mathrm{d}x). \tag{6.33}$$

Using Theorem 6.9 we know that $X^{(4,\varepsilon)}$ has the characteristic exponent

$$\psi^{(4,\varepsilon)}(u) = \int_{\varepsilon<|x|\leq 1}\left(\mathrm{e}^{i\left\langle u, x\right\rangle} - 1 - i\left\langle u, x\right\rangle\right)\nu(\mathrm{d}x). \tag{6.34}$$

Now, according to Theorem 6.12 there exists a Lévy process, denoted by $X^{(4)}$, which is a square integrable, pure jump martingale defined on the space $(\Omega^\sharp, \mathcal{F}^\sharp, \mathbb{P}^\sharp)$, such that $X^{(4,\epsilon)}$ converges to $X^{(4)}$ uniformly on $[0,T]$ along an appropriate subsequence as $\varepsilon \downarrow 0$. Obviously, the characteristic exponent of the latter Lévy process is

$$\psi^{(4)}(u) = \int\limits_{|x|\leq 1} \mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u,x\rangle\,)\nu(\mathrm{d}x). \tag{6.35}$$

Since the sets $\{|x|>1\}$ and $\{|x|\leq 1\}$ are obviously disjoint, the processes $X^{(3)}$ and $X^{(4)}$ are independent. Moreover, they are both independent of $X^{(1)}$ and $X^{(2)}$, which are defined on a different probability space.

Step 4. In order to conclude the proof, we consider the product space

$$(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega^\natural, \mathcal{F}^\natural, \mathbb{P}^\natural) \times (\Omega^\sharp, \mathcal{F}^\sharp, \mathbb{P}^\sharp). \tag{6.36}$$

The process $X = (X_t)_{t\geq 0}$ with

$$\begin{aligned} X_t &= X_t^{(1)} + X_t^{(2)} + X_t^{(3)} + X_t^{(4)} \\ &= bt + \sqrt{c}W_t + \int_0^t\int_{D^c} x\mu^X(\mathrm{d}s,\mathrm{d}x) + \int_0^t\int_D x(\mu^X-\nu^X)(\mathrm{d}s,\mathrm{d}x), \end{aligned} \tag{6.37}$$

is defined on this space, has stationary and independent increments, càdlàg paths, and the characteristic exponent is

$$\begin{aligned} \psi(u) &= \psi^{(1)}(u) + \psi^{(2)}(u) + \psi^{(3)}(u) + \psi^{(4)}(u) \\ &= i\,\langle u,b\rangle - \frac{\langle u,cu\rangle}{2} + \int\limits_{\mathbb{R}^d} \big(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\,\langle u,x\rangle\,1_D(x)\big)\nu(\mathrm{d}x). \end{aligned}$$

□

Remark 6.13 (Truncation function)**.** Assume that the infinitely divisible distribution ρ has the Lévy triplet $(b_h, c, \nu)_h$ relative to the truncation function h, that is, assume that the Fourier transform of ρ is given by (5.5)–(5.6) instead of (5.2). Then, the Lévy–Itô decomposition takes the form

$$X_t = b_h t + \sqrt{c}W_t + \int_0^t\int_{\mathbb{R}^d} h^c(x)\mu^X(\mathrm{d}s,\mathrm{d}x) + \int_0^t\int_{\mathbb{R}^d} h(x)(\mu^X-\nu^X)(\mathrm{d}s,\mathrm{d}x), \tag{6.38}$$

where $h^c(x) = x - h(x)$. This form of the Lévy–Itô decomposition is *consistent* with the choice of the truncation function h; see also Remark 5.10.

Example 6.14. Revisiting the Lévy jump-diffusion process, we can easily see that (3.1) is the Lévy–Itô decomposition of this Lévy process for the truncation function $h(x) = x$, while

$$X_t = b_0 t + \sigma W_t + \sum_{k=1}^{N_t} J_k, \tag{6.39}$$

where $b_0 = b - \lambda\beta$ is the Lévy–Itô decomposition of X relative to the truncation function $h(x) \equiv 0$. See also Example 5.11.

Exercise 12. Suppose X, Y are two independent Lévy processes (on the same probability space). Show that $X + Y$ and $X - Y$ are again Lévy processes. Can $X - Y$ be a Lévy process in case X and Y are not independent?

Exercise 13. Consider the space $(\mathbb{R}_{\geqslant 0} \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_{\geqslant 0}) \times \mathcal{B}(\mathbb{R}^d), \mathrm{Leb} \otimes \nu)$ and denote by μ^X the Poisson random measure with intensity $\mathrm{Leb} \otimes \nu$. Let $f : \mathbb{R}^d \to \mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |f(x)|^2 \nu(\mathrm{d}x) < \infty$. Show that the process $I = (I_t)_{t\geq 0}$ with

$$I_t = \int_0^t \int_{\mathbb{R}^d} f(x) \mu^X(\mathrm{d}s, \mathrm{d}x) - t \int_{\mathbb{R}^d} f(x) \nu(\mathrm{d}x) \tag{6.40}$$

is a square integrable martingale and prove the following simplified version of the Itô isometry

$$\mathbb{E}\big[|I_t|^2\big] = t \int_{\mathbb{R}^d} |f(x)|^2 \nu(\mathrm{d}x). \tag{6.41}$$

Exercise 14. Consider the setting of the previous exercise.
(i) Show that, for each $n \geq 2$ and each $t > 0$,

$$\int_0^t \int_{\mathbb{R}^d} x^n \mu^X(\mathrm{d}s, \mathrm{d}x) < \infty \quad \text{a.s.} \iff \int_{|x|>1} |x|^n \nu(\mathrm{d}x) < \infty. \tag{6.42}$$

(ii) Assuming that the previous condition holds, show that

$$\left(\int_0^t \int_{\mathbb{R}^d} x^n \mu^X(\mathrm{d}s, \mathrm{d}x) - t \int_{\mathbb{R}^d} x^n \nu(\mathrm{d}x) \right)_{t\geq 0} \tag{6.43}$$

is a martingale.

6.7. **Another approach to the basic connections.** We have now proved the basic connections between Lévy processes, infinitely divisible distributions and Lévy triplets, as announced in §3.1. The line of these proofs is diagrammatically represented in Figure 3.2. These relations are useful for the construction of new classes of Lévy processes and for the simulation of Lévy processes.

Naturally, there are other ways to prove these connections. Another approach is diagrammatically represented in Figure 6.4. The steps in these

$$\begin{tikzcd}
(X_t)_{t\geq 0} \arrow[rr, shift left, "SII"] \arrow[dr, "LI"'] & & \mathcal{L}(X_1) = \rho \arrow[ll, shift left, "KET"] \\
& (b, c, \nu) \arrow[ur, "LK"'] &
\end{tikzcd}$$

FIGURE 6.4. Another approach to the basic connections between Lévy processes, infinitely divisible distributions and Lévy triplets.

proofs can be summarized as follows:

(i) show that the law of X_t is infinitely divisible using the stationarity and independence of the increments (cf. Lemma 4.18);
(ii) show that for every Lévy triplet (b, c, ν) that satisfies (5.2) the measure ρ is infinitely divisible (cf. Theorem 5.3, "If" part);
(iii) use Kolmogorov's extension theorem to show that for every infinitely divisible distribution ρ, there exists a Lévy process $X = (X_t)_{t\geq 0}$ such that $\mathbb{P}_{X_1} = \rho$;
(iv) prove the following version of the Lévy–Itô decomposition: every Lévy process admits the path decomposition (6.2). A corollary of the last result is the Lévy–Khintchine formula, cf. (5.27)-(5.29).

This line of proofs is based on the analysis of the jumps of Lévy process and follows in spirit the analysis of the jumps of the compound Poisson process in §6.2. We refer the interested reader to [App09] and [Pro04].

7. The Lévy measure and path properties

The Lévy measure is the most interesting part of a Lévy process and is responsible for the richness of the class of these processes. The behaviour of the sample paths of a Lévy process, as well as many other properties, e.g. existence of moments, smoothness of densities, etc, can be completely characterized based on the Lévy measure and the presence or absence of a Brownian compenent.

Let us recall the definition of a Lévy measure: it is a measure on $\mathbb{R}^d$ that satisfies

$$\nu(\{0\}) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}} (1 \wedge |x|^2)\nu(\mathrm{d}x) < \infty. \tag{7.1}$$

The Lévy measures of certain examples of Lévy processes are presented in Figures 7.5 and 7.6. Using the proporties of the Poisson random measure, we can deduce that the Lévy measure satisfies

$$\mathbb{E}\left[\mu^X([0,1] \times A)\right] = \nu(A), \tag{7.2}$$

for every set $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. In other words, the Lévy measure describes *the expected number of jumps of a certain height in a time interval of length one.* The relation between Poisson random measures and Lévy measures allows us to draw the following conclusion about the sample paths of Lévy processes based on their Lévy measure: the Lévy measure has no mass at the origin while singularities can occur around it, thus a Lévy process can have an infinite number of small jumps — "small" here means bounded by one in absolute value, although we can consider any $\varepsilon > 0$ instead of one. Moreover, the mass away from the origin is bounded, hence only a finite number of big jumps can occur — again, "big" here means greater than one in absolute value.

7.1. **Path properties.** We would like to discuss some finer properties of the paths of a Lévy process, in particular, when are the paths continuous or piecewise constant and when they have finite or infinite variation. Throughout this section we assume that $X = (X_t)_{t\geq 0}$ is a Lévy process with triplet (b, c, ν).

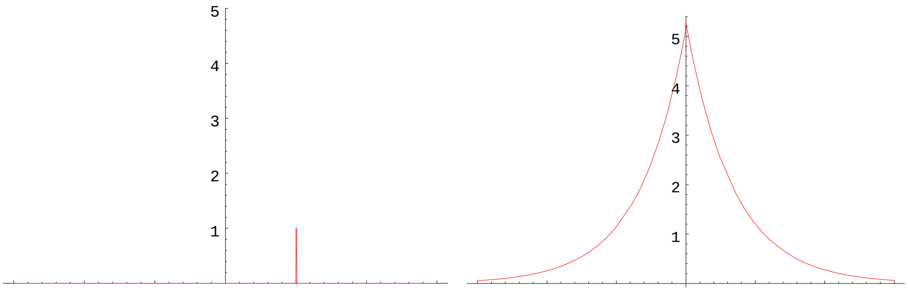

FIGURE 7.5. The distribution function of the Lévy measure of the standard Poisson process (left) and the density of the Lévy measure of a compound Poisson process with double-exponentially distributed jumps.

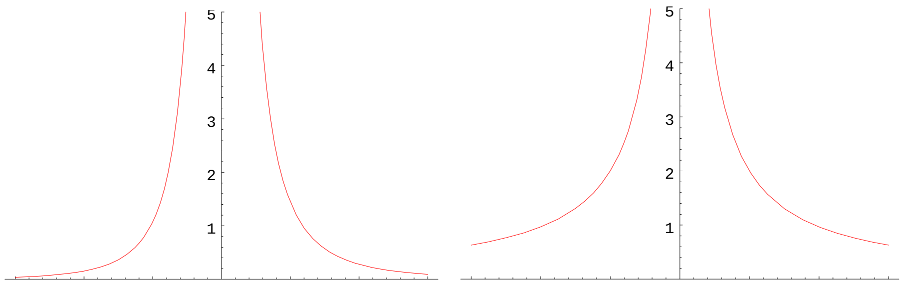

FIGURE 7.6. The density of the Lévy measure of a normal inverse Gaussian (NIG, left) and an α-stable process.

Proposition 7.1. *The paths of $X = (X_t)_{t\geq 0}$ are a.s. continuous if and only if $\nu \equiv 0$.*

Exercise 15. Let X be a Lévy process with Lévy measure ν.

(i) Show that for $a > 0$

$$\mathbb{P}\Big(\sup_{0<s\leq t} |X_s - X_{s-}| > a\Big) = 1 - \mathrm{e}^{-t\nu(\mathbb{R}\backslash(-a,a))}.$$

(ii) Use this to prove Proposition 7.1.

Proposition 7.2. *The paths of $X = (X_t)_{t\geq 0}$ are a.s. piecewise constant if and only if X is a compound Poisson process without drift.*

Exercise 16. Prove Proposition 7.2

Definition 7.3. A Lévy process X has *infinite activity* if the sample paths of X have an a.s. countably infinite number of jumps on every compact time interval $[0,T]$. Otherwise, X has *finite activity.*

Proposition 7.4. *(1) If $\nu(\mathbb{R}^d) = \infty$ then X has infinite activity.*
(2) If $\nu(\mathbb{R}^d) < \infty$ then X has finite activity.

Exercise 17. Prove Proposition 7.4

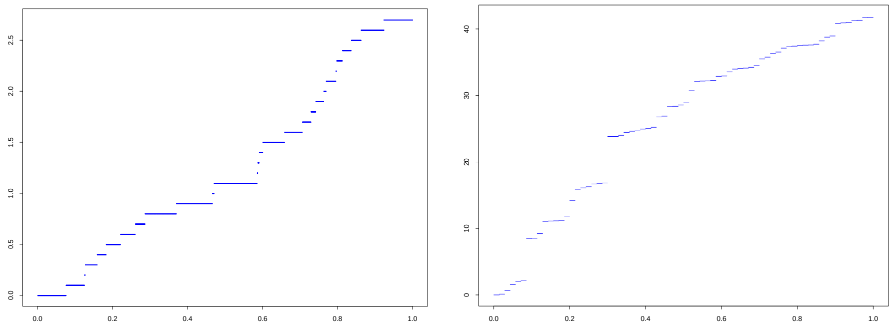

FIGURE 7.7. Simulated paths of a finite activity (left) and an infinite activity subordinator.

Remark 7.5. By the definition of a Lévy measure, cf. Definition 5.1, we get immediately the following equivalences:

$$\begin{aligned} \nu(\mathbb{R}^d) = \infty \quad &\iff \quad \nu(D) = \infty \\ \nu(\mathbb{R}^d) < \infty \quad &\iff \quad \nu(D) < \infty. \end{aligned} \tag{7.3}$$

Remark 7.6. Intuitively speaking, a Lévy process with infinite activity will jump more often than a process with finite activity. This is visually verified by the simulated paths of a compound Poisson and an inverse Gaussian subordinator presented in Figure 7.7.

Remark 7.7. The aforestated results allow us to deduce that if $\nu(D) < \infty$ and $c = 0$ then the Lévy process is actually a compound Poisson process. Since the Lévy measure ν is finite, i.e. $\lambda := \nu(\mathbb{R}) < \infty$, we can define $F(\mathrm{d}x) := \frac{\nu(\mathrm{d}x)}{\lambda}$, which is a probability measure. Thus, λ will be the expected number of jumps and $F(\mathrm{d}x)$ the distribution of the jump size.

7.2. **Variation of the paths.** Next, we want to analyze the variation of the paths of a Lévy process. We will consider a real-valued Lévy process for simplicity, although the main result, Proposition 7.11, is also valid for $\mathbb{R}^d$-valued Lévy processes.

Definition 7.8. Consider a function $f : [a,b] \to \mathbb{R}$. The *total variation* of f over $[a,b]$ is

$$\mathrm{TV}(f) = \sup_{\pi} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| \tag{7.4}$$

where $\pi = \{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of the interval $[a,b]$.

Lemma 7.9. *If $f : [a,b] \to \mathbb{R}$ is càdlàg and has finite variation on $[a,b]$, then*

$$\mathrm{TV}(f) \geq \sum_{t \in [a,b]} |\Delta f(t)|. \tag{7.5}$$

Proof. [App09, Theorem 2.3.14]. □

Definition 7.10. A stochastic process $X = (X_t)_{t\geq 0}$ has *finite variation* if the paths $(X_t(\omega))_{t\geq 0}$ have finite variation for almost all $\omega \in \Omega$. Otherwise, the process has *infinite variation.*

Proposition 7.11. *A Lévy process $X = (X_t)_{t\geq 0}$ with triplet (b, c, ν) has finite variation if and only if*

$$c = 0 \quad \textit{and} \quad \int_{|x|\leq 1} |x|\nu(\mathrm{d}x) < \infty. \tag{7.6}$$

Proof. Assume that $c = 0$, then the Lévy–Itô decomposition of the Lévy process takes the form

$$X_t = bt + \int_0^t \int_{|x|>1} x\mu^X(\mathrm{d}s, \mathrm{d}x) + \underbrace{\int_0^t \int_{|x|\leq 1} x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x)}_{=X_t^{(4)}}. \tag{7.7}$$

We know that the first and second processses have finite variation, hence we will concentrate on the last part. Using the definition we have

$$\begin{aligned}
\mathrm{TV}\big(X_t^{(4)}\big) &= \sup_\pi \sum_{i=1}^n \big|X_{t_i}^{(4)} - X_{t_{i-1}}^{(4)}\big| \\
&= \sup_\pi \sum_{i=1}^n \Big| \int_{t_{i-1}}^{t_i} \int_{|x|\leq 1} x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x) \Big| \\
&\leq \sup_\pi \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \int_{|x|\leq 1} |x|(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x) \\
&= \int_0^t \int_{|x|\leq 1} |x|(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x) \\
&= \int_0^t \int_{|x|\leq 1} |x|\mu^X(\mathrm{d}s, \mathrm{d}x) - t \int_{|x|\leq 1} |x|\nu^X(\mathrm{d}x) < \infty \text{ a.s.,}
\end{aligned} \tag{7.8}$$

since condition (7.6) for the Lévy measure implies that the integral with respect to the Poisson random measure μ^X in (7.8) is well defined and a.s. finite; cf. Theorem 6.9. Hence, we can split the integral wrt to the compensated random measure $\mu^X - \nu^X$ in two a.s. finite parts.

Conversely, assume that X has finite variation; then, we can use estimation (7.5), which yields

$$\infty > \mathrm{TV}(X_t) \geq \sum_{0\leq s\leq t} |\Delta X_s| \geq \sum_{0\leq s\leq t} |\Delta X_s| 1_{\{|\Delta X_s|\leq 1\}} = \int_0^t \int_{|x|\leq 1} |x|\mu^X(\mathrm{d}s, \mathrm{d}x).$$

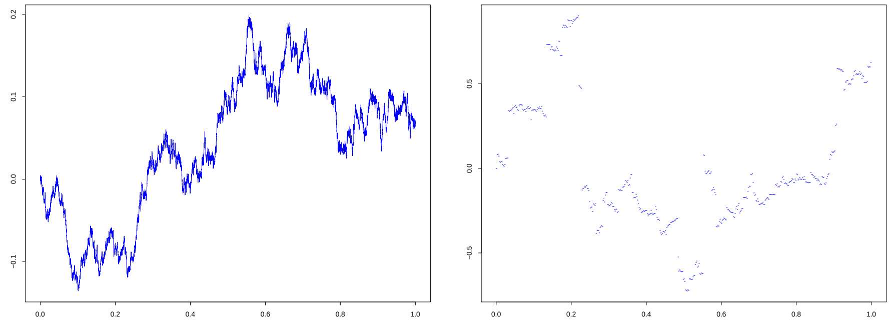

FIGURE 7.8. Simulated paths of two infinite variation Lévy processes: Brownian motion (left) and NIG process.

Using again Theorem 6.9, finiteness of the RHS implies that

$$\int_0^t \int_{|x|\leq 1} |x|\nu^X(\mathrm{d}s, \mathrm{d}x) < \infty \quad \Longrightarrow \quad \int_{|x|\leq 1} |x|\nu(\mathrm{d}x) < \infty, \tag{7.9}$$

which yields the second condition. The Lévy–Itô decomposition of this Lévy process—where the jumps have finite variation—takes the form

$$X_t = b't + \sqrt{c}W_t + \sum_{s\leq t} \Delta X_s. \tag{7.10}$$

However, the paths of a Brownian motion have infinite variation, see e.g. [RY99, Cor. I.2.5], hence X will have paths of finite variation if and only if $c = 0$. □

The simulated sample paths of a continuous Lévy process with infinite variation (i.e. Brownian motion) and a purely discontinuous one are depicted in Figure 7.8. We can observe that, locally, the pure-jump infinite variation process behaves like a Brownian motion, as it proceeds by infinitesimally small movements. However, these small jumps are interlaced with, less frequent, big jumps.

Remark 7.12. Assume that the *jump part* of the Lévy process X has finite variation, i.e. it holds that

$$\int_{|x|\leq 1} |x|\nu(\mathrm{d}x) < \infty. \tag{7.11}$$

Then, the Lévy–Itô decomposition of X takes the form

$$X_t = b_0 t + \sqrt{c}W_t + \int_0^t \int_{\mathbb{R}^d} x\mu^X(\mathrm{d}s, \mathrm{d}x), \tag{7.12}$$

and the Lévy–Khintchine formula can be written as

$$\mathbb{E}\big[\mathrm{e}^{i\langle u, X_1\rangle}\big] = \exp\left(i\,\langle u, b_0\rangle - \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}^d} \big(\mathrm{e}^{i\langle u, x\rangle} - 1\big)\nu(\mathrm{d}x)\right). \tag{7.13}$$

In other words, we can use the truncation function $h(x) = 0$ and the drift term relative to this truncation function (denoted by b_0) is related to the drift term b in (5.2) via

$$b_0 = b - \int_{|x|\leq 1} x\nu(\mathrm{d}x). \tag{7.14}$$

Note that this process is not necessarily a compound Poisson process, as the activity of the process might be infinite (i.e. $\nu(D) = \infty$).

7.3. **Subordinators.** The last part of this section will be devoted to Lévy processes which have a.s. increasing paths. These processes are typically called *subordinators.* Subordinators play an important role in the theory and also in the applications of Lévy processes in various fields, as they constitute a stochastic model for the evolution of time. We start by characterizing the Lévy triplet of a subordinator. We will concentrate on real-valued subordinators, for simplicity; for multivariate subordinators we refer the reader to [BNPS01].

Proposition 7.13. *Let $X = (X_t)_{t\geq 0}$ be a real-valued Lévy process with triplet (b, c, ν). The following are equivalent:*

(1) $X_t \geq 0$ a.s. for some $t > 0$;
(2) $X_t \geq 0$ a.s. for all $t > 0$;
(3) The sample paths of X are a.s. non-decreasing, that is $t \geq s \Longrightarrow X_t \geq X_s$;
(4) The triplet (b, c, ν) satisfies: $\int_0^1 x\nu(\mathrm{d}x) < \infty$, $c = 0$, $\nu(-\infty, 0]) = 0$ and $b \geq \int_0^1 x\nu(\mathrm{d}x)$. In other words, no diffusion component, jumps are only positive and have finite variation and the drift dominates the compensator of the small jumps.

Proof. The statements (3) $\Rightarrow$ (2) (take $s = 0$ in (3)) and (2) $\rightarrow$ (1) are obvious.
(1) $\Rightarrow$ (2) Without loss of generality we may assume that $X_1 \geq 0$ a.s. Since X_1 is the sum of n independent copies of $X_{1/n}$ it follows that $X_{1/n} \geq 0$ a.s. Similarly, it follows that $X_k \geq 0$ a.s. for any $k \in \mathbb{N}$. Combining these two observations leads to $X_q \geq 0$ a.s. for any rational $q > 0$. Statement (2) now follows from the right-continuity of paths of a Lévy process.

(2) $\Rightarrow$ (3) Since $X_t - X_s \overset{\mathrm{d}}{=} X_{t-s}$ it follows from (2) that $X_t - X_s \geq 0$ a.s.

(3) $\Rightarrow$ (4) Suppose that X is a Lévy process with triplet (b, c, ν) satisfying

$$c = 0, \nu(-\infty, 0) = 0, \int_{0<x<1} x\nu(\mathrm{d}x) < \infty \text{ and } b \geq \int_{0<x<1} x\nu(\mathrm{d}x).$$

This means that we can write

$$X_t = \left(b - \int_{0<x<1} x\nu(\mathrm{d}x)\right) t + \int_{[0,t]} \int_{(0,\infty)} xN(ds, dx),$$

where N is the Poisson random measure corresponding to the jumps of the Lévy process (no compensation needed here since X is of finite variation). This is clearly an increasing process.

(4) ⇒ (3) Conversely, if X has increasing paths, then it is of finite variation implying that $c = 0$ and $\int_{|x|<1} |x|\nu(\mathrm{d}x) < \infty$. We can write

$$X_t = \left(b - \int\limits_{0<x<1} x\nu(\mathrm{d}x) \right) t + \int\limits_{[0,t]} \int\limits_{\mathbb{R}} xN(ds, dx).$$

For this to be an increasing process it must hold that $b \geq \int_{0<x<1} x\Pi(\mathrm{d}x)$ and $\Pi(-\infty, 0) = 0$.

□

Corollary 7.14. *The Lévy–Itô decomposition of a subordinator is*

$$X_t = bt + \int\limits_0^t \int\limits_{\mathbb{R}_+} x\mu^X(\mathrm{d}s, \mathrm{d}x), \tag{7.15}$$

while the Lévy–Khintchine representation takes the form

$$\mathbb{E}\left[\mathrm{e}^{iuX_1}\right] = \exp\left(iub + \int\limits_{\mathbb{R}_+} (\mathrm{e}^{iux} - 1)\nu(\mathrm{d}x) \right). \tag{7.16}$$

Remark 7.15. Consider a Lévy process X with triplet (b, c, ν) such that the following proporties holds:

$$b \geq 0, \ c = 0, \ \nu((-\infty, 0]) = 0 \ \text{ and } \int_{(0,1]} |x|\nu(\mathrm{d}x) = \infty.$$

This process has the Lévy–Itô decomposition

$$X_t = bt + \int\limits_0^t \int\limits_{\mathbb{R}_+} x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x), \tag{7.17}$$

its paths are fluctuating but are *not* increasing—the paths have infinite variation—and this process is not a subordinator. The intuitive explanation for this bahavior is that the jump part will converge only if we add an "infinitely strong" deterministic term in the negative direction to compensate for the divergent sum of jumps. This term cannot be negated, however large we choose b.

Exercise 18. Let X be an $\mathbb{R}^d$-valued Lévy process and consider a function $f : \mathbb{R}^d \to \mathbb{R}_+$ such that $f(x) = O(|x|^2)$ as $|x| \to 0$. Show that the process $S = (S_t)_{t\geq 0}$ defined by

$$S_t = \sum_{0\leq s\leq t} f(\Delta X_s) \tag{7.18}$$

is a subordinator.

8. Elementary operations

In this section, we will study the result of certain elementary operations when applied to Lévy processes. The operations we have in mind are linear transformations, projections and subordination. The resulting processes belong to the class of Lévy processes again, and we will see that these operations can be expressed as simple transformations of the Lévy triplet or the Lévy exponent.

8.1. **Linear transformations and projections.** A very simple transformation of a Lévy process is to restrict and project it into a subspace of its state space, or to apply a linear transformation to it. The following result provides a complete characterization of linear transformations of Lévy processes in terms of their Lévy triplet.

Proposition 8.1. *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with triplet $(b, c, \nu)_h$. Let U be an $n \times d$ matrix with real entries ($U \in M_{nd}(\mathbb{R})$). Then, $X^U = (X_t^U)_{t\geq 0}$ with $X_t^U := UX_t$ is an $\mathbb{R}^n$-valued Lévy process with Lévy triplet $(b^U, c^U, \nu^U)_{h'}$, where*

$$b^U = Ub + \int_{\mathbb{R}^d} (h'(Ux) - Uh(x))\nu(\mathrm{d}x) \tag{8.1a}$$

$$c^U = UcU^\top \tag{8.1b}$$

$$\nu^U(E) = \nu(\{x \in \mathbb{R}^d : Ux \in E\}), \quad E \in \mathcal{B}(\mathbb{R}^n \backslash \{0\}). \tag{8.1c}$$

Here $h'(x)$ denotes a truncation function on $\mathbb{R}^n$.

Remark 8.2. The Lévy measure ν^U in (8.1c) is the *push-forward* of the measure ν by the operator U, also denoted $U_*\nu$. We have that, for suitable functions f and $E \in \mathcal{B}(\mathbb{R}^n \backslash \{0\})$, it holds

$$\int_E f(y)\nu^U(\mathrm{d}y) = \int_{\mathbb{R}^d} 1_E(Ux)f(Ux)\nu(\mathrm{d}x).$$

Proof. Since U defines a linear mapping from $\mathbb{R}^d$ to $\mathbb{R}^n$, it is clear that X^U has independent and stationary increments, and is stochastically continuous; moreover, $X_0^U = 0$ a.s. In other words, X^U is an $\mathbb{R}^n$-valued Lévy process.

We will show that ν^U is a Lévy measure and the integral on the RHS of b^U is finite; hence, the triplet (b^U, c^U, ν^U) in (8.1a) is indeed a Lévy triplet. Then we will derive the characteristic function of X_t^U.

Clearly ν^U has no mass at the origin; in addition we have that

$$\begin{aligned}\int_{\mathbb{R}^n} (|y|^2 \wedge 1)\nu^U(\mathrm{d}y) &= \int_{\mathbb{R}^d} (|Ux|^2 \wedge 1)\nu(\mathrm{d}x) \\ &\leq (\|U\|^2 \vee 1) \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\nu(\mathrm{d}x) < \infty,\end{aligned}$$

because the induced norm satisfies $|Ux| \leq \|U\||x|$ for all $U \in M_{nd}(\mathbb{R})$ and $x \in \mathbb{R}^d$.

Next, we restrict ourselves to the canonical truncation function for simplicity, i.e. $h(x) = x1_{\{|x|\leq 1\}}$, and derive the following result for the integral on the RHS of b^U:

$$\begin{aligned}
\int_{\mathbb{R}^d} |h'(Ux) - Uh(x)|\nu(\mathrm{d}x) \\
&\leq \int_{\mathbb{R}^d} |Ux||1_{\{|Ux|\leq 1\}} - 1_{\{|x|\leq 1\}}|\nu(\mathrm{d}x) \\
&= \int_{\mathbb{R}^d} |Ux||1_{\{|Ux|\leq 1<|x|\}} - 1_{\{|x|\leq 1<|Ux|\}}|\nu(\mathrm{d}x) \\
&\leq \int_{\{|Ux|\leq 1<|x|\}} |Ux|\nu(\mathrm{d}x) + \int_{\{|x|\leq 1<|Ux|\}} |Ux|\nu(\mathrm{d}x) \\
&\leq \int_{\{|x|>1\}} \nu(\mathrm{d}x) + \|U\| \int_{\{|x|\leq 1<\|U\||x|\}} |x|\nu(\mathrm{d}x) \\
&\leq \int_{\{|x|>1\}} \nu(\mathrm{d}x) + \|U\|^2 \int_{\{\frac{1}{\|U\|}<|x|\leq 1\}} |x|^2\nu(\mathrm{d}x) < \infty.
\end{aligned}$$

Finally, regarding the characteristic function we have for any $z \in \mathbb{R}^n$

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{i\langle z, X_1^U\rangle}\right] = \mathbb{E}\left[\mathrm{e}^{i\langle z, UX_1\rangle}\right] &= \mathbb{E}\left[\mathrm{e}^{i\langle U^\top z, X_1\rangle}\right] \\
&= \exp\bigg(i\langle U^\top z, b\rangle - \frac{1}{2}\langle U^\top z, cU^\top z\rangle \\
&\qquad + \int_{\mathbb{R}^d} (\mathrm{e}^{i\langle U^\top z, x\rangle} - 1 - i\langle U^\top z, h(x)\rangle)\nu(\mathrm{d}x)\bigg) \\
&= \exp\bigg(i\langle z, Ub\rangle - \frac{1}{2}\langle z, UcU^\top z\rangle \\
&\qquad + \int_{\mathbb{R}^d} (\mathrm{e}^{i\langle z, Ux\rangle} - 1 - i\langle z, Uh(x)\rangle)\nu(\mathrm{d}x)\bigg) \\
&= \exp\bigg(i\langle z, b^U\rangle - \frac{1}{2}\langle z, c^U z\rangle \\
&\qquad + \int_{\mathbb{R}^n} (\mathrm{e}^{i\langle z, y\rangle} - 1 - i\langle z, h'(y)\rangle)\nu^U(\mathrm{d}y)\bigg),
\end{aligned}$$

where b^U is given by (8.1a). Thus, (b^U, c^U, ν^U) is indeed the triplet of the Lévy process X^U. □

8.2. **Subordination.** Subordinators are Lévy processes with a.s. non-decreasing paths; see . . . for a complete characterization. Subordinators can be thought of a stochastic model for the evolution of time. *Subordination* is the tranformation of one stochastic process to a new one through a random time-change by an indepedent subordinator. This idea was introduced by Bochner. Note that one can also subordinate a semigroup of linear operators to create a new semigroup.

In mathematical finance, subordination plays a prominent role. Many popular Lévy modes can be constructed by subordinating Brownian motion, e.g. VG and NIG. In that setting, one often speaks about "calendar" time and "business" time. Subordination is also used to create multidimensional models with dependence structure via a common time-change.

Let $Y = (Y_t)_{t\geq 0}$ be a suborinator, i.e. a Lévy process with a.s. increasing paths. Let ψ_Y denote the characteristic exponent of Y; using ... we know that it has the form

$$\psi_Y(u) = ib_Y u + \int\limits_{(0,\infty)} (\mathrm{e}^{iux} - 1)\nu_Y(\mathrm{d}y). \tag{8.2}$$

Note that $\mathbb{E}[\mathrm{e}^{uY_t}] < \infty$ for all $u \leq 0$ since Y takes only non-negative values; therefore, $\int_{x>1} \mathrm{e}^{uy}\nu_Y(\mathrm{d}y) < \infty$ for all $u \leq 0$. Therefore, the characteristic exponent of Y can be extended to an analytic function for $u \leq 0$, and the moment generating function of Y_t is

$$\mathbb{E}[\mathrm{e}^{\langle u, Y_t\rangle}] = \mathrm{e}^{t\phi_Y(u)} \tag{8.3}$$

where

$$\phi_Y(u) = b_Y u + \int\limits_{(0,\infty)} (\mathrm{e}^{ux} - 1)\nu_Y(\mathrm{d}y). \tag{8.4}$$

Theorem 8.3. *Let X be an $\mathbb{R}^d$-valued Lévy process with characteristic exponent ψ_X. Let Y be a subordinator with cumulant generating function ϕ_Y, where Y is independent of X. Define the process $Z = (Z_t)_{t\geq 0}$ for each $\omega \in \Omega$ via*

$$Z_t(\omega) = X_{Y_t(\omega)}(\omega). \tag{8.5}$$

Then, Z is a Lévy process with characteristic exponent

$$\psi_Z(u) = \phi_Y(\psi_X(u)). \tag{8.6}$$

Proof. ... □

Exercise 19. Show that any Lévy process with finite variation can be written as the difference of two independent subordinators.

9. Moments and Martingales

In this section, we turn out attention to the finiteness of the moments of Lévy processses. As motivation, we consider the Brownian motion and the compound Poisson process. Let B denote a Brownian motion with drift b and variance $\sqrt{c}$; then it is well known that the moments of B_t, for any order, are finite and the moment generating function has the form

$$M_{B_t}(u) = \mathbb{E}\big[\mathrm{e}^{\langle u, B_t\rangle}\big] = \exp\left(t\Big(\langle u, b\rangle + \frac{\langle u, cu\rangle}{2}\Big)\right), \tag{9.1}$$

for any $t \geq 0$ and any $u \in \mathbb{R}^d$.

On the other hand, consider a compound Poisson process X with jump intensity λt, $0 \leq \lambda < \infty$, and jump distribution F, i.e.

$$X_t = \sum_{k=1}^{N_t} J_k.$$

We know from Example ... that

$$\mathbb{E}[X_t] = \lambda t \mathbb{E}[J_k].$$

Thus, X_t does not have a finite moment if $\mathbb{E}[J_k] = \infty$. It turns out that the finiteness of the moments of a Lévy process is closely related to the finiteness of an integral over the Lévy measure of the process, and in particular over the *big* jumps (i.e. $|x| > 1$).

Theorem 9.1. *Let X be an $\mathbb{R}$-valued Lévy process with triplet (b, c, ν) and $u \in \mathbb{R}$. Then, for any $t \geq 0$*

$$\mathbb{E}\big[\mathrm{e}^{uX_t}\big] < \infty \qquad \textit{if and only if} \qquad \int\limits_{|x|>1} \mathrm{e}^{ux}\nu(\mathrm{d}x). \tag{9.2}$$

Proof. "Only If" part. Assume that $\mathbb{E}[\mathrm{e}^{uX_t}] < \infty$ for some $t > 0$. Recall that the Lévy–Itô decomposition of a Lévy process X is

$$X_t = X_t^{(1)} + X_t^{(2)} + X_t^{(3)} + X_t^{(4)}, \tag{9.3}$$

where $X^{(3)}$ is a compound Poisson process with arrival rate $\lambda := \nu(D^c)$ and jump distribution $F(\mathrm{d}x) := \frac{\nu(\mathrm{d}x)}{\nu(D^c)} 1_{\{|x|\geq 1\}}$. Now, using the independence of $X^{(1)}, \dots, X^{(4)}$ we have that

$$\infty > \mathbb{E}\big[\mathrm{e}^{uX_t}\big] = \mathbb{E}\big[\mathrm{e}^{u(X_t^{(1)}+X_t^{(2)}+X_t^{(4)})}\big] \times \mathbb{E}\big[\mathrm{e}^{uX_t^{(3)}}\big]; \tag{9.4}$$

thus, by the assumption we get that

$$\mathbb{E}\big[\mathrm{e}^{uX_t^{(3)}}\big] < \infty. \tag{9.5}$$

Since $X^{(3)}$ is a compound Poisson process we know (see Example ...) that

$$\begin{aligned}
\mathbb{E}[\mathrm{e}^{uX_t^{(3)}}] &= \mathrm{e}^{-\lambda t} \sum_{k\geq 0} \frac{(\lambda t)^k}{k!} \left(\int\limits_{\mathbb{R}} \mathrm{e}^{ux} F(\mathrm{d}x) \right)^k \\
&= \mathrm{e}^{-\lambda t} \sum_{k\geq 0} \frac{t^k}{k!} \left(\int\limits_{\mathbb{R}} \mathrm{e}^{ux} 1_{\{|x|\geq 1\}} \frac{\nu(\mathrm{d}x)}{\nu(D^c)} \right)^k .
\end{aligned}$$

All the summands must be finite, hence the one corresponding to $k = 1$ must also be finite, and we can conclude

$$\mathrm{e}^{-\lambda t} \frac{\lambda t}{\nu(D^c)} \int\limits_{\mathbb{R}} \mathrm{e}^{ux} 1_{\{|x|\geq 1\}} \nu(\mathrm{d}x) < \infty \Longrightarrow \int\limits_{|x|\geq 1} \mathrm{e}^{ux}\nu(\mathrm{d}x) < \infty.$$

"If" part. Conversely, assume that $\int_{|x|\geq 1} \mathrm{e}^{ux}\nu(\mathrm{d}x) < \infty$ for some $u \in \mathbb{R}$. Since $(\nu|_{D^c})^{*n}$ is a finite measure, we have that

$$\int\limits_{\mathbb{R}} \mathrm{e}^{ux}(\nu|_{D^c})^{*n}(\mathrm{d}x) = \Big(\int\limits_{|x|\geq 1} \mathrm{e}^{ux}\nu(\mathrm{d}x)\Big)^n < \infty, \tag{9.6}$$

therefore,

$$\mathbb{E}[\mathrm{e}^{uX_t^{(3)}}] = \mathrm{e}^{-\lambda t} \sum_{k\geq 0} \frac{t^k}{k!} \left(\int\limits_{\mathbb{R}} \mathrm{e}^{ux} 1_{\{|x|\geq 1\}} \frac{\nu(\mathrm{d}x)}{\nu(D^c)} \right)^k < \infty$$

for all $t > 0$. In order to complete the proof, we have to show that

$$\mathbb{E}\big[\mathrm{e}^{u(X_t^{(1)}+X_t^{(2)}+X_t^{(4)})}\big] < \infty \tag{9.7}$$

for all $t > 0$. Note that $X^{(1)}+X^{(2)}+X^{(4)}$ is a Lévy process with characteristic exponent

$$\psi'(u) = iub - \frac{u^2 c}{2} + \int\limits_{|x|\leq 1} \big(\mathrm{e}^{iux} - 1 - iux\big)\nu(\mathrm{d}x). \tag{9.8}$$

The first and second summands clearly admit an analytic expression to the complex plane. Regarding the integral term, we have that

$$\int\limits_{|x|\leq 1} \big(\mathrm{e}^{iux} - 1 - iux\big)\nu(\mathrm{d}x) = \int\limits_{|x|\leq 1} \Big(\sum_{k\geq 0} \frac{(iux)^{k+2}}{(k+2)!} \Big)\nu(\mathrm{d}x), \tag{9.9}$$

and we can exchange the sum and the integral using Fubini's theorem and the estimation

$$\sum_{k\geq 0} \int\limits_{|x|\leq 1} \frac{|ux|^{k+2}}{(k+2)!}\nu(\mathrm{d}x) \leq \sum_{k\geq 0} \frac{|u|^{k+2}}{(k+2)!} \int\limits_{|x|\leq 1} |x|^2 \nu(\mathrm{d}x) < \infty, \tag{9.10}$$

since the Lévy measure has bounded support, namely in $\{|x| \leq 1\}$.

Therefore, ψ' can be extended to an analytic function to the whole complex plane $\mathbb{C}$ and eq. (9.7) holds true. $\square$

Corollary 9.2. *Let ρ be an infinitely divisible distribution on $\mathbb{R}$ whose Lévy measure has bounded support. Then, $\widehat{\rho}$ can be extended to an entire function on $\mathbb{C}$.*

9.1. **Submultiplicative functions.** The previous theorem actually holds for a broader class of functions than just the exponential; these functions are called submultiplicative functions, and we will briefly describe them here.

Definition 9.3. A function $f : \mathbb{R}^d \to \mathbb{R}_{\geqslant 0}$ is called *submultiplicative* if there exists a constant $c > 0$ such that for all $x, y \in \mathbb{R}^d$

$$f(x+y) \leq cf(x)f(y). \tag{9.11}$$

Definition 9.4. A function is called *locally bounded* if it is bounded on every compact set.

Proposition 9.5. *(1) Let f, g be submultiplicative functions then fg is also submultiplicative.*

(2) If f is a submultiplicative function then so is $f(\beta x + \gamma)^\alpha$, where $\alpha > 0$, $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}^d$.

Proof. Direct consequences of the definition. □

Example 9.6. The following are some characteristic examples of submultiplicative functions:

$$|x| \vee 1,\ |x_i| \vee 1,\ x_i \vee 1$$
$$\exp(|x|^\beta),\ \exp(|x_i|^\beta),\ \exp((x_i \vee 0)^\beta)$$
$$\log(|x| \vee \mathrm{e}),\ \log(|x_i| \vee \mathrm{e}),\ \log(x_i \vee \mathrm{e}),$$

where $\beta \in (0,1]$, $x \in \mathbb{R}^d$ and x_i denotes the i-th component of x.

9.2. **f-Moments.** We will call the expectation of X_t with respect to a suitable function f an *f-moment.*

Theorem 9.7. *Let X be an $\mathbb{R}^d$-valued Lévy process with triplet (b, c, ν). Let f be a submultiplicative, locally bounded, measurable function on $\mathbb{R}^d$. Then, for any $t \geq 0$*

$$\mathbb{E}\big[f(X_t)\big] < \infty \qquad \textit{if and only if} \qquad \int\limits_{|x|>1} f(x)\nu(\mathrm{d}x) < \infty. \tag{9.12}$$

Proof. [Sat99, Theorem 25.3]. □

Actually, the proof of this Theorem follows along the same lines as the proof of Theorem ... making also use of the following result on submultiplicative functions.

Lemma 9.8. *Let f be a submultiplicative and locally bounded function. Then*

$$f(x) \leq b\mathrm{e}^{a|x|}, \tag{9.13}$$

for some constants $a, b > 0$.

Proof. Since f is bounded on compact sets, we can choose b such that

$$\sup_{|x|\leq 1} f(x) \leq b, \tag{9.14}$$

and further $bc > 1$ (where c is the constant from the submultiplicative property). Using that f is submultiplicative we get

$$f(x) = f\Big(\sum_{i=1}^{n} \frac{x}{n}\Big) \leq c^{n-1} f\Big(\frac{x}{n}\Big)^n. \tag{9.15}$$

If we choose n such that $n - 1 < |x| \leq n$, then we have

$$\begin{aligned} f(x) &\leq c^{n-1} g\Big(\frac{x}{n}\Big)^n \leq c^{n-1} b^n = b(bc)^{n-1} \\ &\leq b(bc)^{|x|} = b\mathrm{e}^{a|x|}. \end{aligned}$$
□

Remark 9.9. In case the Lévy process X has finite first moment, i.e.

$$\int\limits_{|x|>1} |x|\nu(\mathrm{d}x) < \infty, \tag{9.16}$$

then the Lévy–Itô decomposition of X takes the form

$$L_t = b_1 t + \sqrt{c} W_t + \int_0^t \int_{\mathbb{R}^d} x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x), \tag{9.17}$$

and the Lévy–Khintchine formula can be written as

$$\mathbb{E}\left[\mathrm{e}^{i\langle u, X_1\rangle}\right] = \exp\left(i\langle u, b_1\rangle - \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}^d} \left(\mathrm{e}^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\right)\nu(\mathrm{d}x)\right). \tag{9.18}$$

In other words, we can use the truncation function $h(x) = x$ and the drift term relative to this truncation function (denoted by b_1) is related to the drift term b in ... via

$$b_1 = b + \int_{|x|>1} x\nu(\mathrm{d}x). \tag{9.19}$$

9.3. **Moment generating function.** Finally, we are interested in the domain of definition and the form of the characteristic function of a Lévy process. The next result provides the answer.

Theorem 9.10. *Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}^d$-valued Lévy process with Lévy triplet (b, c, ν) and characteristic exponent ψ. Let*

$$\mathcal{U} = \left\{ u \in \mathbb{R}^d : \int_{\{|x|>1\}} \mathrm{e}^{\langle u,x\rangle}\nu(\mathrm{d}x) < \infty \right\}. \tag{9.20}$$

(i) The set $\mathcal{U}$ is convex and contains the origin.
(ii) $u \in \mathcal{U}$ if and only if $\mathbb{E}[\mathrm{e}^{\langle u, X_t\rangle}] < \infty$ for some (and hence all) $t \geq 0$.
(iii) If $u \in \mathbb{C}^d$ such that $\Re u \in \mathcal{U}$, then

$$\phi(u) = \langle u, b\rangle + \frac{\langle u, cu\rangle}{2} + \int_{\mathbb{R}^d} \left(\mathrm{e}^{\langle u,x\rangle} - 1 - \langle u, h(x)\rangle\right)\nu(\mathrm{d}x) \tag{9.21}$$

is well-defined, $\mathbb{E}|\mathrm{e}^{\langle u, X_t\rangle}| < \infty$, and the moment generating function is

$$\mathbb{E}[\mathrm{e}^{\langle u, X_t\rangle}] = \mathrm{e}^{t\phi(u)} = \mathrm{e}^{t\psi(iu)}. \tag{9.22}$$

Proof. [Sat99, Theorem 25.17]. □

We close this section on moments of Lévy processes with a result concerning the f-moments of the supremum of a Lévy process.

Theorem 9.11. *Let X be an $\mathbb{R}^d$-valued Lévy process. Define*

$$X_t^* = \sup_{0\leq s\leq t} |X_s|. \tag{9.23}$$

Let $f : \mathbb{R}^d \to [0, \infty)$ be a continuous submultiplicative function, increasing to ∞ as $x \to \infty$. The following are equivalent:
(1) $\mathbb{E}[f(X_t^*)] < \infty$ for all $t > 0$,
(2) $\mathbb{E}[f(|X_t|)] < \infty$ for all $t > 0$.

Proof. [Sat99, Theorem 25.18]. □

9.4. **Martingales and Lévy processes.** We are interested in constructing martingales that are driven by Lévy processes.

Proposition 9.12. *Let X be an $\mathbb{R}^d$-valued Lévy process with Lévy triplet (b, c, ν), characteristic exponent ψ and cumulant generating function ϕ.*

(1) If $\int_{|x|>1} |x|\nu(\mathrm{d}x) < \infty$, then X is a martingale if and only if

$$b + \int\limits_{|x|>1} x\nu(\mathrm{d}x) = 0.$$

(2) If $\int_{|x|>1} |x|\nu(\mathrm{d}x) < \infty$, then $(X_t - \mathbb{E}[X_t])_{t\geq 0}$ is a martingale.

(3) If $\int_{|x|>1} \mathrm{e}^{\langle u,x\rangle}\nu(\mathrm{d}x) < \infty$ for some $u \in \mathbb{R}^d$, then $M = (M_t)_{t\geq 0}$ is a martingale, where

$$M_t = \frac{\mathrm{e}^{\langle u,X_t\rangle}}{\mathrm{e}^{t\phi(u)}}. \tag{9.24}$$

(4) The process $N = (N_t)_{t\geq 0}$ is a complex-valued martingale, where

$$N_t = \frac{\mathrm{e}^{i\langle u,X_t\rangle}}{\mathrm{e}^{t\psi(u)}}. \tag{9.25}$$

Proof. . . . □

We can also construct martingales using the conditional expectation; this recipe is, of course, not restricted to Lévy processes

Proposition 9.13. *Let Y be an integrable and measurable random variable. Then, $L = (L_t)_{t\geq 0}$ is a martingale, where*

$$L_t = \mathbb{E}[Y|\mathcal{F}_t]. \tag{9.26}$$

Proof. . . . □

10. Popular examples

In this section, we review some popular models in the mathematical finance literature from the point of view of real-valued Lévy processes. We describe their Lévy triplets and characteristic functions and provide, whenever possible, their – infinitely divisible – laws.

10.1. **Black–Scholes.** The most famous asset price model based on a Lévy process is that of [Sam65], [BS73] and Merton [Mer73]. The log-returns are normally distributed with mean μ and variance σ^2, i.e. $L_1 \sim \text{Normal}(\mu, \sigma^2)$ and the density is

$$f_{L_1}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big[- \frac{(x-\mu)^2}{2\sigma^2}\Big].$$

The characteristic function is

$$\varphi_{L_1}(u) = \exp\Big[i\mu u - \frac{\sigma^2 u^2}{2}\Big],$$

the first and second moments are

$$\mathrm{E}[L_1] = \mu, \qquad \mathrm{Var}[L_1] = \sigma^2,$$

while the skewness and kurtosis are

$$\text{skew}[L_1] = 0, \qquad \text{kurt}[L_1] = 3.$$

The canonical decomposition of L is

$$L_t = \mu t + \sigma W_t$$

and the Lévy triplet is $(\mu, \sigma^2, 0)$.

10.2. **Merton.** [Mer76] was one of the first to use a discontinuous price process to model asset returns. The canonical decomposition of the driving process is

$$L_t = \mu t + \sigma W_t + \sum_{k=1}^{N_t} J_k$$

where $J_k \sim \text{Normal}(\mu_J, \sigma_J^2)$, $k = 1, \ldots$, hence the distribution of the jump size has density

$$f_J(x) = \frac{1}{\sigma_J \sqrt{2\pi}} \exp\Big[- \frac{(x-\mu_J)^2}{2\sigma_J^2} \Big].$$

The characteristic function of L_1 is

$$\varphi_{L_1}(u) = \exp\Big[i\mu u - \frac{\sigma^2 u^2}{2} + \lambda\big(\mathrm{e}^{i\mu_J u - \sigma_J^2 u^2/2} - 1 \big) \Big],$$

and the Lévy triplet is $(\mu, \sigma^2, \lambda \times f_J)$.

The density of L_1 is not known in closed form, while the first two moments are

$$\mathrm{E}[L_1] = \mu + \lambda \mu_J \qquad \text{and} \qquad \mathrm{Var}[L_1] = \sigma^2 + \lambda \mu_J^2 + \lambda \sigma_J^2$$

10.3. **Kou.** [Kou02] proposed a jump-diffusion model similar to Merton's, where the jump size is double-exponentially distributed. Therefore, the canonical decomposition of the driving process is

$$L_t = \mu t + \sigma W_t + \sum_{k=1}^{N_t} J_k$$

where $J_k \sim \text{DbExpo}(p, \theta_1, \theta_2)$, $k = 1, \ldots$, hence the distribution of the jump size has density

$$f_J(x) = p\theta_1 \mathrm{e}^{-\theta_1 x} 1_{\{x<0\}} + (1-p)\theta_2 \mathrm{e}^{\theta_2 x} 1_{\{x>0\}}.$$

The characteristic function of L_1 is

$$\varphi_{L_1}(u) = \exp\Big[i\mu u - \frac{\sigma^2 u^2}{2} + \lambda\Big(\frac{p\theta_1}{\theta_1 - iu} - \frac{(1-p)\theta_2}{\theta_2 + iu} - 1 \Big) \Big],$$

and the Lévy triplet is $(\mu, \sigma^2, \lambda \times f_J)$.

The density of L_1 is not known in closed form, while the first two moments are

$$\mathrm{E}[L_1] = \mu + \frac{\lambda p}{\theta_1} - \frac{\lambda(1-p)}{\theta_2} \qquad \text{and} \qquad \mathrm{Var}[L_1] = \sigma^2 + \frac{\lambda p}{\theta_1^2} + \frac{\lambda(1-p)}{\theta_2^2}.$$

10.4. **Generalized Hyperbolic.** The generalized hyperbolic model was introduced by [EP02] following the seminal work on the hyperbolic model by [EK95]. The class of hyperbolic distributions was invented by O. E. Barndorff-Nielsen in relation to the so-called 'sand project' (cf. [BN77]). The increments of time length 1 follow a generalized hyperbolic distribution with parameters $\alpha, \beta, \delta, \mu, \lambda$, i.e. $L_1 \sim \mathrm{GH}(\alpha, \beta, \delta, \mu, \lambda)$ and the density is

$$\begin{aligned} f_{GH}(x) = c(\lambda, \alpha, \beta, \delta)\big(\delta^2 + (x-\mu)^2\big)^{(\lambda-\frac{1}{2})/2} \\ \times K_{\lambda-\frac{1}{2}}\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)\exp\big(\beta(x-\mu)\big), \end{aligned}$$

where

$$c(\lambda, \alpha, \beta, \delta) = \frac{(\alpha^2-\beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)}$$

and K_λ denotes the Bessel function of the third kind with index λ (cf. [AS68]). Parameter $\alpha > 0$ determines the shape, $0 \leq |\beta| < \alpha$ determines the skewness, $\mu \in \mathbb{R}$ the location and $\delta > 0$ is a scaling parameter. The last parameter, $\lambda \in \mathbb{R}$ affects the heaviness of the tails and allows us to navigate through different subclasses. For example, for $\lambda = 1$ we get the hyperbolic distribution and for $\lambda = -\frac{1}{2}$ we get the normal inverse Gaussian (NIG).

The characteristic function of the GH distribution is

$$\varphi_{GH}(u) = \mathrm{e}^{iu\mu}\Big(\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+iu)^2}\Big)^{\frac{\lambda}{2}}\frac{K_\lambda\big(\delta\sqrt{\alpha^2-(\beta+iu)^2}\big)}{K_\lambda\big(\delta\sqrt{\alpha^2-\beta^2}\big)},$$

while the first and second moments are

$$\mathrm{E}[L_1] = \mu + \frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}$$

and

$$\mathrm{Var}[L_1] = \frac{\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \frac{\beta^2\delta^4}{\zeta^2}\Big(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K_{\lambda+1}^2(\zeta)}{K_\lambda^2(\zeta)}\Big),$$

where $\zeta = \delta\sqrt{\alpha^2-\beta^2}$.

The canonical decomposition of a Lévy process driven by a generalized hyperbolic distribution (i.e. $L_1 \sim \mathrm{GH}$) is

$$L_t = t\mathrm{E}[L_1] + \int_0^t\int_{\mathbb{R}} x(\mu^L - \nu^{GH})((\mathrm{d}s, \mathrm{d}x))$$

and the Lévy triplet is $(E[L_1], 0, \nu^{GH})$. The Lévy measure of the GH distribution has the following form

$$\nu^{GH}(\mathrm{d}x) = \frac{\mathrm{e}^{\beta x}}{|x|}\left(\int_0^\infty \frac{\exp(-\sqrt{2y+\alpha^2}\,|x|)}{\pi^2 y(J_{|\lambda|}^2(\delta\sqrt{2y}) + Y_{|\lambda|}^2(\delta\sqrt{2y}))}\mathrm{d}y + \lambda\mathrm{e}^{-\alpha|x|}1_{\{\lambda\geq 0\}}\right);$$

here J_λ and Y_λ denote the Bessel functions of the first and second kind with index λ. We refer to [Rai00, section 2.4.1] for a fine analysis of this Lévy measure.

The GH distribution contains as special or limiting cases several known distributions, including the normal, exponential, gamma, variance gamma, hyperbolic and normal inverse Gaussian distributions; we refer to Eberlein and v. Hammerstein [EvH04] for an exhaustive survey.

10.5. **Normal Inverse Gaussian.** The normal inverse Gaussian distribution is a special case of the GH for $\lambda = -\frac{1}{2}$; it was introduced to finance in [BN97]. The density is

$$f_{NIG}(x) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2-\beta^2} + \beta(x-\mu)\right) \frac{K_1\left(\alpha\delta\sqrt{1+(\frac{x-\mu}{\delta})^2}\right)}{\sqrt{1+(\frac{x-\mu}{\delta})^2}},$$

while the characteristic function has the simplified form

$$\varphi_{NIG}(u) = \mathrm{e}^{iu\mu} \frac{\exp(\delta\sqrt{\alpha^2-\beta^2})}{\exp(\delta\sqrt{\alpha^2-(\beta+iu)^2})}.$$

The first and second moments of the NIG distribution are

$$\mathrm{E}[L_1] = \mu + \frac{\beta\delta}{\sqrt{\alpha^2-\beta^2}} \qquad \text{and} \qquad \mathrm{Var}[L_1] = \frac{\delta}{\sqrt{\alpha^2-\beta^2}} + \frac{\beta^2\delta}{(\sqrt{\alpha^2-\beta^2})^3},$$

and similarly to the GH, the canonical decomposition is

$$L_t = t\mathrm{E}[L_1] + \int_0^t \int_{\mathbb{R}} x(\mu^L - \nu^{NIG})((\mathrm{d}s, \mathrm{d}x)),$$

where now the Lévy measure has the simplified form

$$\nu^{NIG}(\mathrm{d}x) = \mathrm{e}^{\beta x} \frac{\delta\alpha}{\pi|x|} K_1(\alpha|x|)\mathrm{d}x.$$

The NIG is the only subclass of the GH that is closed under convolution, i.e. if $X \sim \mathrm{NIG}(\alpha, \beta, \delta_1, \mu_1)$ and $Y \sim \mathrm{NIG}(\alpha, \beta, \delta_2, \mu_2)$ and X is independent of Y, then

$$X + Y \sim \mathrm{NIG}(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2).$$

Therefore, if we estimate the returns distribution at some time scale, then we know it – in closed form – for all time scales.

10.6. **CGMY.** The CGMY Lévy process was introduced by Carr, Geman, Madan, and Yor [CGMY02]; another name for this process is (generalized) tempered stable process (see e.g. [CT04]). The characteristic function of L_t, $t \in [0, T]$ is

$$\varphi_{L_t}(u) = \exp\Big(tC\Gamma(-Y)\big[(M-iu)^Y + (G+iu)^Y - M^Y - G^Y\big]\Big).$$

The Lévy measure of this process admits the representation

$$\nu^{CGMY}(\mathrm{d}x) = C\frac{\mathrm{e}^{-Mx}}{x^{1+Y}} 1_{\{x>0\}}\mathrm{d}x + C\frac{\mathrm{e}^{Gx}}{|x|^{1+Y}} 1_{\{x<0\}}\mathrm{d}x,$$

where $C > 0$, $G > 0$, $M > 0$, and $Y < 2$. The CGMY process is a pure jump Lévy process with canonical decomposition

$$L_t = t\mathrm{E}[L_1] + \int_0^t \int_{\mathbb{R}} x(\mu^L - \nu^{CGMY})((\mathrm{d}s, \mathrm{d}x)),$$

and Lévy triplet $(\mathrm{E}[L_1], 0, \nu^{CGMY})$, while the density is not known in closed form.

The CGMY processes are closely related to stable processes; in fact, the Lévy measure of the CGMY process coincides with the Lévy measure of the stable process with index $\alpha \in (0,2)$ (cf. [ST94, Def. 1.1.6]), *but* with the additional exponential factors; hence the name *tempered* stable processes. Due to the exponential tempering of the Lévy measure, the CGMY distribution has finite moments of all orders. Again, the class of CGMY distributions contains several other distributions as subclasses, for example the variance gamma distribution (Madan and Seneta [MS90]) and the bilateral gamma distribution (Küchler and Tappe [KT08]).

10.7. **Meixner.** The Meixner process was introduced by Schoutens and Teugels [ST98], see also [Sch02]. Let $L = (L_t)_{0\le t\le T}$ be a Meixner process with $\mathrm{Law}(H_1|P) = \mathrm{Meixner}(\alpha, \beta, \delta)$, $\alpha > 0$, $-\pi < \beta < \pi$, $\delta > 0$, then the density is

$$f_{\mathrm{Meixner}}(x) = \frac{\left(2\cos\frac{\beta}{2}\right)^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left(\frac{\beta x}{\alpha}\right)\left|\Gamma\left(\delta + \frac{ix}{\alpha}\right)\right|^2.$$

The characteristic function L_t, $t \in [0,T]$ is

$$\varphi_{L_t}(u) = \left(\frac{\cos\frac{\beta}{2}}{\cosh\frac{\alpha u - i\beta}{2}}\right)^{2\delta t},$$

and the Lévy measure of the Meixner process admits the representation

$$\nu^{\mathrm{Meixner}}(\mathrm{d}x) = \frac{\delta\exp\left(\frac{\beta}{\alpha}x\right)}{x\sinh(\frac{\pi x}{\alpha})}.$$

The Meixner process is a pure jump Lévy process with canonical decomposition

$$L_t = t\mathrm{E}[L_1] + \int_0^t\int_{\mathbb{R}} x(\mu^L - \nu^{\mathrm{Meixner}})((\mathrm{d}s, \mathrm{d}x)),$$

and Lévy triplet $(\mathrm{E}[L_1], 0, \nu^{\mathrm{Meixner}})$.

10.8. **Spectrally negative.** A spectrally negative Lévy process has no positive jumps and is not the negative of a subordinator. The lack of positive jumps implies the existence of

Proposition 10.1. *Let X be a spectrally negative Lévy process with Lévy triplet (b, c, ν). Then $\mathbb{E}[\mathrm{e}^{X_t}] < \infty$ for any $t \ge 0$ and*

$$\phi(\lambda) := \log\left(\mathbb{E}\left[\mathrm{e}^{X_1}\right]\right) = b\lambda + \frac{c}{2}\lambda^2 + \int_{(-\infty,0)} (\mathrm{e}^{\lambda x} - 1 - \lambda x 1_{\{x>-1\}})\nu(\mathrm{d}x).$$

The function ϕ is convex, infinitely often differentiable on $(0,\infty)$ with $\phi(0) = 0$ and $\phi(\infty) = \infty$.

Proof. Since X has no positive jumps, $\nu((0,\infty)) = 0$ and thus for any $\lambda \geq 0$

$$\int_{\{|x|>1\}} \mathrm{e}^{\lambda x}\nu(\mathrm{d}x) = \int_{\{x>-1\}} \mathrm{e}^{\lambda x}\nu(\mathrm{d}x) \leq \nu((-\infty,-1)) < \infty.$$

From Theorem 9.1 it then follows that $\mathbb{E}[\mathrm{e}^{\lambda X_t}] < \infty$ for all $t \geq 0$. This implies that the characteristic exponent $\psi(\lambda)$ of X can be extended to complex λ with negative imaginary part, hence the Laplace exponent of X exists and is given by

$$\phi(\lambda) := \log\left(\mathbb{E}\left[\mathrm{e}^{X_1}\right]\right) = \psi(-i\lambda) = b\lambda + \frac{c}{2}\lambda^2 + \int_{(-\infty,0)} (\mathrm{e}^{\lambda x} - 1 - \lambda x 1_{\{x>-1\}})\nu(\mathrm{d}x).$$

Since for any $n \in \mathbb{N}$ the function $x \to x^n e^{\lambda x}$ is smaller than 1 on $(-\infty, 0)$ for some $C < 0$ if follows that ϕ is infinitely often differentiable on $(0,\infty)$ and that

$$\phi''(\lambda) = c + \lambda^2 \int_{(-\infty,0)} e^{\lambda x}\nu(\mathrm{d}x) > 0.$$

From the assumption that X does not have monotone paths it follows that $\mathbb{P}(X_1 > 0) > 0$ which implies that $\mathbb{E}\left[\mathrm{e}^{\lambda X_1}\right] \to \infty$ as $\lambda \to \infty$. Finally, $\phi(0) = 0$ is follows by definition. $\square$

The lack of positive jumps allows us to deduce the following interesting characterization of the first passage times over $x > 0$

$$\tau_x^+ = \inf\{t \geq 0 : X_t > x\}$$

which is a stopping time due to our regularity assumptions on the filtration. Note that

$$\{\tau_x^+ \leq t\} = \{\sup_{0 \leq s \leq t} X_s > x\}.$$

Theorem 10.2. *Let X be a spectrally negative Lévy process with Laplace exponent ϕ and denote by Φ the right inverse of ϕ, i.e. for $q \geq 0$*

$$\Phi(q) = \sup\{\lambda : \phi(\lambda) = q\}.$$

Then τ_x^+ is a (possibly killed) subordinator with

$$\mathbb{E}[\mathrm{e}^{-q\tau_x^+}] = e^{-\Phi(q)x}$$

for $q > 0$. In particular, $\mathbb{P}(\tau_x^+ < \infty) = e^{-\Phi(0)x}$.

Proof. Due to the càdlàg paths of X the process $x \to \tau_x^+$ also has càdlàg paths. A spectrally negative Lévy process is strictly positive immediately (this is obvious when X has paths of infinite variation; in the case of finite variation X must also be strictly positive immediately since otherwise X would have decreasing paths which we ruled out by definition, see also Proposition **??**). This implies that $\tau_0^+ = 0$. Next, consider an increment $\tau_{x+y}^+ - \tau_x^+$ for some $y, x > 0$. On the event $\{\tau_x^+ < \infty\}$ we apply Theorem **??** with $\tau = \tau_x^+$ and deduce that $\tilde{X}_t = X_{\tau_x^+ + t} - X_{\tau_x^+}$ is again a Lévy process with the same law as X on $\{\tau_{x^+} < \infty\}$. As X has no negative jumps, it follows that $X_{\tau_x^+} = x$ on $\{\tau_{x^+} < \infty\}$. We find that

$$\tau_{x+y}^+ = \inf\{t > 0 : X_t > x + y\} = \inf\{t > 0 : \tilde{X}_t > y\} + \tau_x^+$$

from which the independence and stationarity of increments follows. The fact that the paths of τ_x^+ are non-increasing is obvious. The event $\{\tau_x^+ = \infty\}$ occurs with positive probability whenever X drifts to $-\infty$, in which case τ_x^+ is a subordinator killed at rate $\mathbb{P}(\tau_x^+ = \infty)$ for which we find an expression below.

To calculate the Laplace exponent of τ_x^+ we use the martingale from part (3) of Proposition **??**, i.e. $\mathrm{e}^{\lambda X_t - \phi(\lambda)t}$ for some $\lambda \geq 0$. It immediately follows that

$$\mathbb{E}[\mathrm{e}^{\lambda X_t - \phi(\lambda)t}] = 1$$

and hence, for any stopping time τ

$$\mathbb{E}[\mathrm{e}^{\lambda X_{t\wedge\tau} - \phi(\lambda)(t\wedge\tau)}] = 1.$$

For $\tau = \tau_x^+$ it follows that $X_{t\wedge\tau_x^+} \leq x$ and $X_{\tau_x^+ = x}$ on $\{\tau_x^+ < \infty\}$. An application of dominated convergence theorem yields for $\lambda > 0$

$$1 = \lim_{t\to\infty} \mathbb{E}[\mathrm{e}^{\lambda X_{t\wedge\tau_x^+} - \phi(\lambda)(t\wedge\tau_x^+)}] = \mathrm{e}^{\lambda x}\mathbb{E}[\mathrm{e}^{-\phi(\lambda)\tau_x^+} 1_{\{\tau_x^+ < \infty\}}].$$

The properties of ϕ derived in Proposition 10.1 imply that for any $q > 0$ the equation $\phi(\lambda) = q$ has a unique on $(0, \infty)$ which we denote by $\Phi(q)$. The equation $\phi(\lambda) = 0$ with $\lambda \in [0, \infty)$ has one solution $\lambda = 0$ and a second one in case $\phi'(0+) < 0$. Denote by $\Phi(0)$ the largest solution of this equation. Then we find that for any $q \geq 0$

$$\mathbb{E}[\mathrm{e}^{-q\tau_x^+} 1_{\{\tau_x^+ < \infty\}}] = \mathrm{e}^{-\Phi(q)x}$$

and thus $\mathbb{P}(\tau_x^+ < \infty) = e^{-\Phi(0)x}$ showing that τ_x^+ is a subordinator killed at rate $\Phi(0)$. □

11. Simulation of Lévy processes

We shall briefly describe simulation methods for Lévy processes. Our attention is focused on finite activity Lévy processes (i.e. Lévy jump-diffusions) and some special cases of infinite activity Lévy processes, namely the normal inverse Gaussian and the variance gamma processes.

Here, we do not discuss simulation methods for random variables with known density; various algorithms can be found at
`http://cg.scs.carleton.ca/ luc/rnbookindex.html`.

11.1. **Finite activity.**

Assume we want to simulate the Lévy jump-diffusion

$$L_t = bt + \sigma W_t + \sum_{k=1}^{N_t} J_k$$

where $N_t \sim \text{Poisson}(\lambda t)$ and $J \sim F(\mathrm{d}x)$. W denotes a standard Brownian motion, i.e. $W_t \sim \text{Normal}(0, t)$.

We can simulate a discretized trajectory of the Lévy jump-diffusion L at fixed time points $t_1, \ldots, t_n$ as follows:

- generate a standard normal variate and transform it into a normal variate, denoted G_i, with variance $\sigma\Delta t_i$, where $\Delta t_i = t_i - t_{i-1}$;
- generate a Poisson random variate N with parameter λT;
- generate N random variates τ_k uniformly distributed in $[0, T]$; these variates correspond to the jump times;

- simulate the law of jump size J, i.e. simulate random variates J_k with law $F(\mathrm{d}x)$.

The discretized trajectory is

$$L_{t_i} = bt_i + \sum_{j=1}^{i} G_j + \sum_{k=1}^{N} 1_{\{\tau_k < t_i\}} J_k.$$

Alternatively, we can also generate a sequence of exponential variates which are the interarrival times of the compound Poisson process together with the variates from the jump distribution.

11.2. **Infinite activity.** The variance gamma and the normal inverse Gaussian process can be easily simulated because they are time-changed Brownian motions.

Assume we want to simulate a normal inverse Gaussian (NIG) process with parameters $\alpha, \beta, \delta, \mu$; cf. also section 10.5. We can simulate a discretized trajectory at fixed time points $t_1, \dots, t_n$ as follows:

- simulate n independent inverse Gaussian variables I_i with parameters $(\delta \Delta t_i)^2$ and $\alpha^2 - \beta^2$, where $\Delta t_i = t_i - t_{i-1}$, $i = 1, \dots, n$;
- simulate n i.i.d. standard normal variables G_i;
- set $\Delta L_i = \mu \Delta t_i + \beta I_i + \sqrt{I_i} G_i$.

The discretized trajectory is

$$L_{t_i} = \sum_{k=1}^{i} \Delta L_k.$$

Assume we want to simulate a variance gamma (VG) process with parameters σ, θ, κ; we can simulate a discretized trajectory at fixed time points $t_1, \dots, t_n$ as follows:

- simulate n independent gamma variables Γ_i with parameter $\frac{\Delta t_i}{\kappa}$
- set $\Gamma_i = \kappa \Gamma_i$;
- simulate n standard normal variables G_i;
- set $\Delta L_i = \theta \Gamma_i + \sigma \sqrt{\Gamma_i} G_i$.

The discretized trajectory is

$$L_{t_i} = \sum_{k=1}^{i} \Delta L_k.$$

If the Lévy process has jumps of infinite variation, then (under a mild condition) we can approximate the jumps of absolute size smaller than ε by a Brownian motion with variance

$$t \int_{|x|<\varepsilon} x^2 \Pi(\mathrm{d}x).$$

12. Stochastic integration

In this section, we develop the theory of stochastic integration with respect to general semimartingales and Poisson random measures in several steps.

12.1. **Semimartingales and Doob–Meyer decompositions.** We start with some useful definitions and results from stochastic analysis. Let $\mathcal{B} = (\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ denote a stochastic basis. We denote by $\mathcal{M}$ the space of martingales and by $\mathcal{H}^2$ the space of square integrable martingales on $\mathcal{B}$, starting at zero, while $\mathcal{V}$ denotes the space of processes with finite variation. If $\mathcal{C}$ denotes a class of processes, then $\mathcal{C}_{\text{loc}}$ denotes the localized class.

Definition 12.1. A càdlàg, adapted stochastic process X is called a *semimartingale* if it admits a decomposition

$$X = X_0 + M + A, \tag{12.1}$$

where X_0 is $\mathcal{F}_0$-measurable and finite, $M \in \mathcal{M}_{\text{loc}}$ and $A \in \mathcal{V}$. A semimartingale X is called *special* if A is also predictable.

Example 12.2. Every Lévy process is a semimartingale. This follows directly from the Lévy–Itô decomposition (6.2) of a Lévy process.

Remark 12.3. The decomposition (12.1) of a semimartingale X is *not* unique. In order to construct a counterexample, consider a Poisson process N with intensity λ. The compensated Poisson process $\overline{N}$ is a martingale, see Exercise 2, with decomposition

$$\overline{N}_t = N_t - \lambda t. \tag{12.2}$$

This is clearly the decomposition of a semimartingale, however the RHS of (12.2) is both a process of finite variation and a martingale.

On the contrary, the decomposition (12.1) of a special semimartingale is *unique*, since the only predictable local martingale of finite variation is the constant process; cf. [JS03, Cor. I.3.16]. If we consider the Poisson process again—which is also a special semimartingale—, we see that the RHS of (12.2) is not predictable, hence the decomposition is unique.

Definition 12.4. A process $X = (X_t)_{t\geq 0}$ is *of class (D)* if the set of random variables $(X_\tau)_{\tau\in\mathcal{T}}$ is uniformly integrable, where $\mathcal{T}$ denotes the set of all finite-valued stopping times on $\mathcal{B}$.

Theorem 12.5 (Dood–Meyer decomposition)**.**
1. *Every submartingale $X = (X_t)_{t\geq 0}$ admits the unique decomposition* (12.1) *where M is a local martingale with $M_0 = 0$ a.s., and A is a predictable, increasing, locally integrable process, with $A_0 = 0$ a.s.*
2. *Every submartingale $X = (X_t)_{t\geq 0}$ of class (D) admits the unique decomposition* (12.1) *where M is a uniformly integrable martingale with $M_0 = 0$ a.s., and A is a predictable, increasing, integrable process, with $A_0 = 0$ a.s.*

Proof. [Pro04, Chapter 3]. □

Theorem 12.6 (Doob's martingale inequality)**.** *Let $X = (X_t)_{t\geq 0}$ be a nonnegative submartingale, then for any $p > 1$ and any $t \geq 0$*

$$\mathbb{E}[X_t^p] \leq \mathbb{E}\big[\sup_{0\leq s\leq t} X_s^p\big] \leq q^p \mathbb{E}\big[X_t^p\big], \tag{12.3}$$

where p and q are conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. *In particular, if* X *is a square-integrable martingale then*

$$\mathbb{E}\big[\sup_{0\le s\le t} X_s^p\big] \le 4\mathbb{E}\big[X_t^2\big]. \tag{12.4}$$

Proof. [JP03] or [Wil91]. □

Theorem 12.7. *Let* M *be a square integrable martingale with* $M_0 = 0$ *a.s. There exists a unique, increasing, predictable and integrable process* $\langle M\rangle$ *such that* $M^2 - \langle M\rangle$ *is a uniformly integrable martingale.*

Definition 12.8. The process $\langle M\rangle$ is called the *predictable quadratic variation*, or *angle bracket*, of M.

Proof. Let $M \in \mathcal{H}^2$ then, by Jensen's inequality, we have that M^2 is a non-negative submartingale. Applying Doob's inequality and using the square integrability of M we get that $\sup_{0\le s\le t} M_s^2$ is integrable (for any t). Moreover, Doob's stopping theorem yields that the stopped process $M^{2,\tau} = (M^2_{t\wedge\tau})_{t\ge0}$ is a submartingale. Therefore, $M^{2,\tau}$ is a process of class (D) and the Dood–Meyer decomposition yields that there exists a unique, increasing, predictable and integrable process, denoted by $\langle M\rangle$, such that

$$M^2 - \langle M\rangle$$

is a uniformly integrable martingale. □

Remark 12.9. The processes M^2 and $\langle M\rangle$ obviously satisfy

$$\mathbb{E}[M_t^2] = \mathbb{E}[\langle M\rangle_t]. \tag{12.5}$$

Note that a "converse" to this theorem is also true.

Lemma 12.10. *Let* M *be a square-integrable martingale, and assume there exists an increasing, predictable and integrable process* A *such that*

$$\mathbb{E}[(M_t - M_s)^2|\mathcal{F}_s] = \mathbb{E}\left[A_t - A_s|\mathcal{F}_s\right]. \tag{12.6}$$

Then A *is the predictable quadratic variation of* M.

Proof. Using the martingale property of M, we can rearrange (12.6) as follows:

$$\begin{aligned}(12.6) &\iff \mathbb{E}[M_t^2 - M_s^2|\mathcal{F}_s] = \mathbb{E}[A_t - A_s|\mathcal{F}_s]\\ &\iff \mathbb{E}[M_t^2 - A_t|\mathcal{F}_s] = \mathbb{E}[M_s^2 - A_s|\mathcal{F}_s] = M_s^2 - A_s,\end{aligned}$$

hence $M^2 - A$ is a martingale. Now, the uniqueness of the Doob–Meyer decomposition yields that A is the predictable quadratic variation of M, i.e. $A \equiv \langle M\rangle$. □

Example 12.11. Let $X = (X_t)_{t\ge0}$ be a square integrable Lévy process. The process $M = (M_t)_{t\ge0}$, where

$$\begin{aligned}M_t &= X_t - \mathbb{E}[X_t]\\ &= \sqrt{c}W_t + \int_0^t\int_{\mathbb{R}} x(\mu^X - \nu^X)(\mathrm{d}s, \mathrm{d}x),\end{aligned} \tag{12.7}$$

is a martingale, see Proposition 9.12, and it follows directly it is also square-integrable. We have that

$$\mathbb{E}[M_t] = 0 \quad \text{and} \quad \mathrm{Var}[M_t] = \Big(c + \int_{\mathbb{R}} x^2 \nu(\mathrm{d}x)\Big)t. \tag{12.8}$$

Now, M^2 is a submartingale of class (D) and has the Doob–Meyer decomposition

$$M^2 - \langle M \rangle \in \mathcal{M}.$$

Since $M^2 - \langle M \rangle$ and M are zero-mean martingales we get

$$\mathbb{E}[M_t^2 - \langle M \rangle_t] = 0 = \mathbb{E}[M_t^2] - \mathrm{Var}[M_t], \tag{12.9}$$

thus, using Lemma 12.10, we can conclude that

$$\langle M \rangle_t = \Big(c + \int_{\mathbb{R}} x^2 \nu(\mathrm{d}x)\Big)t. \tag{12.10}$$

12.2. **Spaces of integrands.** Here we define the class of simple, predictable processes and the class of square-integrable processes (wrt $\langle M \rangle$). These are the classes of integrands with respect to which we can reasonably define a stochastic integral, when the integrator is a square integrable martingale.

Definition 12.12. Let M be a square integrable martingale with predictable quadratic variation $\langle M \rangle$, and fix $T > 0$. We denote by $L^2(M)$ the space of square integrable processes with respect to $\langle M \rangle$, that is

$$L^2(M) = \Big\{ f : [0,T] \times \Omega \to \mathbb{R} \big| f \text{ predictable and } \mathbb{E}\Big[\int_0^T |f_s|^2 \mathrm{d}\langle M \rangle_s\Big] < \infty \Big\}.$$

Let $f, g \in L^2(M)$, then

$$\langle f, g \rangle_{L^2(M)} = \mathbb{E}\Big[\int_0^T f_s g_s \mathrm{d}\langle M \rangle_s\Big] \tag{12.11}$$

defines an inner product on $L^2(M)$, while $L^2(M)$ endowed with this inner product becomes a real Hilbert space. The norm on $L^2(M)$ is naturally induced by

$$\|f\|_{L^2(M)} = \langle f, f \rangle_{L^2(M)}. \tag{12.12}$$

Definition 12.13. The *stochastic set* $\rrbracket r, s \rrbracket$ is defined as

$$\rrbracket r, s \rrbracket := \{(\omega, t) : \omega \in \Omega, r < t \le s\}. \tag{12.13}$$

Definition 12.14. Let $\Pi = \{0 = s_0 < \cdots < s_n = T\}$ denote a partition of $[0,T]$. The space of *simple* predictable processes is defined as

$$\mathcal{S} = \Big\{ f : [0,T] \times \Omega \to \mathbb{R} \big| f(s) = \sum_{i=1}^n f_i 1_{\rrbracket s_i, s_{i+1} \rrbracket}, f_i \text{ is } \mathcal{F}_{s_i}\text{-adapted \& bounded} \Big\}.$$

Lemma 12.15. *The space $\mathcal{S}$ is dense in $L^2(M)$.*

Proof. This follows analogously to Theorem 8.8. in [Sch05]. □

Exercise 20. Prove Lemma 12.15.

12.3. **Stochastic integration wrt an L^2-martingale, simple processes.** We are now ready to define the stochastic integral of a simple, predictable process f with respect to a square integrable martingale M.

Definition 12.16. Let $f \in \mathcal{S}$ and $M \in \mathcal{H}^2$. The *stochastic integral* of f with respect to M is defined by

$$I_t = \sum_{i=0}^{n-1} f_i \big(M_{s_{i+1}\wedge t} - M_{s_i \wedge t} \big). \tag{12.14}$$

We will denote the stochastic integral by

$$I_t = I_t(f) = \int_0^t f_s \mathrm{d}M_s. \tag{12.15}$$

The idea for the construction of the stochastic integral is due to Kiyoshi Itô, hence it is also called the *Itô stochastic integral*.

Remark 12.17. Let us point out that for each time interval $(s_i, s_{i+1}]$ the integrand f_i is adapted to $\mathcal{F}_{s_i}$, i.e. to the "past", while the integrator $M_{s_{i+1}} - M_{s_i}$ "looks into the future". In mathematical finance, f_i has the natural interpretation as a *trading strategy* for the time interval $[s_i, s_{i+1}]$, which should be based on information up to time s_i. Thus it should be $\mathcal{F}_{s_i}$-adapted.

Lemma 12.18. *The stochastic integral is linear, i.e. if $\alpha, \beta \in \mathbb{R}$ and $f, g \in \mathcal{S}$, then*

$$I_t(\alpha f + \beta g) = \alpha I_t(f) + \beta I_t(g). \tag{12.16}$$

Proof. Immediate from (12.14). □

As a first result, we show that the stochastic integral of a simple process with respect to a square integrable martingale is a square integrable martingale itself, and prove an isometry between $\mathcal{S}$ and $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Lemma 12.19. *Let $f \in \mathcal{S}$ and $M \in \mathcal{H}^2$. Then, the stochastic integral $I = \int_0^\cdot f_s \mathrm{d}M_s$ defined by (12.14) is a square integrable martingale, the predictable quadratic variation of I is*

$$\langle I \rangle = \int_0^\cdot f_s \mathrm{d}\langle M \rangle_s, \tag{12.17}$$

and

$$\mathbb{E}[|I_t|^2] = \mathbb{E}\Big[\int_0^t |f_s|^2 \mathrm{d}\langle M \rangle_s \Big]. \tag{12.18}$$

Proof. Step 1. For the martingale property, we have to show that

$$\mathbb{E}[I_t | \mathcal{F}_s] = I_s. \tag{12.19}$$

Let $0 \le s < t \le T$ and Π be a partition of $[0, T]$. Assume that s and t belong to different subintervals of Π, that is, there exist points s_l and s_k

with $s_{l+1} < s_k$ such that $s \in (s_l, s_{l+1}]$ and $t \in (s_k, s_{k+1}]$. Then, we can decompose I_t as follows:

$$\begin{aligned} I_t &= \sum_{i=0}^{l-1} f_i\big(M_{s_{i+1}} - M_{s_i}\big) + f_l\big(M_{s_{l+1}} - M_{s_l}\big) \\ &\qquad + \sum_{i=l+1}^{k-1} f_i\big(M_{s_{i+1}} - M_{s_i}\big) + f_k\big(M_t - M_{s_k}\big) \\ &= I_1 + I_2 + I_3 + I_4. \end{aligned} \tag{12.20}$$

Every random variable in I_1 is $\mathcal{F}_s$-measurable since $s_l \leq s$, hence

$$\mathbb{E}[I_1|\mathcal{F}_s] = I_1. \tag{12.21}$$

Regarding I_2, since f_l and M_{s_l} are $\mathcal{F}_s$-measurable and using the martingale property of M, we get that

$$\mathbb{E}[f_l\big(M_{s_{l+1}} - M_{s_l}\big)|\mathcal{F}_s] = f_l\mathbb{E}[M_{s_{l+1}}|\mathcal{F}_s] - M_{s_l} = f_l\big(M_s - M_{s_l}\big). \tag{12.22}$$

Looking again at (12.21)–(12.22), we have that

$$\mathbb{E}[I_1 + I_2|\mathcal{F}_s] = I_1 + f_l\big(M_s - M_{s_l}\big) = I_s. \tag{12.23}$$

Hence, for the martingale property to hold we have to show that

$$\mathbb{E}[I_3 + I_4|\mathcal{F}_s] = 0.$$

Next, regarding I_3 and I_4, since $s_{l+1} \geq s$ we will use the following iterated conditioning trick for the proof: for $s_m \geq s_{l+1} \geq s$ we have

$$\begin{aligned} \mathbb{E}[f_m\big(M_{s_{m+1}} - M_{s_m}\big)|\mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[f_m\big(M_{s_{m+1}} - M_{s_m}\big)|\mathcal{F}_{s_m}]|\mathcal{F}_s] \\ &= \mathbb{E}[f_m\big(\underbrace{\mathbb{E}[M_{s_{m+1}})|\mathcal{F}_{s_m}]}_{=M_{s_m}} - M_{s_m}\big)|\mathcal{F}_s] = 0. \end{aligned}$$

Applying this to I_3 and I_4 yields that $\mathbb{E}[I_3+I_4|\mathcal{F}_s] = 0$, hence the martingale property of the stochastic integral $I = \int_0^\cdot f_s \mathrm{d}M_s$ has been proved.

Step 2. Let us now turn to the predictable quadratic variation of I. We will denote the increment of M by Δ, i.e. $\Delta_i := M_{s_{i+1}} - M_{s_i}$, therefore

$$I_t = \sum_{i=0}^{n-1} f_i\big(M_{s_{i+1}} - M_{s_i}\big) = \sum_{i=0}^{n-1} f_i\Delta_i, \tag{12.24}$$

where $\Delta_{n-1} = M_t - M_{s_{n-1}}$, and then

$$I_t^2 = \sum_{i=0}^{n-1} f_i^2\Delta_i^2 + 2 \sum_{0\leq i<j\leq n-1} f_i f_j \Delta_i \Delta_j. \tag{12.25}$$

Let us observe that, for $s = s_0 < \cdots < s_n = t$, we have

$$\begin{aligned}
\mathbb{E}[(I_t - I_s)^2|\mathcal{F}_s] &= \mathbb{E}\Big[\Big(\int_s^t f_u \mathrm{d}M_u\Big)^2\big|\mathcal{F}_s\Big] = \mathbb{E}\Big[\Big(\sum_{i=0}^{n-1} f_i\Delta_i\Big)^2\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[\sum_{i=0}^{n-1} f_i^2\Delta_i^2\big|\mathcal{F}_s\Big] + 2\mathbb{E}\Big[\sum_{0\le i<j\le n-1} f_i f_j \Delta_i\Delta_j\big|\mathcal{F}_s\Big] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\Big[f_i^2\Delta_i^2\big|\mathcal{F}_s\Big] + 2\sum_{0\le i<j\le n-1} \mathbb{E}\Big[f_i f_j \Delta_i\Delta_j\big|\mathcal{F}_s\Big] \\
&= I_5 + I_6.
\end{aligned} \tag{12.26}$$

Regarding I_6, using the martingale property of M, we have that for $i < j$

$$\begin{aligned}
\mathbb{E}\Big[f_i f_j\Delta_i\Delta_j\big|\mathcal{F}_s\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[f_i f_j\Delta_i\Delta_j\big|\mathcal{F}_{s_j}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[f_i f_j\Delta_i\mathbb{E}\Big[M_{s_{j+1}} - M_{s_j}\big|\mathcal{F}_{s_j}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[f_i f_j\Delta_i\Big(\underbrace{\mathbb{E}\big[M_{s_{j+1}}\big|\mathcal{F}_{s_j}\big] - M_{s_j}}_{=0}\Big)\big|\mathcal{F}_s\Big] \\
&= 0,
\end{aligned} \tag{12.27}$$

because f_i, f_j and Δ_i are $\mathcal{F}_{s_j}$-measurable. A similar argument applies for $i > j$, hence $I_6 = 0$. Regarding I_5, we have that

$$\begin{aligned}
\mathbb{E}\Big[f_i^2\Delta_i^2\big|\mathcal{F}_s\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[f_i^2\Delta_i^2\big|\mathcal{F}_{s_i}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[f_i^2\mathbb{E}\Big[\big(M_{s_{i+1}} - M_{s_i}\big)^2\big|\mathcal{F}_{s_i}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[f_i^2\mathbb{E}\Big[\langle M\rangle_{s_{i+1}} - \langle M\rangle_{s_i}\big|\mathcal{F}_{s_i}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[f_i^2\big(\langle M\rangle_{s_{i+1}} - \langle M\rangle_{s_i}\big)\big|\mathcal{F}_{s_i}\Big]\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[f_i^2\big(\langle M\rangle_{s_{i+1}} - \langle M\rangle_{s_i}\big)\big|\mathcal{F}_s\Big],
\end{aligned}$$

therefore, we get

$$\begin{aligned}
I_5 &= \mathbb{E}\Big[\sum_{i=0}^{n-1} f_i^2\big(\langle M\rangle_{s_{i+1}} - \langle M\rangle_{s_i}\big)\big|\mathcal{F}_s\Big] \\
&= \mathbb{E}\Big[\int_s^t f_u^2 \mathrm{d}\langle M\rangle_u\big|\mathcal{F}_s\Big],
\end{aligned} \tag{12.28}$$

hence, we can conclude that

$$\mathbb{E}\Big[\Big(I_t - I_s\Big)^2\big|\mathcal{F}_s\Big] = \mathbb{E}\Big[\int_s^t f_u^2 \mathrm{d}\langle M\rangle_u\big|\mathcal{F}_s\Big]. \tag{12.29}$$

Setting $s = 0$ we arrive at

$$\mathbb{E}\big[I_t^2\big] = \mathbb{E}\Big[\int_0^t f_u^2 \mathrm{d}\langle M\rangle_u\Big], \tag{12.30}$$

which also shows the square-integrability of I, while Lemma 12.10 immediately yields that

$$\langle I\rangle_t = \int_0^t f_u^2 \mathrm{d}\langle M\rangle_u.$$

□

References

[App09] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge University Press, 2nd edition, 2009.

[AS68] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions.* Dover, 5th edition, 1968.

[Bil99] P. Billingsley. *Convergence of Probability Measures.* Wiley, 2nd edition, 1999.

[BN77] O. E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. A*, 353:401–419, 1977.

[BN97] O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.*, 24:1–13, 1997.

[BNPS01] O. E. Barndorff-Nielsen, J. Pedersen, and K. Sato. Multivariate subordination, self-decomposability and stability. *Adv. Appl. Probab.*, 33:160–187, 2001.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Econ.*, 81:637–654, 1973.

[CGMY02] P. Carr, H. Geman, D. B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *J. Business*, 75:305–332, 2002.

[CT04] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman and Hall/CRC Press, 2004.

[Dud02] R. M. Dudley. *Real Analysis and Probability.* Cambridge University Press, 2002.

[EK95] Ernst Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281–299, 1995.

[EP02] E. Eberlein and K. Prause. The generalized hyperbolic model: financial derivatives and risk measures. In H. Geman, D. Madan, S. Pliska, and T. Vorst, editors, *Mathematical Finance – Bachelier Congress 2000*, pages 245–267. Springer, 2002.

[EvH04] E. Eberlein and E. A. v. Hammerstein. Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In R.C. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications IV*, Progress in Probability 58, pages 221–264. Birkhäuser, 2004.

[JP03] J. Jacod and Ph. Protter. *Probability Essentials.* Springer, 2nd edition, 2003.

[JS03] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, 2nd edition, 2003.

[Kou02] Steven G. Kou. A jump diffusion model for option pricing. *Manag. Sci.*, 48:1086–1101, 2002.

[KT08] U. Küchler and S. Tappe. Bilateral gamma distributions and processes in financial mathematics. *Stochastic Process. Appl.*, 118:261–283, 2008.

[Kyp06] A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications.* Springer, 2006.

[Mer73] R. C. Merton. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.*, 4:141–183, 1973.

[Mer76] R. C. Merton. Option pricing with discontinuous returns. *Bell J. Financ. Econ.*, 3:145–166, 1976.

[MS90] D. B. Madan and E. Seneta. The variance gamma (VG) model for share market returns. *J. Business*, 63:511–524, 1990.

[Pro04] Ph. Protter. *Stochastic Integration and Differential Equations*. Springer, 3rd edition, 2004.

[PS06] G. Peskir and A. N. Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Birkhäuser, 2006.

[Rai00] S. Raible. *Lévy processes in finance: Theory, numerics, and empirical facts*. PhD thesis, Univ. Freiburg, 2000.

[RY99] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 3rd edition, 1999.

[Sam65] P. A. Samuelson. Rational theory of warrant pricing. *Indust. Manag. Rev.*, 6:13–31, 1965.

[Sat99] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[Sch02] W. Schoutens. The Meixner process: Theory and applications in finance. In O. E. Barndorff-Nielsen, editor, *Mini-proceedings of the 2nd MaPhySto Conference on Lévy Processes*, pages 237–241, 2002.

[Sch05] R. Schilling. *Measures, Integrals and Martingales*. Cambridge University Press, 2005.

[Shi96] A. N. Shiryaev. *Probability*. Springer, 2nd edition, 1996.

[ST94] Gennady Samorodnitsky and Murad Taqqu. *Stable non-Gaussian Random Processes*. Chapman and Hall, 1994.

[ST98] W. Schoutens and J. L. Teugels. Lévy processes, polynomials and martingales. *Comm. Statist. Stochastic Models*, 14:335–349, 1998.

[Wil91] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.