This course is an introduction to the theory of Lévy processes. It covers definitions, the Lévy–Itô decomposition, and some early results in the direction of fluctuation theory.

1 Lévy processes

Let us begin by recalling the definition of two familiar processes, a Brownian motion and a Poisson process.

A real-valued process $B = \{B_t : t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a Brownian motion if the following hold:

- The paths of $B$ are $\mathbb{P}$-almost surely continuous.
- $\mathbb{P}(B_0 = 0) = 1$.
- For $0 \leq s \leq t$, $B_t - B_s$ is equal in distribution to $B_{t-s}$.
- For $0 \leq s \leq t$, $B_t - B_s$ is independent of $\{B_u : u \leq s\}$.
- For each $t > 0$, $B_t$ is equal in distribution to a normal random variable with variance $t$.

A process valued on the non-negative integers $N = \{N_t : t \geq 0\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is said to be a Poisson process with intensity $\lambda > 0$ if the following hold:

- The paths of $N$ are $\mathbb{P}$-almost surely right continuous with left limits.
- $\mathbb{P}(N_0 = 0) = 1$.
- For $0 \leq s \leq t$, $N_t - N_s$ is equal in distribution to $N_{t-s}$.
- For $0 \leq s \leq t$, $N_t - N_s$ is independent of $\{N_u : u \leq s\}$.
• For each $t > 0$, $N_t$ is equal in distribution to a Poisson random variable with parameter $\lambda t$.

There are clearly a lot of similarities between these two definitions, and these lead us to the definition of a Lévy process.

**Definition 1.1.** A process $X = \{X_t : t \geq 0\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a *Lévy process* if it possesses the following properties:

- [Càdlàg paths] The paths of $X$ are $\mathbb{P}$-almost surely right continuous with left limits.
- [Initial condition] $\mathbb{P}(X_0 = 0) = 1$.
- [Stationary increments] For $0 \leq s \leq t$, $X_t - X_s$ is equal in distribution to $X_{t-s}$.
- [Independent increments] For $0 \leq s \leq t$, $X_t - X_s$ is independent of $\{X_u : u \leq s\}$.

Certain facts follow immediately from this definition. For instance, the property of independent increments implies that any Lévy process is a spatially-homogeneous Markov process; indeed, Lévy processes even possess the strong Markov property, but we will postpone discussion of this.

The class of Lévy processes is rather rich, and the reader may already know plenty of examples which fall into it. To better appreciate this, we first need to discuss infinitely divisible distributions, which are in bijection with Lévy processes.

**Definition 1.2.** We say that a real-valued random variable $U$ possesses an *infinitely divisible distribution* if for each $n = 1, 2, \ldots$ there exists a collection of i.i.d. random variables $U_{1,n}, \ldots, U_{n,n}$ such that

$$U \overset{d}{=} U_{1,n} + \cdots + U_{n,n},$$

where $\overset{d}{=}$ represents equality in distribution.

The fundamental theorem about infinitely divisible distributions is formulated in terms of characteristic exponents, which we now define. Recall that, if $U$ is a random variable, then its characteristic function (Fourier transform) $h : \mathbb{R} \to \mathbb{C}$ is given by

$$h(\theta) = \mathbb{E}[e^{i\theta U}], \quad \theta \in \mathbb{R}.$$ 

We will begin with a technical lemma.

**Lemma 1.3** (and definition). *Let $U$ be an infinitely divisible random variable, and $h$ its characteristic function. Then:*

(i) The function $h$ is continuous and non-zero, and $h(0) = 1$.

(ii) There exists a unique continuous function $f : \mathbb{R} \to \mathbb{C}$, such that $e^{f(\theta)} = h(\theta)$ for all $\theta \in \mathbb{R}$ and $f(0) = 0$. We will denote the function $f$ by $\log h$, and refer to it as the characteristic exponent of $U$. 

2
We make some remarks on the above lemma. Firstly, the converse of (i) is not true, as is shown by the example of the Binomial distribution. Secondly, the function \( \log h \) in (ii) is not formed by composing \( h \) with some branch of the logarithm function (i.e., by cutting the complex plane), and one should be careful not to assume that \( h(\theta_1) = h(\theta_2) \) implies that \( \log h(\theta_1) = \log h(\theta_2) \).

We do not offer a proof of Lemma 1.3, but one may be found in \([8, Lemma 7.6]\). It is perhaps not too surprising—one may visualise the continuous path drawn out by \( h \) as lifting to a continuous path on the Riemann surface of the log function.

It’s worth remarking on the following fact, whose proof is immediate. Suppose that \( U_1 \) and \( U_2 \) are two independent infinitely divisible random variables with characteristic functions \( h_1 \) and \( h_2 \) and characteristic exponents \( f_1 \) and \( f_2 \). Then, the random variable \( U = U_1 + U_2 \) has an infinitely divisible distribution, and if we denote by \( h \) and \( f \) the characteristic function and exponent of \( U \), then \( h = h_1 h_2 \) and \( f = f_1 + f_2 \).

The relationship between infinitely divisible distributions and Lévy processes is given by the following lemma, of which we will only prove one direction. The remaining direction is given in \([8, Theorem 7.10]\).

**Lemma 1.4.** If \( X \) is a Lévy process, then for any \( t \geq 0 \), the random variable \( X_t \) possesses an infinitely divisible distribution. Conversely, if \( U \) is an infinitely divisible random variable, then there exists a (unique in distribution) Lévy process \( X \) such that \( X_1 \overset{d}{=} U \).

**Proof (of easy direction).** Let \( X \) be a Lévy process and \( t \geq 0 \). Then, for any \( n = 1, 2, \ldots, \)

\[
X_t = X_{t/n} + (X_{2t/n} - X_{t/n}) + \cdots + (X_t - X_{(n-1)t/n}),
\]

and the summands on the right-hand side are i.i.d. by the properties of stationary independent increments. Thus, \( X_t \) possesses an infinitely divisible distribution.

We now examine the characteristic exponents of the one-dimensional distributions of Lévy processes. Take a Lévy process \( X \), and write \( \Psi_t \) for the characteristic exponent of \( X_t \), that is,

\[
e^{\Psi_t(\theta)} = \mathbb{E}[e^{i\theta X_t}], \quad \theta \in \mathbb{R}.
\]

Applying (1.1) we obtain that for \( m, n \in \mathbb{N} \),

\[
m\Psi_1(\theta) = \Psi_m(\theta) = n\Psi_{m/n}(\theta),
\]

which is to say that for rational \( t > 0 \),

\[
\Psi_t(\theta) = t\Psi_1(\theta).
\]

We now wish to extend this to all \( t > 0 \). Recall that \( X \) is almost surely right-continuous; by applying bounded convergence in (1.2) the same holds for \( t \mapsto e^{\Psi_t(\theta)} \), for each fixed \( \theta \in \mathbb{R} \). Thus, \( e^{\Psi_t(\theta)} = e^{t\Psi_1(\theta)} \) holds for all real \( t > 0 \) and \( \theta \in \mathbb{R} \), that is,

\[
\mathbb{E}[e^{i\theta X_t}] = e^{t\Phi(\theta)}, \quad \theta \in \mathbb{R},
\]

where \( \Phi := \Psi_1 \) is the characteristic exponent of \( X_1 \). This leads us to:
**Definition 1.5.** Let $X$ be a Lévy process. We refer to $\Psi$, the characteristic exponent of $X_1$, as the characteristic exponent of the Lévy process $X$.

The main representation theorem for Lévy processes, which may of course be viewed as a theorem about infinitely divisible distributions, is the following.

**Theorem 1.6** (Lévy–Khintchine formula). Let $X$ be a Lévy process with characteristic exponent $\Psi$. Then, there exist (unique) $a \in \mathbb{R}$, $\sigma \geq 0$, and a measure $\Pi$, with no atom at zero, satisfying

$$\int_{\mathbb{R}} 1 \wedge x^2 \Pi(dx) < \infty,$$

such that

$$\Psi(\theta) = ia\theta - \frac{1}{2} \sigma^2 \theta^2 + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x 1_{[-1,1]}(x)) \Pi(dx).$$

(1.4)

Conversely, given any admissible choices $(a, \sigma, \Pi)$, there exists a Lévy process $X$ with characteristic exponent given by the above formula.

The first part of this theorem is rather technical, and we will not address it; see §8, Theorem 8.1]. The second part amounts to constructing a Lévy process out of its characteristics, and we will prove it shortly as part of the more detailed Lévy–Itô decomposition.

The names and approximate meanings of the characteristics in the Lévy–Khintchine formula are as follows. The term $a$ is known as the centre of $X$, and incorporates any deterministic drift term; $\sigma$ is the Gaussian coefficient, and represents the volatility of a Brownian component, if present; and the so-called Lévy measure $\Pi$ represents the size and intensity of the jumps of $X$.

## 2 Examples of Lévy processes

As part of our thesis on the variety of Lévy processes, we list a few examples. Several of these will be used later in the text; in particular, the characteristic exponents of compound Poisson processes and linear Brownian motion will be essential in §3, and stable processes will remain a prominent example throughout.

### 2.1 Poisson processes

For each $\lambda > 0$, denote by $\mu_\lambda$ the Poisson distribution, that is, a measure concentrated on $k = 0, 1, 2, \ldots$ such that $\mu_\lambda(k) = e^{-\lambda} \lambda^k / k!$. The characteristic function of this distribution satisfies

$$\sum_{k \geq 0} e^{ik\theta} \mu_\lambda(k) = e^{\lambda(e^{i\theta} - 1)} = \left[e^{\lambda(e^{i\theta} - 1)}\right]^n.$$

The right-hand side is the characteristic function of the sum of $n$ independent Poisson distributions, each with parameter $\lambda/n$, from which it follows that $\mu_\lambda$ is infinitely divisible. In the Lévy–Khintchine decomposition we see that $\sigma = 0$; $\Pi = \lambda \delta_1$, the Dirac measure supported on $\{1\}$; and $a = \lambda$. 
A Poisson process \( \{N_t : t \geq 0\} \) is a Lévy process whose distribution at time \( t > 0 \), is Poisson with parameter \( \lambda t \). From the above calculations we have

\[
\mathbb{E}(e^{i\theta N_t}) = e^{\lambda t(e^{i\theta} - 1)},
\]

and hence its characteristic exponent is given by \( \Psi(\theta) = \lambda(e^{i\theta} - 1) \) for \( \theta \in \mathbb{R} \).

### 2.2 Compound Poisson processes

Suppose now that \( N \) is a Poisson random variable with parameter \( \lambda > 0 \) and that \( \{\xi_i : i \geq 1\} \) is an sequence of i.i.d. random variables (independent of \( N \)), whose common
law \( F \) having no atom at zero. By first conditioning on \( N \), we have for \( \theta \in \mathbb{R} \),

\[
\mathbb{E}(e^{i\theta \sum_{i=1}^{N} \xi_i}) = \sum_{n \geq 0} \mathbb{E}(e^{i\theta \sum_{i=1}^{n} \xi_i})e^{-\lambda n} n!
= \sum_{n \geq 0} \left( \int_{\mathbb{R}} e^{i\theta x} F(dx) \right)^n e^{-\lambda n} n!
= e^{\lambda \int_{\mathbb{R}} (e^{i\theta x} - 1) F(dx)}.
\]

(2.1)

We see from (2.1) that distributions of the form \( \sum_{i=1}^{N} \xi_i \) are infinitely divisible with triple

\[ a = \lambda \int_{0<|x|<1} x F(dx), \quad \sigma = 0 \] and \( \Pi(dx) = \lambda F(dx) \). When \( F \) is simply a Dirac mass at the point 1, we have the Poisson process considered in the previous section.

Suppose now that \( N = \{N_t : t \geq 0\} \) is a Poisson process with intensity \( \lambda \), independent of the sequence \( \{\xi_i : i \geq 1\} \). We say that the process \( \{X_t : t \geq 0\} \) defined by

\[ X_t = \sum_{i=1}^{N_t} \xi_i, \quad t \geq 0, \]

is a **compound Poisson process** with **jump distribution** \( F \). If we write

\[ X_t = X_s + \sum_{i=N_s+1}^{N_t} \xi_i, \]

and recall that \( N \) has independent stationary increments, we see that \( X_t \) is the sum of \( X_s \) and an independent copy of \( X_{t-s} \). Right continuity and left limits of the process \( N \) also ensure right continuity and left limits of \( X \). Thus compound Poisson processes are Lévy processes. From the calculations in the previous paragraph, for each \( t \geq 0 \) we may substitute \( N_t \) for the variable \( N \) to discover that the Lévy–Khintchine formula for a compound Poisson process takes the form \( \Psi(\theta) = \lambda \int_{\mathbb{R}} (e^{i\theta x} - 1) F(dx) \). Note in particular that the Lévy measure of a compound Poisson process is always finite with total mass equal to the rate \( \lambda \) of the underlying process \( N \).

If a drift of rate \( c \in \mathbb{R} \) is added to a compound Poisson process so that now

\[ X_t = \sum_{i=1}^{N_t} \xi_i + ct, \quad t \geq 0 \]

then it is straightforward to see that the resulting process is again a Lévy process. The associated infinitely divisible distribution is nothing more than a shifted compound Poisson distribution with shift \( c \). The Lévy-Khintchine exponent is given by

\[ \Psi(\theta) = \lambda \int_{\mathbb{R}} (e^{i\theta x} - 1) F(dx) + ic\theta. \]

If further the shift is chosen to centre the compound Poisson distribution, that is \( c = \lambda \int_{\mathbb{R}} x F(dx) \), then

\[ \Psi(\theta) = \lambda \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x) F(dx) \]

(2.2)

and the resulting process has mean zero at all times. We call the corresponding Lévy process a **compensated compound Poisson process**.
2.3 Linear Brownian motion

![Figure 3: A sample path of a Brownian motion.](image1)

![Figure 4: A sample path of the independent sum of a Brownian motion and a compound Poisson process.](image2)

Consider the probability law

\[ \mu_{s,\gamma}(dx) := \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(x-\gamma)^2}{2s^2}} \, dx, \quad x \in \mathbb{R}, \]

where \( \gamma \in \mathbb{R} \) and \( s > 0 \). This is the Gaussian distribution with mean \( \gamma \) and variance \( s^2 \).

It is well known that

\[
\int_{\mathbb{R}} e^{i\theta x} \mu_{s,\gamma}(dx) = e^{-\frac{1}{2} s^2 \theta^2 + i\theta \gamma}
= \left[ e^{-\frac{1}{2} \left( \frac{x}{s} \right)^2 + i\theta \frac{x}{s^2}} \right]^n,
\]

7
from which it follows that $\mu_{a,\gamma}$ is an infinitely divisible distribution, this time with $a = -\gamma$, $\sigma = s$ and $\Pi = 0$.

The characteristic exponent $\Psi(\theta) = -s^2\theta^2/2 + i\theta\gamma$ is that of a linear Brownian motion, namely
\[ X_t := sB_t + \gamma t, \quad t \geq 0, \]
where $B = \{B_t : t \geq 0\}$ is a standard Brownian motion. Since (as we have already discussed) $B$ has stationary independent increments and continuous paths, it is simple to deduce that $X$ does as well, making it another example of a Lévy process.

### 2.4 Gamma processes

Let $\alpha, \beta > 0$ and define the probability measure
\[ \mu_{\alpha,\beta}(dx) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} \, dx \]
supported on $[0, \infty)$. This is called the Gamma($\alpha, \beta$) distribution; note that when $\beta = 1$ this is the exponential distribution.

We have
\[
\int_0^\infty e^{\theta x} \mu_{\alpha,\beta}(dx) = \frac{1}{(1 - i\theta/\alpha)^\beta} = \left[ \frac{1}{(1 - i\theta/\alpha)^{\beta/n}} \right]^n
\]
and infinite divisibility follows. In its Lévy–Khintchine representation we have $\sigma = 0$ and $\Pi(dx) = \beta x^{-1} e^{-\alpha x} \, dx$, supported on $[0, \infty)$, and $a = \int_0^1 x \Pi(dx)$. However, this is not immediately obvious. The key to showing this is the so-called Frullani integral, which we give in the following lemma.

**Lemma 2.1** (Frullani integral). For all $\alpha, \beta > 0$ and $z \in \mathbb{C}$ such that $\text{Re} \, z \leq 0$ we have
\[
(1 - z/\alpha)^{-\beta} = \exp \left\{ -\int_0^\infty (1 - e^{zx}) \beta x^{-1} e^{-\alpha x} \, dx \right\}.
\]

To see how this lemma helps note that the Lévy–Khintchine formula for a gamma distribution takes the form
\[ \Psi(\theta) = -\beta \int_0^\infty (1 - e^{i\theta x}) \frac{1}{x} e^{-\alpha x} \, dx = -\beta \log(1 - i\theta/\alpha) \]
for $\theta \in \mathbb{R}$. The choice of $a$ in the Lévy–Khintchine formula is the necessary quantity to cancel the term coming from $i\theta 1_{|x|<1}$ in the integral with respect to $\Pi$ in the general Lévy–Khintchine formula.

The corresponding Lévy process is called a gamma process. Suppose now that $X = \{X_t : t \geq 0\}$ is a gamma process. Stationary independent increments tell us that for all $0 \leq s < t < \infty$, $X_t = X_s + \widetilde{X}_{t-s}$ where $\widetilde{X}_{t-s}$ is an
independent copy of $X_{t-s}$. This increment is strictly positive with probability one, since it is gamma distributed; that is, $X_t > X_s$ almost surely. Hence a gamma process is an example of a Lévy process with almost surely non-decreasing paths. Another such Lévy process is a compound Poisson process where the jump distribution $F$ is concentrated on $(0, \infty)$. However, there are two key differences between these processes. Firstly, the Lévy measure of a gamma process has infinite total mass, unlike the Lévy measure of a compound Poisson process, whose mass is necessarily finite (and equal to the arrival rate of jumps). Secondly, while a compound Poisson process with positive jumps does have paths which are almost surely non-decreasing, it does not have paths that are almost surely strictly increasing.

Lévy processes whose paths are almost surely non-decreasing (or simply non-decreasing for short) are called subordinators, and we will discuss them again in §4.5.

2.5 Stable processes

The reader may already be aware of the so-called Brownian scaling property, which states that if $B$ is a Brownian motion and $c > 0$, then the scaled stochastic process $(cB_{tc^{-2}} : t \geq 0)$ is equal in distribution to $B$. It is interesting to ask which Lévy processes satisfy a similar property, and these are the subject of this section. A Lévy process $X$ is strictly stable if there exists $\alpha \in (0, 2]$ such that for any $c > 0$

$$(cX_{tc^{-\alpha}} : t \geq 0) \overset{d}{=} X.$$ (2.3)

These processes are quite important, so we will say a little more about them immediately.

Strictly stable processes are those Lévy processes whose one-dimensional distributions are strictly stable distributions. A random variable $Y$ is said to have a strictly stable distribution if, for any $n \geq 1$, it observes the distributional equality

$$Y_1 + \cdots + Y_n \overset{d}{=} a_n Y.$$ (2.4)

where $Y_1, \ldots, Y_n$ are independent copies of $Y$ and $a_n > 0$. It is clear that particular this definition implies that any strictly stable random variable is infinitely divisible. It turns out that necessarily $a_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$ (see [4, §VI.1]), and we refer to the parameter $\alpha$ as the index of $Y$. The case $\alpha = 2$ corresponds to zero mean Gaussian random variables and has essentially been dealt with in §2.3, so we exclude it in the remainder of the discussion.

A strictly stable random variable $Y$ with index $\alpha \in (0, 1) \cup (1, 2)$ has characteristic exponent of the form

$$\Psi (\theta) = -c|\theta|^{\alpha} \left(1 - i\beta \tan \frac{\pi \alpha}{2} \text{sgn} \theta\right),$$ (2.5)

where $\beta \in [-1, 1]$ and $c > 0$.

A strictly stable random variable with index $\alpha = 1$ has characteristic exponent of the form

$$\Psi (\theta) = -c|\theta| + i\theta \eta,$$ (2.6)
2 Examples of Lévy processes

Figure 5: Graphs of strictly stable processes for varying choices of $\alpha$. Plots (a) through (c) show symmetric processes, while (d) shows a process which only jumps upwards. Note the scale in plot (a).
where $\eta \in \mathbb{R}$ and $c > 0$.

(The parameters in (2.5) and (2.6) are unique, and every choice of them gives the characteristic exponent of a stable distribution.)

Since we have said that strictly stable random variables are infinitely divisible, let us make the connection with the Lévy–Khintchine formula. We define

$$
\Pi(dx) = \left\{ \begin{array}{ll}
c_1 x^{-1-\alpha} dx, & x \in (0, \infty), \\
c_2 |x|^{-1-\alpha} dx, & x \in (-\infty, 0),
\end{array} \right.
$$

(2.7)

where $c_1, c_2 \geq 0$ and $\beta = (c_1 - c_2)/(c_1 + c_2)$ if $\alpha \in (0, 1) \cup (1, 2)$ and $c_1 = c_2$ if $\alpha = 1$. The relationship between $c$ and $c_1, c_2$ is given by $c = -\Gamma(-\alpha) \cos(\pi\alpha/2)(c_1 + c_2)$, if $\alpha \in (0, 1) \cup (1, 2)$; and $c = \pi/2(c_1 + c_2)$, if $\alpha = 1$. (See [8, proof of Theorem 4.10].) The choice of $a \in \mathbb{R}$ also depends on $\alpha$.

Unlike the previous examples, the distributions that lie behind these characteristic exponents are heavy tailed, in that, if $Y$ is a strictly stable random variable with index $\alpha$, then

$$
E[|Y|^{\gamma}] = \gamma \int_0^\infty t^{\gamma-1} \mathbb{P}(|Y| > t) dt \left\{ \begin{array}{ll}
< \infty, & \gamma < \alpha, \\
= \infty, & \gamma \geq \alpha.
\end{array} \right.
$$

In particular, if $Y$ is not Gaussian, then it does not have finite variance; and if $\alpha \leq 1$, then $Y$ has no mean. The value of the parameter $\beta$ (in the case $\alpha \neq 1$) gives control over the asymmetry in the Lévy measure.

### 2.6 Other examples, and concluding remarks

There are many more known examples of infinitely divisible distributions (and hence Lévy processes). Of the many known proofs of infinitely divisibility for specific distributions, most of them are non-trivial, often requiring intimate knowledge of special functions. A brief list of such distributions might include generalised inverse Gaussian, variance gamma, truncated stable, tempered stable (CGMY), generalised hyperbolic, Meixner, Pareto, $F$-distributions, Gumbel, Weibull, log-normal, Student $t$-distributions, Lamperti stable, and $\beta$- and $\theta$-processes (see [6] §1.2.7 or [8] Remark 8.2] for an extensive list of references.)

Despite being able to identify a large number of infinitely divisible distributions and hence associated Lévy processes, it is not clear at this point what the paths of Lévy processes look like. The task of giving a mathematically precise account of this lies ahead in the next section. In the meantime let us make the following informal remarks concerning paths of Lévy processes.

Any linear combination of a finite number of independent Lévy processes is again a Lévy process. It turns out that one may consider any Lévy process as an independent sum of a Brownian motion with drift and a countable number of independent compound Poisson processes with different jump rates, jump distributions and drifts. The superposition occurs in such a way that the resulting path remains almost surely finite at all times and, for each $\varepsilon > 0$, the process experiences at most a countably infinite number
of jumps of magnitude \( \varepsilon \) or less with probability one and an almost surely finite number of jumps of magnitude greater than \( \varepsilon \), over all fixed finite time intervals. If in the latter description there is always an almost surely finite number of jumps over each fixed time interval then it is necessary and sufficient that one has the linear independent combination of a Brownian motion with drift and a compound Poisson process. Depending on the underlying structure of the jumps and the presence of a Brownian motion in the described linear combination, a Lévy process will either have paths of bounded variation on all finite time intervals or paths of unbounded variation on all finite time intervals.

We include some computer simulations to give a rough sense of how the paths of Lévy processes look. Figures 1 and 2 depict the paths of Poisson process and a compound Poisson process, respectively. Figures 3 and 4 show the paths of a Brownian motion and the independent sum of a Brownian motion and a compound Poisson process, respectively. Figure 5 depicts four sample paths of strictly stable processes with different parameter values. All stable processes experience an infinite number of jumps over any finite time horizon, but when \( \alpha < 1 \), the paths are almost surely of bounded variation, whereas when \( \alpha \geq 1 \), they are of unbounded variation. The reader is cautioned however that, ultimately, computer simulations can only depict a finite number of jumps in any given path. Figures 1–4 were very kindly produced by Antonis Papapantoleon.

3 Lévy–Itô decomposition

We will prove the converse part of Theorem 1.6, the Lévy–Khintchine formula, by means of constructing a Lévy process with a given Lévy triple \((a, \sigma, \Pi)\). To this end, let us first recall, from the examples we saw in the previous section, that the characteristic exponent

\[
i a\theta - \frac{1}{2} \sigma^2 \theta^2
\]

belongs to a linear Brownian motion with volatility \( \sigma \) and drift \( a \), while the characteristic exponent

\[
\lambda \int_{\mathbb{R}} (e^{i\theta x} - 1) \mu(dx)
\]

belongs to a compound Poisson process with jump distribution \( \mu \) and jump rate \( \lambda \).

Recall furthermore that the sum of two independent Lévy processes \( X \) and \( Y \) with characteristic exponents \( \Psi \) and \( \Phi \) is a Lévy process with characteristic exponent \( \Psi + \Phi \).

With this in mind, let \( \Psi \) be the function given in Theorem 1.6 and we give the following decomposition, which appears to make sense at least formally.

Let \( A_\varepsilon = [-1, 1] \setminus (-\varepsilon, \varepsilon) \) be a closed annulus around zero with inner radius \( \varepsilon \). We have:

\[
\Psi(\theta) = i a\theta - \frac{1}{2} \sigma^2 \theta^2
+ \Pi(\mathbb{R} \setminus [-1, 1]) \int_{\mathbb{R}\setminus[-1,1]} (e^{i\theta x} - 1) \frac{\Pi(dx)}{\Pi(\mathbb{R} \setminus [-1, 1])}
+ \lim_{\varepsilon \downarrow 0} \left[ \int_{A_\varepsilon} (e^{i\theta x} - 1) \Pi(dx) - i \theta \int_{A_\varepsilon} x \Pi(dx) \right].
\]  

(3.1)
We therefore claim that a Lévy process with triple \((a, \sigma, \Pi)\) may be obtained by summing three independent simpler Lévy processes: the first term represents a linear Brownian motion; the second term represents a compound Poisson process of ‘large’ jumps (note that this process need not have a mean); and the third term represents a limit of compound Poisson processes compensated via deterministic drift to have zero mean. It is far from obvious that the final limit gives a characteristic exponent of an infinitely divisible random variable, or that this has meaning in terms of convergence of stochastic processes; this will be the main content of the proof.

3.1 Poisson random measures

It is clear by now that compound Poisson processes will play an important role in the proof of the Lévy–Itô decomposition. We will therefore make a short digression into the theory of Poisson point processes in general state spaces.

To motivate our development, and provide a simple example, consider a compound Poisson process with drift, \(X\), which has jump distribution \(\mu\); that is,

\[
X_t = \delta t + \sum_{i=1}^{N_t} \xi_i, \quad t \geq 0,
\]

where \(N\) is a Poisson process with intensity \(\lambda\), and the \(\xi_i\) are i.i.d. random variables with common law \(\mu\). Let us denote the arrival times of the Poisson process by \(T_i\), for \(i \geq 1\), and write \(L = \{(T_i, \xi_i) : i \geq 1\}\) for the (random) set of ‘time-space jump points’.

A sample path of \(X\) is visualised in Figure 6, together with the set \(L\) of its jumps.

Define a measure \(N\) on the Borel sets of \([0, \infty) \times \mathbb{R}\) by

\[
N(A) = \#(L \cap A) = \sum_{i=1}^{\infty} 1_A(T_i, \xi_i).
\]  

(3.2)

\(N\) is a random measure, and the random point set \(L\) is its support. In fact, \(N\) will be our first example of a Poisson random measure; here is the definition.

**Definition 3.1.** Let \((S, \mathcal{S}, \eta)\) be a \(\sigma\)-finite measure space, and \((\Omega, \mathcal{F}, P)\) a probability space. A function \(N : \Omega \times S \to \{0, 1, 2, \ldots, \infty\}\) is called a discrete random measure on \(S\) if, for every \(A \in \mathcal{S}\), \(N(\cdot, A)\) is a random variable; and for every \(\omega \in \Omega\), \(N(\omega, \cdot)\) is a measure. For convenience, we will usually write a random measure \(N\) without the first argument.

A random measure \(N\) is called a Poisson random measure on \(S\) with intensity measure \((\text{or mean measure})\) \(\eta\), if the following are satisfied:

(i) if \(A_1, \ldots, A_n\) are pairwise disjoint elements of \(\mathcal{S}\), then \(N(A_1), \ldots, N(A_n)\) are independent;

(ii) for each \(A \in \mathcal{S}\), \(N(A)\) is distributed as a Poisson random variable with parameter \(\eta(A)\).
Figure 6: The path of the compound Poisson process with drift, $X$ (above), the set of points $L$ generated by its jumps (below), and an evaluation of the Poisson random measure $N$ (below, dashed shape).
With respect to (ii), note that we consider a Poisson distribution with parameter 0 (resp., ∞) to be equal to 0 (resp., ∞) with probability one.

Our first example is as follows.

**Lemma 3.2.** The measure $N$ defined in (3.2) is a Poisson random measure on $[0, \infty) \times \mathbb{R}$ with intensity measure $\lambda \text{Leb} \times \mu$.

We do not offer a proof of this result since it would be so similar to that of the forthcoming Theorem 3.3, which we prove in full.

**Theorem 3.3.** Let $(S, \mathcal{S}, \eta)$ be a σ-finite measure space. Then there exists a Poisson random measure $N$ on $S$ with intensity measure $\eta$.

**Proof.** Suppose first that $\eta$ is finite. Let $(\Omega, \mathcal{F}, P)$ be a probability space supporting independent random variables $\xi, i \geq 1$ such that $\xi$ has a Poisson distribution with parameter $\eta(S)$, and

$$P(\xi \in A) = \frac{\eta(A)}{\eta(S)}, \quad A \in \mathcal{S}, \quad i \geq 1.$$  

Then, let

$$N(A) = \sum_{i=1}^{N} \mathbb{I}_{A}(\xi), \quad A \in \mathcal{S}. \quad (3.3)$$

This $N$ is certainly a random measure on $(S, \mathcal{S})$. Take disjoint measurable sets $A_1, \ldots, A_k$; we must show that these are independent and each $N(A_i)$ is Poisson distributed.

Fixing a set $A \in \mathcal{S}$, under the conditional law $P(\cdot | \mathbb{N} = n)$, the random variable $N(A)$ is the number of ‘independently thrown’ points $\{\xi_i : i = 1, \ldots, n\}$ which happen to land inside $A$; that is, it is Binomial$(n, p)$-distributed, where $p = \eta(A)/\eta(S)$. The generalisation of this is to say that, again under this conditional law, the tuple $(N(A_1), \ldots, N(A_k))$ has multinomial distribution, in that

$$P(N(A_1) = n_1, \ldots, N(A_k) = n_k | \mathbb{N} = n) = \frac{n!}{n_0!n_1! \cdots n_k!} \prod_{i=0}^{k} \left( \frac{\eta(A_i)}{\eta(S)} \right)^{n_i},$$

where $n_0 = n - (n_1 + \cdots + n_k)$ and $A_0 = S \setminus \bigcup_{i=1}^{k} A_i$. We then sum over $n$, finding

$$P(N(A_1) = n_1, \ldots, N(A_k) = n_k)$$

$$= \sum_{n_0} e^{-\eta(S)} \left( \frac{\eta(S)}{n!} \right)^n \frac{n!}{n_0!n_1! \cdots n_k!} \prod_{i=0}^{k} \left( \frac{\eta(A_i)}{\eta(S)} \right)^{n_i}$$

$$= \prod_{i=1}^{k} e^{-\eta(A_i)} \frac{\eta(A_i)}{n_2!} \sum_{n_0} e^{-\eta(A_0)} \frac{\eta(0)}{n_0!} \left( \frac{n_0}{(n_1 + \cdots + n_k)} \right)! \prod_{i=1}^{k} \left( \frac{\eta(A_i)}{\eta(S)} \right)^{n_i}$$

This shows that the $N(A_i)$ are mutually independent and each is Poisson$(\eta(A_i))$ distributed.

15
Suppose instead that $\eta(S) = \infty$. Since we know that $\eta$ is $\sigma$-finite, we can break $S$ up into disjoint subsets $S_1, S_2, \ldots$ such that $S = \bigcup_{i=1}^{\infty} S_i$ and $\eta(S_i) < \infty$. We then set $\eta_i = \eta(\cdot \cap S_i)$, which is a finite measure on $S_i$ (with the $\sigma$-algebra induced from $S$). These measures thus induce Poisson random measures $N_i$ on $S_i$, which we place on one probability space so that they are independent, and we then define

$$N(A) = \sum_{i=1}^{\infty} N_i(A \cap S_i) \quad A \in S.$$  

We claim that $N$ is a random measure on $S$ with intensity $\eta$; this is not difficult to show. We must also show that it is a Poisson random measure. Take $A \in S$; then the random variables $N_i(A)$ are independent and Poisson($\eta(A \cap S_i)$) distributed. Then $N(A) = \sum_{i=1}^{\infty} N_i(A)$, so by the additive property of Poisson distributions, this random variable is Poisson distributed with rate $\sum_{i=1}^{\infty} \eta(A \cap S_i) = \eta(A)$. Now take disjoint sets $A_1, \ldots, A_k$. We need to show that $N(A_1), \ldots, N(A_k)$ are independent, but this follows because the random variables in $\{N_i(A_j) : i = 1, 2, \ldots ; j = 1, \ldots, k\}$ are independent.

We note that, if $S = [0, \infty) \times E$ for some measurable space $E$, and $N$ is a Poisson random measure with intensity measure $\text{Leb} \times \nu$, then the measure $N$ may be constructed as the counting measure of an at most countable random set of points; that is, there exist $(t_i, x_i) \in [0, \infty) \times E$, for $i \in I$, such that

$$N = \sum_{i \in I} \delta_{(t_i, x_i)}.$$  

The set $\text{supp} N := \{(t_i, x_i) : i \in I\}$ is called a Poisson point process on $E$ with characteristic measure $\nu$.

### 3.2 Functionals of Poisson random measures

In this section, we are interested in the law of integrals

$$\int_{[0, \infty) \times E} f(t, x) N(dt, dx),$$

for measurable functions $f : [0, \infty) \times E \to \mathbb{R}$. When this integral has meaning (e.g., $f \geq 0$ or the integral is absolutely convergent) then it is equal to the sum

$$\sum_{(t, x) \in \text{supp} N} f(t, x).$$

**Theorem 3.4.** Let $f : [0, \infty) \times E \to \mathbb{R}$ be a measurable function and $N$ a Poisson random measure on $[0, \infty) \times E$ with intensity measure $\text{Leb} \times \nu$. Then:

(i) The integral

$$X_f := \int_{[0, \infty) \times E} f(t, x) N(dt, dx)$$

is almost surely absolutely convergent if and only if

$$\int_{[0, \infty) \times E} 1 \wedge |f(t, x)| dt \nu(dx) < \infty.$$
(ii) If $X_f$ is a.s. absolutely convergent, then

$$E[e^{i\beta X_f}] = \exp\left\{ -\int_{[0,\infty) \times E} (1 - e^{i\beta f(t,x)}) \, dt \nu(dx) \right\}, \quad \beta \in \mathbb{R}. \quad (3.4)$$

(iii) If

$$\int_{[0,\infty) \times E} |f(t,x)| \, dt \nu(dx) < \infty,$$

then $E|X_f| < \infty$ and

$$E[X_f] = \int_{[0,\infty) \times E} f(t,x) \, dt \nu(dx). \quad (3.5)$$

(iv) If

$$\int_{[0,\infty) \times E} |f(t,x)| \, dt \nu(dx) < \infty \quad \text{and} \quad \int_{[0,\infty) \times E} (f(t,x))^2 \, dt \nu(dx) < \infty,$$

then

$$E[X_f^2] = \int_{[0,\infty) \times E} (f(t,x))^2 \, dt \nu(dx) + \left( \int_{[0,\infty) \times E} f(t,x) \, dt \nu(dx) \right)^2.$$

Here (3.4) is often called the exponential formula, and (3.5) the compensation formula. Both are occasionally called master formulas.

Using the theorem above, we have the following result, part of which is a converse to Lemma 3.2. (The remarks about filtrations may seem rather odd at this point—why not just make $(\mathcal{F}_t)_{t \geq 0}$ the natural filtration generated by $X$? But all will become clear in the next section, where we need a collection of martingales all adapted to the same filtration.)

**Lemma 3.5.** Suppose that $N$ is a Poisson random measure on $[0,\infty) \times \mathbb{R}$, with intensity measure $\text{Leb} \times \Pi$, where $\Pi$ has no atom at zero. Define $\hat{\mathcal{F}}_t = \sigma\left\{ N([0,s] \times C) : s \leq t, C \text{ Borel} \right\}$, and let $(\mathcal{F}_t)_{t \geq 0}$ be the natural enlargement of $(\hat{\mathcal{F}}_t)_{t \geq 0}$. Let $B \subset \mathbb{R}$ be a Borel subset of $\mathbb{R}$ such that $0 < \Pi(B) < \infty$.

(i) The process

$$X_t = \int_{[0,t] \times B} x \, N(du,dx), \quad t \geq 0,$$

is a compound Poisson process with rate $\Pi(B)$ and jump distribution $\Pi(\cdot \cap B)/\Pi(B)$. Indeed, for $s < t$, $X_t - X_s$ is even independent of $\mathcal{F}_s$.

---

1The natural enlargement of $(\hat{\mathcal{F}}_t)$ is given as follows. One first sets $\hat{\mathcal{F}}_t = \sigma\{ A \cup N : A \in \mathcal{F}_t, N \subset M \text{ for some event } M \text{ s.t. } P(M) = 0 \}$. This is called $P$-regularisation. One then defines $\mathcal{F}_t = \bigcap_{s > t} \hat{\mathcal{F}}_s$. The filtration $\mathcal{F}_t$ is now right-continuous and contains all $P$-nullsets.
(ii) If, in addition, $\int_B |x| \Pi(dx) < \infty$, then the compensated Poisson process

$$M_t = \int_{[0,t] \times B} x N(du, dx) - t \int_B x \Pi(dx), \quad t \geq 0,$$

is a zero-mean $P$-martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$.

(iii) If furthermore $\int_B x^2 \Pi(dx) < \infty$, then $M$ is a square-integrable martingale, which is to say that $E[M_t^2] < \infty$ for all $t \geq 0$; indeed,

$$E[M_t^2] = t \int_B x^2 \Pi(dx), \quad t \geq 0.$$

**Proof.** (i) The fact that $X_t$ is well-defined for all $t$ follows from Theorem 3.4 (since $\int_{[0,t] \times B} 1 \land |x| \Pi(dx) \leq t \Pi(B) < \infty$.) One may show that $X$ has càdlàg paths by applying the dominated convergence theorem to functions of the form $1_{[0,t] \times B}$, path-by-path on a set of full measure.

Now, we calculate the increment from $s$ to $t$ to be

$$X_t - X_s = \int_{(s,t] \times B} xN(du, dx), \quad 0 \leq s < t < \infty,$$

and this is independent of $\mathcal{F}_s$ by the independence properties of Poisson random measures, since $(s, t] \times B$ is disjoint from $[0, s] \times C$ for any Borel $C$. Finally, $N(\cdot \cap C)$ restricted to $[0, t - s] \times B$ has the same law as $N$ restricted to $(s, t] \times B$, because the same holds for their intensity measures; this implies that the increment has the same law as $X_{t-s}$.

Thus, $X$ is certainly a Lévy process. We need to show that it is a compound Poisson process. Applying the exponential formula [3.4] we obtain

$$E[e^{i \theta X_t}] = \exp \left\{ t \int_B (e^{i \theta x} - 1) \Pi(dx) \right\}, \quad t \geq 0, \; \theta \in \mathbb{R},$$

and from (2.1) this is precisely the characteristic function of a compound Poisson process with rate $\Pi(B)$ and jump distribution $\Pi(\cdot \cap B)/\Pi(B)$.

(ii) Under the conditions of this part of the lemma, it follows from Theorem 3.4(iii) that $X_t$ is integrable and has expectation

$$E[X_t] = t \int_B x \Pi(dx). \quad \text{(3.6)}$$

Therefore certainly $E[M_t] = 0$; we must also show it is a martingale. $M$ is definitely adapted to $(\mathcal{F}_t)_{t \geq 0}$, and $M_t$ is integrable for each $t$ because the same holds for $X_t$.

To show the martingale property, let $s < t$ and calculate, using the Markov property of $X$,

$$E[M_t - M_s \mid \mathcal{F}_s] = E[M_{t-s}]$$

$$= E \left[ \int_{[0,t-s] \times B} xN(du, dx) \right] - (t - s) \int_B x \Pi(dx)$$

$$= 0,$$

where the final equality follows from (3.6).
(iii) Apply Theorem 3.4(iii) and perform a short computation.

While we are here, let us remind ourselves of our goals. We have already seen, in (2.2),
that the characteristic exponent of the process $M$ appearing in the previous lemma is
given by:
\[ \int_B (e^{i\theta x} - 1 - i\theta x) \Pi(dx), \quad \theta \in \mathbb{R}. \]  

(3.7)

If $\Pi$ is a Lévy measure, i.e. $\int_{\mathbb{R}} 1 \wedge x^2 \Pi(dx) < \infty$, then one may choose for $B$ any of
the annuli $A_\varepsilon = [-1, 1] \setminus (-\varepsilon, \varepsilon)$ considered at the beginning of the section and obtain a
square-integrable martingale $M$ with characteristic exponent given by (3.7). These are
precisely the limands in (3.1)! However, the set $[-1, 1]$ itself is not an acceptable choice
for $B$, because $\int_{[-1,1]} |x| \Pi(dx)$ may be infinite (for example, this is the case for the stable
processes defined in §2.5.) The aim of this section is to find a way to extend $B$ to $[-1, 1]$
via a limiting procedure; and the final point in the above lemma gives a hint about how
we can do this.

3.3 Square-integrable martingales

The last lemma made reference to square-integrable martingales. The space of these
processes has a nice structure which we now discuss. Let us first give (or refresh) some
definitions.

We take as given a probability measure $P$ and a filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfying the
natural conditions. This means that $(\mathcal{F}_t)_{t \geq 0}$ is equal to its natural enlargement, which
was described in footnote 1 on page 17; we reiterate the definition here: the natural
enlargement of $(\mathcal{F}_t)$ is given as follows. One first sets $\hat{\mathcal{F}}_t = \sigma\{A \cup N : A \in \mathcal{F}_t, N \subset M
$ for some event $M$ s.t. $P(M) = 0\}$. This is called $P$-regularisation. One then defines
$\mathcal{F}_t = \bigcap_{s>t} \hat{\mathcal{F}}_s$. The filtration $\mathcal{F}_t$ is now right-continuous and contains all $P$-nullsets.
For a discussion of this definition, see [3, Definition 1.3.38].

**Definition 3.6.** A martingale $M$ is called a square-integrable martingale if $E[M_t^2] < \infty$
for every $t \geq 0$. We will write $\mathcal{M}^2$ for the collection of zero-mean, càdlàg, square-
integrable martingales. For each $t \geq 0$, we will also define $\mathcal{M}_t^2 = \{(M_s, 0 \leq s \leq t) : M \in \mathcal{M}^2\}$.

For each $t \geq 0$, we define a seminorm on $\mathcal{M}_t^2$ by
\[ \|M\|_t = \sqrt{E[M_t^2]}, \]

and we also define the symbol
\[ \|M\| = \sum_{n=1}^{\infty} 2^{-n}(\|M\|_n \wedge 1), \quad M \in \mathcal{M}^2. \]

Recall that a seminorm satisfies all of the conditions of a norm except for the requirement that $\|M\|_t = 0$ implies $M = 0$. In fact, $\|\cdot\|_t$ is about as good a seminorm as you
can get: if $M, N \in \mathcal{M}^2$ and $\|M - N\|_s = 0$, then $M$ and $N$ restricted to $[0, t]$ are indistinguishable as stochastic processes. (This means that $P(M_s = N_s$ for all $s \in [0, t]) = 1$.)

We should also remark that

$$\|M\|_s = \sqrt{E[M^2]} \leq \sqrt{E[M^2]} = \|M\|_t, \quad s < t,$$

since $M^2$ is a submartingale; in particular, if $(M^{(n)})_{n \geq 1}$ converges under $\|\cdot\|_t$, then it will also converge under $\|\cdot\|_s$ when $s < t$; and in particular, every sequence $(M^{(n)})_{n \geq 1}$ converges in $L^2(\Omega, \mathcal{F}_s, P)$.

We warn that $\|\cdot\|$ does not define even a seminorm on $\mathcal{M}^2$; however, $d(M, N) := \|M - N\|$ is a pseudometric, that is, a function satisfying all conditions of a metric except that $d(M, N) = 0$ implies $M = N$. Again, $d$ comes with the nice property that $d(M, N) = 0$ implies that $M$ and $N$ are indistinguishable as stochastic processes on $[0, \infty)$. The point of $\|\cdot\|$ is that $d(M^{(n)}, M) \to 0$ if and only if $\|M^{(n)} - M\|_t \to 0$ for every $t \geq 0$. (This follows from a standard result of metric space theory, as $d$ is essentially a (Hilbert cube-like) product metric on $\bigoplus_{n \geq 1} \mathcal{M}^2_{t,n}$.)

Let us remind the reader that a square-integrable martingale is not necessarily bounded in $L^2$ (i.e., $\sup_{t \geq 0} E[M^2_t] < \infty$). The theory of so-called $L^2$-martingales is rather neater than that of square-integrable martingales, but unfortunately the processes $M$ appearing in Lemma 3.5 do not fall into that class!

The following proposition foreshadows how useful the space $\mathcal{M}^2$ will be.

**Proposition 3.7.** The metric space $(\mathcal{M}^2, d)$ is complete.

**Proof.** Take a sequence $(X^{(n)})_{n \geq 1}$ which is Cauchy in $\mathcal{M}^2$. For each $t \geq 0$, the sequence $(X^{(n)}_t)_{n \geq 1}$ is Cauchy in $L^2(\Omega, \mathcal{F}_t, P)$, by much the same reasoning as in the second remark above. Denote the $L^2$ limit by $\tilde{X}_t$, and define a stochastic process $\tilde{X} = (\tilde{X}_t : t \geq 0)$. Certainly $\tilde{X}$ is adapted to $(\mathcal{F}_t)_{t \geq 0}$, and $E[\tilde{X}_t^2] < \infty$ for every $t$.

We must next show that $\tilde{X}$ is a martingale, and we do this directly using the definition of conditional expectation. (This argument is weirdly low-level, but I haven’t found a better way.) Fix $s < t$ and pick $A \in \mathcal{F}_s$. Firstly,

$$|E[1_A(X^{(n)}_s - \tilde{X}_s)| \leq \sqrt{E[1_A] \sqrt{E[(X^{(n)}_t - \tilde{X}_t)^2]} \to 0},$$

by the Cauchy–Schwarz inequality; and the same argument shows that $E[1_A(X^{(n)}_t - \tilde{X}_t)] \to 0$ also. Now we calculate:

$$E[1_A(\tilde{X}_t - \tilde{X}_s)] = E[1_A(\tilde{X}_t - X^{(n)}_t)] + E[1_A(X^{(n)}_t - \tilde{X}_t)] + E[1_A(\tilde{X}_s - X^{(n)}_s)].$$

The first and third terms on the right-hand side have limit zero, as we established; and the second term is identically zero by the martingale property of the process $X^{(n)}$. We have thus shown that $E[\tilde{X}_s | \mathcal{F}_s] = \tilde{X}_s$, i.e., $\tilde{X}$ is a martingale.

Since $\tilde{X}$ is a martingale, it has a càdlàg modification $X$ (see [3] Proposition 2.5.13.) (This means that $P(X_t = \tilde{X}_t) = 1$ for every $t$.) We now have that $X \in \mathcal{M}^2$, and we only have left to show that $X^{(n)}$ converges to $X$ in $\mathcal{M}^2$. But this is automatic, because the convergence holds in $\mathcal{M}^2_m$ for every $m \geq 1$. 

\[\]
We now have everything in place for the following theorem, which is the missing ingredient in the Lévy–Itô decomposition.

Recall that $A_\varepsilon = [-1, 1] \setminus (-\varepsilon, \varepsilon)$. We use the same filtration as in Lemma 3.5, i.e. define $\hat{\mathcal{F}}_t = \sigma\{N([0, s] \times C) : s \leq t, C \text{ Borel}\}$, and let $(\hat{\mathcal{F}}_t)_{t \geq 0}$ be the natural enlargement of $(\hat{\mathcal{F}}_t)_{t \geq 0}$.

**Theorem 3.8.** Let $\Pi$ be a Lévy measure and $N$ a Poisson random measure on $[0, \infty) \times \mathbb{R}$ with intensity measure $\text{Leb} \times \Pi$. For $\varepsilon \in (0, 1)$, define

$$M^\varepsilon_t = \int_{[0,t] \times A_\varepsilon} x N(ds, dx) - t \int_{A_\varepsilon} x \Pi(dx), \quad t \geq 0,$$

which is an element of $\mathcal{M}^2$. Then, there exists a martingale $M \in \mathcal{M}^2$ satisfying the following properties.

(i) $M^\varepsilon \overset{\#^2}{\rightarrow} M$ as $\varepsilon \downarrow 0$. Furthermore, there exist a nullset $Q$ and a deterministic sequence $(\varepsilon_n)_{n \geq 1}$ with $\varepsilon_n \downarrow 0$, such that for every $t \geq 0$,

$$\lim_{n \to \infty} \sup_{s \in [0,t]} (M^\varepsilon_n - M_s)^2 = 0 \quad (3.8)$$

outside of $Q$.

(ii) $M$ has almost surely càdlàg paths and stationary, independent increments.

In other words, $M$ is a Lévy process which is also a martingale. Furthermore, $M$ has characteristic exponent

$$\int_{[-1,1]} (e^{i\theta x} - 1 - i\theta x) \Pi(dx), \quad \theta \in \mathbb{R}.$$ 

**Proof.** (i) We begin by proving that $M$ exists as a limit in $\mathcal{M}^2$. The first step is to show that $(M^\varepsilon)$ is Cauchy.

Fix $0 < \eta < \varepsilon < 1$, and pick $t \geq 0$. Using Theorem 3.4, we calculate

$$\|M^\eta - M^\varepsilon\|^2 = E\left[ (M^\eta_t - M^\varepsilon_t)^2 \right]$$

$$= E\left[ \left( \int_{[0,t]} \int_{\eta \leq |x| < \varepsilon} x N(ds, dx) - t \int_{\eta \leq |x| < \varepsilon} x \Pi(dx) \right)^2 \right]$$

$$= t \int_{\eta \leq |x| < \varepsilon} x^2 \Pi(dx),$$

from which follows that

$$\|M^\eta - M^\varepsilon\| \leq \int_{\eta \leq |x| < \varepsilon} x^2 \Pi(dx) \sum_{n=1}^{\infty} n 2^{-n}.$$

Using the condition $\int_{[-1,1]} x^2 \Pi(dx) < \infty$, it follows that when $\eta, \varepsilon$ become small, the distance $\|M^\eta - M^\varepsilon\| \to 0$; that is, $(M^\varepsilon)$ is Cauchy in $\mathcal{M}^2$. It therefore follows that a limit exists in $\mathcal{M}^2$, and we shall call it $M$. 

21
We now seek to prove the stronger convergence given in the theorem, which will emerge as a fairly simple consequence of classical martingale theory. To begin with, let $t = 1$ and choose any sequence $(\varepsilon_n^0)$ which approaches zero. According to Doob’s maximal inequality [3, Theorem 2.5.19],

$$\lim_{n \to \infty} E \left[ \sup_{s \in [0,1]} \left( M_s - M_s^{\varepsilon_n^0} \right)^2 \right] \leq 4 \lim_{n \to \infty} \| M - M^{\varepsilon_n^0} \|_1 = 0. \quad (3.9)$$

The $L^1$ convergence of the sequence of suprema on the left-hand side of (3.9) entails the existence of a subsequence $(\varepsilon_n^1) \subset (\varepsilon_n^0)$ and a $P$-nullset $Q_1$ such that

$$\lim_{n \to \infty} \sup_{s \in [0,1]} \left( M_s - M_s^{\varepsilon_n^1} \right)^2 = 0$$

outside of $Q_1$. (This is a fact from measure theory; see [1, Theorem 15.7] or [7, Theorem 3.12]. This link in the PDF works at time of writing.)

To extend this idea to all $t$, we first iterate; the relation

$$\lim_{n \to \infty} E \left[ \sup_{s \in [0,m]} \left( M_s - M_s^{\varepsilon_n^{m-1}} \right)^2 \right] \leq 4 \lim_{n \to \infty} \| M - M^{\varepsilon_n^{m-1}} \|_m = 0$$

holds, and we extract a subsequence $(\varepsilon_n^m) \subset (\varepsilon_n^{m-1})$ along which almost sure convergence holds outside of some nullset $Q_m$. Then we define the diagonal sequence $\varepsilon_n = \varepsilon_n^n$, and claim that for any fixed $t \geq 0$,

$$\lim_{n \to \infty} \sup_{s \in [0,t]} \left( M_s - M_s^{\varepsilon_n} \right)^2 = 0$$

outside of the nullset $Q = \bigcup_m Q_m$. This is easy to believe. Take $m$ to be the next integer after $t$ and use the simple bound $\sup_{s \in [0,t]} \cdots \leq \sup_{s \in [0,m]} \cdots$; then observe that $(\varepsilon_n^n)_{n \geq m} \subset (\varepsilon_n^m)_{n \geq m}$, and use the almost sure convergence along the sequence $(\varepsilon_n^m)$ to obtain the result.

(ii) The fact that $M$ has càdlàg paths is automatic by virtue of its being an element of $\mathcal{M}^2$. Fix $0 \leq s_1 < t_1 \leq s_2 < \cdots < s_n < t_n$ and $\theta_1, \ldots, \theta_n$. The almost sure convergence given in the previous parts, together with the dominated convergence theorem, allows us to make the following calculation.

$$E \left[ \prod_{j=1}^n e^{i \theta_j (M_{t_j} - M_{s_j})} \right] = \lim_{m \to \infty} E \left[ \prod_{j=1}^n e^{i \theta_j (M_{t_j}^{(e_m)} - M_{s_j}^{(e_m)})} \right]$$

$$= \lim_{m \to \infty} \prod_{j=1}^n E \left[ e^{i \theta_j M_{t_j}^{(e_m)} - s_j} \right]$$

$$= \prod_{j=1}^n E \left[ e^{i \theta_j M_{t_j} - s_j} \right].$$
The equality of the first and last expressions is equivalent to $M$ having stationary, independent increments.

The final claim about the characteristic exponent follows simply, since $M_1^\varepsilon$ converges to $M_1$ in $L^2$, and therefore also in probability, and indeed in distribution; and this is enough [see Proposition 2.5(iv)] for the characteristic functions to converge! □

Let us make a short remark, which may be of interest but which we will not need again. The convergence given in (3.8) actually tells us that the paths of $M_1^\varepsilon$ converge to the paths of $M_1$ almost surely in the local uniform topology on the space $D[0, \infty)$ of real càdlàg functions on $[0, \infty)$. This is a rather strong mode of convergence and in particular implies that the paths of $M_1^\varepsilon$ converge to those of $M_1$ almost surely in the Skorokhod topology on $D[0, \infty)$. For definitions and a broader discussion of these concepts, see [5, Chapter VI].

### 3.4 Lévy–Itô decomposition

We now state formally the decomposition we have proved.

**Theorem 3.9** (Lévy–Itô decomposition). Let $a \in \mathbb{R}$, $\sigma \geq 0$ and let $\Pi$ be a measure with no atom at zero satisfying $\int_\mathbb{R} (1 \wedge x^2) \Pi(dx)$. Then there exist three independent Lévy processes: a linear Brownian motion

$$X^{(1)} = \sigma B_t + at;$$

a compound Poisson process $X^{(2)}$ with rate $\Pi(\mathbb{R} \setminus [-1, 1])$ and jump distribution $\Pi(\cdot \cap \mathbb{R} \setminus [-1, 1])/\Pi(\mathbb{R} \setminus [-1, 1])$; and a third Lévy process $X^{(3)}$, the square-integrable martingale $M$ in Theorem 3.8 given by the limit of compensated compound Poisson processes the magnitude of whose jumps is less than 1. The process $X = X^{(1)} + X^{(2)} + X^{(3)}$ is a Lévy process with characteristic exponent

$$\Psi(\theta) = ia\theta - \frac{1}{2}\sigma^2\theta^2 + \int_\mathbb{R} (e^{iax} - 1 - x1_{[-1,1]}(x)) \Pi(dx).$$

We remark that this theorem proves the converse part of Theorem 1.6, the Lévy–Khintchine formula.

**Remark 3.10.** The set

$$\{(s, \Delta X_s) : s \geq 0\}$$

is a Poisson point process with characteristic measure $\Pi$. This is because it may be viewed as the union of: (a) the Poisson point process induced by the compound Poisson process $X^{(2)}$, which has characteristic measure $\Pi(\cdot \cap \{x : |x| > 1\})$; and (b) those Poisson point processes with characteristic measures $\Pi(\cdot \cap \{x : \varepsilon_{n+1} \leq |x| < \varepsilon_n\})$, for $n \geq 1$, which arise from the sequence of martingales $M^{\varepsilon_n}$. This superposition is again a Poisson point process by the same reasoning as in the proof of Theorem 3.3.
4 A collection of results

4.1 Path variation

In this section, we are interested in deciding the question of when the paths of a Lévy process have bounded variation. The answer will emerge fairly simply from the Lévy–Itô decomposition.

We give a brief recap of the notion of path variation. First consider a function $f: [0, \infty) \to \mathbb{R}$. Given any partition $\mathcal{P} = \{a = t_0 < t_2 < \cdots < t_n = b\}$ of the bounded interval $[a, b]$ we define the variation of $f$ over $[a, b]$ with partition $\mathcal{P}$ by

$$V_\mathcal{P}(f, [a, b]) = \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|.$$ 

The variation of $f$ over $[a, b]$ is given by

$$V(f, [a, b]) := \sup_\mathcal{P} V_\mathcal{P}(f, [a, b]),$$

where the supremum is taken over all partitions of $[a, b]$; if this quantity is finite, we say that $f$ is of bounded variation over $[a, b]$. Moreover, $f$ is said to be of bounded variation if $V(f, [a, b]) < \infty$ for all bounded intervals $[a, b]$. If $V(f, [a, b]) = \infty$ for all bounded intervals $[a, b]$ then we say that $f$ is of unbounded variation.

A very useful fact, which we will not prove, is that if $f$ is a càdlàg function (and therefore has at most a countable number of discontinuities), then its variation may be written

$$V(f, [a, b]) = V(f^c, [a, b]) + \sum_{x \in (a, b]} |\Delta f(x)|,$$  \hfill (4.1) 

where $f^c(x) = f(x) - \sum_{y \leq x} \Delta f(y)$ is the continuous part of $f$ and $\Delta f(x) := f(x) - f(x-)$ is the jump of $f$ at $x$.

When considering a stochastic process $X = \{X_t : t \geq 0\}$, we may adopt these notions in the almost sure sense. So, for example, the statement $X$ is a process of bounded variation (or has paths of bounded variation) simply means that $P$-almost every path of $X$ is of bounded variation.

For a Lévy process $X$, we give the following test for bounded variation.

**Lemma 4.1.** The Lévy process $X$ is of bounded variation if

$$\sigma = 0 \quad \text{and} \quad \int_{\mathbb{R}} 1 \wedge |x| \Pi(dx) < \infty,$$

and unbounded variation otherwise.

**Proof.** From the description of the Poisson point process of jumps in [Remark 3.10], as well as [Theorem 3.4], we see that the sum $\sum_{s \leq t} |\Delta X_s|$ converges a.s. if and only if $\int_{\mathbb{R}} 1 \wedge |x| \Pi(dx) < \infty$. Suppose that this does hold. Using (4.1) we see that we then only have to be worried about the variation of the continuous part of $X$. The continuous
part of \(X^{(2)}\) is zero, and under our assumption \(X^{(3)}\) is precisely an absolutely convergent jump part compensated by a deterministic drift, \(-t \int_{[-1,1]} x \Pi(dx)\), so the variation of \(X^{(3)}\) is always finite. Finally, \(X^{(1)}\) has finite variation if and only if \(\sigma = 0\), and this completes the proof.

We end with a nice representation of bounded variation Lévy processes.

**Lemma 4.2.** Let \(X\) be a Lévy process of bounded variation. Then the characteristic exponent may be written

\[
\Psi(\theta) = i d\theta + \int_{\mathbb{R}} (e^{i \theta x} - 1) \Pi(dx),
\]

(4.2)

where \(d = a - \int_{[-1,1]} x \Pi(dx)\). Furthermore,

\[
X_t = dt + \sum_{s \leq t} \Delta X_s, \quad t \geq 0.
\]

(4.3)

**Proof.** The expression for \(d\) follows from the Lévy–Khintchine representation. The expression for \(X_t\) is derived by observing that, when \(X\) is of bounded variation,

\[
X^{(3)}_t = \sum_{s \leq t} \Delta X_s \mathbb{1}_{\{\Delta X_s \leq 1\}} - t \int_{[-1,1]} x \Pi(dx), \quad t \geq 0.
\]

Due to (4.3), the term \(d\) is called the drift coefficient of \(X\).

### 4.2 Behaviour around stopping times

We now offer some results around the theme of ‘Lévy processes as Markov processes’. We will not prove anything in this section. Proofs of all the results may be found in [6, Chapter 3].

Let \(X\) be a Lévy process, and define a filtration via \(\mathcal{F}_t = \sigma(X_s, s \leq t)\), for \(t \geq 0\). This filtration is typically neither right-continuous nor complete. However, if we define \(\mathcal{F}_t\) to be the \(P\)-regularisation\(^2\) of \(\mathcal{F}_t\), then the new filtration \((\mathcal{F}_t)_{t \geq 0}\) is automatically right-continuous, thus, it satisfies the natural conditions. (See [2, Proposition I.4].)

As we have already remarked, a direct consequence of the independent increments property of \(X\) is the (spatially homogeneous) Markov property; that is, for any fixed time \(t \geq 0\), the process \((X_{t+s} - X_t, s \geq 0)\) is independent of \(\mathcal{F}_t\) and is equal in distribution to \(X\).

Lévy processes also possess the strong Markov property, which describes the process after a stopping time; here are the definitions. A \([0, \infty]\)-valued random variable \(T\) is called a stopping time (with respect to the filtration \((\mathcal{F}_t)_{t \geq 0}\)) if, for every \(t \geq 0\), the set \(\{T \leq t\} \in \mathcal{F}_t\). Since \((\mathcal{F}_t)_{t \geq 0}\) is right-continuous, this is equivalent to \(\{T < t\} \in \mathcal{F}_t\) for every \(t \geq 0\). For any stopping time \(T\), we may define the associated \(\sigma\)-algebra

\[
\mathcal{F}_T = \{\mathcal{A} : \text{for all } t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}.
\]

\(^2\)i.e., we adjoin the \(P\)-nullsets by defining \(\mathcal{F}_t = \sigma\{A \cup N : A \in \mathcal{F}_t, \exists M \supset N : P(M) = 0\}\)
We will first give some nice examples of stopping times. Let \( B \) be a Borel subset of \( \mathbb{R} \). The random times
\[
T_B = \inf\{t \geq 0 : X_t \in B\} \quad \text{and} \quad \tau_B = \inf\{t > 0 : X_t \in B\}
\]
are referred to as the first entrance time and first passage time of \( B \), respectively.

**Proposition 4.3.** If \( B \) is either open or closed, then

(i) \( T_B \) is a stopping time and \( P(X_{T_B} \in \overline{B} \mid T_B < \infty) = 1 \), and

(ii) \( \tau_B \) is a stopping time and \( P(X_{\tau_B} \in \overline{B} \mid \tau_B < \infty) = 1 \).

We have the following extension of the Markov property to stopping times.

**Theorem 4.4** (strong Markov property). Let \( T \) be any stopping time. Under the conditional measure \( P(\cdot \mid T < \infty) \), the process \( (X_{T+t} - X_T, t \geq 0) \) is independent of \( \mathcal{F}_T \) and has the same distribution as \( X \).

We can even say something about the behaviour of the Lévy process just before certain stopping times. Recall that the Lévy process is right-continuous with left-limits. It is certainly not left-continuous unless it is a linear Brownian motion. However, we do have the following.

**Theorem 4.5** (quasi-left-continuity). Suppose that \((T_n)\) is an (a.s.) increasing sequence of stopping times such that the almost sure limit \( \lim_{n \to \infty} T_n =: T \) exists. Then, under the conditional measure \( P(\cdot \mid T < \infty) \), \( \lim_{n \to \infty} X_{T_n} = X_T \) a.s..

A closely related result applying to fixed times is the following, which states that \( X \) has ‘no fixed jumps’.

**Proposition 4.6.** Fix a time \( t \geq 0 \). Then, \( X \) is continuous at time \( t \) with probability 1.

This can be easily proved without the need for Theorem 4.5: \( N(\{t\} \times \mathbb{R}) = 0 \) a.s., which implies that, with probability 1, \( X \) has no jump at the fixed time \( t \).

### 4.3 Compensation formula

This section gives a precise statement of a ‘process’ version of Theorem 3.4, which is also often called the compensation formula. It is exactly the same as Theorem 3.4 in the case where \( \phi \) is non-random.

**Theorem 4.7** (Compensation formula, Lévy process). Let \( X \) be a Lévy process and \((\mathcal{F}_t)_{t \geq 0}\) the natural enlargement of its induced filtration, and write \( N \) for the Poisson random measure on \([0, \infty) \times \mathbb{R}\) associated with the jumps of \( X \). Consider a function \( \phi : \Omega \times [0, \infty) \times \mathbb{R} \to [0, \infty) \) satisfying

(i) \( \phi \) is a measurable function.
A collection of results

(ii) For each fixed \( t \geq 0 \), the function \((\omega, x) \mapsto \phi(\omega, t, x)\) is \( \mathcal{F}_t \otimes \mathcal{B}\)-measurable, where \( \mathcal{B} \) is the Borel \( \sigma \)-algebra.

(iii) For each fixed \( x \in \mathbb{R} \), \((\phi(t, x), t \geq 0)\), viewed as a stochastic process, is left-continuous with probability 1.

Then,
\[
E \left[ \int_{(0, \infty) \times \mathbb{R}} \phi(t, x) N(dt, dx) \right] = E \left[ \int_{(0, \infty) \times \mathbb{R}} \phi(t, x) dt \Pi(dx) \right].
\]

This result is stated and proved in [6, Theorem 4.4]. We also remark that both these theorems can be extended to predictable processes \( \phi \) via the monotone class theorem, but we will not need this.

4.4 Existence of moments

Here we are interested in knowing when \( E[g(X_t)] \) is finite, and for a certain, quite wide, class of functions, there is a simple integral test.

**Definition 4.8.** A non-negative function \( g \) on \( \mathbb{R} \) is called submultiplicative if there exists a constant \( a > 0 \) such that
\[
g(x + y) \leq ag(x)g(y), \quad x, y \in \mathbb{R}.
\]
A function \( g \) is called locally bounded if it is bounded on every compact subset of its domain.

A very important fact about moments of the one-dimensional distributions of Lévy processes is that, in almost every important case, their finiteness is determined by the Lévy measure \( \Pi \), as follows.

**Theorem 4.9.** Let \( g \) be a locally bounded, submultiplicative measurable function on \( \mathbb{R} \). Then, for any \( t > 0 \),
\[
E[g(X_t)] < \infty \quad \text{if and only if} \quad \int_{|x| > 1} g(x) \Pi(dx) < \infty.
\]

We point out that the functions \( x \mapsto |x|; \quad x \mapsto \exp(x^\beta) \) (for any \( \beta \in (0, 1) \)); and \( x \mapsto \log_+|x| \) are submultiplicative. In addition, if \( g \) is a submultiplicative function, then the function \( x \mapsto g(cx + \gamma)^{\alpha} \) is also submultiplicative, for any choice of \( c, \gamma \in \mathbb{R} \) and \( \alpha > 0 \); and the product of any two submultiplicative functions is submultiplicative. For a proof of these facts, see [8, Proposition 25.4].

We will not prove Theorem 4.9; a proof may be found in Sato [8, Theorem 25.3]. Let us remark here that this theorem gives the proof of our claim, at the end of §2.5, about the ‘heavy tails’ of stable distributions.
4.5 Subordinators

A subordinator is a Lévy process whose paths are non-decreasing with probability one. It is simple to give a criterion for a Lévy process $X$ to be a subordinator. The paths must be of bounded variation; $X$ may not have any negative jumps; and it must have non-negative drift. That is, we require

$$\sigma = 0, \quad \Pi(-\infty, 0) = 0, \quad \int_{(0,\infty)} 1 \wedge x \Pi(dx) < \infty, \quad \text{and} \quad d \geq 0,$$

(4.4)

where $d$ is as given in Lemma 4.2. Further, it is clear that any Lévy process satisfying these conditions has non-decreasing paths, so (4.4) characterises subordinators.

Since $\mathbb{P}(X_t < 0) = 0$, $X$ possesses every negative exponential moment, i.e. $\mathbb{E}[e^{-\lambda X_t}]$ is finite for every $\lambda, t \geq 0$. We therefore define the Laplace exponent $\kappa$ of a subordinator $X$ to be the function $\kappa : [0, \infty) \to \mathbb{R}$ given by

$$\mathbb{E}[e^{-\lambda X_t}] = e^{-t\kappa(\lambda)}, \quad \lambda \geq 0.$$

Proposition 4.10. If $X$ is a subordinator, then for $\lambda \geq 0$,

$$\kappa(\lambda) = d\lambda + \int_{[0,\infty)} (1 - e^{-\lambda x}) \Pi(dx).$$

Furthermore, $\kappa$ has an analytic extension to $\mathbb{C}_r = \{ \lambda \in \mathbb{C} : \text{Re} \lambda \geq 0 \}$, and $-\kappa(-i\theta) = \Psi(\theta)$.

Let us state and prove a rather nice result about subordinators, which has been proved (in various versions) by Kesten, Horowitz and Bertoin. The result is a first step in the direction of fluctuation theory, and it shows in fairly explicit form how a subordinator crosses a level. Generalisations are possible for any Lévy process, but they require the Wiener–Hopf factorisation; see [6]. Suppose that $X$ is a subordinator. We define the potential measure (or the 0-potential measure) $U$ by:

$$U(dx) = \mathbb{E}\left[ \int_0^\infty \mathbb{I}_{\{X_t \in dx\}} \, dt \right] = \int_0^\infty \mathbb{P}(X_t \in dx) \, dt, \quad x \geq 0,$$

and the first-passage time

$$\tau^+_x := \tau(x, \infty) = \inf\{ t > 0 : X_t > x \},$$

for any $x > 0$. Then we have the following proposition concerning the so-called overshoot and undershoot at first passage.

Proposition 4.11. Suppose that $X$ is a subordinator with Lévy measure $\Pi$ and potential measure $U$, and choose a level $x > 0$. Suppose that $f$ and $g$ are Borel-measurable functions and that $f(0) = 0$. Then,

$$\mathbb{E}[f(X_{\tau^+_x} - x)g(x - X_{\tau^+_x})] = \int_{[0, x]} U(dy) \int_{(x - y, \infty)} \Pi(du)f(u + y - x)g(x - y).$$
Proof. The proof is a typical application of the compensation formula for Lévy processes. We make the following calculation. We use the fact that \( X \) has at most countably many jumps (and the assumption \( f(0) = 0 \)) to write the first sum; then the elementary identity \( X_{t} = X_{t-} + \Delta X_{t} \); then the compensation formula; then the fact that Lebesgue measure is diffuse and the set of jumps is countable. The calculation is as follows:

\[
\begin{align*}
\mathbb{E} \left[ f(X_{t+} - x)g(x - X_{t-}) \right] &= \mathbb{E} \left[ \sum_{t>0} f(X_{t} - x)g(x - X_{t}) \mathbb{I}_{\{X_{t-} \leq x \wedge X_{t} > x \}} \right] \\
&= \mathbb{E} \left[ \int_{[0,\infty) \times \mathbb{R}} f(X_{t-} + z - x)g(x - X_{t-}) \mathbb{I}_{\{X_{t-} \leq x \}} \mathbb{I}_{\{X_{t-} + z > x \}} N(dt, dz) \right] \\
&= \mathbb{E} \left[ \int_{0}^{\infty} dt g(x - X_{t-}) \mathbb{I}_{\{X_{t-} \leq x \}} \int_{(0,\infty)} \Pi(du) f(X_{t-} + u - x) \mathbb{I}_{\{u > x - X_{t-} \}} \right] \\
&= \int_{0}^{\infty} dt \mathbb{P}(X_{t} \in dy, X_{t} \leq x) g(x - y) \int_{(x-y,\infty)} \Pi(du) f(u + y - x),
\end{align*}
\]

which yields the statement we are trying to prove. \( \square \)

The result is often stated as:

\[
\mathbb{P}(X_{t+} - x \in dv, x - X_{t+} \in dz) = U(x - dz) \Pi(z + dv), \quad v > 0, z \in [0, x],
\]

and this is obtained simply by a change of variables.

The assumption on \( f \) means that the proposition will not see the event \( \{X_{t+} = x\} \), even if this is possible. This phenomenon is known as ‘creeping’, and a discussion of it may be found in [6, §5.3].

Here is a nice application of the classical renewal theorem, which we also do not prove.

**Corollary 4.12.** Suppose that \( X \) is a subordinator and that \( \mu := \mathbb{E}[X_1] < \infty \). Suppose further that the support of the renewal measure is not concentrated on any sublattice of \( \mathbb{R} \). Then the following limit in the sense of weak convergence of measures holds.

\[
\lim_{x \to \infty} \mathbb{P}(X_{t+} - x \in dv, x - X_{t+} \in dz) = \frac{1}{\mu} dz \Pi(z + dv).
\]

We finish with an example. Let \( X \) be a strictly stable subordinator with index \( \alpha < 1 \). Then \( \kappa(\lambda) = c\lambda^\alpha \), and \( \Pi(dx) = c_{1} x^{-(\alpha + 1)} \), where \( c_{1} = \frac{c}{\Gamma(-\alpha)\cos(\pi\alpha/2)} \). Moreover, we can calculate \( U \). Observe (and this is a general remark applying to all subordinators) that

\[
\int e^{-\lambda x} U(dx) = \int_{0}^{\infty} \mathbb{E}[e^{-\lambda X_{t}}] dt = \int_{0}^{\infty} e^{-t\kappa(\lambda)} dt = \frac{1}{\kappa(\lambda)}.
\]

Using this Laplace transform identity, it follows that \( U(dx) = \frac{x^{\alpha - 1}}{\Gamma(\alpha)} dx \), so

\[
\mathbb{P}(X_{t+} - x \in dv, x - X_{t+} \in dz) = \frac{c_{1}}{\Gamma(\alpha)} (x-z)^{\alpha-1} (z+v)^{-(\alpha+1)} dz dv, \quad v > 0, z \in [0, x].
\]
4 A collection of results

4.6 Duality lemma

In this section we discuss a simple feature of all Lévy processes which follows as a direct consequence of stationary independent increments. When the path of a Lévy process over a finite time horizon is time reversed (in an appropriate sense) the new path is equal in law to the process reflected about the origin.

Lemma 4.13 (Duality lemma). For each fixed $t > 0$, define the reversed process

$$Y_s := X_{(t-s)} - X_t : 0 \leq s \leq t$$

and the process,

$$\{-X_s : 0 \leq s \leq t\}.$$  

Then the two processes have the same law under $P$, and $-X$ is known as the dual Lévy process.

Proof. First, note that $Y_0 = -\Delta X_t = 0$ almost surely (due to Proposition 4.6). As can be seen from Figure 7 (which is to be understood symbolically), the paths of $Y$ are obtained from those of $X$ by a reflection about the vertical axis with an adjustment of the continuity at the jump times so that its paths are almost surely right continuous with left limits. For fixed $0 \leq s \leq t$, the distribution of $X_{(t-s)} - X_t$ is identical to that of $-X_s$, so $Y_s$ has the same distribution as $-X_s$. Since $X$ has stationary independent increments, the same as is true of $Y$ (prove this by considering the characteristic function of the increments). This then implies that the law of the process $Y$ is determined by its one-dimensional distributions, and it follows that $Y$ and $(-X_s, s \leq t)$ are equal in law. \hfill \Box

The duality lemma is well-known for random walks, which are the discrete time analogue of Lévy processes, and is justified using an identical proof. See, for example, Feller [4, §XII.2].

One consequence of the duality lemma is the relationship between the running supremum, the running infimum, the process reflected in its supremum and the process reflected in its infimum. The last four objects are, respectively,

$$\bar{X}_t = \sup_{0 \leq s \leq t} X_s, \quad \underline{X}_t = \inf_{0 \leq s \leq t} X_s,$$

$$\{\bar{X}_t - X_t : t \geq 0\} \text{ and } \{X_t - \underline{X}_t : t \geq 0\}.$$ 

Lemma 4.14. For each fixed $t > 0$, the pairs $(\bar{X}_t, \bar{X}_t - X_t)$ and $(X_t - \underline{X}_t, -\underline{X}_t)$ have the same distribution under $P$.

Proof. Define $\bar{Y}_s = X_t - X_{(t-s)} = -Y_s$ for $0 \leq s \leq t$ and write $\bar{X}_t = \inf_{0 \leq s \leq t} \bar{Y}_s$. Using right continuity and left limits of paths we may deduce that

$$(\bar{X}_t, \bar{X}_t - X_t) = (\bar{X}_t - \bar{X}_t, -\bar{X}_t)$$

almost surely. One may visualise this in Figure 8. By rotating the picture about by $180^\circ$ one sees the almost sure equality of the pairs $(\bar{X}_t, \bar{X}_t - X_t)$ and $(\bar{X}_t - \bar{X}_t, -\bar{X})$. Now appealing to the duality lemma we have that $\{\bar{X}_s : 0 \leq s \leq t\}$ is equal in law to $\{X_s : 0 \leq s \leq t\}$ under $P$. The result now follows. \hfill \Box
Figure 7: Duality of the processes $X = \{X_s : s \leq t\}$ and $Y = \{X_{(t-s)} - X_t : s \leq t\}$. The path of $Y$ is a reflection of the path of $X$ with an adjustment of continuity at jump times.

Figure 8: Duality of the pairs $(X_t, X_t - X_t)$ and $(X_t - X_t, -X_t)$.
Acknowledgements. This course is based on a graduate course of Andreas Kyprianou. All figures (with the exception of [5]) were adapted from Kyprianou [6].

References


