Topics in Probability: Lévy Processes
Math 7880-1; Spring 2011

Davar Khoshnevisan

155 South 1400 East JWB 233, Department of Mathematics, University of Utah, Salt Lake City UT 84112–0090

E-mail address: davar@math.utah.edu

URL: http://www.math.utah.edu/~davar

Contents

Lecture 1. Introduction 1

What is a Lévy process? 1

Infinite divisibility 2

The Lévy–Khintchine formula 3

On equation (1) 4

Problems for Lecture 1 7

Lecture 2. Some Examples 9

Uniform motion 9

Poisson processes on the real line 9

Nonstandard Brownian motion with drift 10

Isotropic stable laws 10

The asymmetric Cauchy distribution on the line 12

The Gamma distribution on the half line 12

Adding independent Lévy processes 13

Problems for Lecture 2 13

Lecture 3. Continuous-Parameter Martingales 15

Filtrations 15

Martingales 16

Modifications 18

Problems for Lecture 3 19

Lecture 4. Poisson Random Measures 21

A construction of Poisson random measures 21
The Poisson process on the line 24
Problems for Lecture 4 24

Lecture 5. Poisson Point Processes 25
A construction of Poisson point processes 25
Compound Poisson processes 26
Problems for Lecture 5 28

Lecture 6. Lévy Processes 29
The Lévy–Itô construction 29
Problems for Lecture 6 32

Lecture 7. Structure Theory 33
The Lévy–Itô decomposition 33
The Gaussian Part 34
The Compound Poisson Part 35
A strong law of large numbers 38
Symmetry and isotropy 39
Problems for Lecture 7 40

Lecture 8. Subordinators 43
Laplace exponents 44
Stable subordinators 45
Subordination 50
Problems for Lecture 8 51

Lecture 9. The Strong Markov Property 53
Transition measures and the Markov property 53
The strong Markov property 55
Feller semigroups and resolvents 56
The Hille–Yosida theorem 58
The form of the generator 60
Problems for Lecture 9 61

Lecture 10. Potential theory 63
Potential measures 63
The range of a Lévy process 64
Problems for Lecture 10 70

Lecture 11. Recurrence and Transience 71
The recurrence/transience dichotomy 71
The Port–Stone theorem 74
Problems for Lecture 11 76

Lecture 12. Excessive Functions 79
Absolute continuity considerations 79
Excessive functions 81
Lévy processes that hit points 82
Problems for Lecture 12 83

Lecture 13. Energy and Capacity 85
Polar and essentially-polar sets 85
An energy identity 86
Proof of Theorem 3 88
Problems for Lecture 13 90

Bibliography 91

Lecture 1

Introduction

What is a Lévy process?

In a nutshell, Lévy processes are continuous-time random walks that are "mathematically viable." We are about to describe these processes in greater depth, but it might help to keep in mind Brownian motion as a central example of a "mathematically viable" continuous-time random walk.

A stochastic process $X := \{X_t\}_{t\geq 0}$ [with values in $\mathbf{R}^d$] is a *continuous-time random walk* if $X_0 = 0$ [the process starts at the origin at time 0] and X has *i.i.d. increments*. The latter property means that for all $s, t \geq 0$:

(1) $X_{t+s} - X_s$ and X_t have the same distribution; and

(2) $X_{t+s} - X_s$ is independent of $\{X_r\}_{r\in[0,s]}$.

While the preceding definition makes perfect sense in discrete time, it does not lend itself to a rich continuous-time theory as the following example might suggest:

Consider the following deterministic [i.e., nonrandom] equation:

$$f(t+s) = f(t) + f(s) \qquad \text{for all } s, t \geq 0. \tag{1}$$

It is a fundamental fact that all Borel-measurable solutions to equation (1) have the form $f(t) = at$ for some $a \geq 0$; see Theorem 9 below. But it is also known (Hamel, 1905) that, under the axiom of choice, (1) has nonmeasurable nonlinear solutions [which can be shown are nowhere continuous also]; see Theorem 10. Choose and fix one such badly-behaved solution, call it f, and observe that $X_t := f(t)$ is a [nonrandom] continuous-time random walk! The nonlinear solutions to the nonrandom equation (1) have very bad measurability properties, and therefore the class of all continuous-time random walks contains such badly-behaved objects that it is hopeless to study them seriously. Fortunately, there is a fix that is simple to describe:

Definition 1. A *Lévy process* $X := \{X_t\}_{t\geq 0}$ is a continuous-time random walk such that the *trajectories*, or *paths*, of X [i.e., the function $t \mapsto X_t$ as a function of $\omega \in \Omega$] are right-continuous with left limits everywhere.[1] □

This is a course about Lévy processes. Some standard references are the following books: Bertoin (1996); Kyprianou (2006); Sato (1999); see also the survey monograph by Fristedt (1974). The notation of this course is on the whole borrowed from Math. 6040 (Khoshnevisan, 2007).

Infinite divisibility

Suppose $X := \{X_t\}_{t\geq 0}$ is a Lévy process on $\mathbf{R}^d$, and let μ_t denote the distribution [or "law"] of X_t for every $t \geq 0$. For all $n \geq 1, t > 0$,

$$X_t = \sum_{j=1}^{n} \left(X_{jt/n} - X_{(j-1)t/n} \right) \tag{2}$$

is a sum of n i.i.d. random variables. For example, we set $t = 1$ to find that

$$\mu_1 = \mu_{1/n} * \cdots * \mu_{1/n} \qquad (n \text{ times}),$$

where "$*$" denotes convolution. Equivalently, we can write this using the Fourier transform as $\hat{\mu}_1 = \hat{\mu}_{1/n}^n$; in particular, $\hat{\mu}_1^{1/n}$ is the Fourier transform of a probability measure.

We can also apply (2) with $t := n$ to find that

$$\mu_n = \mu_1 * \cdots * \mu_1 \qquad (n \text{ times}),$$

or equivalently, $(\hat{\mu}_1)^n$ is also the Fourier transform of a probability measure. Thus, to summarize, $(\hat{\mu}_1)^t$ is the Fourier transform of a probability measure for all rationals $t \geq 0$. In fact, a little thought shows us that $(\hat{\mu}_1)^t$ is the Fourier transform of μ_t for all rational $t \geq 0$. Thus, we can write probabilistically,

$$\left[\hat{\mu}_1(\xi)\right]^t = \mathrm{E}e^{i\xi\cdot X_t} = \hat{\mu}_t(\xi) \qquad \text{for all } \xi \in \mathbf{R}^d,$$

for all rational $t \geq 0$. And because X has right-continuous trajectories, for all $t \geq 0$ we can take rational $s \downarrow t$ to deduce that the preceding holds for all $t \geq 0$. This shows $(\hat{\mu}_1)^t$ is the Fourier transform of a probability measure for every $t \geq 0$.

[1] In other words, $X_t = \lim_{s\downarrow t} X_s$ and $X_{t-} := \lim_{s\uparrow t} X_s$ exists for all $t > 0$.

Definition 2. A Borel probability measure ρ on $\mathbf{R}^d$ is said to be *infinitely divisible* if $(\hat{\rho})^t$ is the Fourier transform of a probability measure for every $t \geq 0$. □

Thus, if X is a Lévy process then the distribution of X_1 is infinitely divisible. In fact, the very same reasoning shows that the distribution of X_s is infinitely divisible for all $s \geq 0$. A remarkable fact, due to Lévy and then Itô, is that the converse is also true: Every infinitely-divisible measure ρ on $\mathbf{R}^d$ corresponds in a unique way to a Lévy process X in the sense that the law [i.e., the distribution] of X_1 is ρ. Thus, we can see immediately that the standard-normal law on $\mathbf{R}$ corresponds to one-dimensional Brownian motion, and Poiss(λ) to a rate-λ Poisson process on the line. In other words, the study of Lévy processes on $\mathbf{R}^d$ is in principle completely equivalent to the analysis of all infinitely-divisible laws on $\mathbf{R}^d$.

The Lévy–Khintchine formula

We need to introduce some terminology before we can characterize infinitely-divisible laws on $\mathbf{R}^d$.

Definition 3. A *Lévy triple* is a trio $(a\,,\sigma,m)$ where $a \in \mathbf{R}^d$, σ is a $(d \times d)$ matrix, and m is a Borel measure on $\mathbf{R}^d$ such that

$$m(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbf{R}^d} \left(1 \wedge \|x\|^2\right) m(\mathrm{d}x) < \infty. \tag{3}$$

The matrix $\sigma'\sigma$ is called the *diffusion* [*or Gaussian*] *covariance matrix*, and m is called the *Lévy measure*. □

Although $m(\mathbf{R}^d)$ might be infinite, (3) ensures that $m(A) < \infty$ for every open set A that does not contain the origin. Indeed, the fact that $q \mapsto q/(1+q)$ is decreasing on $\mathbf{R}_+$ implies that

$$\begin{aligned} m\left\{z \in \mathbf{R}^d : \|z\| > r\right\} &\leq \frac{1+r^2}{r^2} \cdot \int_{\mathbf{R}^d} \frac{\|x\|^2}{1+\|x\|^2}\, m(\mathrm{d}x) \\ &\leq \frac{1+r^2}{r^2} \cdot \int_{\mathbf{R}^d} \left(1 \wedge \|x\|^2\right) m(\mathrm{d}x) < \infty. \end{aligned} \tag{4}$$

A similar argument shows that $\int_{\mathbf{R}^d}(q \wedge \|x\|^2)\, m(\mathrm{d}x) < \infty$ for all $q > 0$.

Definition 4. A *Lévy exponent* is the function $\Psi : \mathbf{R}^d \to \mathbf{C}$, where

$$\Psi(\xi) := i(a \cdot \xi) + \frac{1}{2}\|\sigma\xi\|^2 + \int_{\mathbf{R}^d} \left[1 - \mathrm{e}^{i\xi\cdot z} + i(\xi \cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right] m(\mathrm{d}z), \tag{5}$$

where $(a\,,\sigma,m)$ is a Lévy triple. □

Lemma 5. *The integral in* (5) *is absolutely convergent, and* Ψ *is continuous with* $\Psi(0) = 0$.

Proof. Because $1-\cos\theta \le \theta^2/2$ and $\theta - \sin\theta \le \theta^2/6$ for all $\theta \in \mathbf{R}$ [Taylor's theorem with remainder],

$$\left|\frac{1-\mathrm{e}^{-i\xi\cdot z} - i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)}{\|\xi\|^2}\right| \le \left(\|z\|^2 \wedge \frac{1}{\|\xi\|^2}\right).$$

The definition of a Lévy measure then tells us that the integral in Ψ is absolutely convergent. And the remaining assertions follow easily from this. □

The following result is called the *Lévy–Khintchine formula*; it provides the reason for introducing all this terminology.

Theorem 6 (Khintchine, 1938; Kolmogorov, 1932; Lévy, 1934)**.** *A Borel probability measure ρ on $\mathbf{R}^d$ is infinitely divisible if and only if $\hat\rho(\xi) = \exp(-\Psi(\xi))$ for all $\xi \in \mathbf{R}^d$, where Ψ is a Lévy exponent. The corresponding triple $(a\,,\sigma, m)$ determines ρ uniquely.*

See also Lévy (1937, pp. 212–220).

The uniqueness portion will follow immediately from Fourier analysis; see also the proof of Theorem 3 below [page 29]. That proof also implies the more important half of the theorem; namely, that if $\hat\rho = \mathrm{e}^{-\Psi}$ for a Lévy exponent Ψ, then ρ is infinitely divisible. The proof of the remaining half is a difficult central-limit-type argument, and does not concern our immediate needs; you can find it in Sato (1999, pages 42–45).

There are many other ways of writing the Lévy–Khintchine formula. Here is one that is used frequently: Suppose $\int_{\mathbf{R}^d}(1\wedge\|x\|)\,m(\mathrm{d}x) < \infty$. Then $\int_{\mathbf{R}^d}(1-\mathrm{e}^{i\xi\cdot z})\,m(\mathrm{d}z)$ and $\int_{\mathbf{R}^d} i(\xi\cdot z)\,m(\mathrm{d}z)$ both converge absolutely; this can be seen from a Taylor expansion similar to the one in the proof of Lemma 5. Therefore, in the case that $\int_{\mathbf{R}^d}(1\wedge\|x\|)\,m(\mathrm{d}z) < \infty$, we can also write

$$\begin{aligned}\Psi(\xi) &= i(b\cdot\xi) + \frac{1}{2}\|\sigma\xi\|^2 + \int_{\mathbf{R}^d}\left[1-\mathrm{e}^{i\xi\cdot z}\right]\,m(\mathrm{d}z),\\ \text{where } b &:= a - \int_{\|z\|<1} z\,m(\mathrm{d}z).\end{aligned} \tag{6}$$

Let us conclude with a few basic properties of Lévy exponents.

Lemma 7. *If Ψ is a Lévy exponent, then $\overline{\Psi}$ and $\mathrm{Re}\Psi$ are also Lévy exponents. Moreover, $\overline{\Psi}(\xi) = \Psi(-\xi)$ and $\mathrm{Re}\Psi(\xi) \ge 0$ for all $\xi \in \mathbf{R}^d$.*

On equation (1)

This section is not covered in the lectures. Nonetheless it would be a shame to say nothing indepth about the functional equation (1). Therefore, we close this chapter with a discussion on (1).

Suppose f solves (1). Then $f(0) = 2f(0)$, whence we have $f(0) = 0$. Therefore, we can extend f to a function F on all of $\mathbf{R}$ as follows: $F(x) = f(x)$ if $x \geq 0$; and $F(x) = -f(-x)$ for all $x < 0$. Note that F solves "Cauchy's functional equation,"

$$F(x+y) = F(x) + F(y) \qquad \text{for every } x, y \in \mathbf{R}. \tag{7}$$

The preceding reduces the analysis of (1) to an analysis of (7). Therefore, we investigate the latter equation from now on. The following is immediate, but important.

Proposition 8. *If F solves (7), then $F(kx) = kF(x)$ for all $x \in \mathbf{R}$ and all integers $k \geq 2$. In particular, $F(y) = yF(1)$ for all rationals y.*

Suppose $F : \mathbf{R} \to \mathbf{R}$ and $G : \mathbf{R} \to \mathbf{R}$ are right-continuous functions that have left limits everywhere and agree on the rationals. Then it is easy to see that $F = G$ on all of $\mathbf{R}$. As a consequence of Proposition 8, we find that if F solves (7) and F is right continuous with left limits everywhere, then $F(x) = xF(1)$ for all x; i.e., F is a linear function. As it turns out, "right continuous with left limits" can be reduced to the seemingly-stronger condition "Lebesgue measurable," without changing the content of the preceding discussion. More precisely, we have the following.

Theorem 9. *Every Lebesgue-measurable solution to (7) is linear.*

This is classical; we follow a more recent proof—due to Fitzsimmons (1995)—that is simpler than the classical one.

Proof. Let F be a Lebesgue-measurable solution to (1), consider the $\mathbf{C}$-valued function $G(x) := \exp(iF(x))$, defined for every $x \in \mathbf{R}$. It is clear that

$$G(x+y) = G(x)G(y) \qquad \text{for all } x, y \in \mathbf{R}; \tag{8}$$

compare with (1). Because $|G(x)| = 1$, G is locally integrable and never vanishes. Therefore, $\int_0^a G(x)\,\mathrm{d}x \neq 0$ for almost every $a \geq 0$. We can now integrate (8) over all $y \in [0\,,a]$ [for almost all $a \geq 0$] to obtain the following: For almost all $a \in \mathbf{R}$, we have

$$\int_x^{a+x} G(y)\,\mathrm{d}y = G(x) \cdot \int_0^a G(y)\,\mathrm{d}y \qquad \text{for every } x \in \mathbf{R}. \tag{9}$$

The dominated convergence theorem implies that the left-hand side is a continuous function of x, and hence so is the right-hand side; i.e., G is continuous. Thanks to Proposition 8 every continuous solution to (8) is a complex exponential. Therefore, there exists $\theta \in \mathbf{R}$ such that $G(x) = \mathrm{e}^{i\theta x}$ for all $x \in \mathbf{R}$. From this it follows that $F(x) = \theta x + 2\pi N(x)$, where $N : \mathbf{R} \to \mathbf{Z}$ is measurable. It suffices to prove that $N(x) = 0$ for all x. But this is not too difficult to establish. Indeed, we note that N solves (7).

Therefore, $N(x)/k = N(x/k)$ for all $x \in \mathbf{R}$ and integers $k \geq 2$ (Proposition 8). Consequently, $N(x)/k$ is an integer for all $x \in \mathbf{R}$ and $k \geq 2$; and this implies readily that $N(x) = 0$ [for otherwise, we could set $k = 2|N(x)|$]. □

Theorem 10 (Hamel, 1905)**.** *Assume the axiom of choice. Then, there are uncountably-many nonmeasurable solutions to* (7)*.*

In fact, the proof will show that there are "many more" nonmeasurable solutions to (7) than there are measurable ones.

Proof (sketch). Let $\mathbf{H}$—a socalled "Hamel basis"—denote the maximal linearly-independent subset of $\mathbf{R}$, where $\mathbf{R}$ is viewed as a vector space over the field $\mathbf{Q}$ of rationals. The existence of $\mathbf{H}$ follows from the axiom of choice (Hewitt and Stromberg, 1965, (3.12), p. 14). And it follows fairly easily from the axiom of choice (Hewitt and Stromberg, 1965, (3.20), p. 18) that for every $x \in \mathbf{R}$ there exists a unique function $\xi_x : \mathbf{H} \to \mathbf{Q}$ such that: (i) $\xi_x(h) = 0$ for all but a finite number of $h \in \mathbf{H}$; and (ii) $x = \sum_{h\in\mathbf{H}} h\xi_x(h)$. Because a countable union of countable sets is itself countable [this follows, for instance, from the axiom of choice], we can deduce that $\mathbf{H}$ has the cardinality $\mathfrak{c}$ of the continuum. [For if the cardinality of $\mathbf{H}$ were $< \mathfrak{c}$, then $\mathbf{H}$ would be countable.]

Now let $\mathcal{F}$ denote the collection of all functions $\phi : \mathbf{H} \to \mathbf{R}$; the cardinality of $\mathcal{F}$ is $2^{\mathfrak{c}} > \mathfrak{c}$.

Define

$$F_\phi(x) := \sum_{h\in\mathbf{H}} \phi(h)\xi_x(h) \qquad \text{for all } x \in \mathbf{R} \text{ and } \phi \in \mathcal{F}. \tag{10}$$

Since $\mathbf{H}$ is linearly independent, it follows that if ϕ and ψ are two different elements of $\mathcal{F}$, then $F_\phi \neq F_\psi$. Consequently, the cardinality of $\{F_\phi\}_{\phi\in\mathcal{F}}$ is $2^{\mathfrak{c}}$ [one for every $\phi \in \mathcal{F}$].

The definition of the ξ_x's implies that F_ϕ solves (7) for every $\phi \in \mathcal{F}$. It follows from Proposition 8 that if F_ϕ were continuous, then F_ϕ would be linear; in fact, $F_\phi(x) = xF_\phi(1)$ for all $x \in \mathbf{R}$. Note that the collection of all numbers $F_\phi(1)$ such that F_ϕ is continuous is at most $\mathfrak{c}$. Therefore, the cardinality of all linear/continuous solutions to (7) that have the form F_ϕ for some $\phi \in \mathcal{F}$ is at most $\mathfrak{c}$. Because $\{F_\phi\}_{\phi\in\mathcal{F}}$ has cardinality $2^{\mathfrak{c}} > \mathfrak{c}$, it follows that there are at least $2^{\mathfrak{c}} - \mathfrak{c} = 2^{\mathfrak{c}}$ different solutions to (7) none of which are continuous. Theorem 9 then implies that every discontinuous solution F_ϕ to (7) is nonmeasurable, and this completes the proof. □

Interestingly enough, Lévy (1961)—after whom the stochastic processes of this course are named—has used the Hamel basis $\mathbf{H}$ of the preceding proof in order to construct an "explicitly constructed" set in $\mathbf{R}$ that is not measurable.

Problems for Lecture 1

1. Verify Lemma 7.

2. Prove that every Lévy process X on $\mathbf{R}^d$ is *continuous in probability*; i.e., if $s \to t$ then $X_s \to X_t$ in probability. [Our later examples show that convergence in probability cannot in general be improved to almost-sure convergence.]

3. Verify that if X is a Lévy process on $\mathbf{R}^d$, then $\{X_{jh}\}_{j=0}^{\infty}$ is a d-dimensional random walk for every fixed $h > 0$.

4. Let μ be an infinitely-divisible distribution on $\mathbf{R}^d$, so that for every integer $n \geq 1$ there exists a Borel probability measure μ_n on $\mathbf{R}^d$ such that $\hat{\mu}^{1/n} = \hat{\mu}_n$. Observe that $\lim_{n\to\infty} |\hat{\mu}_n|^2$ is the indicator of the set $\{\xi \in \mathbf{R}^d : \hat{\mu}(\xi) \neq 0\}$. Use this to prove that $\hat{\mu}$ is never zero. (Hint: You may use the following theorem of P. Lévy without proof: *If $\{\mu_n\}_{n\geq 1}$ is a sequence of probability measures on $\mathbf{R}^d$ such that $q := \lim_{n\to\infty} \hat{\mu}_n$ exists and q is continuous in an open neighborhood of the origin, then there exists a probability measure μ such that $q = \hat{\mu}$, and $\mu_n \Rightarrow \mu$, in particular, $\hat{\mu}_n \to \hat{\mu}$ everywhere.*)

5. Is Unif$\,(a\,,b)$ infinitely divisible? (Hint: Exercise 4!)

6. Verify the following:

(1) Y is infinitely divisible iff its law [or *distribution*] is;

(2) Constants are infinitely-divisible random variables;

(3) $\mathrm{N}(\mu\,,\sigma^2)$ is infinitely divisible for every fixed $\mu \in \mathbf{R}$ and $\sigma^2 > 0$;

(4) Poiss(λ) is infinitely divisible for every $\lambda > 0$ fixed;

(5) Gamma$\,(\alpha\,,\lambda)$ is infinitely divisible for all $\alpha, \lambda > 0$ fixed. Recall that the density function of Gamma$\,(\alpha\,,\lambda)$ is $f(x) := \lambda^\alpha x^{\alpha-1} \mathrm{e}^{-\lambda x} / \Gamma(\alpha) \cdot \mathbb{1}_{(0,\infty)}(x)$. Gamma laws include Exp$\,(\lambda) =$ Gamma$\,(1\,,\lambda)$ and $\xi_k^2 =$ Gamma$\,(k\,,1/2)$ for $k \geq 1$.

7. One can combine Lévy processes to form new ones:

(1) Prove that if $X^1, \ldots, X^k$ are independent Lévy processes with values in $\mathbf{R}^d$, then $Y_t := X_t^1 + \cdots + X_t^k$ defines a Lévy process on $\mathbf{R}^d$ as well. Identify the Lévy triple and exponent of Y in terms of those of X^j's.

(2) Prove that if $X^1, X^2, \cdots, X^k$ are independent Lévy processes with respective values in $\mathbf{R}^{d_1}, \ldots, \mathbf{R}^{d_k}$, then $Z_t := (X_t^1\,, \ldots, X_t^k)$ defines a Lévy process with values in $\mathbf{R}^d$, where $d = d_1 + \cdots + d_k$. Identify the Lévy triple and exponent of Z in terms of those of X^j's. In particular, prove that if X is a Lévy process on $\mathbf{R}^d$, then $Y_t := (t\,, X_t)$ defines a Lévy process on $\mathbf{R}_+ \times \mathbf{R}^d$. Identify the Lévy triple and exponent of Y in terms of those of X^j's.

8. Prove that $\limsup_{\|\xi\|\to\infty} |\Psi(\xi)|/\|\xi\|^2 < \infty$ for all Lévy exponents Ψ.

9 (Time reversal)**.** Verify that if X is a Lévy process on $\mathbf{R}^d$, then $\tilde{X}_t := -X_t$ defines a Lévy process. Identify its Lévy triple and exponent of $\tilde{X}$ in terms of those of X. Furthermore, prove that for every fixed time $t \geq 0$, the processes $\{X_{(t-s)-} - X_t\}_{s\in[0,t]}$ and $\{\tilde{X}_s\}_{s\in[0,t]}$ have the same finite-dimensional distributions.

Some Examples

Our immediate goal is to see some examples of Lévy processes, and/or infinitely-divisible laws on $\mathbf{R}^d$.

Uniform motion

Choose and fix a nonrandom $a \in \mathbf{R}^d$ and define

$$X_t := at \qquad \text{for all } t \geq 0. \tag{1}$$

Then, $\{X_t\}_{t\geq 0}$ is a [nonrandom] Lévy process with Lévy triple $(a\,,0\,,0)$. The process $\{X_t\}_{t\geq 0}$ denotes uniform motion in the direction of a.

Poisson processes on the real line

If $N = \text{Poiss}(\lambda)$ for some $\lambda > 0$, then

$$\mathrm{E}e^{i\xi N} = \sum_{n=0}^{\infty} e^{i\xi n} \cdot \frac{e^{-\lambda}\lambda^n}{n!} = \exp\left(-\lambda\left(1 - e^{i\xi}\right)\right). \tag{2}$$

That is, $\mathrm{E}\exp(i\xi N) = \exp(-\Psi(\xi))$, where

$$\Psi(\xi) = \int_{\mathbf{R}} \left(1 - e^{i\xi z} - i\xi z \mathbb{1}_{(0,1)}(|z|)\right) m(dz),$$

and $m(dz) := \lambda\delta_{\{1\}}(dz)$. The corresponding Lévy process is called the *Poisson process* with *intensity parameter* λ.

Nonstandard Brownian motion with drift

The Lévy triple $(0\,,\,I\,,0)$, where "I" denotes the $(d \times d)$ identity matrix, belongs to a vector of d i.i.d. standard-normal random variables, and the corresponding Lévy process is [standard] d-dimensional Brownian motion. We can generalize this example easily: Choose and fix a vector $a \in \mathbf{R}^d$, and a $(d\times d)$ matrix σ, and consider the Lévy triple $(a\,,\sigma\,,0)$. The corresponding Lévy exponent is

$$\Psi(\xi) = i(a \cdot \xi) + \frac{1}{2}\|\sigma\xi\|^2.$$

Therefore, Ψ is the Lévy exponent of random vector X in $\mathbf{R}^d$ if and only if $X = -a + \sigma Z$ where Z is a vector of d i.i.d. standard-normal random variables. The corresponding Lévy process is described by $W_t := -at + \sigma B_t$, where B is standard Brownian motion [check!] . The jth coordinate of W is a Brownian motion with mean $-a_j$ and variance $v_j^2 := (\sigma'\sigma)_{j,j}$, and the coordinates of W are not in general independent. Since $\lim_{t\to\infty}(W_t/t) = -a$ a.s. by the law of large numbers for Brownian motion, $-a$ is called the "drift" of the nonstandard Brownian motion W.

Isotropic stable laws

Choose and fix a number α. An *isotropic stable law of index α* is the infinitely-divisible distribution with Lévy exponent $\Psi(\xi) = c\|\xi\|^\alpha$, where $c \in \mathbf{R}$ is a fixed constant. The corresponding Lévy process is called the isotropic stable process with index α. We consider only random vectors with Lévy exponent $\exp(-c\|\xi\|^\alpha)$ in this discussion.

Of course, $c = 0$ leads to $\Psi \equiv 0$, which is the exponent of the infinitely divisible, but trivial, random variable $X \equiv 0$. Therefore, we study only $c \neq 0$. Also, we need $|\exp(-\Psi(\xi))| = \exp\{-c\|\xi\|^\alpha\} \leq 1$, and this means that c cannot be negative. Finally, $\Psi(0) = 0$, and hence we have $\alpha > 0$.

Lemma 1 (Lévy)**.** *$\Psi(\xi) = c\|\xi\|^\alpha$ is a Lévy exponent iff $\alpha \in (0\,,2]$.*

The case $\alpha = 2$ is the Gaussian case we just saw. And $\alpha = 1$ is also noteworthy; the resulting distribution is the "isotropic [or symmetric, in dimension one] Cauchy distribution" whose density is

$$f(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}c^{d/2}}\left(1 + \frac{\|x\|^2}{c^2}\right)^{-(d+1)/2} \qquad \text{for all } x \in \mathbf{R}^d. \tag{3}$$

Proof of Lemma 1. Exercise 8 below shows that $|\Psi(\xi)| = O(\|\xi\|^2)$ as $\|\xi\| \to \infty$; therefore, $\alpha \leq 2$. Since $\alpha = 2$ is Gaussian, we limit ourselves to $\alpha < 2$.

Next let us consider $\alpha \in (0\,,2)$ [and of course $c > 0$]. In that case, a change of variables shows that for all $\xi \in \mathbf{R}^d$,

$$\int_{\mathbf{R}^d} \left(1 - e^{i\xi\cdot z} - i(\xi \cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right) \frac{dz}{\|z\|^{d+\alpha}} = \int_{\mathbf{R}^d} (1 - \cos(\xi \cdot z)) \frac{dz}{\|z\|^{d+\alpha}} \propto \|\xi\|^{\alpha}. \tag{4}$$

[The first identity is justified because the left-most integral is a radial function of ξ, and hence real.] Therefore, we can choose $C \in (0\,,\infty)$ so that $m(dx) = C\|x\|^{-(d+\alpha)}\,dx$ is the Lévy measure with exponent $\Psi(\xi) = \exp(-c\|\xi\|^{\alpha})$ iff (3) on page 3 holds. □

Sometimes, a reasonable knowledge of the Lévy exponent of an infinitely-divisible law yields insight into its structure. Here is a first example; Skorohod (1961) contains much more precise [and very useful] estimates of the tails of stable laws.

Proposition 2. *If X has an isotropic stable distribution with $\alpha \in (0\,,2)$, then for all $\beta > 0$, $\mathrm{E}(\|X\|^{\beta}) < \infty$ iff $\beta < \alpha$.*

In other words, except in the case that $\alpha = 2$, the decay of the tail of an isotropic stable law is slow. It is also possible to say more about the tails of the distrbution (Theorem 15, page 50). But that requires a more sophisticated analysis.

Proof. Because $\mathrm{E}\exp(iz \cdot X) = \exp(-c\|z\|^{\alpha}) = \mathrm{E}\cos(z \cdot X)$ for some $c > 0$, we may apply (4) in the following form:

$$\int_{\mathbf{R}^d} \left(1 - e^{i\xi\cdot z}\right) \frac{dz}{\|z\|^{d+\alpha}} = \int_{\mathbf{R}^d} (1 - \cos(i\xi \cdot z)) \frac{dz}{\|z\|^{d+\alpha}} \propto \|\xi\|^{\alpha}. \tag{5}$$

Replace ξ by X and take expectations to obtain

$$\mathrm{E}\left(\|X\|^{\beta}\right) \propto \mathrm{E}\int_{\mathbf{R}^d} (1 - \cos(z \cdot X)) \frac{dz}{\|z\|^{d+\beta}} = \int_{\mathbf{R}^d} \left(1 - e^{-c\|z\|^{\alpha}}\right) \frac{dz}{\|z\|^{d+\beta}} \propto \int_0^{\infty} \left(1 - e^{-r^{\alpha}}\right) \frac{dr}{r^{\beta+1}}.$$

Now

$$\int_1^{\infty} \left(1 - e^{-r^{\alpha}}\right) \frac{dr}{r^{\beta+1}} \le \int_1^{\infty} \frac{dr}{r^{\beta+1}} < \infty \qquad \text{for every } \beta > 0.$$

Therefore, $\mathrm{E}(\|X\|^{\beta}) < \infty$ iff $\int_0^1 (1 - e^{-r^{\alpha}})dr/r^{\beta+1} < \infty$. The result follows from this and the fact that $(\theta/2) \le 1 - e^{-\theta} \le \theta$ for all $\theta \in (0\,,1)$. □

The asymmetric Cauchy distribution on the line

We can specialize the preceding example to $d = \alpha = 1$ and obtain the symmetric Cauchy law μ on the line. Of course the density of μ is known as well.[1] But more significantly, we have learned that the Lévy triple of μ is $(0\,,0\,,m)$, where $m(\mathrm{d}z) \propto z^{-2}\,\mathrm{d}z$. This suggests that perhaps we can create an asymmetric variation of the Cauchy law by considering a Lévy triple of the form $(c_0\,,0\,,m)$ for a Lévy measure of the form

$$\frac{m(\mathrm{d}z)}{\mathrm{d}z} = \frac{c_1}{z^2}\mathbb{1}_{(0,\infty)}(z) + \frac{c_2}{z^2}\mathbb{1}_{(-\infty,0)}(z),$$

where $c_1 \neq c_2$ are both positive and c_0 is selected carefully. This can in fact be done, as we will see next.

Theorem 3. *For every $c, c_1, c_2 > 0$ and $\theta \in [-2/\pi\,,2/\pi]$ there exists $c_0 \in \mathbf{R}$ and an infinitely-divisible Borel probability measure μ on $\mathbf{R}$ such that: (i) The Lévy triple of μ has the form $(c_0\,,0\,,m)$ for m as above; and (ii) $\hat{\mu}(\xi) = \exp(-c|\xi| - ic\theta\xi\log|\xi|)$ for $\xi \in \mathbf{R}\setminus\{0\}$.*[2]

The measure μ given above is called the *Cauchy distribution* on $\mathbf{R}$ with *asymmetry parameter* θ [and scale parameter c]. When $\theta = 0$, μ is the symmetric Cauchy law. When $|\theta| = 2/\pi$, μ is called the *completely asymmetric Cauchy law* on $\mathbf{R}$.

The Gamma distribution on the half line

It is easy to see that the Gamma $(\alpha\,,\lambda)$ distribution on $(0\,,\infty)$ is infinitely divisible for every $\alpha, \lambda > 0$. Next we identify its Lévy triple.

Proposition 4. *If μ is a* Gamma $(\alpha\,,\lambda)$ *distribution on $\mathbf{R}_+$ for some $\alpha, \lambda > 0$, then μ is infinitely divisible with $a := 0$, $\sigma := 0$, and Lévy measure m and Lévy exponent Ψ respectively given by*

$$m(\mathrm{d}x) = \frac{\alpha \mathrm{e}^{-\lambda x}}{x}\,\mathrm{d}x \cdot \mathbb{1}_{(0,\infty)}(x), \quad \Psi(\xi) = \alpha\log\left(1 - \frac{i\xi}{\lambda}\right),$$

where "log" denotes the principle branch of the logarithm.

---

[1] It is $f(x) = 1/\{\pi(1+x^2)\}$ for $-\infty < x < \infty$.

[2] I am making a little fuss about $\xi \in \mathbf{R}\setminus\{0\}$, since the function $\xi\log|\xi|$ is defined only for $\xi \in \mathbf{R}\setminus\{0\}$. But of course $\hat{\mu}(0) = 1$ because μ is a probability measure. Alternatively, we can define the function $\xi\log|\xi|$ cotinuously on all of $[0\,,\infty)$ by letting $0\log|0|$ be $\lim_{\xi\to 0, \xi\neq 0}\xi\log|\xi| = 0$. If so, then the stated formula for $\hat{\mu}(\xi)$ is valid for all $\xi \in \mathbf{R}$.

Adding independent Lévy processes

Finally, let me mention the following device which can be used to generate new Lévy processes from old ones.

Lemma 5. *If $\{X_t\}_{t\geq 0}$ and $\{\bar{X}_t\}_{t\geq 0}$ are independent Lévy processes on $\mathbf{R}^d$ with respective triples $(a\,,\,\sigma\,,m)$ and $(\bar{a}\,,\,\bar{\sigma}\,,\bar{m})$, then $\{X_t + Y_t\}_{t\geq 0}$ is a Lévy process on $\mathbf{R}^d$ with Lévy triple $(A\,,\Sigma\,,M)$, where $A := a + \bar{a}$ and $M := m + \bar{m}$, and Σ can be chosen in any fashion as long as $\Sigma'\Sigma = \sigma'\sigma + \bar{\sigma}'\bar{\sigma}$.*

This is elementary, and can be checked by directly verifying the defining properties of Lévy processes. However, I emphasize that: (i) There is always such a Σ;[3] and (ii) $\Sigma'\Sigma$ is defined uniquely even though Σ might not be.

Problems for Lecture 2

1. Prove that if $q := \int_0^1 z^{-2}(z - \sin z)\,\mathrm{d}z - \int_1^\infty z^{-2}\sin z\,\mathrm{d}z$, then for all $\xi > 0$,

$$\int_0^\infty \left(1 - \mathrm{e}^{i\xi z} + i\xi z \mathbb{1}_{(0,1)}(z)\right)\frac{\mathrm{d}z}{z^2} = \frac{\pi\xi}{2} + i\xi\ln\xi + iq\xi.$$

Deduce Theorem 3 from this identity.

2. This problem outlines the proof of Proposition 4.

(1) Suppose $f : (0\,,\infty) \to \mathbf{R}$ has a continuous derivative $f' \in L^1(\mathbf{R})$, and $f(0+) := \lim_{x\downarrow 0} f(x)$ and $f(\infty-) := \lim_{x\uparrow\infty} f(x)$ exist and are finite. Prove that for all $\lambda, \rho > 0$,

$$\int_0^\infty \left[\frac{f(\lambda x) - f((\lambda+\rho)x)}{x}\right]\mathrm{d}x = \left[f(0+) - f(\infty-)\right]\ln\left(1 + \frac{\rho}{\lambda}\right).$$

This is the *Frullani integral identity.*

(2) Prove that for all $\lambda > 0$ and $\xi \in \mathbf{R}$,

$$1 - \frac{i\xi}{\lambda} = \exp\left(\int_0^\infty \mathrm{e}^{-\lambda x}\left(1 - \mathrm{e}^{i\xi x}\right)\frac{\mathrm{d}x}{x}\right),$$

and deduce Proposition 4 from this identity.

3 (Stable scaling)**.** Prove that if X is an isotropic stable process in $\mathbf{R}^d$ with index $\alpha \in (0\,,2]$, then $Y_t := R^{-1/\alpha}X_{Rt}$ defines a Lévy process for every $R > 0$ fixed. Explicitly compute the Lévy exponent of Y. Is the same result true when R depends on t? What happens if you consider instead an asymmetric Cauchy process on the line?

[3] This follows readily from the fact that $\sigma'\sigma + \tilde{\sigma}'\tilde{\sigma}$ is a nonnegative-definite matrix.

Continuous-Parameter Martingales

Here and throughout, $(\Omega, \mathcal{F}, \mathrm{P})$ denotes a complete probability space. [Recall that "completeness" is a property of $\mathcal{F}$; namely, that all subsets of P-null sets are $\mathcal{F}$-measurable and P-null.]

Filtrations

Definition 1. A *filtration* $\{\mathcal{F}_t\}_{t\geq 0}$ is a family of sub-sigma-algebras of $\mathcal{F}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$.

Definition 2. A filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies the *usual conditions* ["conditions habituelles"] if:

(1) $\mathcal{F}_t$ is P-complete for every $t \geq 0$; and
(2) $\{\mathcal{F}_t\}_{t\geq 0}$ is *right continuous*; i.e., $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ for all $t \geq 0$.

Given a filtration $\{\mathcal{G}_t\}_{t\geq 0}$, there exists a smallest filtration $\{\mathcal{F}_t\}_{t\geq 0}$ that satisfies the usual conditions. We can construct the latter filtration in a few steps as follows:

(1) Let $\mathcal{F}_0^0$ denote the completion of $\mathcal{G}_0$;
(2) Define $\mathcal{F}_t^0$ to be the smallest sigma-algebra that contains both $\mathcal{F}_0^0$ and $\mathcal{G}_t$ [for all $t \geq 0$];
(3) Define $\mathcal{F}_t := \cap_{s>t}\mathcal{F}_s^0$. Then, $\{\mathcal{F}_t\}_{t\geq 0}$ is the desired filtration.

From now on, we assume tacitly that all filtrations satisfy the usual conditions, unless it is stated explicitly otherwise.

Martingales

Let $X := \{X_t\}_{t\geq 0}$ be a real-valued stochastic process on $(\Omega, \mathcal{F}, \mathrm{P})$, and $\{\mathcal{F}_t\}_{t\geq 0}$ a filtration on the same space.

Definition 3. X is *adapted* to $\{\mathcal{F}_t\}_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.

Definition 4. X is a *martingale* [with respect to $\{\mathcal{F}_t\}_{t\geq 0}$] if:

(1) X is *adapted* to $\{\mathcal{F}_t\}_{t\geq 0}$;
(2) $X_t \in L^1(\mathrm{P})$ for all $t \geq 0$;
(3) $\mathrm{E}(X_{t+s} \,|\, \mathcal{F}_s) = X_s$ almost surely for all $s, t \geq 0$.

Thus, continuous-time martingales are defined just as in the discrete-time theory. However, there is a notable technical matter that arises: The last part of the definition of martingales has to be understood carefully. It states that for all $s, t \geq 0$ there exists a P-null set $N_{s,t}$ such that

$$\mathrm{E}(X_{t+s} \,|\, \mathcal{F}_s) = X_s \qquad \text{a.s. on } N_{s,t}^c.$$

Definition 5. The filtration $\{\mathcal{X}_t\}_{t\geq 0}$ *generated* by the stochastic process X is defined as the smallest filtration such that: (a) X is adapted to $\{\mathcal{X}_t\}_{t\geq 0}$; and (b) $\{\mathcal{X}_t\}_{t\geq 0}$ satisfies the usual conditions. We might refer to $\{\mathcal{X}_t\}_{t\geq 0}$ also as the *natural filtration* of X.

It can be verified directly that it X is a martingale with respect to some filtration $\{\mathcal{F}_t\}_{t\geq 0}$, then X is certainly a martingale in its own natural filtration $\{\mathcal{X}_t\}_{t\geq 0}$. Therefore, unless we need explicit information about the filtrations involved, we say that X is a martingale without mentioning the filtrations explicitly. [If this happens, then we are assuming that the underlying filtration is the natural filtration of X.]

Definition 6. A stochastic process $\{X_t\}_{t\geq 0}$ with values in a Euclidean space is *cadlag* ["continue á droite, limitée à gauche"] if $t \mapsto X_t$ is right continuous and the left limits

$$X_{t-} := \lim_{s\uparrow t} X_s \qquad \text{exist for all } t > 0.$$

Some authors use "rcll" in place of "cadlag." But let's not do that here ☺.

The continuous-time theory of general processes is quite complicated, but matters simplify greatly for cadlag processes, as the next theorem shows. First, let us recall

Definition 7. A map $T : \Omega \to \mathbf{R}_+ \cup \{\infty\}$ is a *stopping time* [for a given filtration $\{\mathcal{F}_t\}_{t\geq 0}$] if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. The sigma-algebra $\mathcal{F}_T$ is defined as in discrete-parameter theory, viz.,

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

Theorem 8. *Suppose X is a process that takes values in $\mathbf{R}^d$, is cadlag, and is adapted to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Then for all stopping times T, $X_T \mathbb{1}_{\{T<\infty\}}$ is a random variable. Moreover, $T_A := \inf\{s > 0 : X_s \in A\}$ is a stopping time for every $A \in \mathcal{B}(\mathbf{R}^d)$, provided that we define $\inf \varnothing := \infty$.*

Theorem 9 (The optional stopping theorem)**.** *Suppose X is a cadlag martingale and T is a stopping time that is bounded a.s. That is, suppose there exists a nonrandom $k > 0$ such that $\mathrm{P}\{T < k\} = 1$. Then, $\mathrm{E}(X_T) = \mathrm{E}(X_0)$.*

I will not prove this here, but suffice it to say that the idea is to follow a discretization scheme, which enables us to appeal to the optional stopping theorem of discrete-parameter martingale theory. See the proof of Theorem 11 below for this sort of argument.

Definition 10. Choose and fix an real number $p > 1$. X is said to be a *cadlag L^p martingale* if it is a cadlag martingale and $X_t \in L^p(\mathrm{P})$ for all $t \geq 0$.

Theorem 11 (Doob's maximal inequality)**.** *If X is a cadlag L^p martingale for some $p > 1$, then*

$$\mathrm{E}\left(\sup_{s\in[0,t]} |X_s|^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathrm{E}\left(|X_t|^p\right) \qquad \textit{for every } t \geq 0.$$

In other words, the L^p norm of the maximum process is at most q times the L^p norm of the terminal value of the process, where q is the conjugate to p; i.e., $p^{-1} + q^{-1} = 1$.

Sketch of Proof. Notice that if F is an arbitrary finite set in $[0\,,t]$, then $\{X_s\}_{s\in F}$ is a discrete-time martingale [in its own filtration]. Therefore, discrete-time theory tells us that

$$\mathrm{E}\left(\max_{s\in F} |X_s|^p\right) \leq \left(\frac{p}{p-1}\right)^p \max_{s\in F} \mathrm{E}(|X_s|^p) \leq \left(\frac{p}{p-1}\right)^p \mathrm{E}(|X_t|^p).$$

(Why the last step?) Now replace F by $F_n := \{jt/2^n;\ 0 \leq j \leq 2^n\}$ and take limits [$n \uparrow \infty$]: By the dominated convergence theorem,

$$\lim_{n\to\infty} \mathrm{E}\left(\max_{s\in F_n} |X_s|^p\right) = \mathrm{E}\left(\sup_{n\geq 1} \max_{s\in F_n} |X_s|^p\right),$$

and the supremum is equal to $\sup_{s\in[0,t]} |X_s|^p$ because X [and hence $s \mapsto |X_s|^p$] is cadlag. □

Similarly, one can derive the following from the discrete-parameter theory of martingales:

Theorem 12 (The martingale convergence theorem). *Let X be a cadlag martingale such that either: (a) $X_t \geq 0$ a.s. for all $t \geq 0$; or (b) $\sup_{t\geq 0} \mathrm{E}(|X_t|) < \infty$. Then, $\lim_{t\to\infty} X_t$ exists a.s. and is finite a.s.*

In like manner, we can define continuous-parameter supermartingales, submartingales, and reverse martingales. In the case that those processes are cadlag, the discrete-parameter theory extends readily to the continuous-parameter setting. I will leave the numerous details and variations to you.

Modifications

Now we address briefly what happens if we have a quite general continuous-parameter martingale that is not cadlag.

Definition 13. The *finite-dimensional distributions* of a stochastic process X are the collection of all joint probabilities of the form

$$\mathrm{P}\left\{X_{t_1} \in A_1\,, \ldots, X_{t_k} \in A_k\right\},$$

as $t_1, \ldots, t_k$ range over all possible numbers in $\mathbf{R}_+ := [0\,,\infty)$, and $A_1, \ldots, A_k$ over all possible measurable subsets of the state space where X takes its values.

It is important to remember that, *a priori*, the finite-dimensional distributions of X are the only hard piece of information that we have available on a process X. [Think, for example, about how we learned about Brownian motion in Math. 6040.]

Definition 14. Let $X := \{X_t\}_{t\geq 0}$ and $Y := \{Y_t\}_{t\geq 0}$ be two stochastic processes with values in a common space. We say that X is a *modification* of Y if $\mathrm{P}\{X_t = Y_t\} = 1$ for all $t \geq 0$.

We can make some elementary observations: First, if X is a modification of Y, then Y is also a modification of X; second—and this is important—if X and Y are modifications of one another, then their finite-dimensional distributions are the same. In other words, if X and Y are modifications of one another, then they are "stochastically the same." However, we next see that not all modifications are created equal; some are clearly better than others.

Example 15. Let B denote a one-dimensional Brownian motion. Let us introduce an independent positive random variable T with an absolutely continuous distribution [say, $T = \mathrm{Unif}\,[0\,,1]$, or $T = \mathrm{Exp}\,(1)$, etc.]. And now we can define a new process X by setting $X_t(\omega) := B_t(\omega)$ if $t \neq T(\omega)$, and $X_t(\omega) := 5$ if $T(\omega) = t$ for all $t \geq 0$ and $\omega \in \Omega$. Since $\mathrm{P}\{T = t\} = 0$ for all

$t \geq 0$, it follows that X and B are modifications of one another. Therefore, X is a Brownian motion in the sense that X has i.i.d. increments with $X_t = \mathrm{N}(0\,, t)$ for all $t \geq 0$. However, $t \mapsto X_t$ is a.s. discontinuous [with probability one, X has a jump at T]. □

The following is an important result in the general theory of processes. In words, it states that $\{X_t\}_{t\geq 0}$ always has a cadlag modification, which has the same finite-dimensional distributions as $\{X_t\}_{t\geq 0}$.

Theorem 16. *Every martingale $\{X_t\}_{t\geq 0}$ has a cadlag modification. That modification is a cadlag martingale.*

Therefore we can, and will, always consider only cadlag martingales.

The proof of the preceding theorem is not particularly hard, but it takes us too far afield. Therefore, we will skip it. You can find the details of a proof in (Khoshnevisan, 2002, p. 225). However, here is an important consequence, which we will use from now on.

If Y is an integrable random variable and $\{\mathcal{F}_t\}_{t\geq 0}$ a filtration, then $M_t := \mathrm{E}(Y \,|\, \mathcal{F}_t)$ defines a martingale. If T is a simple stopping time with possible values in a finite [nonrandom] set F, then for all $A \in \mathcal{F}_T$,

$$\mathrm{E}\left(\mathrm{E}\left(Y \,\middle|\, \mathcal{F}_T\right); A\right) = \mathrm{E}\left(Y; A\right) = \sum_{t\in F} \mathrm{E}\left(Y; A \cap \{T = t\}\right).$$

Because $A \cap \{T = t\} \in \mathcal{F}_t$, it follows that

$$\mathrm{E}\left(\mathrm{E}\left(Y \,\middle|\, \mathcal{F}_T\right); A\right) = \sum_{t\in F} \mathrm{E}\left(M_t; A \cap \{T = t\}\right) = \mathrm{E}\left(M_T\,; A\right) \quad \text{for all } A \in \mathcal{F}_T.$$

Therefore, $M_T = \mathrm{E}(Y \,|\, \mathcal{F}_T)$ a.s. for all simple stopping times T. If T is a bounded stopping time, then we can find simple stopping times $T_n \downarrow T$ [as in Math. 6040]. Therefore, the cadlag version of M satisfies $M_T = \mathrm{E}(Y \,|\, \mathcal{F}_T)$ a.s. for all bounded, hence a.s.-finite, stopping times T.

Problems for Lecture 3

1. Let $\{\mathcal{F}_t\}_{t\geq 0}$ denote a filtration [that satisfies the usual conditions]. Then prove that for all $Y \in L^1(\mathrm{P})$, $t \mapsto \mathrm{E}(Y \,|\, \mathcal{F}_t)$ has a cadlag modification. Use this to prove that if $\{X_t\}_{t\geq 0}$ is a cadlag martingale, then there is a version of modification expectations that leads to: $\mathrm{E}(X_{t+s} \,|\, \mathcal{F}_t) = X_t$ for all $s, t \geq 0$, a.s. [Note the order of the quantifiers.] In other words, there exists one null set off which the preceding martingale identity holds simultaneously for all $s, t \geq 0$.

2. Prove Theorem 8, but for the second part [involving T_A] restrict attention to only sets A that are either open or closed. The result for general Borel sets A is significantly harder to prove, and requires the development of a great deal more measure theory [specifically, Choquet's celebrated "capacitability theorem"]. You can learn about Choquet's theorem, as well as the measurability of T_A for a Borel set A, in Chapters 3 and 4 of the definitive account by Dellacherie and Meyer (1978).

3. Prove that the process X of Example 15 is a.s. discontinuous at T.

4. Prove that if $\sup_{t\geq 0} \mathrm{E}(|X_t|^p) < \infty$ for a martingale X and some $p \in (1\,,\infty)$, then $\lim_{t\to\infty} X_t$ exists a.s. and in $L^p(\mathrm{P})$. [This is false for $p = 1$.]

5 (Change of measure)**.** Let M be a nonnegative cadlag mean-one martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Define

$$\widehat{\mathrm{P}}(A) := \mathrm{E}\,(M_t\,; A) \qquad \text{for all } A \in \mathcal{F}_t.$$

Then, prove that $\hat{\mathrm{P}}$ defines consistently a probability measure on the measurable space $(\Omega\,, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty := \vee_{t\geq 0}\mathcal{F}_t$.

(1) Let $\hat{\mathrm{E}}$ denote the expectation operator for $\hat{\mathrm{P}}$. Then prove that $\hat{\mathrm{E}}(Y) = \mathrm{E}(M_t Y)$ for all nonnegative $\mathcal{F}_t$-measurable random variables Y;

(2) Suppose $\{\mathcal{F}_t\}_{t\geq 0}$ denotes the natural filtration of a d-dimensional Brownian motion $B := \{B_t\}_{t\geq 0}$. Show that for all $\lambda \in \mathbf{R}^d$ fixed, M is a nonnegative cadlag mean-one martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$, where

$$M_t := \exp\left(-\lambda \cdot B_t - \frac{t\|\lambda\|^2}{2}\right) \qquad (t \geq 0);$$

(3) Prove that $X_t := B_t + \lambda t$ defines a d-dimensional Brownian motion on the probability space $(\Omega\,, \mathcal{F}_\infty, \hat{\mathrm{P}})$. That is, if we start with an ordinary Brownian motion B under P, then we obtain a Brownian motion with drift λ if we change our measure to $\hat{\mathrm{P}}$. This is called the *Girsanov and/or Cameron–Martin transformation* of Brownian motion [to Brownian motion with a drift].

Poisson Random Measures

Throughout, let $(S, \mathcal{S}, m)$ denote a sigma-finite measure space with $m(S) > 0$, and $(\Omega, \mathcal{F}, \mathrm{P})$ the underlying probability space.

A construction of Poisson random measures

Definition 1. A *Poisson random measure* Π *with intensity* m is a collection of random variables $\{\Pi(A)\}_{A\in\mathcal{S}}$ with the following properties:

(1) $\Pi(A) = \mathrm{Poiss}(m(A))$ for all $A \in \mathcal{S}$;

(2) If $A_1, \ldots, A_k \in \mathcal{S}$ are disjoint, then $\Pi(A_1), \ldots, \Pi(A_k)$ are independent.

We sometimes write "PRM(m)" in place of "Poisson random measure with intensity m."

Theorem 2. PRM(m) *exists and is a.s. purely atomic.*

Proof. The proof proceeds in two distinct steps.

Step 1. First consider the case that $m(S) < \infty$.

Let $N, X_1, X_2, \ldots$ be a collection of independent random variables with $N = \mathrm{Poiss}(m(S))$, and $\mathrm{P}\{X_j \in A\} = m(A)/m(S)$ for all $j \geq 1$ and $A \in \mathcal{S}$. Define,

$$\Pi(A) := \sum_{j=1}^{N} \mathbb{1}_A(X_j) \qquad \text{for all } A \in \mathcal{S}.$$

Clearly, Π is almost surely a purely-atomic measure with a random number [i.e., N] atoms. Next we compute the finite-dimensional distributions of Π.

If we condition first on N, then we find that for every disjoint $A_1, \ldots, A_k \in \mathcal{S}$ and $\xi_1, \ldots, \xi_k \in \mathbf{R}$,

$$\begin{aligned} \mathrm{E}e^{i\sum_{j=1}^k \xi_j \Pi(A_j)} &= \mathrm{E}\left(\prod_{\ell=1}^N \exp\left\{ i\sum_{j=1}^k \xi_j \mathbb{1}_{A_j}(X_\ell)\right\}\right) \\ &= \mathrm{E}\left[\left(\mathrm{E}\exp\left\{ i\sum_{j=1}^k \xi_j \mathbb{1}_{A_j}(X_1)\right\}\right)^N\right]. \end{aligned}$$

Because the A_j's are disjoint, the indicator function of $(A_1 \cup \cdots \cup A_k)^c$ is equal to $1 - \sum_{j=1}^k \mathbb{1}_{A_j}$, and hence

$$\begin{aligned} \exp\left\{ i\sum_{j=1}^k \xi_j \mathbb{1}_{A_j}(x)\right\} &= \sum_{j=1}^k \mathbb{1}_{A_j}(x)e^{i\xi_j} + 1 - \sum_{j=1}^k \mathbb{1}_{A_j}(x) \\ &= 1 + \sum_{j=1}^k \mathbb{1}_{A_j}(x)\left(e^{i\xi_j} - 1\right) \quad \text{for all } x \in S. \end{aligned}$$

Consequently,

$$\mathrm{E}\exp\left\{ i\sum_{j=1}^k \xi_j \mathbb{1}_{A_j}(X_1)\right\} = 1 + \sum_{j=1}^k \frac{m(A_j)}{m(S)}\left(e^{i\xi_j} - 1\right),$$

and hence,

$$\mathrm{E}\exp\left(i\sum_{j=1}^k \xi_j \Pi(A_j)\right) = \mathrm{E}\left(\left\{1 + \sum_{j=1}^k \frac{m(A_j)}{m(S)}\left(e^{i\xi_j} - 1\right)\right\}^N\right).$$

Now it is easy to check that if $r \in \mathbf{R}$, then $\mathrm{E}(r^N) = \exp\{-m(S)(1-r)\}$. Therefore,

$$\mathrm{E}\exp\left(i\sum_{j=1}^k \xi_j \Pi(A_j)\right) = e^{-\sum_{j=1}^k m(A_j)\left(1 - e^{i\xi_j}\right)}. \tag{1}$$

This proves the result, in the case that $m(S) < \infty$, thanks to the uniqueness of Fourier transforms.

Step 2. In the general case we can find disjoint sets $S_1, S_2, \ldots \in \mathcal{S}$ such that $S = \cup_{k=1}^\infty S_k$ and $m(S_j) < \infty$ for all $j \geq 1$. We can construct independent PRM's $\Pi_1, \Pi_2, \ldots$ as in the preceding, where Π_j is defined solely based on subsets of S_j. Then, define $\Pi(A) := \sum_{j=1}^\infty \Pi_j(A \cap S_j)$ for all

$A \in \mathcal{S}$. Because a sum of independent Poisson random variables has a Poisson law, it follows that $\Pi = \mathrm{PRM}(m)$. □

Theorem 3. *Let $\Pi :=\ \mathrm{PRM}(m)$, and suppose $\varphi : S \to \mathbf{R}^k$ is measurable and satisfies $\int_{\mathbf{R}^d} \|\varphi(x)\|\, m(\mathrm{d}x) < \infty$. Then, $\int_{\mathbf{R}^d} \varphi\, \mathrm{d}\Pi$ is finite a.s., $\mathrm{E} \int_{\mathbf{R}^d} \varphi\, \mathrm{d}\Pi = \int \varphi\, \mathrm{d}m$, and for every $\xi \in \mathbf{R}^k$,*

$$\mathrm{E}e^{i\xi\cdot\int \varphi\, \mathrm{d}\Pi} = \exp\left(-\int \left(1 - e^{i\xi\cdot\varphi(x)}\right) m(\mathrm{d}x) \right). \tag{2}$$

The preceding holds also if m is a finite measure, and φ is measurable. If, in addition, $\int_{\mathbf{R}^d} \|\varphi(x)\|^2\, m(\mathrm{d}x) < \infty$, then also

$$\mathrm{E}\left(\left\| \int_{\mathbf{R}^d} \varphi\, \mathrm{d}\Pi - \int_{\mathbf{R}^d} \varphi\, \mathrm{d}m \right\|^2 \right) \le 2^{k-1} \int_{\mathbf{R}^d} \|\varphi(x)\|^2\, m(\mathrm{d}x).$$

Proof. By a monotone-class argument it suffices to prove the theorem in the case that $\varphi = \sum_{j=1}^n c_j \mathbb{1}_{A_j}$, where $c_1, \ldots, c_n \in \mathbf{R}^k$ and $A_1, \ldots, A_n \in \mathcal{S}$ are disjoint with $m(A_j) < \infty$ for all $j = 1, \ldots, n$. In this case, $\int \varphi\, \mathrm{d}\Pi = \sum_{j=1}^n c_j \Pi(A_j)$ is a finite weighted sum of independent Poisson random variables, where the weights are k-dimensional vectors $c_1, \ldots, c_n$. The formula for the characteristic function of $\int \varphi\, \mathrm{d}\Pi$ follows readily from (1). And the mean of $\int \varphi\, \mathrm{d}\Pi$ is elementary. Finally, if φ^j denotes the jth coordinate of φ, then

$$\mathrm{Var} \int \varphi^j\, \mathrm{d}\Pi = \sum_{i=1}^n c_i^2 \mathrm{Var}\Pi(A_i) = \sum_{i=1}^n c_i^2 m(A_i) = \int |\varphi^j(x)|^2\, m(\mathrm{d}x). \tag{3}$$

The L^2 computation follows from adding the preceding over $j = 1, \ldots, k$, using the basic fact that for all random [and also nonrandom] mean-zero variables $Z_1, \ldots, Z_k \in L^2(\mathrm{P})$,

$$|Z_1 + \cdots + Z_k|^2 \le 2^{k-1} \sum_{j=1}^k |Z_j|^2. \tag{4}$$

Take expectations to find that $\mathrm{Var} \sum_{j=1}^k Z_j \le 2^{k-1} \sum_{j=1}^k \mathrm{Var}(Z_j)$. We can apply this in (3) with $Z_j := \int \varphi^j\, \mathrm{d}\Pi$ to finish. □

The Poisson process on the line

In the context of the present chapter let $S := \mathbf{R}_+$, $\mathcal{S} := \mathcal{B}(\mathbf{R}_+)$, and consider the intensity $m(A) := \lambda|A|$ for all $A \in \mathcal{B}(\mathbf{R}_+)$, where $|\cdots|$ denotes the one-dimensional Lebesgue measure on $(\mathbf{R}_+\,, \mathcal{B}(\mathbf{R}_+))$, and $\lambda > 0$ is a fixed finite constant. If Π denotes the corresponding $\mathrm{PRM}(m)$, then we can define

$$N_t := \Pi((0\,,t]) \qquad \text{for all } t \geq 0.$$

That is, N is the cumulative distribution function of the random measure Π. It follows immediately from Theorem 2 that:

(1) $N_0 = 0$ a.s., and N has i.i.d. increments; and
(2) $N_{t+s} - N_s = \mathrm{Poiss}(\lambda t)$ for all $s, t \geq 0$.

That is, N is a classical Poisson process with intensity parameter λ in the same sense as in Math. 5040.

Problems for Lecture 4

Throughout let N denote a Poisson process with intensity $\lambda \in (0\,,\infty)$.

1. Check that N is cadlag and prove the following:

(1) $N_t - \lambda t$ and $(N_t - \lambda t)^2 - \lambda t$ define mean-zero cadlag martingales;
(2) (The strong law of large numbers) $\lim_{t\to\infty} N_t/t = \lambda$ a.s.

2. Let $\tau_0 := 0$ and then define iteratively for all $k \geq 1$,

$$\tau_k := \inf\left\{s > \tau_{k-1} :\ N_s > N_{s-}\right\}.$$

Prove that $\{\tau_k - \tau_{k-1}\}_{k=1}^{\infty}$ is an i.i.d. sequence of $\mathrm{Exp}(\lambda)$ random variables.

3. Let τ_k be defined as in the previous problem. Prove that $N_{\tau_k} - N_{\tau_k -} = 1$ a.s.

Poisson Point Processes

Let $(S, \mathcal{S}, m)$ be as in the preceding lectures. The goal of this lecture is to learn quickly about Poisson point processes. The book by Kingman (1972) contains a more detailed treatment, as well as a more extensive bibliography.

A construction of Poisson point processes

Definition 1. A stochastic process $\Pi := \{\Pi_t(A)\}_{t\geq 0, A\in\mathcal{S}}$ is a *Poisson point process with intensity* m [written as PPP(m)] if:

(1) For all $t, s \geq 0$, $\Pi_{s+t} - \Pi_s$ is a PRM(tm) that is independent of $\{\Pi_s(A)\}_{A\in\mathcal{S}}$;
(2) $\{\Pi_t(A)\}_{t\geq 0}$ is a Poisson process with intensity $m(A)$ for all $A \in \mathcal{S}$.

Theorem 2. PPP(m) *exists.*

Once you learn why PPP(m) exists, you should convince yourself that the finite-dimensional distributions are determined uniquely.

Proof. The proof is easy: Let

$$S^* := \mathbf{R}_+ \times S, \quad \mathcal{S}^* := \mathcal{B}(S^*), \quad m^* := \text{Leb} \times m, \tag{1}$$

where "Leb" denotes the Lebesgue measure on $\mathbf{R}_+$. Then, let $\Pi^* = \text{PRM}(m^*)$ on $(S^*, \mathcal{S}^*, m^*)$ and define $\Pi_t(A) := \Pi^*((0, t] \times A)$ for all $t \geq 0$ and $A \in \mathcal{S}$. A direct computation or two shows that the process $\{\Pi_t\}_{t\geq 0}$ does the job. □

Theorem 3 on page 23 yields the following byproduct for PPP's.

Proposition 3. *The following are valid:*

(1) *For all fixed* $A \in \mathcal{B}(\mathbf{R}^d)$, $\{\Pi_t(A)\}_{t\geq 0}$ *is a Poisson process with rate* $m(A)$ *[this is true even if* $m(A) = \infty$*];*

(2) *If* $A_1, \ldots, A_k \in \mathcal{B}(\mathbf{R}^d)$ *are disjoint and* $m(A_j) < \infty$ *for all* $1 \leq j \leq k$*, then* $\{\Pi_t(A_1)\}_{t\geq 0}, \ldots, \{\Pi_t(A_k)\}_{t\geq 0}$ *are independent processes;*

(3) *For every fixed* $t \geq 0$, $\Pi_t = \mathrm{PRM}(tm)$ *on* $(\mathbf{R}^d, \mathcal{B}(\mathbf{R}^d))$.

And here is a little more.

Theorem 4. *Let* $\{\Pi_t\}_{t\geq 0} = \mathrm{PPP}(m)$*, and suppose* $\varphi : \mathbf{R}^d \to \mathbf{R}^k$ *is measurable and satisfies* $\int_{\mathbf{R}^d} \|\varphi(x)\|\, m(\mathrm{d}x) < \infty$*. Then,* $\int_{\mathbf{R}^d} \varphi(x)\, \Pi_t(\mathrm{d}x)$ *is finite a.s. and* $t \mapsto \int_{\mathbf{R}^d} \varphi\, \mathrm{d}\Pi_t - t\int_{\mathbf{R}^d} \varphi\, \mathrm{d}m$ *is a k-dimensional mean-zero martingale. And for every* $\xi \in \mathbf{R}^k$,

$$\mathrm{E}e^{i\xi\cdot\int \varphi\, \mathrm{d}\Pi_t} = e^{-t\int_{\mathbf{R}^d}\left(1-e^{i\xi\cdot\varphi(x)}\right) m(\mathrm{d}x)}. \tag{2}$$

The preceding holds also if m is a finite measure, and φ *is measurable. If, in addition,* $\int_{\mathbf{R}^d} \|\varphi(x)\|^2\, m(\mathrm{d}x) < \infty$*, then for all* $T > 0$,

$$\mathrm{E}\left(\sup_{t\in[0,T]} \left\| \int_{\mathbf{R}^d} \varphi\, \mathrm{d}\Pi_t - t\int_{\mathbf{R}^d} \varphi\, \mathrm{d}m \right\|^2\right) \leq 2^{k+1}T \int_{\mathbf{R}^d} \|\varphi(x)\|^2\, m(\mathrm{d}x).$$

Proof. I will describe only the two parts that differ from Theorem 3; the rest follows from Theorem 3. Namely:

(1) It is enough to check the martingale property by working only with φ of the form $\varphi(x) = c\mathbb{1}_A(x)$, where $c \in \mathbf{R}$ and $A \in \mathcal{B}(S)$. In this case, we wish to prove that $X_t := c\Pi_t(A) - ctm(A)$ $(t \geq 0)$ defines a mean zero martingale in the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ generated by $\{\Pi_s\}_{s\geq 0}$. But this follows from Exercise 1 [you have to pay some attention to the filtration though].

(2) It remains to check the L^2 maximal inequality. Without loss of generality we may assume that $k = 1$; otherwise we work with individual coordinates of φ separately, and then add, using (4). Also, it suffices to consider only the case that $\varphi(x) = c\mathbb{1}_A(x)$ as in **(1)**. According to Doob's maximal inequality,

$$\mathrm{E}\left(\sup_{t\in[0,T]} |X_t|^2\right) \leq 4\mathrm{E}(X_T^2) = 4c^2\mathrm{Var}\left(\Pi_T(A)\right) = 4\int_{-\infty}^{\infty} |\varphi(x)|^2\, m(\mathrm{d}x).$$

The result follows. □

Compound Poisson processes

Compound Poisson processes are a generalization of Poisson processes.

Definition 5. Let $X_1, X_2, \ldots$ be i.i.d. random variables in $\mathbf{R}^d$ with common law m. Let N denote an independent Poisson process with rate $\lambda \in (0\,,\infty)$. Then, $C := \{C_t\}_{t\geq 0}$ is a *compound Poisson process* [with parameters m and λ], where

$$C_t := \sum_{j=1}^{N_t} X_j \qquad (t \geq 0),$$

where $\sum_{j=1}^{0} X_j := 0$. If $\mathrm{E}\|X_1\| < \infty$, then

$$C_t - \mathrm{E}C_t = C_t - t\mathrm{E}(X_1)$$

$(t \geq 0)$ is called a *compensated compound Poisson process with parameters m and λ.*

Remark 6. Compound Poisson processes are also [sometimes] called *continuous-time random walks.* □

In the case that $d = 1$ and $m := \delta_1$, $C = N$ is a Poisson process. Note that, in general, C behaves much like N: It jumps at i.i.d. $\mathrm{Exp}(\lambda)$ times; the difference now is that the jump sizes are themselves i.i.d., independent of the jump times, with jumping distribution m.

Proposition 7. *If C is a compound Poisson process with parameters m and λ, then C is cadlag, and has i.i.d. increments with incremental distribution governed by the following characteristic function:*

$$\mathrm{E}e^{i\xi\cdot(C_{t+s}-C_s)} = \exp\left\{-\lambda t \int_{\mathbf{R}^d} \left(1 - e^{i\xi\cdot z}\right)\, m(dz)\right\} \qquad \text{for } \xi \in \mathbf{R}^d,\ s, t \geq 0.$$

Theorem 8 (The strong Markov property)**.** *The following are valid:*

(1) *(The strong Markov property, part 1) For all finite stopping times T [with respect to the natural filtration of C], all nonrandom $t_1, \ldots t_k \geq 0$, and $A_1, \ldots, A_k \in \mathcal{B}(\mathbf{R}_+)$,*

$$\mathrm{P}\left(\bigcap_{j=1}^{k} \left\{C_{T+t_j} - C_T \in A_j\right\} \,\middle|\, \mathcal{F}_T\right) = \mathrm{P}\left(\bigcap_{j=1}^{k} \left\{C_{t_j} \in A_j\right\}\right) \quad \text{a.s.;} \quad \text{and}$$

(2) *(The strong Markov property, part 2) For all finite stopping times T, all nonrandom $t_1, \ldots t_k \geq 0$, and measurable $\varphi_1, \ldots, \varphi_k : \mathbf{R}_+ \to \mathbf{R}_+$,*

$$\mathrm{E}\left(\prod_{j=1}^{k} \varphi_j(C_{T+t_j} - C_T) \,\middle|\, \mathcal{F}_T\right) = \mathrm{E}\left(\prod_{j=1}^{k} \varphi_j(C_{t_j})\right) \quad \text{a.s.;} \quad \text{and}$$

Next we construct compound Poisson processes using Poisson point processes. If $\{\Pi_t\}_{t\geq 0}$ denote a PPP(λm) on $\mathbf{R}^d$ where $\lambda \in (0\,,\infty)$ is fixed, then

$$Y_t := \int_{\mathbf{R}^d} x\, \Pi_t(\mathrm{d}x) \qquad (t \geq 0)$$

defines a cadlag process with i.i.d. increments. Moreover, $\int_{\mathbf{R}^d} \|x\|\, \Pi_t(\mathrm{d}x) < \infty$ a.s. for all $t \geq 0$. This is because each Π_t has at most a finite number of atoms.[1] And (2) applies to yield

$$\mathrm{E}e^{i\xi\cdot(Y_{t+s}-Y_s)} = \exp\left\{-\lambda t \int_{\mathbf{R}^d} \left(1 - e^{i\xi\cdot x}\right)\, m(\mathrm{d}x)\right\}.$$

We can compare this with Exercise 2 to find that Y is a compound Poisson process with parameters m and λ. In order to better understand this construction of compound Poisson processes [using PPP's], note that Π_t has $N_t := \Pi_t(\mathbf{R}^d)$ atoms, where N is a Poisson process of rate λ. If we denote those atoms by $X_1, \ldots, X_{N_t}$, then $\int_{\mathbf{R}^d} x\, \Pi_t(\mathrm{d}x) = \sum_{j=1}^{N_t} X_j$ is compound Poisson, as desired.

Problems for Lecture 5

1. Prove Proposition 3.

2. Prove Proposition 7.

3. Prove Theorem 8.

[1] In fact, $\mathrm{E}\Pi_t(\mathbf{R}^d) = \lambda m(\mathbf{R}^d)t = \lambda t$.

Lévy Processes

Recall that a *Lévy process* $\{X_t\}_{t\ge0}$ on $\mathbf{R}^d$ is a cadlag stochastic process on $\mathbf{R}^d$ such that $X_0 = 0$ and X has i.i.d. increments. We say that X is *continuous* if $t \mapsto X_t$ is continuous. On the other hand, X is *pure jump* if $t \mapsto X_t$ can move only when it jumps [this is not a fully rigorous definition, but will be made rigorous en route the Itô–Lévy construction of Lévy processes].

Definition 1. If X is a Lévy process, then its *tail sigma-algebra* is $\mathcal{T} := \cap_{t\ge0}\sigma(\{X_{r+t} - X_t\}_{r\ge0})$. □

The following is a continuous-time analogue of the Kolmogorov zero-one law for sequences of i.i.d. random variables.

Proposition 2 (Kolmogorov zero-one law)**.** *The tail sigma algebra of a Lévy process is trivial; i.e.,* $\mathrm{P}(A) \in \{0\,,1\}$ *for all* $A \in \mathcal{T}$*.*

The Lévy–Itô construction

The following is the starting point of the classification of Lévy processes, and is also known as the *Lévy–Khintchine formula*; compare with the other Lévy–Khintchine formula (Theorem 6).

Theorem 3 (The Lévy–Khintchine formula; Itô, 1942; Lévy, 1934)**.** *For every Lévy exponent* Ψ *on* $\mathbf{R}^d$ *there exists a Lévy process* X *such that for all* $t \ge 0$ *and* $\xi \in \mathbf{R}^d$*,*

$$\mathrm{E}e^{i\xi\cdot X_t} = e^{-t\Psi(\xi)}. \tag{1}$$

Conversely, if X *is a Lévy process on* $\mathbf{R}^d$ *then* (1) *is valid for a Lévy exponent* Ψ*.*

In words, the collection of all Lévy processes on $\mathbf{R}^d$ is in one-to-one correspondence with the family of all infinitely-divisible laws on $\mathbf{R}^d$.

We saw already that if X is a Lévy process, then X_1 [in fact, X_t for every $t \geq 0$] is infinitely divisible. Therefore, it remains to prove that if Ψ is a Lévy exponent, then there is a Lévy process X whose exponent is Ψ. The proof follows the treatment of Itô (1942), and is divided into two parts.

Isolating the pure-jump part. Let $B := \{B_t\}_{t\geq 0}$ be a d-dimensional Brownian motion, and consider the Gaussian process defined by

$$W_t := \sigma B_t - at. \qquad (t \geq 0).$$

A direct computation shows that $W := \{W_t\}_{t\geq 0}$ is a continuous Lévy process with Lévy exponent

$$\Psi^{(c)}(\xi) = ia'\xi + \frac{1}{2}\|\sigma\xi\|^2 \qquad \text{for all } \xi \in \mathbf{R}^d.$$

[W is a Brownian motion with drift $-a$, where the coordinates of W are possibly correlated, unless σ is diagonal.] Therefore, it suffices to prove the following:

Proposition 4. *There exists a Lévy process Z with exponent*

$$\Psi^{(d)}(\xi) := \int_{\mathbf{R}^d} \left(1 - \mathrm{e}^{i\xi\cdot z} + i(\xi \cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right) m(\mathrm{d}z),$$

for all $\xi \in \mathbf{R}^d$.

Indeed, if this were so, then we could construct W and Z independently from one another, and set

$$X_t = W_t + Z_t \qquad \text{for all } t \geq 0.$$

This proves Theorem 3, since $\Psi = \Psi^{(c)} + \Psi^{(d)}$. In fact, together with Theorem 6, this implies the following:

Theorem 5. *(1) The only continuous Lévy processes are Brownian motions with drift, and; (2) The Gaussian and the non-Gaussian parts of an arbitrary Lévy process are independent from one another.*

Therefore, it suffices to prove Proposition 4.

Proof of Proposition 4. Consider the measurable sets

$$A_{-1} := \left\{z \in \mathbf{R}^d : \ \|z\| \geq 1\right\}, \quad \text{and} \quad A_n := \left\{z \in \mathbf{R}^d : 2^{-n+1} \leq \|z\| < 2^{-n}\right\},$$

as n varies over all nonnegative integers. Now we can define stochastic processes $\{X^{(n)}\}_{n=-1}^{\infty}$ as follows: For all $t \geq 0$,

$$X_t^{(-1)} := \int_{A_{-1}} x\, \Pi_t(\mathrm{d}x), \qquad X_t^{(n)} := \int_{A_n} x\, \Pi_t(\mathrm{d}x) - tm(A_n) \qquad (n \geq 0).$$

Thanks to the construction of Lecture 5 (pp. 26 and on), $\{X^{(n)}\}_{n=-1}^{\infty}$ are independent Lévy processes, and for all $n \geq 0$, $t \geq 0$, and $\xi \in \mathbf{R}^d$,

$$\mathrm{E}e^{i\xi\cdot X_t^{(n)}} = \exp\left\{-t\int_{A_n}\left(1 - e^{i\xi\cdot z} + i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right) m(dz)\right\}.$$

Moreover, $X^{(-1)}$ is a compound Poisson process with parameters $m(\bullet\cap A_{-1})/m(A_{-1})$ and $\lambda = m(A_{-1})$, for all $n \geq 0$, $X^{(n)}$ is a compensated compound Poisson process with parameters $m(\bullet \cap A_n)/m(A_n)$ and $\lambda = m(A_n)$.

Now $Y_t^{(n)} := \sum_{k=0}^{n} X_t^{(k)}$ defines a Lévy process with exponent

$$\psi_n(\xi) := \int_{1>\|z\|\geq 2^{-n+1}}\left(1 - e^{i\xi\cdot z} + i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right) m(dz),$$

valid for all $\xi \in \mathbf{R}^d$ and $n \geq 1$. Our goal is to prove that there exists a process Y such that for all nonrandom $T > 0$,

$$\sup_{t\in[0,T]}\left\|Y_t^{(n)} - Y_t\right\| \to 0 \qquad \text{in } L^2(\mathrm{P}). \tag{2}$$

Because $Y^{(n)}$ is cadlag for all n, uniform convergence shows that Y is cadlag for all n. In fact, the jumps of $Y^{(n+1)}$ contain those of $Y^{(n)}$, and this proves that Y is pure jump. And because the finite-dimensional distributions of $Y^{(n)}$ converge to those of Y, it follows then that Y is a Lévy process, independent of $X^{(-1)}$, and with characteristic exponent

$$\psi_\infty(\xi) = \lim_{n\to\infty}\psi_n(\xi) = \int_{1>\|z\|}\left(1 - e^{i\xi\cdot z} + i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)\right) m(dz).$$

[The formula for the limit holds by the dominated convergence theorem.] Sums of independent Lévy processes are themselves Lévy. And their exponents add. Therefore, $X_t^{(-1)} + Y_t$ is Lévy with exponent $\Psi^{(d)}$.

It remains to prove the existence of Y. Let us choose and fix some $T > 0$, and note that for all $j, k \geq 1$ and $t \geq 0$,

$$Y_t^{(n+k)} - Y_t^{(n)} = \sum_{j=k+1}^{n+k}\left(\int_{A_j} x\,\Pi_t(dx) - tm(A_j)\right),$$

and the summands are independent because the A_j's are disjoint. Since the left-hand side has mean zero, it follows that

$$\begin{aligned}\mathrm{E}\left(\left\|Y_t^{(n+k)} - Y_t^{(n)}\right\|^2\right) &= \sum_{j=k+1}^{n+k}\mathrm{E}\left(\left\|\int_{A_j} x\,\Pi_t(dx) - tm(A_j)\right\|^2\right)\\ &\leq 2^{d-1}t\sum_{j=k+1}^{n+k}\int_{A_j}\|x\|^2\,m(dx) = 2^{d-1}t\int_{\cup_{j=k+1}^{n+k}A_j}\|x\|^2\,m(dx);\end{aligned}$$

see Theorem 3. Every one-dimensional mean-zero Lévy process is a mean-zero martingale [in the case of Brownian motion we have seen this in Math. 6040; the reasoning in the general case is exactly the same]. Therefore, $Y^{(n+k)} - Y^{(n)}$ is a mean-zero cadlag martingale (coordinatewise). Doob's maximal inequality tells us that

$$\mathrm{E}\left(\sup_{t\in[0,T]}\left\|Y_t^{(n+k)} - Y_t^{(n)}\right\|^2\right) \leq 2^{d+1}T\int_{2^{-k}\leq\|z\|<2^{n-k+1}} \|x\|^2\, m(\mathrm{d}x).$$

This and the definition of a Lévy measure (p. 3) together imply (2), whence the result. □

Problems for Lecture 6

1. Prove the Kolmogorov 0-1 law (page 29).

2. Prove that every Lévy process X on $\mathbf{R}^d$ is a strong Markov process. That is, for all finite stopping times T [in the natural filtration of X], $t_1, \ldots, t_k \geq 0$, and $A_1, \ldots, A_k \in \mathcal{B}(\mathbf{R}^d)$,

$$\mathrm{P}\left(\bigcap_{j=1}^{k}\left\{X_{T+t_j} - X_T \in A_j\right\} \,\middle|\, \mathcal{F}_T\right) = \mathrm{P}\left(\bigcap_{j=1}^{k}\left\{X_{t_j} \in A_j\right\}\right) \quad \text{a.s.}$$

(Hint: Follow the Math. 6040 proof of the strong Markov property of Brownian motion.)

Structure Theory

The Lévy–Itô decomposition

The Lévy–Itô proof of the Lévy–Khintchine formula (Theorem 3, page 29, and the proof of the part that we have not discussed) has also consequences that reach beyond issues of existence etc. Indeed, that proof shows among other things that if X is a Lévy process with triple (a, σ, m) and

$$(\Delta X)_t := X_t - X_{t-} \qquad (t \geq 0),$$

then

$$\Pi_t(A) := \sum_{s\in[0,t]} \mathbb{1}_{\{(\Delta X)_s \in A\}} \qquad (t \geq 0,\, A \in \mathcal{B}(\mathbf{R}^d))$$

defines a PPP(m). And we have the process-wise decomposition

$$X_t = W_t + C_t + D_t \qquad (t \geq 0), \tag{1}$$

called the *Lévy–Itô decomposition* of X, where:

- $W_t := \sigma B_t - at$, where B is standard Brownian motion on $\mathbf{R}^d$;
- C is a compound Poisson process with parameters $m(\bullet \cap A_{-1})/m(A_{-1})$ and $\lambda = m(A_{-1})$, where $A_{-1} := \{z \in \mathbf{R}^d : \|z\| \geq 1\}$. And $\|C_t - C_{t-}\| \geq 1$ for all $t \geq 0$;
- D is a mean-zero Lévy process that is an L^2 martingale and satisfies:
  (a) $\|D_t - D_{t-}\| \leq 1$ for all $t \geq 0$ [a.s.]; and
  (b) For all $T > 0$,

$$\mathrm{E}\left(\sup_{t\in[0,T]} \|D_t\|^2\right) \leq 2^{d+1} T m(A_{-1}^{\mathrm{c}});$$

- B, C, and D are independent processes.

The preceding decomposition teaches us a great deal about the behavior of Lévy processes. Next we make some remarks along these directions.

The Gaussian Part

The following gives us an interpretation of the matrix σ: X has a nontrivial Gaussian component if and only if $\sigma'\sigma$ has a nontrivial spectrum.

Theorem 1. *We have*

$$\limsup_{\|\xi\|\to\infty} \frac{\mathrm{Re}\Psi(\xi)}{\|\xi\|^2} = \limsup_{\|\xi\|\to\infty} \frac{|\Psi(\xi)|}{\|\xi\|^2} = \max_{z\in\mathbf{R}^d\setminus\{0\}} \left(\frac{z'\sigma'\sigma z}{\|z\|^2}\right).$$

Consequently, X has a nontrivial Gaussian part iff

$$\limsup_{\|\xi\|\to\infty} \frac{\mathrm{Re}\Psi(\xi)}{\|\xi\|^2} = \limsup_{\|\xi\|\to\infty} \frac{|\Psi(\xi)|}{\|\xi\|^2} > 0.$$

Remark 2. $\lambda_{\max} := \max_{a\in\mathbf{R}^d\setminus\{0\}}(a'\sigma'\sigma a/\|a\|^2)$ is none other than the largest eigenvalue of σ. And of course $\lambda_{\max} \geq 0$, since σ is nonnegative definite. □

Proof of Theorem 1. By the Lévy–Khintchine formula,

$$\left|\Psi(\xi) - \frac{1}{2}\|\sigma\xi\|^2\right| \leq |a'\xi| + \int_{\mathbf{R}^d} \left|1 - \mathrm{e}^{-i\xi\cdot z} + i(\xi\cdot z)\right| m(\mathrm{d}z),$$

$$\mathrm{Re}\Psi(\xi) = \frac{1}{2}\|\sigma\xi\|^2 + \int_{\mathbf{R}^d} (1 - \cos(\xi\cdot z))\, m(\mathrm{d}z).$$

Therefore, it suffices to prove that

$$\lim_{\|\xi\|\to\infty} \int_{\mathbf{R}^d} \left|\frac{1 - \mathrm{e}^{-i\xi\cdot z} + i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)}{\|\xi\|^2}\right| m(\mathrm{d}z) = 0, \tag{2}$$

$$\lim_{\|\xi\|\to\infty} \int_{\mathbf{R}^d} \left(\frac{1 - \cos(\xi\cdot z)}{\|\xi\|^2}\right) m(\mathrm{d}z) = 0. \tag{3}$$

We saw earlier that

$$\left|\frac{1 - \mathrm{e}^{-i\xi\cdot z} + i(\xi\cdot z)\mathbb{1}_{(0,1)}(\|z\|)}{\|\xi\|^2}\right| \leq \left(\|z\|^2 \wedge \frac{1}{\|\xi\|^2}\right). \tag{4}$$

This and the dominated convergence theorem together imply (2). And (3) is proved similarly. □

The Compound Poisson Part

Recall from (7) on page 35 that if $\int_{\mathbf{R}^d}(1 \wedge \|x\|)\, m(\mathrm{d}x) < \infty$, then we have the following form of the Lévy–Khintchine formula:

$$\Psi(\xi) = i(b \cdot \xi) + \frac{1}{2}\|\sigma\xi\|^2 + \int_{\mathbf{R}^d} \left(1 - \mathrm{e}^{i\xi\cdot z}\right) m(\mathrm{d}z), \tag{5}$$

where

$$b := a - \int_{\|z\|<1} z\, m(\mathrm{d}z). \tag{6}$$

It follows readily from structure theory that in this case, that is when $\int_{\mathbf{R}^d}(1 \wedge \|x\|)\, m(\mathrm{d}x) < \infty$, we have the decomposition

$$X_t = \sigma B_t - bt + C_t, \tag{7}$$

where B is a standard Brownian motion, and C is an independent pure-jump process. If, in addition, m is a finite measure [this condition implies that $\int_{\mathbf{R}^d}(1 \wedge \|x\|)\, m(\mathrm{d}x)$ is finite] then we can recognise $\int_{\mathbf{R}^d}(1 - \mathrm{e}^{i\xi\cdot z})\, m(\mathrm{d}z)$ as the characteristic exponent of a compound Poisson process with parameter $\lambda := m(\mathbf{R}^d)$ and jump distribution $m(\bullet)/m(\mathbf{R}^d)$. This proves the first half of the following theorem.

Theorem 3. *Suppose $m(\mathbf{R}^d) < \infty$, and $b = 0$ and $\sigma = 0$. Then X is compound Poisson. Conversely, if X is compound Poisson then $\sigma = 0$, $m(\mathbf{R}^d) < \infty$, and $b = 0$.*

Proof. We have proved already the first half. Therefore, we assume from here on that X is compound Poisson.

Because X is compound Poisson, it is a pure-jump process. Therefore, (7) tells us readily that X_t must equal C_t; i.e., $\sigma = 0$ and $b = 0$. It remains to demonstrate that $m(\mathbf{R}^d) < \infty$.

Define $A_\epsilon := \{z \in \mathbf{R}^d : \|z\| > \epsilon\}$, where $\epsilon > 0$ is arbitrary. Note that the total number of jumps, during the time interval $[0\,, t]$, whose magnitude is in A_ϵ is $\Pi_t(A_\epsilon) = \sum_{s\in[0,t]} \mathbb{1}_{\{\|(\Delta X)_s\|>\epsilon\}}$, which—by properties of Poisson processes—has mean and variance both equal to $tm(A_\epsilon)$. Thanks to Chebyshev's inequality,

$$\begin{aligned} \mathrm{P}\left\{\Pi_t(A_\epsilon) \le \frac{1}{2}tm(A_\epsilon)\right\} &\le \mathrm{P}\left\{|\Pi_t(A_\epsilon) - \mathrm{E}\Pi_t(A_\epsilon)| > \frac{1}{2}tm(A_\epsilon)\right\} \\ &\le \frac{4}{tm(A_\epsilon)}. \end{aligned}$$

Now, if $m(\mathbf{R}^d) = \infty$ then $m(A_\epsilon) \to \infty$ as $\epsilon \to 0$. It follows readily from the preceding that $\sup_{\epsilon>0} \Pi_t(A_\epsilon) = \infty$. But that supremum is the total number of jumps of X during the time interval $[0\,, t]$, and this would contradict the assumption that X is compound Poisson. Therefore, $m(\mathbf{R}^d) < \infty$. $\square$

One can also characterize when X is compound Poisson purely in terms of the exponent Ψ.

Theorem 4. *X is compound Poisson if and only if Ψ is bounded.*

Proof. Suppose first that X is compound Poisson. Then m is a finite measure, $b = 0$ and $\sigma = 0$ (Theorem 3), and

$$\sup_{\xi\in\mathbf{R}^d} |\Psi(\xi)| \le \sup_{\xi\in\mathbf{R}^d} \left| \int_{\mathbf{R}^d} \left(1 - e^{i\xi\cdot z}\right) m(dz) \right| \le 2m(\mathbf{R}^d) < \infty,$$

thanks to (5). This proves that if X is compound Poisson then Ψ is bounded.

We now assume that Ψ is bounded, and prove the converse.

Note that

$$|\Psi(\xi)| \ge \mathrm{Re}\Psi(\xi) \ge \int_{\mathbf{R}^d} (1 - \cos(\xi\cdot z))\, m(dz).$$

Now let us introduce a d-dimensional standard Brownian motion $Y := \{Y_t\}_{t\ge 0}$, and note that

$$\sup_{\xi\in\mathbf{R}^d} |\Psi(\xi)| \ge \mathrm{E}\mathrm{Re}\Psi(Y_t) \ge \int_{\mathbf{R}^d} \left(1 - e^{-t\|z\|^2/2}\right) m(dz).$$

Send $t \to \infty$ and apply Fatou's lemma to deduce that m is finite. And Theorem 1 implies that $\sigma = 0$. It remains to prove that $b = 0$. But it follows readily from (5) and the dominated convergence theorem that

$$\limsup_{\|\xi\|\to\infty} \frac{|\Psi(\xi)|}{\|\xi\|} = \limsup_{\|\xi\|\to\infty} \frac{|b\cdot\xi|}{\|\xi\|} = \max_{1\le j\le d} |b_j|.$$

Therefore, the boundedness of Ψ implies that $b = 0$, as asserted. □

Theorem 5. *If X is a Lévy process on $\mathbf{R}^d$ with Lévy measure m then for all $\gamma > 0$ the following are equivalent:*

(1) $\int_{\mathbf{R}^d} (1 \wedge \|z\|^\gamma)\, m(dz) < \infty$; *and*
(2) $V_t(\gamma) := \sum_{s\le t} \|(\Delta X)_s\|^\gamma < \infty$ *for one hence all* $t \ge 0$ *a.s.*

Consequently, the random function $t \mapsto X_t$ has bounded variation [a.s.] iff $\sigma = 0$ and $\int_{\mathbf{R}^d} (1 \wedge \|x\|)\, m(dx) < \infty$.

A great deal more is known about the variations of a Lévy process (Millar, 1971; Monroe, 1972).

Before we set out to prove Theorem 5, let us study some examples.

Example 6. Poisson processes are of bounded variation. Indeed, if X is a Poisson process with intensity $\lambda \in (0\,,\infty)$, then $a = \sigma = 0$, $m = \lambda\delta_{\{1\}}$, and $\int_{-\infty}^{\infty} (1 \wedge |x|)\, m(dx) = \lambda$. □

Example 7. Let X denote an isotropic stable process with index $\alpha \in (0\,,2)$. Then, $a = 0$, $\sigma = 0$, and $m(\mathrm{d}x) \propto \|x\|^{-(d+\alpha)}\,\mathrm{d}x$; see Lemma 1. Thus, we can integrate in spherical coordinates to find that

$$\int_{\mathbf{R}^d} (1 \wedge \|x\|)\, m(\mathrm{d}x) \propto \int_0^\infty (1 \wedge r) \frac{\mathrm{d}r}{r^{\alpha+1}}$$

is finite if and only if $\alpha \in (0\,,1)$. In particular, the isotropic Cauchy process [that is, $\alpha = 1$ here] is the borderline case which has unbounded variation. Note also that this example shows that there exist Lévy processes which are *not* compound Poisson [$m(\mathbf{R}^d) = \infty$], and yet have bounded variation almost surely. □

Proof of Theorem 5. First, let us handle the matter of bounded variation, assuming the first portion of the theorem. Then we address the first portion.

We apply (1) to see that the compound Poisson component of X does not contribute to the question of whether or not X has bounded variation. That is, we can assume, without loss of generality, that $m\{z \in \mathbf{R}^d : \|z\| > 1\} = 0$. Since Brownian motion has quadratic variation, it has infinite variation a.s. This implies that unless $\sigma = 0$, X cannot have bounded variation. And since $t \mapsto -at$ [in the process W] has bounded variation, a too does not contribute to our problem. In summary, we need only study the case that $a = 0$, $\sigma = 0$, and $m\{z \in \mathbf{R}^d : \|z\| > 1\} = 0$. In that case, X is pure jump and the first portion of the theorem does the job. It remains to prove the equivalence of (1) and (2).

The convergence of $V_t(\gamma)$ does not depend on the continuous [Gaussian] component W in the decomposition (1). Therefore, we can assume without loss of generality that $a = 0$ and $\sigma = 0$. Also, since C is compound Poisson, $\sum_{s\in[0,t]} \|(\Delta C)_s\|^\gamma$ is a finite sum of a.s.-finite random variables. Therefore, we can assume without loss of generality that $m\{z \in \mathbf{R}^d : \|z\| > 1\} = 0$, whence $C_t \equiv 0$ for all t. Thus, $\int_{\mathbf{R}^d}(1 \wedge \|z\|^\gamma)\, m(\mathrm{d}z) = \int_{\mathbf{R}^d} \|z\|^\gamma\, m(\mathrm{d}z)$. If this integral is finite, then by Theorem 4 [page 26],

$$\mathrm{E} \sum_{s\in[0,t]} \|(\Delta X)_s\|^\gamma = \mathrm{E} \int \|x\|^\gamma\, \Pi_t(\mathrm{d}x) = t \int_{\mathbf{R}^d} \|x\|^\gamma\, m(\mathrm{d}x) < \infty.$$

Therefore, for every $t \geq 0$ there exists a null set off which $V_t(\gamma) < \infty$. Since $t \mapsto V_t(\gamma)$ is nondecreasing, we can choose the null set to be independent of t, and this shows that (1)$\Longrightarrow$(2).

Conversely, if (1) fails, then $\int_{\|x\|\leq 1} \|x\|^\gamma\, m(\mathrm{d}x) = \infty$, under the present reductions. Let $A_\epsilon := \{z \in \mathbf{R}^d : \epsilon < \|z\| \leq 1\}$ and note that $\Pi_t(A_\epsilon) = \sum_{s\leq t} \mathbb{1}_{\{\epsilon < \|(\Delta X)_s\| \leq 1\}}$ and $V_t(\gamma) \geq V_t(\gamma\,,\epsilon) := \int_{A_\epsilon} \|x\|^\gamma\, \Pi_t(\mathrm{d}x)$. From Theorem

4 [page 26] we know that $\mathrm{E}V_t(\gamma\,,\epsilon) = t\int_{A_\epsilon} \|x\|^\gamma\, m(\mathrm{d}x)$ and $\mathrm{Var}V_t(\gamma\,,\epsilon) \le \mathrm{E}V_t(\gamma\,,\epsilon)$ [in fact this is an identity]. Therefore, Chebyshev's inequality shows that

$$\mathrm{P}\left\{V_t(\gamma) \le \frac{t}{2}\int_{A_\epsilon} \|x\|^\gamma\, m(\mathrm{d}x)\right\} \le 4\left(t\int_{A_\epsilon} \|x\|^\gamma\, m(\mathrm{d}x)\right)^{-1}.$$

This shows that if $\int \|x\|^\gamma\, m(\mathrm{d}x) = \infty$ then $V_t(\gamma) = \infty$ a.s. for all $t \ge 0$. By monotonicity, this implies (2). □

Conclusion: To us, the most interesting Lévy processes are those that have a pure-jump component with unbounded variation. The rest are basically Brownian motion with drift, plus a compound Poisson process.

The preceding conclusion is highlighted in the following section.

A strong law of large numbers

Recall that if $x_1, x_2, \ldots$ are i.i.d. random variables with values in $\mathbf{R}^d$, and if $s_n := x_1 + \cdots + x_n$, then:

(1) $\limsup_{n\to\infty} \|s_n/n\|$ and $\liminf_{n\to\infty} \|s_n/n\|$ are a.s. constants;

(2) $\limsup_{n\to\infty} \|s_n/n\| < \infty$ [a.s.] if and only if $\mathrm{E}\|x_1\| < \infty$; and

(3) If and when $\mathrm{E}\|x_1\| < \infty$, then we have $\lim_{n\to\infty}(s_n/n) = \mathrm{E}[x_1]$ a.s.

The preceding is Kolmogorov's strong law of large numbers.

Since we may think of Lévy processes as continuous-time random walks, we might wish to know if the strong law of large numbers has a continuous-time analogue. It does, as the following shows.

Theorem 8. *Let X be a Lévy process in $\mathbf{R}^d$ with triple $(a\,,\sigma\,,m)$. Then, $p := \mathrm{P}\{\limsup_{t\to\infty} \|X_t/t\| < \infty\}$ is zero or one. And $p = 1$ if and only if $\int_{\|x\|\ge 1} \|x\|\, m(\mathrm{d}x) < \infty$. And when $p = 1$,*

$$\lim_{t\to\infty} \frac{X_t}{t} = -a + \int_{\|x\|\ge 1} x\, m(\mathrm{d}x) \quad \text{a.s.}$$

The preceding yields information about the global [i.e., large time] growth of a Lévy process. We will soon see an example of a "local" version [i.e., one that is valid for small t] in Theorem 5 on page 45. For further information on such properties see the paper by Štatland (1965).

Proof. If $\tau > 0$ is fixed and nonrandom, then

$$\limsup_{t\to\infty} \left\|\frac{X_t}{t}\right\| = \limsup_{t\to\infty} \left\|\frac{X_t - X_\tau}{t}\right\| \qquad \text{a.s.}$$

Therefore, $p \in \{0, 1\}$, by the Kolmogorov zero-one law (Proposition 2 on page 29).

We use structure theorem (1) and write $X_t = \sigma B_t - at + C_t + D_t$. By the strong law of large numbers for Brownian motion, $B_t/t \to 0$ a.s. as $t \to \infty$. And

$$\mathrm{E}\left(\sup_{s\in[2^n,2^{n+1}]} \frac{\|D_s\|}{s}\right) \le 2^{-n-1}\sqrt{\mathrm{E}\left(\sup_{s\in[2^n,2^{n+1}]} \|D_s\|^2\right)} = O(2^{-n/2}).$$

Therefore, the Borel–Cantelli lemma implies that $D_t = o(t)$ a.s. as $t \to \infty$. It therefore suffices to prove that $\limsup_{\to\infty} \|C_t/t\| < \infty$ a.s. if and only if $\int_{\|x\|\ge 1} x\, m(\mathrm{d}x) < \infty$, and in the case that the latter condition holds,

$$\lim_{t\to\infty} \frac{C_t}{t} = \int_{\|x\|\ge 1} x\, m(\mathrm{d}x) \quad \text{a.s.} \tag{8}$$

But this is not hard, because we can realize the process C_t as $\sum_{j=1}^{N_t} x_j$, where: (i) $\{N_t\}_{t\ge 0}$ is a Poisson process with rate $\lambda = m\{\|z\| \ge 1\}$; (ii) $x_1, x_2, \ldots$ are i.i.d. with distribution $m(\bullet \cap \{\|z\| \ge 1\})/m\{\|z\| \ge 1\}$; and the x_j's and N are independent. Because $N_t \to \infty$ a.s., we can condition on the process N and apply the Kolmogorov strong law of large numbers to deduce that $\lim_{t\to\infty} \|C_t/N_t\| < \infty$ a.s. if and only if

$$\mathrm{E}\|x_1\| = \frac{1}{m\{\|z\| \ge 1\}} \int_{\|z\|\ge 1} \|z\|\, m(\mathrm{d}z) < \infty.$$

And if the latter holds, then

$$\lim_{t\to\infty} \frac{C_t}{N_t} = \mathrm{E}x_1 = \frac{1}{m\{\|z\| \ge 1\}} \int_{\|z\|\ge 1} z\, m(\mathrm{d}z) < \infty.$$

The result follows because $N_t/t \to \lambda = m\{\|z\| \ge 1\}$ a.s. as $t \to \infty$. □

Symmetry and isotropy

Definition 9. We say that a Lévy process X is *symmetric* if Ψ is real values; X is *isotropic* if there exists a nonrandom $(d \times d)$ orthogonal matrix O such that $\{OX_t\}_{t\ge 0}$ has the same distribution as X. □

Lemma 10. *The process $-X$ is a Lévy process with exponent $\overline{\Psi}$. Therefore, X is symmetric if and only if the finite-dimensional distributions of X and $-X$ are the same. On the other hand, X is isotropic if and only if Ψ is an isotropic—or radial–function; that is, $\Psi(\xi)$ depends on $\xi \in \mathbf{R}^d$ only on $\|\xi\|$.*

The preceding implies, among many other things, that every isotropic Lévy process is symmetric. For one-dimensional Lévy processes, the two notions of symmetry and isotropy are, of course, the same. But the converse is not true in general, viz.,

Example 11 (Processes with stable components)**.** Let $X^1, \ldots, X^d$ denote d independent one-dimensional symmetric stable processes with respective indices $\alpha_1, \ldots, \alpha_d \in (0\,,2]$, where $d \geq 2$. Consider the d-dimensional Lévy process defined by $X_t := (X_t^1\,, \ldots, X_t^d)$ and note that $\Psi(\xi) = \sum_{j=1}^d |\xi_j|^\alpha$ for all $\xi \in \mathbf{R}^d$. Consequently, even though X is manifestly symmetric, it is isotropic if and only if $\alpha = 2$ [i.e., Brownian motion]. □

And just to be sure, let me remind you of an interesting example of an asymmetric Lévy process.

Example 12 (The asymmetric Cauchy process on $\mathbf{R}$)**.** Let X be a Lévy process with exponent $\Psi(\xi) = -|\xi| - i\theta\xi\log|\xi|$, where $|\theta| \leq 2/\pi$ and $0\log|0| := \infty$. The process X is a *Cauchy process* on the line. It is called *symmetric* if $\theta = 0$, *completely asymmetric* if $|\theta| = 2/\pi$, and *asymmetric* otherwise. □

Problems for Lecture 7

1. Prove that every Lévy process X on $\mathbf{R}^d$ is a strong Markov process. That is, for all finite stopping times T [in the natural filtration of X], $t_1, \ldots, t_k \geq 0$, and $A_1, \ldots, A_k \in \mathcal{B}(\mathbf{R}^d)$,

$$\mathrm{P}\left(\bigcap_{j=1}^k \left\{X_{T+t_j} - X_T \in A_j\right\} \,\middle|\, \mathcal{F}_T\right) = \mathrm{P}\left(\bigcap_{j=1}^k \left\{X_{t_j} \in A_j\right\}\right) \quad \text{a.s.}$$

(Hint: Follow the Math. 6040 proof of the strong Markov property of Brownian motion.)

2. Consider the degenerate two-dimensional Brownian motion $X_t := (B_t\,, 0)$ where B is Brownian motion in dimension one. Compute Ψ, and verify that

$$0 = \liminf_{\|\xi\|\to\infty} \frac{|\Psi(\xi)|}{\|\xi\|^2} < \limsup_{\|\xi\|\to\infty} \frac{|\Psi(\xi)|}{\|\xi\|^2} = \frac{1}{2}.$$

Thus, it is possible that the lim sup in Theorem 1 is not a bona fide limit.

3. Derive Lemma 10.

4. Let X denote a Lévy process on $\mathbf{R}^d$ with stable components that respectively have indices $\alpha_1, \ldots, \alpha_d$. Find a necessary and sufficient condition for X to have bounded-variation paths.

5 (The Skorohod–Ottaviani inequality)**.** Let $Y_1, Y_2, \ldots$ be a sequence of independent random variables with values in $\mathbf{R}^d$, and define $S_k := Y_1 + \cdots + Y_k$ for all $k \geq 1$. Prove that for all $n \geq 1$ and $\lambda > 0$,

$$\min_{1\le j\le n} \mathrm{P}\left\{\|S_n - S_j\| \le \lambda\right\} \cdot \mathrm{P}\left\{\max_{1\le j\le n} \|S_j\| \ge 2\lambda\right\} \le \mathrm{P}\left\{\|S_n\| \ge \lambda\right\}.$$

Conclude that if X is a Lévy process on $\mathbf{R}^d$, then for all $t, \lambda > 0$,

$$\inf_{0\le s\le t} \mathrm{P}\left\{\|X_s\| \le \lambda\right\} \cdot \mathrm{P}\left\{\sup_{s\in[0,t]} \|X_s\| \ge 2\lambda\right\} \le \mathrm{P}\left\{\|X_t\| \ge \lambda\right\}.$$

(Hint: Consider the smallest integer j such that $\|S_j\| \geq \lambda$.)

6. Suppose $f : (0\,,\infty) \to \mathbf{R}_+$ is increasing and measurable with $f(0) = 0$. Suppose also that X is a d-dimensional Lévy process such that $\int_0^1 t^{-1}\mathrm{P}\{\|X_t\| > f(t)\}\,\mathrm{d}t < \infty$.

(1) Prove that

$$\sum_{n=1}^{\infty} \int_0^1 \mathrm{P}\left\{\|X_{a^n}\| > f(a^n)\right\} \log(1/a)\,\mathrm{d}a < \infty.$$

Conclude that $\sum_{n=1}^{\infty} \mathrm{P}\{\|X_{a^n}\| > f(a^n)\} < \infty$ for almost every $a \in (0\,,1)$.

(2) Use the Skorohod–Ottaviani inequality (previous problem) to prove that

$$\limsup_{t\downarrow 0} \frac{\|X_t\|}{f(t)} < \infty \qquad \text{a.s.}$$

(Khintchine, 1939).

Subordinators

Definition 1. A *subordinator* $T := \{T_t\}_{t\geq 0}$ is a one-dimensional Lévy process such that $t \mapsto T_t$ is nondecreasing. □

Since $T_0 = 0$ all subordinators take nonnegative values only.

Proposition 2. *A Lévy process T on $\mathbf{R}$ is a subordinator iff its Lévy triple has the form $(a\,,0\,,m)$, where $m((-\infty\,,0)) = 0$ and $\int_0^\infty (1 \wedge x)\, m(\mathrm{d}x) < \infty$.*

Proof. T is a subordinator iff its paths are monotone and its jumps are nonnegative. The monotonicity of the paths is equivalent to their bounded variation; i.e., $\sigma = 0$ and $\int_{-\infty}^{\infty}(1 \wedge |x|)\, m(\mathrm{d}x) < \infty$ [Theorem 5 on page 36]. And the jumps are nonnegative iff $J_t := \sum_{0\leq s\leq t} \mathbb{1}_{(-\infty,0)}((\Delta T)_s) = 0$ for all $t \geq 0$. But $J_t = 0$ iff $\mathrm{E}(J_t) = 0$, by basic facts about PPP's, and $E(J_t) = tm((-\infty\,,0))$. □

One can think of Lévy processes as an extension of the classical family of random walks [sums of i.i.d. random variables]. In the same way, we think of subordinators as an extension of random walks that have nonnegative increments. Such objects are the central pieces of renewal theory, regenerative times, etc., as the following example might suggest.

Example 3. Suppose $Z_1, Z_2, \ldots$ are i.i.d. integer-valued random variables, and consider the random walk defined by $S_0 := 0$ and

$$S_n := Z_1 + \cdots + Z_n \qquad \text{for } n \geq 1.$$

Let τ_n denote the nth return time of S to zero; that is, $\tau_0 := 0$, and iteratively define

$$\tau_n := \inf\{j > \tau_{n-1} : S_j = 0\}.$$

If we define $\inf\varnothing := \infty$, then we find that the τ_n's are increasing stopping times. But also, because S has the strong Markov property [in the sense of Math. 6040], $\{\tau_j - \tau_{j-1}\}_{j=1}^{\infty}$ is a sequence of i.i.d. nonnegative random variables. Thus, $\tau_n := \sum_{j=1}^{n}(\tau_j - \tau_{j-1})$ is a nondecreasing random walk. In other words, the successive return times of S to zero form a discrete-time analogue of a subordinator.

Laplace exponents

It follows that if T is a subordinator, then there exists $a \in \mathbf{R}$ and a Lévy measure m, supported on $(0\,,\infty)$, such that $\int_0^1 x\, m(\mathrm{d}x) < \infty$ and the Lévy exponent of X is given by

$$\Psi(\xi) = -ia\xi + \int_0^\infty \left[1 - \mathrm{e}^{iz\xi} + i(z\xi)\mathbb{1}_{(0,1)}(z)\right] m(\mathrm{d}z) \qquad \text{for all } \xi \in \mathbf{R}.$$

In the present setting, $\int_0^\infty (z\xi)\mathbb{1}_{(0,1)}(z)\, m(\mathrm{d}z) < \infty$. Therefore, we can write

$$\mathrm{E}\exp(i\xi T_t) = \exp\left\{ib\xi t - t\int_0^\infty \left(1 - \mathrm{e}^{iz\xi}\right) m(\mathrm{d}z)\right\} \qquad (\xi \in \mathbf{R}), \tag{1}$$

where

$$b := a - \int_0^1 x\, m(\mathrm{d}x).$$

Because $\mathrm{P}\{X_t \geq 0\} = 1$, both sides of (1) are analytic functions of ξ for ξ in $\{z \in \mathbf{C} : \mathrm{Re}\, z \geq 0\}$. In particular, we consider (1) for $\xi := i\lambda$, where $\lambda \geq 0$, and obtain the following: For all $t, \lambda \geq 0$,

$$\mathrm{E}\mathrm{e}^{-\lambda T_t} = \mathrm{e}^{-t\Phi(\lambda)}, \quad \text{where} \quad \Phi(\lambda) := b\lambda + \int_0^\infty \left(1 - \mathrm{e}^{-\lambda z}\right) m(\mathrm{d}z). \tag{2}$$

By the uniqueness theorem for Laplace transforms, the function Φ determines the law of T uniquely.

Definition 4. The function Φ is called the *Laplace exponent* of the subordinator T. The constant b is called the *drift* of T. □

Most people who study subordinators prefer to work with the Laplace exponent rather than the Lévy exponent because the latter yields a more natural "parametrization." The following is an instance of this; it also explains why b is called the *drift* of T. In the next section we will see an even more compelling instance of why we prefer to work with the Laplace, rather than the Lévy, exponent of subordinators.

The following describes the local behavior of a subordinator; you can find a good deal more information about this topic in the paper by Štatland (1965).

Theorem 5 (A law of large numbers)**.** *If T is a subordinator with drift* b *and Lévy measure m, then $(T_t/t) \xrightarrow{\text{P}} \text{b}$ as $t \downarrow 0$.*

Proof. Thanks to (2) and the Lévy–Khintchine formula, we can write $T_t = \text{b}t + S_t$, where $\{S_t\}_{t\geq 0}$ is a subordinator with Laplace exponent

$$\Phi(\lambda) = \int_0^\infty \left(1 - \mathrm{e}^{-\lambda z}\right) m(\mathrm{d}z).$$

It suffices to prove that $S_t = o(t)$ a.s. as $t \to \infty$. But

$$\lim_{t\to 0} \mathrm{E}\mathrm{e}^{i\xi(S_t/t)} = \lim_{t\to 0} \exp\left\{-t\int_0^\infty \left(1 - \mathrm{e}^{-i\xi z/t}\right) m(\mathrm{d}z)\right\} = 1 \quad \text{for all } \xi \in \mathbf{R},$$

thanks to the dominated convergence theorem. This proves that S_t/t converges to 0 weakly, and hence in probability [since the limit 0 is nonrandom]. □

Stable subordinators

We have seen already one example of a subordinator [called the "Gamma subordinator"] that has no drift and Lévy measure of the form $m(\mathrm{d}x)/\mathrm{d}x = \alpha x^{-1}\exp(-\lambda x)\mathbb{1}_{(0,\infty)}(x)$ [see page 12]. Next we introduce another.

Consider, a Borel measure m_α on $(0\,,\infty)$ with density

$$\frac{m_\alpha(\mathrm{d}x)}{\mathrm{d}x} := \frac{c}{x^{1+\alpha}}\mathbb{1}_{(0,\infty)}(x),$$

where $c > 0$ is a constant. Then, m_α is the Lévy measure of some Lévy process iff $\alpha \in (0\,,2)$, yet m_α is the Lévy measure of some subordinator iff $\alpha \in (0\,,1)$.

Definition 6. A *stable subordinator* with index $\alpha \in (0\,,1)$ is a subordinator with zero drift and Lévy measure m_α. □

In particular, if T is a stable subordinator with index $\alpha \in (0\,,1)$, then $\Phi(\lambda) \propto \lambda^\alpha$ for an arbitrary [but strictly-positive] constant of proportionality. By changing the notation slightly, we can assume hereforth that T_1 is normalized so that

$$\mathrm{E}\mathrm{e}^{-\lambda T_t} = \mathrm{e}^{-t\lambda^\alpha} \qquad \text{for all } \lambda, t > 0.$$

We now study the local behavior of stable subordinators.

If T is a stable subordinator of index $\alpha \in (0\,,1)$, then $T_t = o(t)$ in probability as $t \to 0^+$. This finding can be sharpened in two different ways; they are explained by our next two theorems.

Theorem 7 (Fristedt, 1967). *Let T be a stable subordinator with index $\alpha \in (0\,,1)$. Then whenever $f : (0\,,\infty) \to \mathbf{R}_+$ is measurable and nonincreasing with $t \mapsto t^{-1/\alpha} f(t)$ nondecreasing near $t = 0$,*

$$\limsup_{t\downarrow 0} \frac{T_t}{f(t)} = \begin{cases} 0 & \text{if } \int_1^\infty [f(t)]^{-\alpha}\,\mathrm{d}t/t < \infty, \\ \infty & \text{if } \int_1^\infty [f(t)]^{-\alpha}\,\mathrm{d}t/t = \infty. \end{cases}$$

Theorem 8 (Fristedt, 1964; Khintchine, 1939). *Let T be a stable subordinator with index $\alpha \in (0\,,1)$. Then there exists a positive and finite constant K such that*

$$\liminf_{t\downarrow 0} \frac{T_t}{f(t)} = K \qquad \text{a.s., where} \quad f(t) := \frac{t^{1/\alpha}}{(\ln\ln(1/t))^{(1-\alpha)/\alpha}}.$$

Example 9. Thus, for example [a.s.],

$$\limsup_{t\downarrow 0} \frac{T_t}{t^{1/\alpha}\cdot[\ln(1/t)]^{1/\alpha}} = \infty \quad \text{whereas} \quad \lim_{t\downarrow 0} \frac{T_t}{t^{1/\alpha}\cdot[\ln(1/t)]^{q/\alpha}} = 0,$$

for every $q > 1$. Note also that Theorems 8 and 7 together imply that with probability one, $T_t = t^{(1+o(1))/\alpha}$ a.s. as $t \downarrow 0$. □

Now we start the groundwork needed to prove Theorems 8 and 7. The arguments rely on two probability estimates that are interesting in their own right; namely, we need sharp estimates for $\mathrm{P}\{T_1 \geq z\}$ and $\mathrm{P}\{T_1 \leq \epsilon\}$ when z is large and $\epsilon > 0$ is small.

Theorem 10. *There exist positive and finite constants c_1 and c_2—depending on $\alpha \in (0\,,1)$ and $c > 0$—such that*

$$c_1\left(1 \wedge z^{-\alpha}\right) \leq \mathrm{P}\left\{T_1 > z\right\} \leq c_2\left(1 \wedge z^{-\alpha}\right) \qquad \text{for all } z > 0.$$

Proof. Because $1 - \mathrm{e}^{-\lambda T_1} \geq (1 - \mathrm{e}^{-\lambda z})\mathbb{1}_{\{T_1 > z\}}$ for all $\lambda > 0$ and $z > 1$, we can take expectations to find that

$$\mathrm{P}\left\{T_1 > z\right\} \leq \inf_{\lambda > 0} \frac{1 - \mathrm{e}^{-\lambda^\alpha}}{1 - \mathrm{e}^{-\lambda z}} \leq \inf_{\lambda > 0} \frac{\lambda^\alpha}{1 - \mathrm{e}^{-\lambda z}} = \frac{1}{z^\alpha} \inf_{q > 0} \frac{q^\alpha}{1 - \mathrm{e}^{-q}}.$$

This proves one bound. For the other, we note that $T_1 = \sum_{s\in[0,1]} (\Delta T)_s$, since T is monotonic, hence of bounded variation. In particular, $T_1 \geq \sup_{s\in[0,1]} (\Delta T)_s$. Therefore,

$$\mathrm{P}\left\{T_1 > z\right\} \geq \mathrm{P}\left\{\sum_{s\in[0,1]} \mathbb{1}_{\{(\Delta T)_s > z\}} \geq 1\right\} = 1 - \mathrm{P}\left\{\sum_{s\in[0,1]} \mathbb{1}_{\{(\Delta T)_s > z\}} = 0\right\}.$$

The sum of the indicators is Poisson with mean $m([z\,,\infty))$, which is proportional to $\int_z^\infty x^{-(1+\alpha)}\,\mathrm{d}x \propto z^{-\alpha}$ for $z > 0$. Therefore, there exists a constant C such that $\mathrm{P}\{T_1 > z\} \geq 1 - \mathrm{e}^{-Cz^{-\alpha}}$ for all $z > 0$. This proves the theorem. □

The preceding estimates the distribution of T_1 near infinity. The following estimates it near zero.

Theorem 11. *There exists a positive and finite constant N_α such that*

$$\mathrm{P}\{T_1 \leq \epsilon\} = \exp\left(-\frac{N_\alpha + o(1)}{\epsilon^{\alpha/(1-\alpha)}}\right) \qquad \text{as } \epsilon \downarrow 0.$$

Proof. Note that $T_1 \leq \epsilon$ if and only if $\exp(-\lambda T_1) \geq \exp(-\lambda\epsilon)$, where $\lambda > 0$ is arbitrary. Therefore, Chebyshev's inequality tells us that

$$\mathrm{P}\{T_1 \leq \epsilon\} \leq \inf_{\lambda>0} \mathrm{e}^{\lambda\epsilon} \mathrm{E}\mathrm{e}^{-\lambda T_1} = \inf_{\lambda>0} \mathrm{e}^{\lambda\epsilon - \lambda^\alpha} = \exp\left(-\frac{\nu_\alpha}{\epsilon^{1/(1-\alpha)}}\right),$$

where

$$\nu_\alpha := \alpha^{\alpha/(1-\alpha)}(1-\alpha),$$

and this is positive since $\alpha \in (0\,,1)$. This proves that if N_α exists then it is certainly bounded below by $\nu_\alpha > 0$.

In order to derive the more interesting lower bound, I will apply an elegant rescaling argument (Griffin, 1985). Let us first note that if $T_{(j+1)/n} - T_{j/n} \leq \epsilon/n$ for all $0 \leq j < n$, then certainly

$$T_1 = \sum_{0\leq j<n} \left(T_{(j+1)/n} - T_{j/n}\right) \leq \epsilon.$$

Therefore, by the independence of the increments of the process T,

$$\mathrm{P}\{T_1 \leq \epsilon\} \geq \prod_{0\leq j<n} \mathrm{P}\left\{T_{(j+1)/n} - T_{j/n} \leq \frac{\epsilon}{n}\right\} = \left[\mathrm{P}\left\{T_{1/n} \leq \frac{\epsilon}{n}\right\}\right]^n.$$

Because $T_{1/n}$ has the same distribution as $n^{-1/\alpha}T_1$, we can deduce the recursive inequality,

$$\mathrm{P}\{T_1 \leq \epsilon\} \geq \left[\mathrm{P}\left\{T_1 \leq \epsilon \cdot n^{(1-\alpha)/\alpha}\right\}\right]^n.$$

Now we select n by setting $n := [(\gamma/\epsilon)^{\alpha/(1-\alpha)}]$, where $[\bullet] :=$ the greatest-integer function. It follows easily from this that

$$\mathrm{P}\{T_1 \leq \epsilon\} \geq \left[\mathrm{P}\{T_1 \leq \gamma\}\right]^{[(\gamma/\epsilon)^{\alpha/(1-\alpha)}]}.$$

Consequently,

$$\liminf_{\epsilon\downarrow 0} \epsilon^{\alpha/(1-\alpha)} \ln \mathrm{P}\{T_1 \leq \epsilon\} \geq \sup_{\gamma>0} \gamma^{\alpha/(1-\alpha)} \ln \mathrm{P}\{T_1 \leq \gamma\}.$$

In particular, the lim inf is a genuine limit and equal to $-N_\alpha$, which is strictly greater than $-\infty$ [since $\mathrm{P}\{T_1 \le \gamma\} > 0$ for γ sufficiently large]. This completes the proof. □

Remark 12. We showed in the proof that $N_\alpha \ge \nu_\alpha$. Large-deviations methods can be used to show that this inequality is in fact an identity; and the ensuing proof will show that the constant K in Theorem 7 is $K = \alpha \cdot \beta^{\beta/\alpha}$ for $\beta := 1 - \alpha$. Exercise 3 below outlines the starting point of this approach. □

Proof of Theorem 7. Let $t_n := 2^{-n}$ and note that

$$I(f) := \int_1^\infty \frac{\mathrm{d}t}{[f(t)]^\alpha} < \infty \quad \text{iff} \quad \sum_{n=1}^\infty \frac{t_n}{[f(t_n)]^\alpha} < \infty.$$

[Cauchy's test.] It is easy to check that T_t has the same distribution as $t^{1/\alpha} T_1$; we simply check the characteristic functions. With this fact in mind, we apply Theorem 10 to find that

$$\sum_{n=1}^\infty \mathrm{P}\left\{T_{t_{n-1}} > f(t_n)\right\} = \sum_{n=1}^\infty \mathrm{P}\left\{T_1 > \frac{f(t_n)}{t_n^{1/\alpha}}\right\} \le \text{const} \cdot \sum_{n=1}^\infty \frac{t_n}{[f(t_n)]^\alpha}.$$

Therefore, whenever $I(f) < \infty$, the Borel–Cantelli lemma ensures that, with probability one, $T_{t_{n-1}} \le f(t_n)$ for all but a finite number of n's. Now we apply a monotonicity/sandwich argument [as in the proof of the strong law of large numbers in Math. 6040]: If $t \in [t_n\,, t_{n-1}]$, then $T_t \le T_{t_{n-1}} \le f(t_n) \le f(t)$ for all t sufficiently small; consequently we have $\lim_{t\downarrow 0}(T_t/f(t)) \le 1$ a.s. Because $I(f) < \infty$ implies that $I(\kappa f) < \infty$ for arbitrarily small $\kappa > 0$, it follows that $T_t/f(t) \to 0$ a.s.

The converse is proved similarly:

$$\begin{aligned}\sum_{n=1}^\infty \mathrm{P}\left\{T_{t_{n-1}} - T_{t_n} \le f(t_{n-1})\right\} &= \sum_{n=1}^\infty \mathrm{P}\left\{T_1 \le \frac{f(t_{n-1})}{(t_{n-1} - t_n)^{1/\alpha}}\right\} \\ &= \sum_{n=1}^\infty \mathrm{P}\left\{T_1 \le \frac{f(t_{n-1})}{(2t_n)^{1/\alpha}}\right\} \\ &\ge \text{const} \cdot \sum_{n=2}^\infty \frac{t_n}{[f(t_n)]^\alpha}.\end{aligned}$$

Therefore, $I(f) = \infty$ implies that a.s., $T_{t_{n-1}} - T_{t_n} > f(t_{n-1})$ infinitely often [Borel–Cantelli lemma for independent events]. Since $T_{t_n} \ge 0$, this does the job. □

Now we prove the limit theorems mentioned earlier.

Proof of Theorem 8. Since T_t has the same distribution as $t^{1/\alpha}T_1$, we obtain the following from Theorem 11: For all $\gamma > 0$,

$$\begin{aligned} \mathrm{P}\left\{T_t \le \gamma f(t)\right\} &= \mathrm{P}\left\{T_1 \le \frac{\gamma}{\left[\ln\ln(1/t)\right]^{(1-\alpha)/\alpha}}\right\} \\ &= \exp\left(-N_\alpha \gamma^{-\alpha/(1-\alpha)} \ln\ln t(1+o(1)\right), \\ &= \left[\ln(1/t)\right]^{-\nu+o(1)} \qquad \text{as } t \downarrow 0, \end{aligned}$$

where $\nu := N_\alpha\gamma^{-\alpha/(1-\alpha)}$. This sums along $t = a^n$, provided that $\nu > 1$ and $a \in (0\,,1)$. Because $t \mapsto T_t$ is nondecreasing, a monotonicity/sandwich argument [as in the 6040 proof of the LIL] proves that

$$\liminf_{t\downarrow 0} \frac{T_t}{f(t)} \ge N_\alpha^{(1-\alpha)/\alpha} \qquad \text{a.s.} \tag{3}$$

For the converse we continue using the notation $\nu := N_\alpha\gamma^{-\alpha/(1-\alpha)}$. But now we consider the case that $\nu < 1$. Let us also redefine $t_n := \exp(-n^q)$—where $q > 1$ is to be chosen in a little bit—and then note that

$$\begin{aligned} \mathrm{P}\left\{T_{t_{n-1}} - T_{t_n} \le \gamma f(t_{n-1})\right\} &= \mathrm{P}\left\{T_1 \le \gamma \frac{f(t_{n-1})}{(t_{n-1}-t_n)^{1/\alpha}}\right\} \\ &= \exp\left(-(\nu + o(1))\ln\ln(1/t_{n-1})\right) \\ &= n^{-(\nu q + o(1))}, \end{aligned}$$

since $t_{n-1} - t_n \sim t_n$ as $n \to \infty$. It follows that

$$\sum_n \mathrm{P}\left\{T_{t_{n-1}} - T_{t_n} \le \gamma f(t_{n-1})\right\} = \infty \quad \text{provided that } 1 < q < \nu^{-1}.$$

The Borel–Cantelli lemma for independent events implies that if $1 < q < \nu^{-1}$, then

$$\liminf_{n\to\infty} \frac{T_{t_{n-1}} - T_{t_n}}{f(t_{n-1})} \le \gamma \quad \text{a.s.} \tag{4}$$

At the same time, it is possible to apply Theorem 10 to obtain the following: For all $\epsilon > 0$ and n large,

$$\begin{aligned} \mathrm{P}\left\{T_{t_n} \ge \epsilon f(t_{n-1})\right\} &= \mathrm{P}\left\{T_1 \ge \epsilon \frac{(t_{n-1}/t_n)^{1/\alpha}}{(\ln\ln(1/t_{n-1}))^{(1-\alpha)/\alpha}}\right\} \\ &\le \text{const}\cdot \frac{(\ln\ln(1/t_{n-1}))^{(1-\alpha)}}{t_{n-1}/t_n} \\ &= \frac{n^{(1-\alpha)}}{\exp\left\{qn^{q-1}(1+o(1))\right\}}. \end{aligned}$$

Because $q > 1$ and $\epsilon > 0$ is arbitrary, the preceding estimate and the Borel–Cantelli lemma together imply that $T_{t_n} = o(f(t_{n-1}))$ a.s. as $n \to \infty$. Consequently, (4) tells us that

$$\liminf_{t \downarrow 0} \frac{T_t}{f(t)} \leq \liminf_{n \to \infty} \frac{T_{t_{n-1}}}{f(t_{n-1})} \leq \gamma \quad \text{a.s.},$$

as long as $\nu < 1$; i.e., $\gamma < N_\alpha^{(1-\alpha)/\alpha}$. It follows that the inequality in (3) can be replaced by an identity; this completes our proof. □

Subordination

Let X denote a symmetric Lévy process on $\mathbf{R}^d$ with Lévy exponent Ψ, and suppose T is an independent subordinator [on $\mathbf{R}_+$, of course] with Laplace exponent Φ. Because Ψ is real [and hence nonnegative], it is not hard to see that the process $Y := \{Y_t\}_{t \geq 0}$, defined as

$$Y_t := X_{T_t} \qquad \text{for } t \geq 0,$$

is still a symmetric Lévy process. Moreover, by conditioning we find that

$$\mathrm{E}e^{i\xi \cdot Y_t} = \mathrm{E}\left(e^{-T_t \Psi(\xi)}\right) = e^{-t\Phi(\Psi(\xi))} \qquad \text{for all } t \geq 0 \text{ and } \xi \in \mathbf{R}^d.$$

We say that Y is *subordinated to* X via [the subordinator] T. Let us summarize our findings.

Proposition 13 (Subordination)**.** *If X is a symmetric Lévy process on $\mathbf{R}^d$ with Lévy exponent Ψ and T is an independent subordinator with Laplace exponent Φ, then $Y := X \circ T$ is a symmetric Lévy process on $\mathbf{R}^d$ with Lévy exponent $\Phi \circ \Psi$. If X is isotropic, then so is $X \circ T$.*

If T is a stable subordinator with index $\alpha \in (0\,,1)$, then $\Phi(\lambda) \propto \lambda^\alpha$. Since the Lévy exponent of standard Brownian motion is $\Psi(\xi) = \frac{1}{2}\|\xi\|^2$, we immediately obtain the following.

Theorem 14 (Bochner)**.** *Let X denote standard d-dimensional Brownian motion, and T an independent stable subordinator with index $\alpha \in (0\,,1)$. Then $X \circ T$ is an isotropic stable process with index 2α.*

Because $2\alpha \in (0\,,2)$ whenever $\alpha \in (0\,,1)$, the preceding tells us that we can always realize any isotropic stable process via a subordination of Brownian motion!

Theorem 15. *Let X be a d-dimensional random variable whose law is isotropic stable with index $\alpha \in (0\,,2)$. Then there exist $c_1, c_2 \in (0\,,\infty)$ such that*

$$\frac{c_1}{z^\alpha} \leq \mathrm{P}\{\|X\| > z\} \leq \frac{c_2}{z^\alpha} \qquad \text{for all } z > 1.$$

This theorem improves on Proposition 2 (page 11), which asserted that $\mathrm{E}(\|X\|^\beta) < \infty$ if $\beta < \alpha$, but not if $\beta \geq \alpha$.

Proof. We know that X has the same law as Y_1, where $Y := \{Y_t\}_{t\geq 0}$ is an isotropic stable process on $\mathbf{R}^d$. We can realize Y as follows: $Y_t = B_{T_t}$ where B is d-dimensional Brownian motion, and T an independent stable subordinator with index $\alpha/2$. Therefore, by scaling,

$$\mathrm{P}\{\|X\| > z\} = \mathrm{P}\{\|B_{T_1}\| > z\} = \mathrm{P}\left\{T_1 > \frac{z^2}{\|B_1\|^2}\right\}.$$

It follows easily from Theorem 10 [after conditioning on B] that there are constants c_1 and c_2 such that for all $z > 0$,

$$c_1 \mathrm{E}\left[1 \wedge \frac{\|B_1\|^\alpha}{z^\alpha}\right] \leq \mathrm{P}\{\|X\| > z\} \leq c_2 \mathrm{E}\left[1 \wedge \frac{\|B_1\|^\alpha}{z^\alpha}\right].$$

And by the dominated convergence theorem, the two expectations are equal to $z^{-\alpha}\mathrm{E}(\|B_1\|^\alpha)(1 + o(1))$ as $z \to \infty$. □

Problems for Lecture 8

Throughout these problems, $T := \{T_t\}_{t\geq 0}$ denotes a subordinator with drift b, Lévy measure m, and Laplace exponent Φ.

1. Prove that $\mathrm{E}(T_t/t) = \lim_{\lambda\downarrow 0} \Phi(\lambda)/\lambda = \int_0^\infty x\, m(\mathrm{d}x)$ for all $t \geq 0$. Construct an example where $\mathrm{E}(T_t) = \infty$ for all $t > 0$.

2. Let B denote one-dimensional Brownian motion, and define

$$T_t := \inf\{s > 0 : B_s > t\} \qquad (t \geq 0,\ \inf\varnothing := \infty).$$

(1) Prove that $\{T_t\}_{t\geq 0}$ is a subordinator, and for all $s, t, q \geq 0$,

$$\mathrm{P}\left(T_{t+s} - T_s \leq q \,\middle|\, \mathcal{F}_s\right) = \frac{t}{\sqrt{2\pi}} \int_0^q \frac{\mathrm{e}^{-t^2/(2u)}}{u^{3/2}}\,\mathrm{d}u \qquad \text{a.s.}$$

(2) Conclude that $T_t = \int_0^\infty x\, \Pi_t(\mathrm{d}x)$, for a Poisson point process $\{\Pi_t\}_{t\geq 0}$ with intensity $\mathrm{d}t \times \rho(\mathrm{d}x)$, where $\rho(A) := (2\pi)^{-1/2} \int_A x^{-3/2}\,\mathrm{d}x$.

(3) Show that the Laplace exponent of T is

$$\Phi(\lambda) := \int_0^\infty \left(1 - \mathrm{e}^{-\lambda x}\right) \frac{\mathrm{d}x}{x^{3/2}}.$$

Conclude that T is a stable subordinator of index 1/2 (see Lévy, 1992, Théorème 46.1, p. 221).

3 (Girsanov transformations)**.** Let T be a subordinator with drift b and Laplace exponent Φ. Define $\{\mathcal{F}_t\}_{t\geq 0}$ to be the natural filtration of T.

(1) Prove that $M_t^{(\lambda)} := \exp(-\lambda T_t + t\Phi(\lambda))$ defines a nonnegative mean-one cadlag martingale with respect to $\{\mathcal{F}_t\}_{t\geq 0}$;

(2) For every $\lambda > 0$ and $t \geq 0$ define $\mathrm{P}^{(\lambda)}(A) := \mathrm{E}[M_t^{(\lambda)}\,; A]$ for all $A \in \mathcal{F}_t$. Prove that $\mathrm{P}^{(\lambda)}$ is defined consistently as a probability measure on the measurable space $(\Omega\,, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty := \vee_{t\geq 0}\mathcal{F}_t$;

(3) Prove that T is a subordinator under the measure $\mathrm{P}^{(\lambda)}$. Compute the drift, and more generally Laplace exponent, of T under the new measure $\mathrm{P}^{(\lambda)}$;

(4) Prove that $\mathrm{P}^{(\lambda)}\{\lim_{t\downarrow 0}(T_t/t) = \Phi'(\lambda)\} = 1$. Compare with Theorem 5 on page 45;

(5) Prove that when $t, \lambda > 0$, $\mathrm{E}^{(\lambda)}(T_t) = t\Phi'(\lambda)$ and $\mathrm{Var}^{(\lambda)}(T_t) = -t\Phi''(\lambda)$, where $\mathrm{E}^{(\lambda)}$ and $\mathrm{Var}^{(\lambda)}$ respectively denote the expectation and variance operators for $\mathrm{P}^{(\lambda)}$;

(6) Examine all of the preceding in the special case that T is a stable subordinator with index $\alpha \in (0\,, 1)$.

4. Prove that if the Lévy measure m of a Lévy process X is supported in a compact set and $\int_{\mathbf{R}^d} \|x\|^\gamma\, m(\mathrm{d}x) < \infty$ for some $\gamma \in (0\,, 1]$, then we can write $X_t = -at + \sigma B_t + T_t - S_t$, where T and S are independent subordinators, and B is an independent Brownian motion (Millar, 1971).

The Strong Markov Property

Throughout, $X := \{X_t\}_{t\ge0}$ denotes a Lévy process on $\mathbf{R}^d$ with triple $(a\,,\sigma, m)$, and exponent Ψ. And from now on, we let $\{\mathcal{F}_t\}_{t\ge0}$ denote the natural filtration of X, all the time remembering that, in accord with our earlier convention, $\{\mathcal{F}_t\}_{t\ge0}$ satisfies the usual conditions.

Transition measures and the Markov property

Definition 1. The *transition measures* of X are the probability measures

$$P_t(x\,,A) := \mathrm{P}\,\{x + X_t \in A\}$$

defined for all $t \ge 0$, $x \in \mathbf{R}^d$, and $A \in \mathcal{B}(\mathbf{R}^d)$. In other words, each $P_t(x\,,\bullet)$ is the law of X_t started at $x \in \mathbf{R}^d$. We single out the case $x = 0$ by setting $\mu_t(A) := P_t(0\,,A)$; thus, μ_t is the distribution of X_t for all $t > 0$. □

Note, in particular, that $\mu_0 = \delta_0$ is the point mass at $0 \in \mathbf{R}^d$.

Proposition 2. *For all $s, t \ge 0$, and measurable $f : \mathbf{R}^d \to \mathbf{R}_+$,*

$$\mathrm{E}[f(X_{t+s})\,|\,\mathcal{F}_s] = \int_{\mathbf{R}^d} f(y)\,P_t(X_s\,,\mathrm{d}y) \qquad a.s.$$

Consequently, for all $x_0 \in \mathbf{R}^d$, $0 < t_1 < t_2 < \cdots < t_k$, *and measurable* $f_1, \ldots, f_k : \mathbf{R}^d \to \mathbf{R}_+$,

$$\begin{aligned}
&\mathrm{E}\left(\prod_{j=1}^k f_j(x_0 + X_{t_j})\right) \qquad (1)\\
&= \int_{\mathbf{R}^k} P_{t_1}(x_0\,, \mathrm{d}x_1) \int_{\mathbf{R}^k} P_{t_2-t_1}(x_1\,, \mathrm{d}x_2) \cdots \int_{\mathbf{R}^d} P_{t_k-t_{k-1}}(x_{k-1}\,, \mathrm{d}x_k) \prod_{j=1}^k f_j(x_j).
\end{aligned}$$

Property (1) is called the *Chapman–Kolmogorov equation*. That property has the following ready consequence: Transition measures determine the finite-dimensional distributions of X uniquely.

Definition 3. Any stochastic process $\{X_t\}_{t\ge 0}$ that satisfies the Chapman-Kolmogorov equation is called a *Markov process*. This definition continues to make sense if we replace $(\mathbf{R}^d\,, \mathcal{B}(\mathbf{R}^d))$ by any measurable space on which we can construct infinite families of random variables. □

Thus, Lévy processes are cadlag Markov processes that have special "addition" properties. In particular, as Exercise below 1 shows, Lévy processes have the important property that the finite-dimensional distributions of X are described not only by $\{P_t(x\,,\cdot)\}_{t\ge 0, x\in\mathbf{R}^d}$ but by the much-smaller family $\{\mu_t(\cdot)\}_{t\ge 0}$.

Note, in particular, that if $f : \mathbf{R}^d \to \mathbf{R}_+$ is measurable, $t \ge 0$, and $x \in \mathbf{R}^d$, then

$$\mathrm{E}f(x + X_t) = \int_{\mathbf{R}^d} f(y)\, P_t(x\,, \mathrm{d}y) = \int_{\mathbf{R}^d} f(x + y)\, \mu_t(\mathrm{d}y).$$

Therefore, if we define

$$\tilde{\mu}_t(A) := \mu_t(-A) \qquad \text{for all } t \ge 0 \text{ and } A \in \mathcal{B}(\mathbf{R}^d),$$

where $-A := \{-a : a \in A\}$, then we have the following convolution formula, valid for all measurable $f : \mathbf{R}^d \to \mathbf{R}_+$, $x \in \mathbf{R}^d$, and $t \ge 0$:

$$\mathrm{E}f(x + X_t) = (f * \tilde{\mu}_t)(x).$$

And, more generally, for all measurable $f : \mathbf{R}^d \to \mathbf{R}_+$, $x \in \mathbf{R}^d$, and $s, t \ge 0$

$$\mathrm{E}[f(X_{t+s}) \,|\, \mathcal{F}_s] = (f * \tilde{\mu}_t)(X_s) \qquad \text{a.s.}$$

[Why is this more general?]

Proposition 4. *The family* $\{\mu_t\}_{t\ge 0}$ *of Borel probability measure on* $\mathbf{R}^d$ *is a "convolution semigroup" in the sense that* $\mu_t * \mu_s = \mu_{t+s}$ *for all* $s, t \ge 0$. *Moreover,* $\hat{\mu}_t(\xi) = \exp(-t\Psi(\xi))$ *for all* $t \ge 0$ *and* $\xi \in \mathbf{R}^d$. *Similarly,* $\{\tilde{\mu}_t\}_{t\ge 0}$ *is a convolution semigroup with* $\hat{\tilde{\mu}}_t(\xi) = \exp(-t\Psi(-\xi))$.

Proof. The assertion about $\tilde{\mu}$ follows from the assertion about μ [or you can repeat the following with $\tilde{\mu}$ in place of μ].

Since μ_t is the distribution of X_t, the characteristic function of X_t is described by $\hat{\mu}_t(\xi) = \exp(-t\Psi(\xi))$. The proposition follows immediately from this, because $\hat{\mu}_t(\xi) \cdot \hat{\mu}_s(\xi) = \exp(-(t+s)\Psi(\xi)) = \hat{\mu}_{t+s}(\xi)$. □

The strong Markov property

Theorem 5 (The strong Markov property)**.** *Let T be a finite stopping time. Then, the process $X^T := \{X_t^T\}_{t\geq 0}$, defined by $X_t^T := X_{T+t} - X_T$ is a Lévy process with exponent Ψ and independent of $\mathcal{F}_T$.*

Proof. X^T is manifestly cadlag [because X is]. In addition, one checks that whenever $0 < t_1 < \cdots < t_k$ and $A_1, \ldots, A_k \in \mathcal{B}(\mathbf{R}^d)$,

$$\mathrm{P}\left(\bigcap_{j=1}^{k} \left\{X_{T+t_j} - X_T \in A_j\right\} \,\middle|\, \mathcal{F}_T\right) = \mathrm{P}\left(\bigcap_{j=1}^{k} \left\{X_{t_j} \in A_j\right\}\right) \qquad \text{a.s.;}$$

see Exercise 2 on page 32. This readily implies that the finite-dimensional distributions of X^T are the same as the finite-dimensional distributions of X, and the result follows. □

Theorem 5 has a number of deep consequences. The following shows that Lévy processes have the following variation of strong Markov property. The following is attractive, in part because it can be used to study processes that do not have good additivity properties.

Corollary 6. *For all finite stopping times T, every $t \geq 0$, and all measurable functions $f : \mathbf{R}^d \to \mathbf{R}_+$.*

$$\mathrm{E}\left[f(X_{T+t}) \mid \mathcal{F}_T\right] = \int_{\mathbf{R}^d} f(y) P_t(X_T\,, \mathrm{d}y) \qquad \textit{a.s.}$$

Let T be a finite stopping time, and then define $\mathcal{F}^{(T)} = \{\mathcal{F}_t^{(T)}\}_{t\geq 0}$ to be the natural filtration of the Lévy process X^T. The following is a useful corollary of the strong Markov property.

Corollary 7 (Blumenthal's zero-one law; Blumenthal, 1957)**.** *Let T be a finite stopping time. Then $\mathcal{F}_0^{(T)}$ is trivial; i.e., $\mathrm{P}(A) \in \{0\,, 1\}$ for all $A \in \mathcal{F}_0^{(T)}$.*

The following are nontrivial examples of elements of $\mathcal{F}_0^{(T)}$:

$$\Upsilon_1 := \left\{ \liminf_{t\downarrow 0} \frac{\|X_{T+t} - X_T\|}{t^{1/\alpha}} = 0 \right\} \qquad \text{where } \alpha > 0 \text{ is fixed;}$$

$$\Upsilon_2 := \left\{ \limsup_{t\downarrow 0} \frac{\|X_{T+t} - X_T\|}{\sqrt{2t \ln\ln(1/t)}} = 1 \right\}; \text{ or}$$

$$\Upsilon_3 := \left\{ {}^{\exists} t_n \downarrow 0 \text{ such that } X_{T+t_n} - X_T > 0 \text{ for all } n \geq 1 \right\} \text{ in dimension one, etc.}$$

Proof of Blumenthal's zero-one law. The strong Markov property [Corollary 6] reduces the problem to $T \equiv 0$. And of course we do not need to write $\mathcal{F}^{(0)}$ since $\mathcal{F}_t^{(0)}$ is the same object as $\mathcal{F}_t$.

For all $n \geq 1$ define $\mathcal{A}_n$ to be the completion of the sigma-algebra generated by the collection $\{X_{t+2^{-n}} - X_{2^{-n}}\}_{t\in[0,2^{-n}]}$. By the Markov property, $\mathcal{A}_1, \mathcal{A}_2, \ldots$ are independent sigma-algebras. Their tail sigma-algebra $\mathcal{T}$ is the smallest sigma-algebra that contains $\cup_{i=N}^{\infty}\mathcal{A}_i$ for all $N \geq 1$. Clearly $\mathcal{T}$ is complete, and Kolmogorov's zero-one law tells us that $\mathcal{T}$ is trivial. Because $\cup_{i=N}^{\infty}\mathcal{A}_i$ contains the sigma-algebra generated by all increments of the form $X_{u+v} - X_u$ where $u, v \in [2^{-m}, 2^{-m+1}]$ for some $m \geq N$, and since $X_u \to 0$ as $u \downarrow 0$, it follows that $\mathcal{T}$ contains $\cap_{s\geq 0}\mathcal{X}_s$, where $\mathcal{X}_s$ denotes the sigma-algebra generated by $\{X_r\}_{r\in[0,s]}$. Since $\mathcal{T}$ is complete, this implies $\mathcal{F}_0 \subseteq \mathcal{T}$ [in fact, $\mathcal{T} = \mathcal{F}_0$] as well, and hence $\mathcal{F}_0$ is trivial because $\mathcal{T}$ is. □

Consider, for example, the set Υ_1 introduced earlier. We can apply the Blumenthal zero-one, and deduce the following:

$$\text{For every } \alpha > 0, \qquad \mathrm{P}\left\{ \liminf_{t\downarrow 0} \frac{\|X_{T+t} - X_T\|}{t^{1/\alpha}} = 0 \right\} = 0 \text{ or } 1.$$

You should construct a few more examples of this type.

Feller semigroups and resolvents

Define a collection $\{P_t\}_{t\geq 0}$ of linear operators by

$$(P_t f)(x) := \mathrm{E}f(x + X_t) = \int_{\mathbf{R}^d} f(y) P_t(x\,, dy) = (f * \tilde{\mu}_t)(x) \quad \text{for } t \geq 0,\ x \in \mathbf{R}^d.$$

[Since $X_0 = 0$, $P_0 = \delta_0$ is point mass at zero.] The preceding is well defined for various measurable functions $f : \mathbf{R}^d \to \mathbf{R}$. For instance, everything is fine if f is nonnegative, and also if $(P_t|f|)(x) < \infty$ for all $t \geq 0$ and $x \in \mathbf{R}^d$ [in that case, we can write $P_t f = P_t f^+ - P_t f^-$].

The Markov property of X [see, in particular, Proposition 4] tells us that $(P_{t+s}f)(x) = (P_t(P_s f))(x)$. In other words,

$$P_{t+s} = P_t P_s = P_s P_t \qquad \text{for all } s, t \geq 0, \tag{2}$$

where P_tP_sf is shorthand for $P_t(P_sf)$ etc. Since P_t and P_s commute, in the preceding sense, there is no ambiguity in dropping the parentheses.

Definition 8. The family $\{P_t\}_{t\geq 0}$ is the *semigroup* associated with the Lévy process X. The *resolvent* $\{R_\lambda\}_{\lambda>0}$ of the process X is the family of linear operators defined by

$$(R_\lambda f)(x) := \int_0^\infty \mathrm{e}^{-\lambda t}(P_tf)(x)\,\mathrm{d}t = \mathrm{E}\int_0^\infty \mathrm{e}^{-\lambda t} f(x+X_t)\,\mathrm{d}t \quad (\lambda > 0).$$

This can make sense also for $\lambda = 0$, and we write R in place of R_0. Finally, R_λ is called the *λ-potential of f* when $\lambda > 0$; when $\lambda = 0$, we call it the potential of f instead. □

Remark 9. It might be good to note that we can cast the strong Markov property in terms of the semigroup $\{P_t\}_{t\geq 0}$ as follows: For all $s \geq 0$, finite stopping times T, and $f : \mathbf{R}^d \to \mathbf{R}_+$ measurable, $\mathrm{E}[f(X_{T+s}) \mid \mathcal{F}_T] = (P_sf)(X_T)$ almost surely. □

Formally speaking,

$$R_\lambda = \int_0^\infty \mathrm{e}^{-\lambda t} P_t\,\mathrm{d}t \qquad (\lambda \geq 0)$$

defines the Laplace transform of the [infinite-dimensional] function $t \mapsto P_t$. Once again, $R_\lambda f$ is defined for all Borel measurable $f : \mathbf{R}^d \to \mathbf{R}$, if either $f \geq 0$; or if $R_\lambda|f|$ is well defined.

Recall that $C_0(\mathbf{R}^d)$ denotes the collection of all continuous $f : \mathbf{R}^d \to \mathbf{R}$ that vanish at infinity [$f(x) \to 0$ as $\|x\| \to \infty$]; $C_0(\mathbf{R}^d)$ is a Banach space in norm $\|f\| := \sup_{x\in\mathbf{R}^d} |f(x)|$.

The following are easy to verify:

(1) Each P_t is a contraction [more precisely nonexpansive] on $C_0(\mathbf{R}^d)$. That is, $\|P_tf\| \leq \|f\|$ for all $t \geq 0$;

(2) $\{P_t\}_{t\geq 0}$ is a *Feller semigroup*. That is, each P_t maps $C_0(\mathbf{R}^d)$ to itself and $\lim_{t\downarrow 0} \|P_tf - f\| = 0$;

(3) If $\lambda > 0$, then λR_λ is a contraction [nonexpansive] on $C_0(\mathbf{R}^d)$;

(4) If $\lambda > 0$, then λR_λ maps $C_0(\mathbf{R}^d)$ to itself.

The preceding describe the smoothness behavior of P_t and R_λ for fixed t and λ. It is also not hard to describe the smoothness properties of them as functions of t and λ. For instance,

Proposition 10. *For all $f \in C_0(\mathbf{R}^d)$,*

$$\lim_{t\downarrow 0}\sup_{s\geq 0} \|P_{t+s}f - P_sf\| = 0 \quad \textit{and} \quad \lim_{\lambda\uparrow\infty} \|\lambda R_\lambda f - f\| = 0.$$

Proof. We observe that

$$\|P_tf - f\| = \sup_{x\in\mathbf{R}^d} |\mathrm{E}f(x + X_t) - f(x)| \le \mathrm{E}\left(\sup_{x\in\mathbf{R}^d} |f(x + X_t) - f(x)|\right).$$

Now every $f \in C_0(\mathbf{R}^d)$ is uniformly continuous and bounded on all of $\mathbf{R}^d$. Since X is right continuous and f is bounded, it follows from the bounded convergence theorem that $\lim_{t\downarrow 0} \|P_tf - f\| = 0$. But the semigroup property implies that $\|P_{t+s}f - P_sf\| = \|P_s(P_tf - f)\| \le \|P_tf - f\|$, since P_s is a contraction on $C_0(\mathbf{R}^d)$. This proves the first assertion. The second follows from the first, since $\lambda R_\lambda = \int_0^\infty \mathrm{e}^{-t} P_{t/\lambda}\,\mathrm{d}t$ by a change of variables. □

Proposition 11. *If $f \in C_0(\mathbf{R}^d) \cap L^p(\mathbf{R}^d)$ for some $p \in [1\,,\infty)$, then $\|P_tf\|_{L^p(\mathbf{R}^d)} \le \|f\|_{L^p(\mathbf{R}^d)}$ for all $t \ge 0$ and $\|\lambda R_\lambda f\|_{L^p(\mathbf{R}^d)} \le \|f\|_{L^p(\mathbf{R}^d)}$ for all $\lambda > 0$.*

In words, the preceding states that P_t and λR_λ are contractions on $L^p(\mathbf{R}^d)$ for every $p \in [1\,,\infty)$ and $t, \lambda > 0$.

Proof. If $f \in C_0(\mathbf{R}^d) \cap L^p(\mathbf{R}^d)$, then for all $t \ge 0$,

$$\begin{aligned}\int_{\mathbf{R}^d} |(P_tf)(x)|^p\,\mathrm{d}x = \int_{\mathbf{R}^d} |\mathrm{E}f(x + X_t)|^p\,\mathrm{d}x &\le \int_{\mathbf{R}^d} \mathrm{E}\left(|f(x + X_t)|^p\right)\mathrm{d}x\\ &= \int_{\mathbf{R}^d} |f(y)|^p\,\mathrm{d}y.\end{aligned}$$

This proves the assertion about P_t; the one about R_λ is proved similarly. □

The Hille–Yosida theorem

One checks directly that for all $\mu, \lambda \ge 0$,

$$R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu. \tag{3}$$

This is called the *resolvent equation*, and has many consequences. For instance, the resolvent equation implies readily the commutation property $R_\mu R_\lambda = R_\lambda R_\mu$. For another consequence of the resolvent eqution, suppose $g = R_\mu f$ for some $f \in C_0(\mathbf{R}^d)$ and $\mu > 0$. Then, $g \in C_0(\mathbf{R}^d)$ and by the resolvent equation, $R_\lambda f - g = (\mu - \lambda)R_\lambda g$. Consequently, $g = R_\lambda h$, where $h := f + (\lambda - \mu)R_\lambda g \in C_0(\mathbf{R}^d)$. In other words, $R_\mu(C_0(\mathbf{R}^d)) = R_\lambda(C_0(\mathbf{R}^d))$, whence

$$\mathrm{Dom}[L] := \left\{R_\mu f : f \in C_0(\mathbf{R}^d)\right\} \quad \text{does not depend on } \mu > 0.$$

And $\mathrm{Dom}[L]$ is dense in $C_0(\mathbf{R}^d)$ [Proposition 10].

For yet another application of the resolvent equation, let us suppose that $R_\lambda f \equiv 0$ for some $\lambda > 0$ and $f \in C_0(\mathbf{R}^d)$. Then the resolvent equation implies that $R_\mu f \equiv 0$ for all μ. Therefore, $f = \lim_{\mu\uparrow 0} \mu R_\mu f = 0$. This implies that every R_λ is a one-to-one and onto map from $C_0(\mathbf{R}^d)$ to Dom[L]; i.e., it is invertible!

Definition 12. The [infinitesimal] *generator* of X is the linear operator $L : \text{Dom}[L] \to C_0(\mathbf{R}^d)$ that is defined uniquely by

$$L := \lambda I - R_\lambda^{-1},$$

where $If := f$ defines the identity operator I on $C_0(\mathbf{R}^d)$. The space Dom[L] is the *domain of* L. □

The following is perhaps a better way to think about L; roughly speaking, it asserts that $P_t f - f \simeq tLf$ for t small, and $\lambda R_\lambda f - f \simeq \lambda^{-1} Lf$ for λ large.

Theorem 13 (Hille XXX, Yosida XXX)**.** *If* $f \in \text{Dom}[L]$, *then*

$$\lim_{\lambda\uparrow\infty} \sup_{x\in\mathbf{R}^d} \left| \frac{\lambda(R_\lambda f)(x) - f(x)}{1/\lambda} - (Lf)(x) \right| = \lim_{t\downarrow 0} \sup_{x\in\mathbf{R}^d} \left| \frac{(P_t f)(x) - f(x)}{t} - (Lf)(x) \right| = 0.$$

Because $f = P_0 f$, the Hille–Yosida theorem implies, among other things, that $(\partial/\partial t)P_t|_{t=0} = L$, where the partial derivative is really a right derivative. See Exercise 4 for a consequence in partial integro-differential equations.

Proof. Thanks to Proposition 10 and the definition of the generator, $Lf = \lambda f - R_\lambda^{-1} f$ for all $f \in \text{Dom}[L]$, whence

$$\lambda R_\lambda Lf = \frac{\lambda R_\lambda f - f}{1/\lambda} \to Lf \qquad \text{in } C_0(\mathbf{R}^d) \text{ as } \lambda \uparrow \infty.$$

This proves half of the theorem. For the other half recall that Dom[L] is the collection of all functions of the form $f = R_\lambda h$, where $h \in C_0(\mathbf{R}^d)$ and $\lambda > 0$. By the semigroup property, for such λ and h we have

$$\begin{aligned} P_t R_\lambda h &= \int_0^\infty \mathrm{e}^{-\lambda s} P_{t+s} h \, \mathrm{d}s = \mathrm{e}^{\lambda t} \int_t^\infty \mathrm{e}^{-\lambda s} P_s h \, \mathrm{d}s \\ &= \mathrm{e}^{\lambda t} \left(R_\lambda h - \int_0^t \mathrm{e}^{-\lambda s} P_s h \, \mathrm{d}s \right). \end{aligned}$$

Consequently, for all $f = R_\lambda h \in \text{Dom}[L]$,

$$\begin{aligned} \frac{P_t f - f}{t} &= \left(\frac{\mathrm{e}^{\lambda t} - 1}{t} \right) R_\lambda h - \frac{\mathrm{e}^{\lambda t}}{t} \int_0^t \mathrm{e}^{-\lambda s} P_s h \, \mathrm{d}s \\ &\to \lambda R_\lambda h - h \qquad \text{in } C_0(\mathbf{R}^d) \text{ as } t \downarrow 0. \end{aligned}$$

But $\lambda R_\lambda h - h = \lambda f - R_\lambda^{-1} f = Lf$. □

The form of the generator

Let $\mathcal{S}$ denote the collection of all rapidly-decreasing test functions $f : \mathbf{R}^d \to \mathbf{R}$. That is, $f \in \mathcal{S}$ if and only if $f \in C^\infty(\mathbf{R}^d)$, and f and all of its partial derivatives vanish faster than any polynomial. In other words, if D is a differential operator [of finite order] and $n \geq 1$, then $\sup_{x\in\mathbf{R}^d}(1 + \|x\|^n)|(Df)(x)| < \infty$. It is easy to see that $\mathcal{S} \subset L^1(\mathbf{R}^d) \cap C_0(\mathbf{R}^d)$ and $\mathcal{S}$ is dense in $C_0(\mathbf{R}^d)$. And it is well known that if $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$ as well, and vice versa.

It is possible to see that if $f, \hat{f} \in L^1(\mathbf{R}^d)$, then for all $t \geq 0$ and $\lambda > 0$,

$$\widehat{P_t f}(\xi) = \mathrm{e}^{-t\Psi(-\xi)}\hat{f}(\xi), \quad \widehat{R_\lambda f}(\xi) = \frac{\hat{f}(\xi)}{\lambda + \Psi(-\xi)} \qquad \text{for all } \xi \in \mathbf{R}^d. \tag{4}$$

Therefore, it follows fairly readily that when $f \in \mathrm{Dom}[L] \cap L^1(\mathbf{R}^d)$, $Lf \in L^1(\mathbf{R}^d)$, and $\hat{f} \in L^1(\mathbf{R}^d)$, then we have

$$\widehat{Lf}(\xi) = -\Psi(-\xi)\hat{f}(\xi) \qquad \text{for every } \xi \in \mathbf{R}^d. \tag{5}$$

It follows immediately from these calculations that: (i) Every P_t and R_λ map $\mathcal{S}$ to $\mathcal{S}$; and (ii) Therefore, $\mathcal{S}$ is dense in $\mathrm{Dom}[L]$. Therefore, we can try to understand L better by trying to compute Lf not for all $f \in \mathrm{Dom}[L]$, but rather for all f in the dense subcollection $\mathcal{S}$. But the formula for the Fourier transform of Lf [together with the estimate $|\Psi(\xi)| = O(\|\xi\|^2)$] shows that $L : \mathcal{S} \to \mathcal{S}$ and

$$(Lf)(x) = -\frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \mathrm{e}^{-i\xi\cdot x}\Psi(-\xi)\hat{f}(\xi)\,\mathrm{d}\xi \qquad \text{for all } x \in \mathbf{R}^d \text{ and } f \in \mathcal{S}.$$

Consider the simplest case that the process X satisfies $X_t = at$ for some $a \in \mathbf{R}^d$; i.e., $\Psi(\xi) = -i(a \cdot \xi)$. In that case, we have

$$\begin{aligned}(Lf)(x) = \frac{1}{(2\pi)^d} \cdot \int_{\mathbf{R}^d}(a \cdot i\xi)\mathrm{e}^{-i\xi\cdot x}\hat{f}(\xi)\,\mathrm{d}\xi &= -\frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \mathrm{e}^{-i\xi\cdot x}(a \cdot \widehat{\nabla f}(\xi))\,\mathrm{d}\xi \\ &= -a \cdot (\nabla f)(x),\end{aligned}$$

thanks to the inversion formula. The very same computation works in the more general setting, and yields

Theorem 14. *If $f \in \mathcal{S}$, then $Lf = Cf + Jf$, where*

$$(Cf)(x) = -a \cdot (\nabla f)(x) + \frac{1}{2}\sum\sum_{1\leq i,j\leq d}(\sigma'\sigma)_{ij}\frac{\partial^2}{\partial x_i \partial x_j}f(x),$$

and

$$(Jf)(x) := \int_{\mathbf{R}^d}\left[f(x + z) - f(x) - z \cdot (\nabla f)(x)\mathbf{1}_{[0,1)}(\|z\|)\right] m(\mathrm{d}z) \quad \text{for all } x \in \mathbf{R}^d,$$

Moreover, J is the generator of the non-Gaussian component; and $C = -a \cdot \nabla + \frac{1}{2}\nabla'\sigma'\sigma\nabla$ is the generator of the Gaussian part.

Here are some examples:

- If X is Brownian motion on $\mathbf{R}^d$, then $L = \frac{1}{2}\Delta$ is one-half of the Laplace operator [on $\mathcal{S}$];
- If X is the Poisson process on $\mathbf{R}$ with intensity $\lambda \in (0\,,\infty)$, then $(Lf)(x) = \lambda[f(x+1) - f(x)]$ for $f \in \mathcal{S}$ [might be easier to check Fourier transforms];
- If X is the isotropic stable process with index $\alpha \in (0\,,2)$, then for all $f \in \mathcal{S}$,

$$(Lf)(x) = \text{const} \cdot \int_{\mathbf{R}^d} \left[\frac{f(x+z) - f(x) - z \cdot (\nabla f)(x) \mathbb{1}_{[0,1]}(\|z\|)}{\|z\|^{d+\alpha}} \right] \mathrm{d}z.$$

  Since $\widehat{Lf}(\xi) \propto -\hat{f}(\xi) \cdot \|\xi\|^\alpha$, L is called the "fractional Laplacian" with fractional power $\alpha/2$. It is sometimes written as $L = -(-\Delta)^{\alpha/2}$; the notation is justified [and explained] by the symbolic calculus of pseudo-differential operators.

Problems for Lecture 9

1. Prove that $P_t(x\,,A) = P_t(A - x)$ for all $t \geq 0$, $x \in \mathbf{R}^d$, and $A \in \mathcal{B}(\mathbf{R}^d)$, where $A - x := \{a - x : a \in A\}$. Conclude that the Chapman–Kolmogorov equation is equivalent to the following formula for $\mathrm{E}\prod_{j=1}^k f_j(x_0 + X_{t_j})$:

$$\int_{\mathbf{R}^d} P_{t_1}(\mathrm{d}x_1) \int_{\mathbf{R}^d} P_{t_2 - t_1}(\mathrm{d}x_2) \cdots \int_{\mathbf{R}^d} P_{t_k - t_{k-1}}(\mathrm{d}x_k) \prod_{j=1}^k f_j(x_0 + \cdots + x_j),$$

using the same notation as Proposition 2.

2. Suppose $Y \in L^1(\mathrm{P})$ is measurable with respect to $\sigma(\{X_r\}_{r \geq t})$ for a fixed non-random $t \geq 0$. Prove that $\mathrm{E}(Y \mid \mathcal{F}_t) = \mathrm{E}(Y \mid X_t)$ a.s.

3. Verify that $-X := \{-X_t\}_{t \geq 0}$ is a Lévy process; compute its transition measures $\tilde{P}_t(x\,,\mathrm{d}y)$ and verify the following duality relationship: For all measurable $f, g : \mathbf{R}^d \to \mathbf{R}_+$ and $z \in \mathbf{R}^d$,

$$\int_{\mathbf{R}^d} f(x)\,\mathrm{d}x \int_{\mathbf{R}^d} g(y)\, P_t(x\,,\mathrm{d}y) = \int_{\mathbf{R}^d} g(y)\,\mathrm{d}y \int_{\mathbf{R}^d} f(x)\, \tilde{P}_t(y\,,\mathrm{d}x).$$

4. Prove that $u(s\,,x) := (P_s f)(x)$ solves [weakly] the partial integro-differential equation

$$\frac{\partial u}{\partial s}(s\,,x) = (Lu)(s\,,x) \quad \text{for all } s > 0 \text{ and } x \in \mathbf{R}^d,$$

subject to $u(0\,,x) = f(x)$.

5. Derive the resolvent equation (3).

6. Verify (4) and (5).

7. First, improve Proposition 11 in the case $p = 2$ as follows: Prove that there exists a unique continuous extension of P_t to all of $L^2(\mathbf{R}^d)$. Denote that by P_t still. Next, define

$$\mathrm{Dom}_2[L] := \left\{ f \in L^2(\mathbf{R}^d) : \int_{\mathbf{R}^d} |\Psi(\xi)|^2 \cdot |\hat{f}(\xi)|^2 \, \mathrm{d}\xi < \infty \right\}.$$

Then prove that $\lim_{t \downarrow 0} t^{-1}(P_t f - f)$ exists, as a limit in $L^2(\mathbf{R}^d)$, for all $f \in \mathrm{Dom}_2[L]$. Identify the limit when $f \in C_c(\mathbf{R}^d)$.

Potential theory

Throughout, $X := \{X_t\}_{t\geq 0}$ denotes a Lévy process on $\mathbf{R}^d$, and $\{\mathcal{F}_t\}_{t\geq 0}$ denotes its natural filtration. We also write $\{P_t\}_{t\geq 0}$ and $\{R_\lambda\}_{\lambda\geq 0}$ for the semigroup and resolvent of X respectively.

Potential measures

For all $\lambda \geq 0$ and $A \in \mathcal{B}(\mathbf{R}^d)$ define

$$U_\lambda(A) := \mathrm{E}\int_0^\infty \mathrm{e}^{-\lambda s}\mathbb{1}_A(X_s)\,\mathrm{d}s = \int_0^\infty \mathrm{e}^{-\lambda s}\mathrm{P}\{X_s \in A\}\,\mathrm{d}s. \tag{1}$$

It is easy to see that when $\lambda > 0$, λU_λ is a Borel probability measure on $\mathbf{R}^d$. Moreover,

$$U_\lambda(A) = (R_\lambda \mathbb{1}_A)(0),$$

where $\{R_\lambda\}_{\lambda\geq 0}$ denotes the resolvent of X. Consequently, (4) (p. 60) implies that

$$\hat{U}_\lambda(\xi) = \frac{1}{\lambda + \Psi(-\xi)} \qquad \text{for all } \xi \in \mathbf{R}^d,\ \lambda > 0. \tag{2}$$

Definition 1. U_λ is called the *λ-potential measure* of X; the 0-potential measure of X is denoted by U instead of U_0. □

Remark 2. In general, $U\,[= U_0]$ is only sigma finite. For example, let X denote Brownian motion on $\mathbf{R}^d$, where $d \geq 3$. Then, for all measurable functions $\varphi : \mathbf{R}^d \to \mathbf{R}_+$,

$$\int_{\mathbf{R}^d} \varphi \, \mathrm{d}U = \mathrm{E} \int_0^\infty \varphi(X_s)\, \mathrm{d}s = \frac{1}{(2\pi)^d} \int_0^\infty \mathrm{d}s \int_{\mathbf{R}^d} \mathrm{d}x \; \varphi(x) \mathrm{e}^{-\|x\|^2/(2s)}$$
$$= \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \mathrm{d}x \; \varphi(x) \int_0^\infty \mathrm{d}s \; \mathrm{e}^{-\|x\|^2/(2s)} \propto \int_{\mathbf{R}^d} \frac{\varphi(x)}{\|x\|^{d-2}}\, \mathrm{d}x.$$

That is, $U(A) \propto \int_A \|x\|^{-d+2}\, \mathrm{d}x$ when $d \geq 3$; note that $U(\mathbf{R}^d) = \infty$. When $d = 3$, $U(A)$ is the socalled "Newtonian potential" of A; and for general $d \geq 3$ it is the "$(d-2)$-dimensional Riesz potential of A." The other λ-potentials can be represented [in all dimensions] in terms of Bessel functions. □

The range of a Lévy process

Define $T(x\,, r)$ to be the first hitting time of $B(x\,, r)$; i.e.,

$$T(x\,, r) := \inf\{s > 0 : X_s \in B(x\,, r)\} \qquad (\inf \varnothing := \infty).$$

Proposition 3. *$U_\lambda(A) > 0$ for all open sets $A \subseteq \mathbf{R}^d$ that contain the origin and all $\lambda \geq 0$. Moreover, for all $x \in \mathbf{R}^d$ and $\lambda, r > 0$,*

$$\frac{U_\lambda(B(x\,, r))}{U_\lambda(B(0\,, 2r))} \leq \mathrm{E}\left(\mathrm{e}^{-\lambda T(x\,,r)};\, T(x\,, r) < \infty\right) \leq \frac{U_\lambda(B(x\,, 2r))}{U_\lambda(B(0\,, r))}.$$

Remark 4. Let τ_λ denote an independent $\mathrm{Exp}(\lambda)$ random variable; that is, $\mathrm{P}\{\tau_\lambda > z\} = \mathrm{e}^{-\lambda z}$ for $z > 0$. Then we can interpret the expectation in Proposition 3 as follows:

$$\mathrm{E}\left(\mathrm{e}^{-\lambda T(x\,,r)};\, T(x\,, r) < \infty\right) = \mathrm{P}\{T(x\,, r) < \tau_\lambda\},$$

or equivalently, as the Laplace transform $\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{P}\{T(x\,, r) < t\}\, \mathrm{d}t$. □

Proof. We can write T in place of $T(x\,, r)$, and note that: (i)

$$U_\lambda(B(x\,, r)) = \mathrm{E}\left(\int_0^\infty \mathrm{e}^{-\lambda(s+T)} \mathbb{1}_{B(x,r)}(X_{s+T} - X_T + X_T)\, \mathrm{d}s\,;\, T < \infty\right);$$

and (ii) $\|X_T - x\| \leq r$ a.s. on $\{T < \infty\}$ because X is cadlag. Therefore, the strong Markov property tells us that

$$U_\lambda(B(x\,, r)) \leq \mathrm{E}\left(\int_0^\infty \mathrm{e}^{-\lambda(s+T)} \mathbb{1}_{B(0,2r)}(X_{s+T} - X_T)\, \mathrm{d}s\,;\, T < \infty\right)$$
$$= \mathrm{E}\left(\mathrm{e}^{-\lambda T};\, T < \infty\right) \cdot U_\lambda(B(0\,, 2r)).$$

This implies the first inequality provided that we prove that $U_\lambda(B(0\,, q))$ is never zero; but that is easy. Indeed, the preceding display tells us that $\sup_{x\in\mathbf{R}^d} U_\lambda(B(x\,, r)) \leq U_\lambda(B(0\,, 2r))$. Therefore, if $U_\lambda(B(0\,, q)) = 0$ for some

$q > 0$, it follows that $U_\lambda(B(x\,,q/2)) = 0$ for all $x \in \mathbf{R}^d$. We can integrate the latter λ-potential over all x to find that

$$0 = \int_{\mathbf{R}^d} \mathrm{d}x \int_{\mathbf{R}^d} U_\lambda(\mathrm{d}z)\; \mathbb{1}_{B(x,q/2)}(z) = \int_{\mathbf{R}^d} U_\lambda(\mathrm{d}z) \int_{\mathbf{R}^d} \mathrm{d}x\; \mathbb{1}_{B(z,q/2)}(x),$$

and this is impossible because the right-most term is simply the volume of $B(0\,,q/2)$ times the total mass of U_λ [which is λ^{-1}].

It remain to establish the second inequality. We continue to write T in place of $T(x\,,r)$, and note that

$$\begin{aligned} U_\lambda(B(x\,,2r)) &\geq \mathrm{E}\left(\int_0^\infty \mathrm{e}^{-\lambda(s+T)} \mathbb{1}_{B(x,2r)}(X_{s+T} - X_T + X_T)\,\mathrm{d}s\,;\; T < \infty\right) \\ &\geq \mathrm{E}\left(\int_0^\infty \mathrm{e}^{-\lambda(s+T)} \mathbb{1}_{B(x,r)}(X_{s+T} - X_T)\,\mathrm{d}s\,;\; T < \infty\right), \end{aligned}$$

because $\|X_T\| \leq r$ a.s. on $\{T < \infty\}$ and hence $\|X_{T+s} - X_T + X_T\| \leq \|X_{s+T} - X_T\| + \|X_T\| \leq 2r$ a.s. on $\{\|X_{s+T} - X_T\| \leq r\}$. Another application of the strong Markov property does the job. □

Proposition 3 has many uses; I mention one next.

Theorem 5 (Hawkes, 1986; Kesten, 1969)**.** *The following are equivalent:*

(1) $X(\mathbf{R}_+)$ *a.s. has positive Lebesgue measure;*

(2) *Every* U_λ *is absolutely continuous with a bounded density;*

(3) $\kappa \in L^1(\mathbf{R}^d)$, *where*

$$\kappa(\xi) := \mathrm{Re}\left(\frac{1}{1 + \Psi(\xi)}\right) \qquad \textit{for all } \xi \in \mathbf{R}^d.$$

Moreover, if $\kappa \notin L^1(\mathbf{R}^d)$, *then* $X(\mathbf{R}_+)$ *has zero Lebesgue measure a.s.*

Example 6. Let X be an isotropic stable process in $\mathbf{R}^d$ with index $\alpha \in (0\,,2]$. Then, the range of X has positive Lebesgue measure iff

$$\int_{\mathbf{R}^d} \frac{\mathrm{d}\xi}{1 + \|\xi\|^\alpha} \propto \int_0^\infty \frac{r^{d-1}}{1 + r^\alpha}\,\mathrm{d}r < \infty \qquad \Leftrightarrow \qquad \alpha > d.$$

In particular, the range of planar Brownian motion has zero Lebesgue measure [theorem of Lévy]. And the range of a one-dimensional isotropic Cauchy process has zero Lebesgue measure as well. By contrast, let X be a one-dimensional asymmetric Cauchy process, so that $\Psi(\xi) = |\xi| + i\theta\xi \ln|\xi|$, for some $\theta \in [-2/\pi\,,2/\pi]$. Note that if X is asymmetric (i.e., $\theta \neq 0$), then

$$\mathrm{Re}\left(\frac{1}{1 + \Psi(\xi)}\right) = \frac{\mathrm{Re}\Psi(\xi)}{|1 + \Psi(\xi)|^2} = \frac{1 + o(1)}{\theta^2|\xi|(\ln|\xi|)^2} \qquad \text{as } |\xi| \to \infty.$$

Therefore, while the range of the symmetric Cauchy process on the line has zero Lebesgue measure, the range of an asymmetric Cauchy process has positive Lebesgue measure. □

Corollary 7. *If X is a d-dimensional Lévy process where $d \geq 2$, then $X(\mathbf{R}_+)$ has zero Lebesgue measure a.s.*

Proof of Theorem 5. Let τ_λ denote an independent $\mathrm{Exp}(\lambda)$ random variable and apply Proposition 3 to find that

$$\frac{U_\lambda(B(x\,,r))}{U_\lambda(B(0\,,2r))} \leq \mathrm{P}\{T(x\,,r) < \tau_\lambda\} \leq \frac{U_\lambda(B(x\,,2r))}{U_\lambda(B(0\,,r))}.$$

Note that $T(x\,,r) < \tau_\lambda$ if and only if $X((0\,,\tau_\lambda)) \cap B(x\,,r) \neq \varnothing$. Equivalently, $T(x\,,r) < \tau_\lambda$ if and only if

$$x \in S(\epsilon) := \left\{ z \in \mathbf{R}^d : \mathrm{dist}(x\,, X((0\,,\tau_\lambda)) < \epsilon \right\}.$$

Consequently,

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{T(x\,,r) < \tau_\lambda\right\} \mathrm{d}x = \int_{\mathbf{R}^d} \mathrm{P}\left\{x \in S(r)\right\} \mathrm{d}x = \mathrm{E}\int_{\mathbf{R}^d} \mathbb{1}_{S(r)}(x)\,\mathrm{d}x = \mathrm{E}|S(r)|,$$

where $|\cdots|$ denotes the Lebesgue measure. Also,

$$\begin{aligned}\int_{\mathbf{R}^d} U_\lambda(B(x\,,r))\,\mathrm{d}x &= \int_{\mathbf{R}^d} \mathrm{d}x \int_{\mathbf{R}^d} U_\lambda(\mathrm{d}z)\, \mathbb{1}_{B(x,r)}(z)\\ &= \int_{\mathbf{R}^d} U_\lambda(\mathrm{d}z) \int_{\mathbf{R}^d} \mathrm{d}x\; \mathbb{1}_{B(z,r)}(x) = \frac{cr^d}{\lambda},\end{aligned}$$

where c denotes the volume of a ball of radius one in $\mathbf{R}^d$. To summarize, we obtain the following: For all $r > 0$,

$$\frac{cr^d}{\lambda U_\lambda(B(0\,,2r))} \leq \mathrm{E}|S(r)| \leq \frac{2^d cr^d}{\lambda U_\lambda(B(0\,,r))} \quad \text{for all } r > 0.$$

Note that $S(r)$ decreases to the closure of $X((0\,,\tau_\lambda))$ as $r \downarrow 0$. Because X is cadlag it has at most countably-many jumps. Therefore, the difference between $X((0\,,\tau_\lambda))$ and its closure is at most countable, hence has zero Lebesgue measure. Therefore, the monotone convergence theorem implies that

$$\frac{c}{\lambda \underline{L}_\lambda} \leq \mathrm{E}\,|X((0\,,\tau_\lambda))| \leq \frac{2^d c}{\lambda \overline{L}_\lambda}, \tag{3}$$

where

$$\underline{L}_\lambda := \liminf_{r \downarrow 0} \frac{U_\lambda(B(0\,,r))}{r^d}, \quad \overline{L}_\lambda := \limsup_{r \downarrow 0} \frac{U_\lambda(B(0\,,r))}{r^d}.$$

Note that

$$\mathrm{E}|X((0\,,\tau_\lambda))| = \lambda \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{E}|X((0\,,t))|\,\mathrm{d}t.$$

[This can be obtained by conditioning first on X.] It follows from the monotonicity of $t \mapsto \mathrm{E}|X((0\,,t))|$ that $\mathrm{E}|X((0\,,\tau_\lambda))| > 0$ if and only if $\mathrm{E}|X((0\,,t))| > 0$ for all $t > 0$. And, in particular, $\mathrm{E}|X((0\,,\tau_\lambda))| > 0$ for some $\lambda > 0$ iff $\mathrm{E}|X((0\,,\tau_\lambda))| > 0$ for all $\lambda > 0$. And this implies that $\underline{L}_\lambda < \infty$ for some $\lambda > 0$ iff $\overline{L}_\lambda < \infty$ for all $\lambda > 0$ iff $\mathrm{E}|X(\mathbf{R}_+)| = \lim_{t\uparrow\infty} \mathrm{E}|X((0\,,t))| > 0$.

Now we begin to prove the theorem.

Suppose $|X(\mathbf{R}_+)| > 0$ a.s. Then, $\mathrm{E}|X(\mathbf{R}_+)| > 0$, whence $\overline{L}_\lambda < \infty$ for all $\lambda > 0$. Recall that $U_\lambda(B(x\,,r)) \le U_\lambda(B(0\,,2r))$, and that the latter is $O(r^d)$. Thus, $U_\lambda(B(x\,,r)) \le \text{const}\cdot|B(0\,,r)|$ uniformly in $x \in \mathbf{R}^d$ and $r > 0$, whence it follows from a covering argument that $U_\lambda(A) \le \text{const}\cdot|A|$ for all $A \in \mathscr{B}(\mathbf{R}^d)$. In other words, (1)⇒(2).

Conversely, if (2) holds, then $\overline{L}_\lambda < \infty$ and hence $\mathrm{E}|X(\mathbf{R}_+)| > 0$ by (3). Choose and fix $R > 0$ so large that $\mathrm{E}|X((0\,,R))| > 0$. By the Markov property, $Z_n := |X((nR\,,(n+1)R])|$ are i.i.d. random variables (why?). Because $|X(\mathbf{R}_+)| \ge \sup_{n\ge1} Z_n$, it follows from the Borel–Cantelli lemma for independent events that $|X(\mathbf{R}_+)| \ge \mathrm{E}|X((0\,,R))| > 0$ a.s. Thus, (2)⇒(1). It remains to prove the equivalence of (3) with (1) and (2).

The key computation is the following: For all uniformly-continuous nonnegative $\phi \in L^1(\mathbf{R}^d)$ such that $\hat\phi \ge 0$,

$$\int_{\mathbf{R}^d} \phi\,\mathrm{d}U_1 = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \hat\phi(\xi)\kappa(\xi)\,\mathrm{d}\xi. \tag{4}$$

Let us first complete the proof of the theorem, assuming (4); and then we establish (4).

Define, for every $r > 0$,

$$f_r(x) := \frac{1}{(2r)^d}\mathbb{1}_{C(0,r)}(x), \qquad \phi_r(x) := (f_r * f_r)(x) \qquad \text{for all } x \in \mathbf{R}^d,$$

where $C(0\,,r) := \{z \in \mathbf{R}^d : \max_{1\le j\le d}|z_j| \le r\}$ is a cube of side $2r$ about the origin. One can check directly that: (i) Every ϕ_r is uniformly continuous and in $L^1(\mathbf{R}^d)$; (ii) $\int_{\mathbf{R}^d}\phi_r(x)\,\mathrm{d}x = 1$ so that $|\hat\phi_r(\xi)| \le 1$; and (iii) $\hat\phi_r(\xi) = |\hat f_r(\xi)|^2 \ge 0$. Moreover,

$$\hat f_r(\xi) = \frac{1}{(2r)^d}\prod_{j=1}^d \int_{-r}^r e^{iz\xi_j}\,\mathrm{d}z = \frac{1}{r^d}\prod_{j=1}^d\int_0^r \sin(z\xi_j)\,\mathrm{d}z = \prod_{j=1}^d \frac{1-\cos(r\xi_j)}{r\xi_j}.$$

Two applications of the triangle inequality show that

$$\frac{1}{(4r)^d}\mathbb{1}_{C(0,r)}(x) \le \phi_r(x) \le \frac{1}{(2r)^d}\mathbb{1}_{C(0,2r)}(x).$$

Therefore,

$$\frac{U_\lambda(C(0\,,r))}{(4r)^d} \le \int \phi_r\,\mathrm{d}U_\lambda \le \frac{U_\lambda(C(0\,,2r))}{(2r)^d}.$$

By the triangle inequality, $B(0\,,r) \subset C(0\,,r)$ and $C(0\,,2r) \subset B(2\sqrt{d}\,r)$. Therefore, it follows that U_λ has a bounded density iff $\limsup_{r\downarrow 0}\int \phi_r\,\mathrm{d}U_\lambda < \infty$ iff $\liminf_{r\downarrow 0}\int \phi_r\,\mathrm{d}U_\lambda < \infty$. If $\kappa \in L^1(\mathbf{R}^d)$, then $\int_{\mathbf{R}^d}\hat{\phi}_r(\xi)\kappa(\xi)\,\mathrm{d}\xi \to \int_{\mathbf{R}^d}\kappa(\xi)\,\mathrm{d}\xi$ as $r \downarrow 0$, since $\hat{\phi}_r(\xi)$ is bounded uniformly in r and ξ, and converges to 1 as $r \downarrow 0$. Thus, (3)⟹(2). And conversely, if (2) holds, then by (4) and Fatou's lemma,

$$\int_{\mathbf{R}^d}\kappa(\xi)\,\mathrm{d}\xi \le \liminf_{r\downarrow 0}\int \hat{\phi}_r(\xi)\kappa(\xi)\,\mathrm{d}\xi = (2\pi)^d \liminf_{r\downarrow 0}\int \phi_r\,\mathrm{d}U_1$$
$$\le (2\pi)^d \liminf_{r\downarrow 0}\frac{U_1(C(0\,,2r))}{(2r)^d} < \infty.$$

It remains to verify the truth of (4). Indeed, we first note that the left-hand side is $(R_1\phi)(0)$ so that whenever $\phi, \hat{\phi} \in L^1(\mathbf{R}^d)$,

$$\int_{\mathbf{R}^d}\phi\,\mathrm{d}U_1 = \mathrm{E}\int_0^\infty \mathrm{e}^{-t}\phi(X_t)\,\mathrm{d}t = \mathrm{E}\int_0^\infty \mathrm{e}^{-t}\,\mathrm{d}t\int_{\mathbf{R}^d}\mathrm{d}\xi\;\mathrm{e}^{-i\xi\cdot X_t}\hat{\phi}(\xi)$$
$$= \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d}\frac{\overline{\hat{\phi}(\xi)}}{1+\Psi(-\xi)}\,\mathrm{d}\xi = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d}\frac{\overline{\hat{\phi}(\xi)}}{1+\Psi(\xi)}\,\mathrm{d}\xi.$$

This proves (4) for all nonnegative $\phi \in L^1(\mathbf{R}^d)$ such that $\hat{\phi} \in L^1(\mathbf{R}^d)$ is nonnegative.[1]

In order to prove (4) in full generality, suppose $\phi \in L^1(\mathbf{R}^d)$ is uniformly continuous and $\hat{\phi} \ge 0$, and let φ_ϵ denote the density of B_ϵ where B is Brownian motion. Clearly, $\phi * \varphi_\epsilon \in L^1(\mathbf{R}^d)$, and its Fourier transform is $\hat{\phi}(\xi)\exp(-\epsilon\|\xi\|^2/2)$ is both nonnegative and in $L^1(\mathbf{R}^d)$. What we have proved so far is enough to imply that (4) holds with ϕ replaced by $\phi * \varphi_\epsilon$; i.e.,

$$\int_{\mathbf{R}^d}(\phi * \varphi_\epsilon)\,\mathrm{d}U_1 = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d}\mathrm{Re}\left(\frac{\overline{\hat{\phi}(\xi)}}{1+\Psi(-\xi)}\right)\mathrm{e}^{-\epsilon\|\xi\|^2/2}\,\mathrm{d}\xi$$

Now we let $\epsilon \downarrow 0$; the left-hand side converges to $\int \phi\,\mathrm{d}U_1$ by Fejer's theorem; and the right-hand side converges to the right-hand side of (4) by the monotone convergence theorem. □

Proof of Corollary 7. Suppose there exists a Lévy process X on $\mathbf{R}^d$, with $d \ge 2$, whose range $X(\mathbf{R}_+)$ has positive Lebesgue measure. Let Ψ denote the Lévy exponent of X and apply Theorem 5 to find that $\int_{\mathbf{R}^d}\mathrm{Re}(1+\Psi(\xi))^{-1}\,\mathrm{d}\xi < \infty$. We will derive a contradiction from this.

Denote by $(a\,,\sigma\,,m)$ the Lévy triple of X. Suppose first that σ is not zero. Because σ is nonnegative definite we can find an orthogonal matrix

[1] If $\phi, \hat{\phi} \in L^1(\mathbf{R}^d)$ then the inversion formula holds. Therefore, ϕ is—up to a null set—uniformly continuous. In fact, $\phi \in C_0(\mathbf{R}^d)$ by the Riemann–Lebesgue lemma.

O and a diagonal matrix Λ of eigenvalues of σ such that $\sigma = O\Lambda O'$. Let $\lambda_1 \le \cdots \le \lambda_d$ denote the ordered eigenvalues of σ and suppose λ_k is the first first strictly positive eigenvalue of σ. Then,

$$\mathrm{Re}\Psi(\xi) \ge \frac{1}{2}\xi'\sigma\xi = \frac{1}{2}\left\|\sqrt{\Lambda}\, O'\xi\right\|^2 = \frac{1}{2}\sum_{j=1}^{d}\lambda_j \left|(O'\xi)_j\right|^2 \ge \frac{\lambda_k}{2}\sum_{j=k}^{d}\left|(O'\xi)_j\right|^2.$$

In particular,

$$\mathrm{Re}\Psi(O\xi) \ge \frac{\lambda_k}{4}\|\xi\|^2 = \frac{\lambda_k}{4}\|O\xi\|^2 \qquad \text{for all } \xi \in \mathcal{C}_k,$$

where $\mathcal{C}_1 := \mathbf{R}^d$ and for all $k = 2, \ldots, d$, $\mathcal{C}_k$ is the cone

$$\mathcal{C}_k := \left\{ z = (z_1\,, \ldots, z_d) \in \mathbf{R}^d : \sum_{j=1}^{k-1} z_j^2 \le 2\sum_{j=k}^{d} z_j^2 \right\}.$$

Because $\mathrm{Re}\Psi(\xi) \ge 0$, this shows that

$$\mathrm{Re}\left(\frac{1}{1+\Psi(O\xi)}\right) = \frac{\mathrm{Re}\Psi(O\xi)}{|1+\Psi(O\xi)|^2} \ge \frac{\mathrm{Re}\Psi(O\xi)}{1+|\Psi(O\xi)|^2} \ge \frac{\lambda_k}{4}\frac{\|O\xi\|^2}{1+|\Psi(O\xi)|^2},$$

and hence,

$$\begin{aligned}\int_{\mathbf{R}^d}\kappa(\xi)\,\mathrm{d}\xi &\ge \frac{\lambda_k}{4}\cdot\int_{\mathcal{C}_k}\frac{\|O\xi\|^2}{1+|\Psi(O\xi)|^2}\,\mathrm{d}\xi \\ &\ge \mathrm{const}\cdot\int_{\mathcal{C}_k}\frac{\mathrm{d}\xi}{1+\|\xi\|^2},\end{aligned}$$

since $|\Psi(\xi)| \le \mathrm{const}\cdot(1+\|\xi\|^2)$. Integrate in spherical coordinates to find that $\kappa \notin L^1(\mathbf{R}^d)$. Theorem 5 tells us that $|X(\mathbf{R}_+)| = 0$ a.s. in this case.

It remains to consider the case that $\sigma = 0$. But then $|\Psi(\xi)| = o(\|\xi\|^2)$ by Theorem 1 (page 34), and therefore,

$$\mathrm{E}\exp\left(i\xi\cdot\frac{X_t}{\sqrt{t}}\right) = \mathrm{e}^{-t\Psi(\xi/\sqrt{t})} \to 1 \qquad \text{as } t \downarrow 0.$$

Therefore, $X_t/\sqrt{t} \to 0$ in probability as $t \downarrow 0$. But then

$$U_\lambda(B(0\,,r)) = \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{P}\left\{\|X_t\| \le r\right\}\mathrm{d}t = r^2\int_0^\infty \mathrm{e}^{-\lambda r^2 s}\mathrm{P}\left\{\|X_{r^2 s}\| \le r\right\}\mathrm{d}s.$$

Therefore, $\liminf_{r\downarrow 0} r^{-d}U_\lambda(B(0\,,r)) = \infty$ by Fatou's lemma. Theorem 5 implies that $|X(\mathbf{R}_+)| = 0$ a.s., and hence the result. □

Problems for Lecture 10

Throughout, X is a Lévy process with exponent Ψ, and let $P_t(x\,,\mathrm{d}y)$, P_t, and U_λ respectively denote the transition measures, semigroup, and λ-potential of X.

1. Suppose ρ is a Borel measure on $\mathbf{R}^d$. Choose and fix some finite $c > 0$. We say that ρ is *c-weakly unimodal* if $\rho(B(x\,,r)) \le c\rho(B(0\,,2r))$ for all $r > 0$. Prove that there exists a finite constant $K := K(c\,,d)$, such that

$$\rho(B(x\,,2r)) \le K\rho(B(0\,,r)) \qquad \text{for all } r > 0.$$

Consequently, $\rho(B(x\,,r)) \le K\rho(B(0\,,r))$ for all $r > 0$. Verify that U^λ is c-weakly unimodal for some $c \in (0\,,\infty)$ (Khoshnevisan and Xiao, 2003).

2. Suppose $U(A \ominus A) < \infty$ and $U(A) > 0$, where U denotes the potential of X, $A \in \mathcal{B}(\mathbf{R}^d)$, and $A \ominus A := \{a - b : a\,, b \in A\}$. Then prove that for all $t > 0$,

$$\sup_{x\in\mathbf{R}^d} \mathrm{P}\left\{x + X_s \in A \text{ for some } s \ge t\right\} \le \frac{\int_t^\infty \mathrm{P}\left\{X_s \in A \ominus A\right\} \mathrm{d}s}{U(A)}.$$

In particular, $\sup_{x\in\mathbf{R}^d} \mathrm{P}\{x + X_s \in A \text{ for some } s \ge t\} \to 0$ as $t \to \infty$.

Recurrence and Transience

The recurrence/transience dichotomy

Definition 1. We say that X is *recurrent* if $\liminf_{t\to\infty} \|X_t\| = 0$ a.s. We say that X is *transient* if $\liminf_{t\to\infty} \|X_t\| = \infty$ a.s. □

In other words, X is recurrent when for all $\epsilon > 0$ we can find possibly random times $0 < t_1 < t_2 < \cdots$, tending to infinity, such that X_{t_n} lies in the ball $B(0\,,\epsilon) := \{z \in \mathbf{R}^d : \|z\| < \epsilon\}$.[1] And X is transient means that for every compact set K there exists a random finite time τ such that $X_{t+\tau} \notin K$ for all $t > 0$. As it turns out, recurrence and transience are dichotomous: Either X is recurrent, or it is transient.

Theorem 2. *The following are equivalent:*

(1) *X is transient;*

(2) *X is not recurrent;*

(3) $U_0(B(0\,,r)) < \infty$ *for all* $r > 0$.

The proof relies on a convenient series of equivalences.

Proposition 3. *The following are equivalent:*

(1) $\sup_{x\in\mathbf{R}^d} U_0(B(x\,,r)) < \infty$ *for all* $r > 0$*;*

(2) $U_0(B(0\,,r)) < \infty$ *for all* $r > 0$*;*

[1] Note that I omit writing "a.s." when it is clear from the context. This is likely to happen in the future as well.

(3) $U_0(B(0\,,r)) < \infty$ *for some* $r > 0$;

(4) $\int_0^\infty \mathbb{1}_{B(0,r)}(X_s)\,\mathrm{d}s < \infty$ *a.s. for all* $r > 0$;

(5) $\lim_{z\uparrow\infty} \sup_{x\in B(0,2r)} \mathrm{P}\{\int_0^\infty \mathbb{1}_{B(x,r)}(X_s)\,\mathrm{d}s > z\} < 1$ *for all* $r > 0$.

Proof. Let

$$J(x\,,r) := \int_0^\infty \mathbb{1}_{B(x,r)}(X_s)\,\mathrm{d}s.$$

Since $U_0(B(a\,,r)) = \mathrm{E}J(a\,,r)$, we readily obtain (1)⇒(2) and (2)⇒(3). Also (2)⇒(1), because $U_\lambda(B(x\,,r)) \le U_\lambda(B(0\,,2r))$ uniformly in x, thanks to Proposition 3 (page 64); we can let $\lambda \downarrow 0$ to obtain (2) from (1).

Suppose (3) holds; i.e., $U_0(B(0\,,r)) < \infty$ for some $r > 0$. It is not hard to see that there exists a number $N \ge 1$ such that $B(0\,,2r)$ is a union of N balls of radius r. [The key observation is that N does not depend on r by scaling.] Consequently, $U_0(B(0\,,2r)) \le N \sup_{x\in\mathbf{R}^d} U_0(B(x\,,r/2)) \le N U_0(B(0\,,r))$ thanks to Proposition 3. This shows that (3)⇒(2), and hence (1)–(3) are equivalent.

Next we prove that (1) and (5) are equivalent: Chebyshev's inequality tells us that (1)⇒(5). Therefore, we are concerned with the complementary implication.

Suppose (5) holds, and fix some $r > 0$. We can find $\gamma > 0$ and $\delta \in (0\,,1)$ such that

$$\sup_{x\in B(0,2r)} \mathrm{P}\left\{J(x\,,r) > \gamma\right\} \le \delta.$$

Choose and fix $a \in B(0\,,2r)$, and define

$$T := \inf\left\{s > 0 : \int_0^t \mathbb{1}_{B(a,r)}(X_s)\,\mathrm{d}s > \gamma\right\} \qquad (\inf\varnothing := \infty).$$

For every integer $n \ge 0$, it is not hard to see that if $\int_0^\infty \mathbb{1}_{B(a,r)}(X_r)\,\mathrm{d}t > (n+1)\gamma$, then certainly $T < \infty$; this follows because the process X has cadlag paths and $B(a\,,r)$ is open. Moreover, $\int_0^T \mathbb{1}_{B(a,r)}(X_s)\,\mathrm{d}s = \gamma$ a.s. on $\{T < \infty\}$. Therefore, we can write

$$\mathrm{P}\left\{\int_0^\infty \mathbb{1}_{B(a,r)}(X_s)\,\mathrm{d}s > (n+1)\gamma\right\} = \mathrm{P}\left\{T < \infty\,, \int_T^\infty \mathbb{1}_{B(a,r)}(X_s)\,\mathrm{d}s > n\gamma\right\}.$$

Because X is cadlag and $B(a\,,r)$ is open, it follows that $X_T \in B(a\,,r)$ a.s. on $\{T < \infty\}$. Therefore, the strong Markov property implies that a.s.,

$$\mathrm{P}\left(\int_0^\infty \mathbb{1}_{B(a,r)}(X_{s+T})\,\mathrm{d}s > n\gamma \;\middle|\; \mathcal{F}_T\right) \le \sup_{x\in B(0,2r)} \mathrm{P}\left\{J(x\,,r) > n\gamma\right\}$$

We can iterate this to find that

$$\sup_{a\in B(0,2r)} \mathrm{P}\left\{J(a\,,r) > (n+1)\gamma\right\} \le \left(\sup_{x\in B(0,2r)} \mathrm{P}\left\{J(x\,,r) > \gamma\right\}\right)^{n+1} \le \delta^{(n+1)}.$$

This shows, in particular, that (5)⇒(1); in fact,

$$\sup_{a\in B(0,2r)} \mathrm{E}\left[\frac{J(a\,,r)}{\gamma}\right] \le \sum_{n=0}^{\infty} \sup_{x\in B(0,2r)} \mathrm{P}\left\{J(x\,,r) > n\gamma\right\} < \infty.$$

And (4)⇒(5) because $J(x\,,r) \le J(0\,,3r)$ uniformly for all $x \in B(0\,,2r)$. Since (1)⇒(4), this proves the equivalence of (1), (2), (3), (4), and (5). □

We now derive the recurrence–transience dichotomy.

Proof of Theorem 2. Clearly, (1) ⇒ (2). And if X is transient, then the last hitting time $L := \sup\{t > 0 : X_t \in B(0\,,r)\}$ of the ball $B(0\,,r)$ is a.s. finite. Therefore, $J(0\,,r) \le L < \infty$ a.s., and Proposition 3 implies that (1) ⇒ (3).

Next, we suppose that (3) holds, so that $J(0\,,r) < \infty$ a.s. for all $r > 0$. If (2) did not hold, that is if X were recurrent, then

$$T_n := \inf\left\{s > n : X_s \in B(0\,,r/2)\right\} \qquad (\inf\varnothing := \infty)$$

would be finite a.s. for all $n \ge 1$. And by the Markov property,

$$\begin{aligned}\mathrm{P}\left\{\int_n^\infty \mathbb{1}_{B(0,r)}(X_t)\,\mathrm{d}t \ge z\right\} &\ge \mathrm{P}\left\{\int_{T_n}^\infty \mathbb{1}_{B(0,r)}(X_t)\,\mathrm{d}t \ge z\right\}\\ &= \mathrm{P}\left\{\int_0^\infty \mathbb{1}_{B(0,r)}(X_{T_n+t})\,\mathrm{d}t \ge z\right\}.\end{aligned}$$

Because $X_{T_n} \in B(0\,,r/2)$ a.s., it would follow from the strong Markov property that

$$\mathrm{P}\left\{\int_n^\infty \mathbb{1}_{B(0,r)}(X_t)\,\mathrm{d}t \ge z\right\} \ge \mathrm{P}\left\{\int_0^\infty \mathbb{1}_{B(0,r/2)}(X_t)\,\mathrm{d}t \ge z\right\} = \mathrm{P}\{J(0\,,r/2) \ge z\}.$$

The left-hand side tends to zero as n goes to ∞, for all $z > 0$. Therefore, $J(0\,,r/2) = 0$ a.s. Because $B(0\,,r/2)$ is open, $X_0 = 0$, and X has cadlag paths, this leads us to a contradiction; i.e., (3) ⇒ (2). It remains to prove that (3) ⇒ (1).

Let us assume that (2) holds; i.e., that X is recurrent. Then the following are all a.s.-finite stopping times:

$$
\begin{aligned}
T_1 &:= \inf\{s > 0 : \|X_s\| > r\}, \\
T_2 &:= \inf\{s > T_1 : \|X_s\| < r/2\}, \\
&\qquad\qquad S_2 := \inf\{s > 0 : \|X_{T_2+t} - X_{T_2}\| > r/2\}, \\
T_3 &:= \inf\{s > T_2 + S_2 : \|X_s\| < r/2\}, \\
&\qquad\qquad S_3 := \inf\{s > 0 : \|X_{T_3+s} - X_{T_3}\| > r/2\},
\end{aligned}
$$

etc. Because X is assumed to be recurrent, these are all a.s.-finite stopping times. And it is easy to see that

$$J(0\,, r) = \int_0^\infty \mathbb{1}_{B(0,r)}(X_s)\,\mathrm{d}s \geq T_1 + \sum_{j=1}^\infty S_j \geq \sum_{j=1}^\infty S_j.$$

By the strong Markov property, the S_j's are i.i.d. And since X is cadlag, the S_j's are a.s. strictly positive. From this we can deduce that $\sum_{j=1}^\infty S_j = \infty$ a.s., [2] and hence (3) $\Rightarrow$ (2). This completes the proof. □

The Port–Stone theorem

The following well-known result of Port and Stone (1967, 1971) characterizes recurrence in terms of the Lévy exponent Ψ.[3]

Theorem 4 (Port and Stone, 1967, 1971)**.** *X is transient iff $\mathrm{Re}(1/\Psi)$ is locally integrable near the origin.*

Example 5. Consider standard Brownian motion in $\mathbf{R}^d$, so that $\Psi(\xi) = \frac{1}{2}\|\xi\|^2$. Then, for every $R > 0$,

$$\int_{B(0,R)} \mathrm{Re}\left(\frac{1}{\Psi(\xi)}\right) \mathrm{d}\xi = \int_{B(0,R)} \frac{2}{\|\xi\|^2}\,\mathrm{d}\xi \propto \int_0^R r^{d-3}\,\mathrm{d}r$$

is infinite if and only if $d \leq 2$. Thus, standard Brownian motion is recurrent iff $d \leq 2$. On the other hand, Brownian motion with nonzero drift is transient in all dimensions $d \geq 1$ [why?].

---

[2] Indeed, we can find $\delta > 0$ such that $p := \mathrm{P}\{S_j \in [\delta, 1/\delta]\} > 0$, and note that $\sum_{j=1}^n S_j \geq \sum_{j=1}^n S_j \mathbb{1}_{[\delta,1/\delta]}(S_j) \sim np$ as $n \to \infty$, thanks to the strong law of large numbers.

[3] Port and Stone prove this for [discrete-time] random walks in 1967. The continuous-time version is proved similarly, in great generality, in 1971.

Somewhat more generally, if X denotes an isotropic stable process in $\mathbf{R}^d$ with index $\alpha \in (0\,,2]$, then for every $R > 0$,

$$\int_{B(0,R)} \text{Re}\left(\frac{1}{\Psi(\xi)}\right) \mathrm{d}\xi \propto \int_{B(0,R)} \frac{1}{\|\xi\|^\alpha} \mathrm{d}\xi \propto \int_0^R r^{d-\alpha-1}\,\mathrm{d}r$$

is infinite if and only if $d > \alpha$. Thus, X is recurrent iff $d \le \alpha$. In particular, $\alpha \ge 1$ is the criterion for recurrence in dimension $d = 1$. And in dimension $d = 2$, only Brownian motion [$\alpha = 2$] is recurrent. In dimensions three of higher, all isotropic stable processes are transient. □

A partial proof of Theorem 4. Define

$$G(r) := \int_{C(2r)} \text{Re}\left(\frac{1}{\Psi(\xi)}\right) \mathrm{d}\xi \qquad \text{for all } r > 0, \tag{1}$$

where $C(t) := \{z \in \mathbf{R}^d : \max_{1\le j\le d} |z_j| \le t\}$ for $t > 0$. Because $B(0\,,R) \subset C(R) \subset B(0\,,R\sqrt{d})$ for all $R > 0$, Theorem 4 is equivalent to the statement that X is recurrent iff $G(r) < \infty$ for all $r > 0$. I will prove half of this, and only make some remarks on the harder half.

Consider probability density functions $\{\varphi_r\}_{r>0}$ defined by

$$\varphi_r(x) := \prod_{j=1}^d \left(\frac{1-\cos(2rx_j)}{2\pi r x_j^2}\right) \qquad \text{for all } x \in \mathbf{R}^d. \tag{2}$$

Then, $\hat{\varphi}_r$ is the normalized Pólya kernel,

$$\hat{\varphi}_r(\xi) = \prod_{j=1}^d \left(1 - \frac{|\xi_j|}{2r}\right)^+ \qquad \text{for every } r > 0 \ (\xi \in \mathbf{R}^d), \tag{3}$$

where $z^+ := \max(z\,,0)$, as usual. Since $1 - \cos z \ge z^2/4$ for all $z \in [-2\,,2]$, we conclude that for every $x, \xi \in \mathbf{R}^d$ and $r > 0$,

$$\varphi_r(x) \ge \left(\frac{r}{2\pi}\right)^d \mathbb{1}_{C(1/r)}(x), \quad \text{and} \quad \hat{\varphi}_r(\xi) \le \mathbb{1}_{C(2r)}(\xi). \tag{4}$$

Because $\varphi_r, \hat{\varphi}_r \in L^1(\mathbf{R}^d)$ are both real-valued functions, and $\hat{U}_\lambda(\xi) = (1+\Psi(\xi))^{-1}$, we may apply Parseval's identity and (4) to find that

$$\begin{aligned} U_\lambda(C(2r)) \ge \int_{\mathbf{R}^d} \hat{\varphi}_r\,\mathrm{d}U_\lambda &= \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \varphi_r(\xi)\text{Re}\left(\frac{1}{\lambda+\Psi(\xi)}\right)\mathrm{d}\xi \\ &\ge \frac{r^d}{(2\pi)^{2d}}\int_{C(1/r)} \text{Re}\left(\frac{1}{\lambda+\Psi(\xi)}\right)\mathrm{d}\xi. \end{aligned}$$

Therefore, Fatou's lemma implies that $U_0(C(2r)) \ge \text{const}\cdot r^d G(1/(2r))$ for all $r > 0$. Theorem 2 then tells us that if X is transient then $G(t) < \infty$ for all $t > 0$. This proves the easier half of the theorem. For the other half we start similarly: By Parseval's identity,

$$\int_{\mathbf{R}^d} \varphi_r \, \mathrm{d}U_\lambda = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \hat{\varphi}_r(\xi) \mathrm{Re}\left(\frac{1}{\lambda + \Psi(\xi)}\right) \mathrm{d}\xi$$
$$\leq \frac{1}{(2\pi)^d} \int_{C(2r)} \mathrm{Re}\left(\frac{1}{\lambda + \Psi(\xi)}\right) \mathrm{d}\xi.$$

The left-most term is at least $(r/(2\pi))^d U_\lambda(C(1/r))$; see (4). Therefore, we can let $\lambda \downarrow 0$ and apply Theorem 2, which tells us that if X is recurrent, then $\liminf_{\lambda \downarrow 0} \int_{C(2r)} \mathrm{Re}(\lambda + \Psi(\xi))^{-1} \, \mathrm{d}\xi = \infty$ for all $r > 0$ From here, the remaining difficulty is to prove that one can always "take the limit inside the expectation." See Port and Stone (1967, 1971) for the [difficult] details in the context of [discrete-time] random walks; the extension to continuous time is performed similarly. □

Problems for Lecture 11

1. Describe exactly when X is recurrent when:

(1) X denote an isotropic stable process in $\mathbf{R}^d$ with index $\alpha \in (0\,,2)$;

(2) X is a nonstandard Brownian motion with exponent $\Psi(\xi) = \frac{1}{2}\|\sigma\xi\|^2$;

(3) X is a process with stable components; i.e., $\Psi(\xi) = \sum_{j=1}^d |\xi_j|^{\alpha_j}$ for $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_d$ all in $(0\,,2]$;

(4) X is a the Poisson process on the line with intensity $\lambda \in (0\,,\infty)$.

2. A Lévy process X is said to satisfy the *sector condition* (Kanda, 1976) if $|\Psi(\xi)| = O(\mathrm{Re}\Psi(\xi))$ as $\|\xi\| \to \infty$.

(1) Verify that every symmetric Lévy process satisfies the sector condition, and find an example that does not.

(2) Prove Theorem 4 [completely!] in the case that X satisfies the sector condition.

3. Let X and X' be two i.i.d. Lévy processes on $\mathbf{R}^d$. Prove that if X is recurrent, then $Y_t := X_t - X'_t$ [the socalled "symmetrization of X"] is also recurrent. Construct an example that shows that the converse is not always true.

4. Let B denote a d-dimensional Brownian motion, and T an independent subordinator with Laplace exponent Φ. Prove that $X \circ T$ is recurrent if and only if $\int_0^1 s^{(d-2)/2} \, \mathrm{d}s/\Phi(s) = \infty$.

5. Suppose X is a transient Lévy process in $\mathbf{R}^d$, and $V : \mathbf{R}^d \to \mathbf{R}_+$ is measurable, bounded, and has compact support. Prove that there exists $\theta > 0$ such that $\mathrm{E}(\mathrm{e}^{\theta \int_0^\infty V(X_s)\,\mathrm{d}s}) < \infty$.

6. Let $B := \{B_t\}_{t \geq 0}$ denote Brownian motion in $\mathbf{R}^d$, where $d \geq 3$. Prove that if $\epsilon > 0$ is fixed but arbitrary, then a.s.:

$$\liminf_{r \downarrow 0} \frac{1}{r^{2+\epsilon}} \int_0^\infty \mathbb{1}_{B(0,r)}(B_s)\,\mathrm{d}s = \infty, \qquad \limsup_{r \downarrow 0} \frac{1}{r^{2-\epsilon}} \int_0^\infty \mathbb{1}_{B(0,r)}(B_s)\,\mathrm{d}s = 0.$$

7. Let X denote a Lévy process on $\mathbf{R}^d$ with exponent Ψ. A point $x \in \mathbf{R}^d$ is *possible* if for all $r > 0$ there exists $t > 0$ such that $\mathrm{P}\{X_t \in B(x\,,r)\} > 0$. Let $\mathcal{P}$ denote the collection of all possible points of X. Demonstrate the following assertions:

(1) $\mathcal{P}$ is a closed additive subsemigroup of $\mathbf{R}^d$;

(2) U_λ is supported on $\mathcal{P}$ for all $\lambda > 0$;

(3) X is recurrent if and only if $\lim_{t\to\infty} \|X_t - x\| = 0$ for all $x \in \mathcal{P}$.

Excessive Functions

Absolute continuity considerations

Definition 1. A Lévy process is said to have *transition densities* if there exists a measurable function $p_t(x) : (0\,,\infty)\times\mathbf{R}^d \to \mathbf{R}_+$ such that $P_t(x\,,dy) = p_t(y-x)\,dy$; equivalently, for all $t > 0$, $x \in \mathbf{R}^d$, and measurable $f : \mathbf{R}^d \to \mathbf{R}_+$,

$$(P_t f)(x) = \mathrm{E}f(x + X_t) = (\tilde{p}_t * \mu_t)(x) = (p_t * \tilde{\mu}_t)(x) = \int_{\mathbf{R}^d} f(y)p_t(y - x)\,dx,$$

where $\tilde{g}(x) := g(-x)$. The function $p_t(x)$ is called the *transition density* of X. □

Example 2. If X is d-dimensional Brownian motion, then it has transition densities,

$$p_t(x) = \frac{1}{(2\pi)^d}\mathrm{e}^{-\|x\|^2/(2t)}.$$

And, more generally, if X is an isotropic stable process with index $\alpha \in (0\,,2]$ with exponent $\Psi(\xi) = c\|\xi\|^\alpha$, then X has transition densities given by the inversion formula, viz.,

$$p_t(x) = \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \mathrm{e}^{-i\xi\cdot x - ct\|\xi\|^\alpha}\,d\xi.$$

This integral can be computed when $\alpha = 2$ [see above], and also when $\alpha = 1$, in which case,

$$p_t(x) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\pi^{(d+1)/2}t^{d/2}}\left(1 + \frac{\|x\|^2}{t^2}\right)^{-(d+1)/2}.$$

See also the remarks that follow Lemma 10 on page 1. □

It is quite natural to ask, "what are the necessary and sufficient conditions on Ψ that guarantee that X has transition densities"? This is an open problem that has been unresolved for a very long time. However it is possible to construct simple sufficient conditions (see Exercise 2 on page 83).

Definition 3. A Lévy process is said to have a λ-*potential density* for some $\lambda \geq 0$ if there exists a measurable function $u_\lambda(x) : \mathbf{R}^d \to \mathbf{R}_+ \cup \{\infty\}$ such that $U_\lambda(\mathrm{d}x) = u_\lambda(x)\,\mathrm{d}x$; equivalently, $(R_\lambda f)(x) = \int_{\mathbf{R}^d} f(y) u_\lambda(y-x)\,\mathrm{d}x$ for all $x \in \mathbf{R}^d$ and measurable $f : \mathbf{R}^d \to \mathbf{R}_+$. If this property holds for all $\lambda > 0$, then we say that X has *potential densities*. □

Remark 4. (1) If it exists, then λu_λ is a [probability] density for $\lambda > 0$ fixed;

(2) If X has transition densities, then X has potential densities, and we can write

$$u_\lambda(x) = \int_0^\infty \mathrm{e}^{-\lambda t} p_t(x)\,\mathrm{d}t.$$

The converse is not true in general; see Exercise 1 below. □

Recall that if $f \in C_0(\mathbf{R}^d)$ then $R_\lambda f \in C_0(\mathbf{R}^d)$; this is called the "Feller property." As it turns out, Lévy processes that have λ-potential densities satisfy a stronger property.

Proposition 5 (Hawkes, 1979)**.** *If X has potential densities $\{u_\lambda\}_{\lambda>0}$, then X has the "strong Feller property;" that is, $R_\lambda f \in C_b(\mathbf{R}^d)$ for all bounded and measurable $f : \mathbf{R}^d \to \mathbf{R}$ and every $\lambda > 0$.*

Proof. Because

$$(R_\lambda f)(x) = \int_{\mathbf{R}^d} f(y) u_\lambda(y-x)\,\mathrm{d}y = (\tilde{f} * u_\lambda)(x) \qquad [\tilde{f}(x) := f(-x)],$$

this proposition follows from a general property of convolutions; namely, $h * g \in C_b(\mathbf{R}^d)$ whenever $h, g \in L^1(\mathbf{R}^d)$ and g is bounded.[1] □

---

[1]Indeed we first find, for all $\epsilon > 0$, a function $k \in C_c(\mathbf{R}^d)$ such that $\|k - h\|_{L^1(\mathbf{R}^d)} \leq \epsilon$, which implies that $\sup_x |(h*g)(x) - (k*g)(x)| \leq \sup_x |g(x)|\epsilon$. Because $k * g$ is continuous, it follows that $h * g$ is uniformly within $(1 + \sup_x |g(x)|)\epsilon$ of a continuous function. This proves the continuity of $h * g$; boundedness is trivial.

Excessive functions

Definition 6. A function $f : \mathbf{R}^d \to \mathbf{R}$ is said to be *λ-excessive* if $f \geq 0$ is measurable and $\exp(-\lambda t)P_t f \uparrow f$ as $t \downarrow 0$ [pointwise]. □

Potentials of nonnegative measurable functions are the archetypal examples. Indeed, if $g : \mathbf{R}^d \to \mathbf{R}_+$ is measurable, then

$$e^{-\lambda t} P_t(R_\lambda g) = \int_t^\infty e^{-\lambda s} P_s g \, ds \uparrow R_\lambda g \quad \text{as } t \downarrow 0.$$

Proposition 7. *If $f : \mathbf{R}^d \to \mathbf{R}_+$ is λ-excessive, then $\alpha R_\alpha f \uparrow f$ as $\alpha \uparrow \infty$.*

Proof. Because of a change of variables,

$$\mu R_{\mu+\lambda} f = \mu \int_0^\infty e^{-(\mu+\lambda)t} P_t f \, dt = \int_0^\infty e^{-s} \, e^{-\lambda(s/\mu)} P_{s/\mu} f \, ds.$$

The monotone convergence theorem implies that $\mu R_{\mu+\lambda} f \uparrow f$ as $\mu \uparrow \infty$. This is equivalent to the statement of the result. □

Proposition 8. *Choose and fix some $\lambda > 0$ and suppose X has potential densities. If f and g are λ-excessive for the same $\lambda > 0$, and $f \leq g$ a.e., then $f(x) \leq g(x)$ for all $x \in \mathbf{R}^d$.*

Proof. We have $\mu(R_{\mu+\lambda} f)(x) = \mu \int_{\mathbf{R}^d} f(y) u_{\mu+\lambda}(y - x) \, dy$ and a similar expression for $\mu(R_{\mu+\lambda} g)(x)$. Therefore, $\mu(R_{\mu+\lambda} f)(x) \leq \mu(R_{\mu+\lambda} g)(x)$ for all $x \in \mathbf{R}^d$ and $\mu > 0$. Now let $\mu \uparrow \infty$. □

Proposition 9 (Hawkes)**.** *If X has a potential densities $\{\pi_\lambda\}_{\lambda>0}$, then for every $\lambda > 0$ there exist a lower semicontinuous modification u_λ of π_λ, and u_λ is itself a λ-potential density.*

From now on, we always choose a lower semicontinuous version of the λ-potential densities when they exist.

Proof. By the Lebesgue differentiation theorem, for every $\lambda > 0$,

$$\pi_\lambda(-x) = \lim_{\epsilon \downarrow 0} \frac{U_\lambda(B(-x\,,\epsilon))}{|B(0\,,\epsilon)|} = \lim_{\epsilon \downarrow 0} \frac{(R_\lambda \mathbb{1}_{B(0,\epsilon)})(x)}{|B(0\,,\epsilon)|} \quad \text{for almost all } x \in \mathbf{R}^d.$$

Therefore, the resolvent equation (3, page 58) tells us that

$$\pi_\lambda(-x) - \pi_\mu(-x) = (\mu - \lambda) \left(\pi_\lambda * \pi_\mu\right)(-x) \qquad \text{for almost all } x \in \mathbf{R}^d.$$

Because $R_\lambda \mathbb{1}_{B(0,\epsilon)}$ is λ-excessive, it follows from Fatou's lemma that: (i) $\mu R_{\mu+\lambda} \pi_\lambda \leq \pi_\lambda$ a.e.; and (ii) $\mu \mapsto \mu R_{\mu+\lambda} \pi_\lambda$ is nondecreasing. Define $u_\lambda := \lim_{\mu \uparrow \infty} \mu(R_{\mu+\lambda} \pi_\lambda)$, and observe that u_λ is lower semicontinuous by the strong Feller property. Also, $u_\lambda \le \pi_\lambda$ a.e. It remains to prove that this a.e.-inequality is an a.e.-equality.

For all bounded measurable $\varphi : \mathbf{R}^d \to \mathbf{R}_+$,

$$\begin{aligned}\int_{\mathbf{R}^d} u_\lambda(x)\varphi(x)\,\mathrm{d}x &= \lim_{\mu\uparrow\infty} \int_{\mathbf{R}^d} \mu\left(R_{\mu+\lambda}\pi_\lambda\right)(x)\varphi(x)\,\mathrm{d}x \\ &= \lim_{\mu\uparrow\infty} \mu \int_{\mathbf{R}^d} \varphi(x)\,\mathrm{d}x \int_{\mathbf{R}^d} U_{\mu+\lambda}(\mathrm{d}y)\ \pi_\lambda(y-x) \\ &= \lim_{\mu\uparrow\infty} \mu \int_{\mathbf{R}^d} (R_\lambda\varphi)(y)\, U_{\mu+\lambda}(\mathrm{d}y) = \lim_{\mu\uparrow\infty} \mu\left(R_{\mu+\lambda}R_\lambda\varphi\right)(0).\end{aligned}$$

Thanks to the resolvent equation (3, page 58), $\mu(R_{\mu+\lambda}R_\lambda\varphi)(0) = (R_\lambda\varphi)(0) - (R_\mu\varphi)(0)$, and $(R_\mu\varphi)(0) = \mathrm{E}\int_0^\infty \mathrm{e}^{-\mu s}\varphi(X_s)\,\mathrm{d}s \to 0$ as $\mu \uparrow \infty$, by the monotone convergence theorem. Consequently,

$$\int_{\mathbf{R}^d} u_\lambda(x)\varphi(x)\,\mathrm{d}x = (R_\lambda\varphi)(0) = \int_{\mathbf{R}^d} \pi_\lambda(x)\varphi(x)\,\mathrm{d}x,$$

which implies that $\pi_\lambda = u_\lambda$ a.e. □

Lévy processes that hit points

Theorem 10. *Let X be a Lévy process on $\mathbf{R}$ that has a bounded and positive λ-potential density. Then $\mathrm{P}\{T_x < \infty\} > 0$ for all $x \in \mathbf{R}$, where $T_x := \inf\{s > 0 : X_s = x\}$.*

Remark 11.
(1) We are restricting attention to one dimension because there are no Lévy processes in $\mathbf{R}^d$ with $d \ge 2$ that have a bounded λ-potential; see Corollary 7 (page 66).
(2) We will see later on [Lemma **??**, page **??**] that if u_λ is continuous for all $\lambda > 0$, then it is positive everywhere. And a sufficient criterion for the continuity of u_λ [for all λ] is that $\int_{-\infty}^{\infty} |1 + \Psi(\xi)|^{-1}\,\mathrm{d}\xi < \infty$. □

Proof. First of all, we can note that $T_z < \infty$ if and only if $z \in X(\mathbf{R}_+)$. Therefore, by the Tonelli theorem,

$$0 < \mathrm{E}|X(\mathbf{R}_+)| = \int_{\mathbf{R}^d} \mathrm{P}\left\{T_z < \infty\right\}\,\mathrm{d}z.$$

Therefore, Theorem 5 tells us that

$$|\mathscr{Z}| > 0, \quad \text{where} \quad \mathscr{Z} := \{z \in \mathbf{R} : \mathrm{P}\{T_z < \infty\} > 0\}.$$

By the Markov property, for all $x \in \mathbf{R}$ and $s > 0$,

$$\mathrm{P}\left\{X_t = x \text{ for some } t > s\right\} = \int_{-\infty}^{\infty} \mathrm{P}\left\{T_{x-y} < \infty\right\} \mu_s(\mathrm{d}y),$$

where μ_s denotes the law of X_s. Consequently, we multiply both sides by $\exp(-\lambda s)$ and integrate [ds] to find that

$$\int_0^\infty e^{-\lambda s}\mathrm{P}\left\{X_t = x \text{ for some } t > s\right\} \mathrm{d}s = \int_{-\infty}^\infty \mathrm{P}\left\{T_{x-y} < \infty\right\} u_\lambda(y)\,\mathrm{d}y.$$

Since $\mathrm{P}\{T_{x-y} < \infty\} > 0$ and $u_\lambda(y) > 0$ for all $y \in x - \mathscr{Z}$, the left-hand side is positive. But the left-hand side is at most $\lambda^{-1}\mathrm{P}\{T_x < \infty\}$. □

Problems for Lecture 12

1. Let $N := \{N_t\}_{t\geq 0}$ denote a rate-α Poisson process on $\mathbf{R}$, where $\alpha \in (0\,,\infty)$. Prove that $X_t := N_t - \alpha t$ is a Lévy process which does not have transition densities, but X has λ-potential densities for every $\lambda \geq 0$.

2. Let X be a Lévy process in $\mathbf{R}^d$ with exponent Ψ such that $e^{-t\mathrm{Re}\Psi} \in L^1(\mathbf{R}^d)$ for all $t > 0$. Then prove that X has transition densities given by

$$p_t(x) := \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} e^{-i\xi\cdot x - t\Psi(\xi)}\,\mathrm{d}\xi.$$

Prove, in addition, that $(t\,,x) \mapsto p_t(x)$ is uniformly continuous for $(t\,,x) \in (\eta\,,\infty)\times\mathbf{R}^d$, for every $\eta > 0$ fixed.

3. Let X denote a Lévy process on $\mathbf{R}^d$ with Lévy exponent Ψ, and T an independent subordinator with Laplace exponent Φ.

(1) Verify that the subordinated process $Y_t := X_{T_t}$ is a Lévy process; compute its Lévy exponent in terms of Ψ and Φ.

(2) Prove that if T and X both have transition densities, then so does Y. Do this by expressing the transition densities of Y in terms of those of X and T.

(3) Suppose T has transition densities and X has a λ-potential density for all $\lambda > 0$. Then prove that Y has a λ-potential density for all $\lambda > 0$; do this by expressing the latter potential densities in terms of the corresponding one for X, and the transition densities of T.

4. Prove that if $\{P_t\}_{t\geq 0}$ is the semigroup of a Lévy process on $\mathbf{R}^d$, then:

(1) $P_t : C_b(\mathbf{R}^d) \to C_b(\mathbf{R}^d)$ for all $t \geq 0$;

(2) If the process has transition densities, then P_t maps bounded and measurable functions to $C_b(\mathbf{R}^d)$ for all $t \geq 0$.

5 (Hawkes' estimates)**.** Consider a Lévy process X on $\mathbf{R}^d$ that has transition densities $p_t(x)$. Prove that the following are equivalent:

(1) p_t is bounded for all $t > 0$;

(2) $p_t \in L^2(\mathbf{R}^d)$ for all $t > 0$;

(3) $\exp(-t\mathrm{Re}\Psi) \in L^1(\mathbf{R}^d)$ for all $t > 0$, where Ψ denotes the Lévy exponent of X.

(Hint: Prove first that $p_t * p_s = p_{t+s}$ a.e.)

6 (Meyer–type inequalities)**.** Let X denote a Lévy process on $\mathbf{R}^d$ and $\{R_\lambda\}_{\lambda>0}$ its resolvent. The main goal of this problem is to establish the following *Meyer–type inequalities*: For all bounded and measurable $f : \mathbf{R}^d \to \mathbf{R}$:

$$\mathrm{E}\left(\sup_{t\geq 0}\left|\int_0^t \mathrm{e}^{-\lambda s} f(X_s)\,\mathrm{d}s\right|^2\right) \leq 10\mathrm{E}\left(\sup_{t\geq 0}\left|\mathrm{e}^{-\lambda t}(R_\lambda f)(X_t)\right|^2\right),$$

and

$$\mathrm{E}\left(\sup_{t\geq 0}\left|\int_0^t \mathrm{e}^{-\lambda s} f(X_s)\,\mathrm{d}s\right|^2\right) \geq \frac{1}{10}\mathrm{E}\left(\sup_{t\geq 0}\left|\mathrm{e}^{-\lambda t}(R_\lambda f)(X_t)\right|^2\right).$$

(1) First prove that for all bounded and measurable $f : \mathbf{R}^d \to \mathbf{R}$,

$$\mathrm{E}\left(\left|\int_0^\infty \mathrm{e}^{-\lambda s} f(X_s)\,\mathrm{d}s\right|^2\right) \leq 4\mathrm{E}\left(\sup_{t\geq 0}\left|\mathrm{e}^{-\lambda t}(R_\lambda f)(X_t)\right|^2\right).$$

(Hint: Expand the left-hand side as a double integral.)

(2) Prove that for all bounded and measurable $f : \mathbf{R}^d \to \mathbf{R}$,

$$\mathrm{e}^{-\lambda t}(R_\lambda f)(X_t) = \mathrm{E}\left(\int_0^\infty \mathrm{e}^{-\lambda s} f(X_s)\,\mathrm{d}s \;\middle|\; \mathcal{F}_t\right) - \int_0^t \mathrm{e}^{-\lambda s} f(X_s)\,\mathrm{d}s,$$

where $\{\mathcal{F}_t\}_{t\geq 0}$ denotes the natural filtration of X.

(3) Conclude by proving the Meyer–type inequalities mentioned above.

Energy and Capacity

Polar and essentially-polar sets

Choose and fix a Borel set $G \subset \mathbf{R}^d$, and define the stopping time T_G to be the *entrance time* of G:

$$T_G := \inf\{s > 0 : X_s \in G \text{ or } X_{s-} \in G\} \qquad (\inf \varnothing := \infty). \tag{1}$$

In other words, T_G is the first time, if ever, that the closure of the range of the process X enters the set G.

Definition 1. A Borel set $G \subseteq \mathbf{R}^d$ is called *polar* if $\mathrm{P}\{T_G < \infty\} = 0$; otherwise G is said to be *nonpolar*. Similarly, G is called *essentially polar* if $\mathrm{P}\{T_{G-x} < \infty\} = 0$ for almost all $x \in \mathbf{R}^d$; otherwise G is deemed *essentially nonpolar*. □

We are abusing notation slightly; "essentially nonpolar" is being treated as an equivalent to "not essentially polar."

We can note that

$$\int_{\mathbf{R}^d} \mathrm{P}\{T_{G-x} < \infty\}\, \mathrm{d}x = \int_{\mathbf{R}^d} \mathrm{P}\left\{\overline{X(\mathbf{R}_+)} \cap (G - x) \neq \varnothing\right\} \mathrm{d}x.$$

But $\overline{X(\mathbf{R}_+)} \cap (G-x)$ is nonempty if and only if x is an element of $G \ominus \overline{X(\mathbf{R}_+)}$. Therefore, Fubini's theorem tells us that

$$G \text{ is essentially polar iff } \mathrm{E}\left|G \ominus \overline{X(\mathbf{R}_+)}\right| = 0.$$

Or equivalently,

$$G \text{ is essentially polar iff } \mathrm{E}\left|\overline{X(\mathbf{R}_+)} \ominus G\right| = 0.$$

(Why?) In particular, set $G := \{x\}$ to see that a singleton is essentially polar if and only if the range of $X(\mathbf{R}_+)$ has positive Lebesgue measure with positive probability. [This ought to seem familiar!]

Our goal is to determine all essentially-polar sets, and relate them to polar sets in most interesting cases. To this end define for all $\lambda > 0$ and Borel probability measures ν and μ on $\mathbf{R}^d$ the following:

$$\mathcal{E}_\lambda(\mu\,,\nu) := \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \hat\mu(\xi)\,\overline{\hat\nu(\xi)}\,\mathrm{Re}\left(\frac{1}{\lambda+\Psi(\xi)}\right)\mathrm{d}\xi. \tag{2}$$

And if $\mu(\mathrm{d}x) = f(x)\,\mathrm{d}x$ and $\nu(\mathrm{d}x) = g(x)\,\mathrm{d}x$, then we may write $\mathcal{E}_\lambda(f\,,g)$ in place of $\mathcal{E}_\lambda(\mu\,,\nu)$ as well. Also define

$$\mathrm{Cap}_\lambda(G) := \left[\inf_{\mu\in M_1(G)} \mathcal{E}_\lambda(\mu\,,\mu)\right]^{-1}, \tag{3}$$

where $M_1(G)$ denotes the collection of all probability measures μ such that $\mu(G^c) = 0$, $\inf\varnothing := \infty$, and $\infty^{-1} := 0$.

Definition 2. $\mathcal{E}_\lambda(\mu\,,\nu)$ is called the *mutual λ-energy between μ and ν*, and $\mathrm{Cap}_\lambda(G)$ the *λ-capacity* of G. □

Our goal is to prove the following:

Theorem 3. *If $\mathrm{Cap}_\lambda(G) > 0$ then G is essentially nonpolar. And if $\mathrm{Cap}_\lambda(G) = 0$, then G is polar.*

Because of the preceding, we care mostly whether or not a given set G has positive λ-capacity. Therefore, let me remind you that $\mathrm{Cap}_\lambda(G) > 0$ if and only if there exists a probability measure μ, supported in G, such that $\int_{\mathbf{R}^d} |\hat\mu(\xi)|^2 \mathrm{Re}(1+\Psi(\xi))^{-1}\,\mathrm{d}\xi < \infty$.

Note that $\mathrm{Cap}_\lambda(G) = \mathrm{Cap}_\lambda(G+x)$ for all $x \in \mathbf{R}^d$. As a consequence of Theorem 3 we find then that G is polar if and only if $\mathrm{P}\{T_{G-x} < \infty\} = 0$ for all $x \in \mathbf{R}^d$. That is: (a) All polar sets are essentially polar; and (b) The difference between polarity and essential polarity is about at most a Lebesgue-null set of shifts of G. As the following shows, there is in fact no difference in almost all cases of interest.

Proposition 4. *Suppose U_λ is absolutely continuous for some $\lambda > 0$. Then, a Borel set G is essentially polar if and only if it is polar.*

An energy identity

Theorem 5 (Foondun and Khoshnevisan, 2010, Corollary 3.7)**.** *If f is a probability density on $\mathbf{R}^d$, then*

$$\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x = \mathcal{E}_\lambda(f\,,f) \qquad \textit{for all } \lambda > 0. \tag{4}$$

Proof. If $f \in C_0(\mathbf{R}^d)$ with $\hat{f} \in L^1(\mathbf{R}^d)$, then (4) follows from direct computations. Indeed, we can use the fact that $\hat{u}_\lambda(\xi) = \mathrm{Re}(\lambda + \Psi(\xi))^{-1} \geq 0$ [see (2, p. 63)] together with Fubini's theorem and find that

$$\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x = \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\hat{f}(\xi)|^2 \mathrm{Re}\left(\frac{1}{\lambda + \Psi(\xi)}\right) \mathrm{d}\xi. \tag{5}$$

But in the present case, Fubini's theorem is not applicable. Instead, we proceed in two steps: First we prove that

$$\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x \geq \mathcal{E}_\lambda(f\,,f). \tag{6}$$

This holds trivially unless the left-hand side is finite, which we now assume is the case. Because f is a density function, Lusin's theorem tells us that for all $\delta > 0$ there exists a compact set $K_\delta \subset \mathbf{R}^d$ such that

$$\int_{K_\delta^c} f(x)\,\mathrm{d}x \leq \delta, \quad \text{and} \quad R_\lambda f \text{ is continuous on } K_\delta.$$

In particular,

$$\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x \geq \int_{K_\delta} (R_\lambda f)(x) f(x)\,\mathrm{d}x = \lim_{\epsilon \downarrow 0} \int_{K_\delta} \left((R_\lambda f) * \varphi_\epsilon\right)(x) f(x)\,\mathrm{d}x,$$

where φ_ϵ denotes the density of B_ϵ for a d-dimensional Brownian motion B. Let $f_\delta := f \mathbb{1}_{K_\delta}$ and note that $\hat{f}_\delta \to \hat{f}$, pointwise, as $\delta \downarrow 0$.

Since $(R_\lambda f) * \varphi_\epsilon = R_\lambda(f * \varphi_\epsilon) \geq R_\lambda(f_\delta * \varphi_\epsilon)$ and $\varphi_\epsilon = \varphi_{\epsilon/2} * \varphi_{\epsilon/2}$, we can apply Tonelli's theorem to find that

$$\begin{aligned}
\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x &\geq \liminf_{\epsilon \downarrow 0} \int_{\mathbf{R}^d} \left(R_\lambda(f_\delta * \varphi_\epsilon)\right)(x) f_\delta(x)\,\mathrm{d}x \\
&= \liminf_{\epsilon \downarrow 0} \int_{\mathbf{R}^d} \left(R_\lambda(f_\delta * \varphi_{\epsilon/2})\right)(x) \left(f_\delta * \varphi_{\epsilon/2}\right)(x)\,\mathrm{d}x \\
&= \frac{1}{(2\pi)^d} \liminf_{\epsilon \downarrow 0} \int_{\mathbf{R}^d} |\hat{f}_\delta(\xi)|^2 \mathrm{e}^{-\epsilon\|\xi\|^2/2} \mathrm{Re}\left(\frac{1}{\lambda + \Psi(-\xi)}\right) \mathrm{d}\xi,
\end{aligned}$$

thanks to (5). This proves that

$$\int_{\mathbf{R}^d} (R_\lambda f)(x) f(x)\,\mathrm{d}x \geq \frac{1}{(2\pi)^d} \liminf_{\delta \downarrow 0} \int_{\mathbf{R}^d} |\hat{f}_\delta(\xi)|^2 \mathrm{Re}\left(\frac{1}{1 + \Psi(\xi)}\right) \mathrm{d}\xi,$$

and Fatou's lemma proves (6). The converse bound is much easier: We merely note that, as above,

$$\begin{aligned}
\int_{\mathbf{R}^d} \left(R_\lambda(f * \varphi_\epsilon)\right)(x) f(x)\,\mathrm{d}x &= \int_{\mathbf{R}^d} \left(R_\lambda(f * \varphi_{\epsilon/2})\right)(x) \left(f * \varphi_{\epsilon/2}\right)(x)\,\mathrm{d}x \\
&= \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} |\hat{f}(\xi)|^2 \mathrm{e}^{-\epsilon\|\xi\|^2/2} \mathrm{Re}\left(\frac{1}{\lambda + \Psi(\xi)}\right) \mathrm{d}\xi.
\end{aligned}$$

Then we let $\epsilon \downarrow 0$; the right-most term converges to $\mathcal{E}_\lambda(f\,,f)$ by the dominated convergence theorem, and the lim inf of the left-most term is at most $\int_{\mathbf{R}^d}(R_\lambda f)(x)f(x)\,dx$ by Fatou's lemma. □

Proof of Theorem 3

Theorem 3 will follow immediately from Lemmas 7 and 9 below.

Define

$$(J_\lambda f)(x) := \int_0^\infty e^{-\lambda s} f(x + X_s)\,ds.$$

Lemma 6. *For all $f \in L^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d)$ and $\lambda > 0$,*

$$\int_{\mathbf{R}^d} E\left[(J_\lambda f)(x)\right]\,dx = \frac{1}{\lambda}, \qquad \int_{\mathbf{R}^d} E\left(|(J_\lambda f)(x)|^2\right)\,dx = \frac{1}{\lambda}\mathcal{E}_\lambda(f\,,f).$$

Proof. The first computation follows because f is a probability density and hence $\int_{\mathbf{R}^d}(J_\lambda f)(x)\,dx = \lambda^{-1}$. Now we begin with the second computation of the lemma:

$$\begin{aligned}
E\left(|(J_\lambda f)(x)|^2\right) &= 2\int_0^\infty e^{-\lambda s}\,ds \int_s^\infty e^{-\lambda t}\,dt\; E\left[f(x+X_s)\cdot f(x+X_t)\right] \\
&= 2\int_0^\infty e^{-\lambda s}\,ds \int_s^\infty e^{-\lambda t}\,dt\; E\left[f(x+X_s)\cdot (P_{t-s}f)(x+X_s)\right] \\
&= 2\int_0^\infty e^{-2\lambda s} E\left[f(x+X_s)\cdot (R_\lambda f)(x+X_s)\right]\,ds,
\end{aligned}$$

thanks to the Markov property. Therefore,

$$\int_{\mathbf{R}^d} E\left(|(J_\lambda f)(x)|^2\right)\,dx = \frac{1}{\lambda}\int_{\mathbf{R}^d} f(y)\cdot(R_\lambda f)(y)\,dy.$$

And the lemma follows from Theorem 5. □

Lemma 7. *Regardless of the value of $\lambda > 0$,*

$$E\left(|G \ominus X(\mathbf{R}_+)|\right) = \int_{\mathbf{R}^d} P\left\{T_{G-x} < \infty\right\}\,dx \geq \frac{1}{\lambda}\cdot \mathrm{Cap}_\lambda(G).$$

Remark 8. It is important to note that $T_{G-x} < \infty$ if and only if the Lévy process$x + X_t$ [which starts at $x \in \mathbf{R}^d$ at time zero] everh hits G; more precisely, there exists $t > 0$ such that $x + X_t \in G$ or $x + X_{t-} \in G$. Therefore, the preceding states that if G has positive λ-capacity, then X hits G, starting from almost every starting point $x \in \mathbf{R}^d$. In fact, this property is one way of thinking about the essential nonpolarity of G. □

Proof. Let us begin with a simple fact from classical function theory.

The Paley–Zygmund inequality.[1] *Suppose* $Y : \mathbf{R}^d \times \Omega \to \mathbf{R}_+$ *is nonnegative and measurable,* $Y \not\equiv 0$*,* $Y \in L^2(\mathbf{R}^d \times \Omega)$*, and* $\int_{\mathbf{R}^d} \mathrm{E}Y(x)\,\mathrm{d}x = c > 0$*. Then,*

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{Y(x) > 0\right\} \mathrm{d}x \geq \frac{c^2}{\int_{\mathbf{R}^d} \mathrm{E}\left(|Y(x)|^2\right) \mathrm{d}x},$$

where $1/\infty := 0$.

Let f be a probability density that is supported on the closed [say] ϵ-enlargment G^ϵ of G. We apply Lemma 6 together with the Paley-Zygmund inequality [with $Y(x) := (J_\lambda f)(x)$] and obtain

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{(J_\lambda f)(x) > 0\right\} \mathrm{d}x \geq \frac{1}{\lambda \cdot \mathscr{E}_\lambda(f\,,f)}.$$

If $(J_\lambda f)(x) > 0$, then certainly $x + X_s \in G^\epsilon$ for some $s > 0$; i.e., $T_{G^\epsilon - x} < \infty$. Therefore,

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{T_{G^\epsilon - x} < \infty\right\} \mathrm{d}x \geq \frac{1}{\lambda} \cdot \sup_g \frac{1}{\mathscr{E}_\lambda(g\,,g)},$$

where the supremum is taken over all probability densities g that are supported on G^ϵ. Let h_ϵ be a probability density, supported on $B(0\,,\epsilon)$, and observe that $\rho * h_\epsilon$ is a probability density supported on G^ϵ whenever $\rho \in M_1(G)$. Because of (2), $\mathscr{E}_\lambda(\rho * h_\epsilon\,,\rho * h_\epsilon) \leq \mathscr{E}_\lambda(\rho\,,\rho)$, and hence

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{T_{G^\epsilon - x} < \infty\right\} \mathrm{d}x \geq \frac{1}{\lambda} \cdot \mathrm{Cap}_\lambda(G).$$

Note that

$$\bigcap_{\epsilon>0} \left\{T_{G^\epsilon - x} < \infty\right\} = \bigcap_{\epsilon>0} \left\{x + X(\mathbf{R}_+) \cap G^\epsilon \neq \varnothing\right\} = \left\{x + \overline{X(\mathbf{R}_+)} \cap G \neq \varnothing\right\}.$$

Therefore,

$$\int_{\mathbf{R}^d} \mathrm{P}\left\{x + \overline{X(\mathbf{R}_+)} \cap G \neq \varnothing\right\} \mathrm{d}x \geq \frac{1}{\lambda} \cdot \mathrm{Cap}_\lambda(G).$$

Now the left-hand side is the expectation of the Lebesgue measure of the random set $G \ominus \overline{X(\mathbf{R}_+)}$ [check!]. Because X is cadlag, the set difference between $X(\mathbf{R}_+)$ and its closure has zero measure (in fact, is countable).

[1] Here is the proof: By the Cauchy–Schwarz inequality,

$$\begin{aligned} c = \int_{\mathbf{R}^d} \mathrm{E}f(x)\,\mathrm{d}x &= \int_{\mathbf{R}^d \times \Omega} \mathbb{1}_{\{f>0\}}(x\,,\omega) \cdot f(x\,,\omega)\,\mathrm{d}x\,\mathrm{P}(\mathrm{d}\omega) \\ &\leq \left(\int_{\mathbf{R}^d \times \Omega} \mathbb{1}_{\{f>0\}}(x\,,\omega)\,\mathrm{d}x\,\mathrm{P}(\mathrm{d}\omega) \cdot \int_{\mathbf{R}^d \times \Omega} |f(x\,,\omega)|^2\,\mathrm{d}x\,\mathrm{P}(\mathrm{d}\omega)\right)^{1/2} \\ &= \left(\int_{\mathbf{R}^d} \mathrm{P}\{f(x) > 0\}\,\mathrm{d}x\right)^{1/2} \cdot \left(\int_{\mathbf{R}^d} \mathrm{E}\left(|f(x)|^2\right) \mathrm{d}x\right)^{1/2}. \end{aligned}$$

□

Therefore, the Lebesgue measure of $G\ominus\overline{X(\mathbf{R}_+)}$ is the same as the Lebesgue measure of $G \ominus X(\mathbf{R}_+)$. This proves the result. □

Lemma 9. $\mathrm{P}\{T_G \le n\} \le \mathrm{e}^{\lambda n} \cdot \mathrm{Cap}_\lambda(G)$ *for all* $n, \lambda > 0$.

Proof. This is trivial unless

$$\mathrm{P}\left\{T_G \le n\right\} > 0, \tag{7}$$

which we assume is the case.

For all measurable $f : \mathbf{R}^d \to \mathbf{R}_+$,

$$\begin{aligned}
\mathrm{E}\left((J_\lambda f)(0) \mid \mathcal{F}_{T_G\wedge n}\right) &\ge \int_{T_G}^\infty \mathrm{e}^{-\lambda s}\mathrm{E}\left[f(X_s) \,\middle|\, \mathcal{F}_{T_G\wedge n}\right] \mathrm{d}s \cdot \mathbb{1}_{\{T_G \le n\}}\\
&= \mathrm{e}^{-\lambda n}\int_0^\infty \mathrm{e}^{-\lambda s}\mathrm{E}\left[f(X_{s+T_G\wedge n}) \,\middle|\, \mathcal{F}_{T_G}\right] \mathrm{d}s \cdot \mathbb{1}_{\{T_G \le n\}}\\
&= \mathrm{e}^{-\lambda n}\int_0^\infty \mathrm{e}^{-\lambda s}(P_s f)(X_{T_G\wedge n})\, \mathrm{d}s \cdot \mathbb{1}_{\{T_G \le n\}},
\end{aligned}$$

thanks to the strong Markov property. Therefore,

$$\mathrm{E}\left((J_\lambda f)(0) \mid \mathcal{F}_{T_G\wedge n}\right) \ge \mathrm{e}^{-\lambda n}(R_\lambda f)(X_{T_G}) \cdot \mathbb{1}_{\{T_G \le n\}}.$$

The expectation of the term on the left-hand side is 1, thanks to Lemma 6 and the optional stopping theorem. Therefore,

$$\begin{aligned}
1 &\ge \mathrm{e}^{-\lambda n}\mathrm{E}\left[(R_\lambda f)(X_{T_G}) \,\middle|\, T_G \le n\right] \cdot \mathrm{P}\{T_G \le n\}\\
&= \mathrm{e}^{-\lambda n}\int_{\mathbf{R}^d} (R_\lambda f)\, \mathrm{d}\rho \cdot \mathrm{P}\{T_G \le n\},
\end{aligned}$$

where $\rho(A) := \mathrm{P}(X_{T_G} \in A \mid T_G \le n)$. In accord with (7), $\rho \in M_1(G)$.

We apply the preceding with $f := \rho * \varphi_\epsilon$, where φ_ϵ denotes the density of B_ϵ for a d-dimensional Brownian motion. Because

$$\int_{\mathbf{R}^d} (R_\lambda f)\, \mathrm{d}\rho = \int_{\mathbf{R}^d} \left(R_\lambda(\rho * \varphi_{\epsilon/2})\right)(x) \left(\rho * \varphi_{\epsilon/2}\right)(x)\, \mathrm{d}x,$$

it follows from Theorem 5 that

$$\mathrm{e}^{\lambda n} \ge \frac{1}{(2\pi)^d}\int_{\mathbf{R}^d} \mathrm{e}^{-\epsilon\|\xi\|^2/2}|\hat\rho(\xi)|^2 \mathrm{Re}\left(\frac{1}{\lambda + \Psi(\xi)}\right) \mathrm{d}\xi \cdot \mathrm{P}\left\{T_G \le n\right\}.$$

Let $\epsilon \downarrow 0$ and appeal to the monotone convergence theorem to finish. □

Problems for Lecture 13

1. Prove that $\mathrm{Cap}_\lambda(G) > 0$ for some $\lambda > 0$ iff $\mathrm{Cap}_\lambda(G) > 0$ for all $\lambda > 0$.

2. Prove Proposition 4. (Hint: Inspect the proof of Theorem 10 on page 82.)

Bibliography

Jean Bertoin. *Lévy Processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996. ISBN 0-521-56243-0.

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

R. M. Blumenthal. An extended Markov property. *Trans. Amer. Math. Soc.*, 85: 52–72, 1957. ISSN 0002-9947.

Claude Dellacherie and Paul-André Meyer. *Probabilities and Potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1978. ISBN 0-7204-0701-X.

S. N. Evans. The range of a perturbed Lévy process. *Probab. Theory Related Fields*, 81(4):555–557, 1989. ISSN 0178-8051. doi: 10.1007/BF00367302. URL `http://dx.doi.org/10.1007/BF00367302`.

P. J. Fitzsimmons. On Cauchy's functional equation, 1995. preprint (available at `http://math.ucsd.edu/~pfitz/preprints.html`).

Mohammud Foondun and Davar Khoshnevisan. On the stochastic heat equation with spatially-colored random forcing, 2010. Preprint can be found at `http://www.math.utah.edu/~davar/PPT/FK_COLOREDNOISE`.

Bert Fristedt. The behavior of increasing stable processes for both small and large times. *J. Math. Mech.*, 13:849–856, 1964.

Bert Fristedt. Sample functions of stochastic processes with stationary, independent increments. In *Advances in probability and related topics, Vol. 3*, pages 241–396. Dekker, New York, 1974.

Bert E. Fristedt. Sample function behavior of increasing processes with stationary, independent increments. *Pacific J. Math.*, 21:21–33, 1967. ISSN 0030-8730.

S. E. Graversen. "Polar"-functions for Brownian motion. *Z. Wahrsch. Verw. Gebiete*, 61(2):261–270, 1982. ISSN 0044-3719. doi: 10.1007/BF01844636. URL `http://dx.doi.org/10.1007/BF01844636`.

Philip S. Griffin. Laws of the iterated logarithm for symmetric stable processes. *Z. Wahrsch. Verw. Gebiete*, 68(3):271–285, 1985. ISSN 0044-3719. doi: 10.1007/BF00532641. URL http://dx.doi.org/10.1007/BF00532641.

Georg Hamel. Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x + y) = f(x) + f(y)$. *Math. Ann.*, 60 (3):459–462, 1905. ISSN 0025-5831. doi: 10.1007/BF01457624. URL http://dx.doi.org/10.1007/BF01457624.

J. Hawkes. Local times as stationary processes. In *From local times to global geometry, control and physics (Coventry, 1984/85)*, volume 150 of *Pitman Res. Notes Math. Ser.*, pages 111–120. Longman Sci. Tech., Harlow, 1986.

John Hawkes. Potential theory of Lévy processes. *Proc. London Math. Soc. (3)*, 38(2):335–352, 1979. ISSN 0024-6115. doi: 10.1112/plms/s3-38.2.335. URL http://dx.doi.org/10.1112/plms/s3-38.2.335.

Edwin Hewitt and Karl Stromberg. *Real and Abstract Analysis*. Springer-Verlag, New York, 1965.

Joseph Horowitz. The Hausdorff dimension of the sample path of a subordinator. *Israel J. Math.*, 6:176–182, 1968. ISSN 0021-2172.

Kiyosi Itô. On stochastic processes. I. (Infinitely divisible laws of probability). *Jap. J. Math.*, 18:261–301, 1942.

Mamuro Kanda. Two theorems on capacity for Markov processes with stationary independent increments. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, pages 159Ð–165, 1976.

Harry Kesten. *Hitting Probabilities of Single Points for Processes with Stationary Independent Increments*. Memoirs of the American Mathematical Society, No. 93. American Mathematical Society, Providence, R.I., 1969.

A. Ya. Khintchine. Zwei Sätze über stochastische Prozesse mit stablien Verteilungen. *Mat. Sbornik*, 3:577–584, 1938. in Russian, with German summary.

A. Ya. Khintchine. Sur la croissance locale des processus stochastique homogènes ś acroissements indépendants. *Izv. Akad. Nauk SSSR, Ser. Math.*, pages 487–508, 1939.

Davar Khoshnevisan. *Multiparameter Processes*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2002. ISBN 0-387-95459-7. An introduction to random fields.

Davar Khoshnevisan. *Probability*, volume 80 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007. ISBN 978-0-8218-4215-7; 0-8218-4215-3.

Davar Khoshnevisan and Yimin Xiao. Weak unimodality of finite measures, and an application to potential theory of additive Lévy processes. *Proc. Amer. Math. Soc.*, 131(8):2611–2616 (electronic), 2003. ISSN 0002-9939. doi: 10.1090/S0002-9939-02-06778-3. URL http://dx.doi.org/10.1090/S0002-9939-02-06778-3.

Davar Khoshnevisan and Yimin Xiao. Harmonic analysis of additive Lévy processes. *Probab. Theory Related Fields*, 145(3-4):459–515, 2009. ISSN 0178-8051. doi: 10.1007/s00440-008-0175-5. URL http://dx.doi.org/10.1007/s00440-008-0175-5.

Davar Khoshnevisan, Yimin Xiao, and Yuquan Zhong. Measuring the range of an additive Lévy process. *Ann. Probab.*, 31(2):1097–1141, 2003. ISSN 0091-1798. doi: 10.1214/aop/1048516547. URL `http://dx.doi.org/10.1214/aop/1048516547`.

J. F. C. Kingman. *Regenerative Phenomena.* John Wiley & Sons Ltd., London-New York-Sydney, 1972. Wiley Series in Probability and Mathematical Statistics.

A. N. Kolmogorov. Sulla forma generale di un processo stocastico omogeneo. *R. C. Acad. Lincei*, 15(6):805–806, 866–869, 1932.

Andreas E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications.* Universitext. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-31342-7; 3-540-31342-7.

Jean-François Le Gall. Sur les fonctions polaires pour le mouvement brownien. In *Séminaire de Probabilités, XXII*, volume 1321 of *Lecture Notes in Math.*, pages 186–189. Springer, Berlin, 1988. doi: 10.1007/BFb0084136. URL `http://dx.doi.org/10.1007/BFb0084136`.

Paul Lévy. Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Pisa*, 3:337–366; ibid. 4, 217–218, 1934. Reprinted in *Euvre de Paul Lévy*, Vol. 4, Gauthier–Villars, Paris, 1980.

Paul Lévy. *Théorie de l'addition des variables aléatoires.* Gauthier–Villars, Paris, 1937.

Paul Lévy. An extension of the Lebesgue measure of linear sets. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II*, pages 273–287. Univ. California Press, Berkeley, Calif., 1961.

Paul Lévy. *Processus stochastiques et mouvement brownien.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Éditions Jacques Gabay, Sceaux, 1992. ISBN 2-87647-091-8. Followed by a note by M. Loève, Reprint of the second (1965) edition.

P. W. Millar. Path behavior of processes with stationary independent increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 17:53–73, 1971.

Itrel Monroe. On the γ-variation of processes with stationary independent increments. *Ann. Math. Statist.*, 43:1213–1220, 1972. ISSN 0003-4851.

T. S. Mountford. Time inhomogeneous Markov processes and the polarity of single points. *Ann. Probab.*, 17(2):573–585, 1989. ISSN 0091-1798. URL `http://links.jstor.org/sici?sici=0091-1798(198904)17:2<573:TIMPAT>2.0.CO;2-N&origin=MSN`.

Sidney C. Port and Charles J. Stone. Hitting time and hitting places for non-lattice recurrent random walks. *J. Math. Mech.*, 17:35–57, 1967.

Sidney C. Port and Charles J. Stone. Infinitely divisible processes and their potential theory. *Ann. Inst. Fourier (Grenoble)*, 21(2):157–275; ibid. 21 (1971), no. 4, 179–265, 1971. ISSN 0373-0956.

Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999. ISBN 0-521-55302-4. Translated from the 1990 Japanese original, Revised by the author.

A. V. Skorohod. Asymptotic formulas for stable distribution laws. In *Select. Transl. Math. Statist. and Probability, Vol. 1*, pages 157–161. Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., 1961.

E. S. Štatland. On local properties of processes with independent increments. *Teor. Verojatnost. i Primenen.*, 10:344–350, 1965. ISSN 0040-361x.

Ming Yang. On a theorem in multi-parameter potential theory. *Electron. Comm. Probab.*, 12:267–275 (electronic), 2007. ISSN 1083-589X.