Advanced Courses in Mathematics CRM Barcelona

Davar Khoshnevisan
René Schilling

From Lévy-Type Processes to Parabolic SPDEs

CENTRE DE RECERCA MATEMÀTICA

Birkhäuser

Birkhäuser

Advanced Courses in Mathematics
CRM Barcelona

Centre de Recerca Matemàtica

Managing Editor:
Enric Ventura

More information about this series at http://www.springer.com/series/5038

Davar Khoshnevisan • René Schilling

From Lévy-Type Processes to Parabolic SPDEs

Editors for this volume:
Lluís Quer-Sardanyons, Universitat Autònoma de Barcelona
Frederic Utzet, Universitat Autònoma de Barcelona

Birkhäuser

Davar Khoshnevisan
Department of Mathematics
University of Utah
Salt Lake City, Utah, USA

René Schilling
Institut für Mathematische Stochastik
Fachrichtung Mathematik
TU Dresden
Dresden, Germany

ISSN 2297-0304 ISSN 2297-0312 (electronic)
Advanced Courses in Mathematics - CRM Barcelona
ISBN 978-3-319-34119-4 ISBN 978-3-319-34120-0 (eBook)
DOI 10.1007/978-3-319-34120-0

Library of Congress Control Number: 2016962054

Mathematics Subject Classification (2010): 60H15, 60G51, 35K57, 60H20, 60J25, 60J35

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Part I An Introduction to Lévy and Feller Processes

Part II Invariance and Comparison Principles for Parabolic Stochastic Partial Differential Equations

In fond memory of Don Burkholder (1927–2013),
Wenbo V. Li (1963–2013), and Marc Yor (1949–2014).

Part I

An Introduction to Lévy and Feller Processes

By René L. Schilling

Preface

These lecture notes are an extended version of my lectures on Lévy and Lévy-type processes given at the *Second Barcelona Summer School on Stochastic Analysis* organized by the *Centre de Recerca Matemàtica* (CRM). The lectures are aimed at advanced graduate and PhD students. In order to read these notes, one should have sound knowledge of measure theoretic probability theory and some background in stochastic processes, as it is covered in my books *Measures, Integals and Martingales* [54] and *Brownian Motion* [56].

My purpose in these lectures is to give an introduction to Lévy processes, and to show how one can extend this approach to space inhomogeneous processes which behave locally like Lévy processes. After a brief overview (Chapter 1) I introduce Lévy processes, explain how to characterize them (Chapter 2) and discuss the quintessential examples of Lévy processes (Chapter 3). The Markov (loss of memory) property of Lévy processes is studied in Chapter 4. A short analytic interlude (Chapter 5) gives an introduction to operator semigroups, resolvents and their generators from a probabilistic perspective. Chapter 6 brings us back to generators of Lévy processes which are identified as pseudo-differential operators whose symbol is the characteristic exponent of the Lévy process. As a by-product we obtain the Lévy–Khintchine formula.

Continuing this line, we arrive at the first construction of Lévy processes in Chapter 7. Chapter 8 is devoted to two very special Lévy processes: (compound) Poisson processes and Brownian motion. We give elementary constructions of both processes and show how and why they are special Lévy processes, indeed. This is also the basis for the next chapter (Chapter 9) where we construct a random measure from the jumps of a Lévy process. This can be used to provide a further construction of Lévy processes, culminating in the famous Lévy–Itô decomposition and yet another proof of the Lévy–Khintchine formula.

A second interlude (Chapter 10) embeds these random measures into the larger theory of random orthogonal measures. We show how we can use random orthogonal measures to develop an extension of Itô's theory of stochastic integrals for square-integrable (not necessarily continuous) martingales, but we restrict ourselves to the bare bones, i.e., the L^2-theory. In Chapter 11 we introduce Feller processes as the proper spatially inhomogeneous brethren of Lévy processes, and we show how our proof of the Lévy–Khintchine formula carries over to this setting. We will see, in particular, that Feller processes have a **symbol** which is the state-space-dependent analogue of the characteristic exponent of a Lévy process. The symbol describes the process and its generator. A probabilistic way to calculate the symbol and some first consequences (in particular the semimartingale decomposition of Feller processes) is discussed in Chapter 12; we also show that the symbol contains information on global properties of the process, such as conservativeness. In the final Chapter 13, we summarize (mostly without proofs) how other path properties of a Feller process can be obtained via the symbol. In order to make these notes self-contained, we collect in the appendix some material which is not always included in standard graduate probability courses.

It is now about time to thank many individuals who helped to bring this enterprise on the way. I am grateful to the scientific board and the organizing committee for the kind invitation to deliver these lectures at the *Centre de Recerca Matemàtica* in Barcelona. The CRM is a wonderful place to teach and to do research, and I am very happy to acknowledge their support and hospitality. I would like to thank the students who participated in the CRM course as well as all students and readers who were exposed to earlier (temporally & spatially inhomogeneous. . .) versions of my lectures; without your input these notes would look different!

I am greatly indebted to Ms. Franziska Kühn for her interest in this topic; her valuable comments pinpointed many mistakes and helped to make the presentation much clearer.

And, last and most, I thank my wife for her love, support and forbearance while these notes were being prepared.

Dresden, September 2015 *René L. Schilling*

Symbols and Notation

This index is intended to aid cross-referencing, so notation that is specific to a single chapter is generally not listed. Some symbols are used locally, without ambiguity, in senses other than those given below; numbers following an entry are page numbers.

Unless otherwise stated, functions are real-valued and binary operations between functions such as $f \pm g$, $f \cdot g$, $f \wedge g$, $f \vee g$, comparisons $f \leqslant g$, $f < g$ or limiting relations $f_n \xrightarrow{n\to\infty} f$, $\lim_n f_n$, $\liminf_n f_n$, $\limsup_n f_n$, $\sup_n f_n$ or $\inf_n f_n$ are understood pointwise.

General notation: analysis

positive	always in the sense $\geqslant 0$
negative	always in the sense $\leqslant 0$
$\mathbb{N}$	$1, 2, 3, \dots$
$\inf \emptyset$	$\inf \emptyset = +\infty$
$a \vee b$	maximum of a and b
$a \wedge b$	minimum of a and b
$\lfloor x \rfloor$	largest integer $n \leqslant x$
$\lvert x\rvert$	norm in $\mathbb{R}^d$: $\lvert x\rvert^2 = x_1^2 + \dots + x_d^2$
$x \cdot y$	scalar product in $\mathbb{R}^d$: $\sum_{j=1}^d x_j y_j$
$\mathbb{1}_A$	$\mathbb{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$
δ_x	point mass at x
$\mathcal{D}$	domain
Δ	Laplace operator
∂_j	partial derivative $\frac{\partial}{\partial x_j}$
∇, ∇_x	gradient $\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d}\right)^\top$
$\mathcal{F} f$, $\widehat{f}$	Fourier transform $(2\pi)^{-d} \int e^{-i\,x\cdot\xi} f(x)\,dx$
$\mathcal{F}^{-1} f$, $\check{f}$	inverse Fourier transform $\int e^{i\,x\cdot\xi} f(x)\,dx$
$e_\xi(x)$	$e^{-i\,x\cdot\xi}$

General notation: probability

$\sim$	'is distributed as'
$\perp$	'is stochastically independent'
a.s.	almost surely (w. r. t. $\mathbb{P}$)
iid	independent and identically distributed
N, Exp, Poi	normal, exponential, Poisson distribution
$\mathbb{P}$, $\mathbb{E}$	probability, expectation
$\mathbb{V}$, $\mathbb{C}$ov	variance, covariance
(L0)–(L3)	definition of a Lévy process, 7
(L2′)	13

Sets and σ-algebras

A^c	complement of the set A
$\overline{A}$	closure of the set A
$A \mathbin{\dot\cup} B$	disjoint union, i.e., $A \cup B$ for disjoint sets $A \cap B = \emptyset$
$B_r(x)$	open ball, centre x, radius r

$\operatorname{supp} f$	support, $\overline{\{f \neq 0\}}$
$\mathscr{B}(E)$	Borel sets of E
$\mathscr{F}_t^X$	canonical filtration $\sigma(X_s : s \leqslant t)$
$\mathscr{F}_\infty$	$\sigma\left(\bigcup_{t\geqslant 0} \mathscr{F}_t\right)$
$\mathscr{F}_\tau$	88
$\mathscr{F}_{\tau+}$	32
$\mathscr{P}$	predictable σ-algebra, 119

Stochastic processes

$\mathbb{P}^x, \mathbb{E}^x$	law and mean of a Markov process starting at x, 28
X_{t-}	left limit $\lim_{s\uparrow t} X_s$
ΔX_t	jump at time t: $X_t - X_{t-}$
σ, τ	stopping times: $\{\sigma \leqslant t\} \in \mathscr{F}_t$, $t \geqslant 0$
τ_r^x, τ_r	$\inf\{t > 0 : \lvert X_t - X_0\rvert \geqslant r\}$, first exit time from the open ball $B_r(x)$ centered at $x = X_0$
càdlàg	right continuous on $[0, \infty)$ with finite left limits on $(0, \infty)$

Spaces of functions

$\mathcal{B}(E)$	Borel functions on E
$\mathcal{B}_b(E)$	– –, bounded
$\mathcal{C}(E)$	continuous functions on E
$\mathcal{C}_b(E)$	– –, bounded
$\mathcal{C}_\infty(E)$	– –, $\lim_{\lvert x\rvert\to\infty} f(x) = 0$
$\mathcal{C}_c(E)$	– –, compact support
$\mathcal{C}^n(E)$	n times continuously diff'ble functions on E
$\mathcal{C}_b^n(E)$	– –, bounded (with all derivatives)
$\mathcal{C}_\infty^n(E)$	– –, 0 at infinity (with all derivatives)
$\mathcal{C}_c^n(E)$	– –, compact support
$L^p(E, \mu), L^p(\mu), L^p(E)$	L^p space w. r. t. the measure space $(E, \mathscr{A}, \mu)$
$\mathcal{S}(\mathbb{R}^d)$	rapidly decreasing smooth functions on $\mathbb{R}^d$, 41

Chapter 1

Orientation

Stochastic processes with stationary and independent increments are classical examples of Markov processes. Their importance both in theory and for applications justifies to study these processes and their history.

The origins of processes with independent increments reach back to the late 1920s and they are closely connected with the notion of infinite divisibility and the genesis of the Lévy–Khintchine formula. Around this time, the limiting behaviour of sums of independent random variables

$$X_0 := 0 \quad \text{and} \quad X_n := \xi_1 + \xi_2 + \cdots + \xi_n, \quad n \in \mathbb{N},$$

was well understood through the contributions of Borel, Markov, Cantelli, Lindeberg, Feller, de Finetti, Khintchine, Kolmogorov and, of course, Lévy; two new developments emerged, on the one hand the study of dependent random variables and, on the other, the study of continuous-time analogues of sums of independent random variables. In order to pass from $n \in \mathbb{N}$ to a continuous parameter $t \in [0, \infty)$ we need to replace the steps ξ_k by increments $X_t - X_s$. It is not hard to see that X_t, $t \in \mathbb{N}$, with iid (independent and identically distributed) steps ξ_k enjoys the following properties:

$$X_0 = 0 \quad \text{a.s.} \tag{L0}$$

stationary increments $\qquad X_t - X_s \sim X_{t-s} - X_0 \qquad \forall s \leqslant t \qquad$ (L1)

independent increments $\qquad X_t - X_s \perp \sigma(X_r, r \leqslant s) \quad \forall s \leqslant t \qquad$ (L2)

where '$\sim$' stands for 'same distribution' and '$\perp$' for stochastic independence. In the non-discrete setting we will also require a mild regularity condition

continuity in probability $\qquad \lim_{t\to 0} \mathbb{P}(|X_t - X_0| > \epsilon) = 0 \qquad \forall \epsilon > 0 \qquad$ (L3)

which rules out fixed discontinuities of the path $t \mapsto X_t$. Under (L0)–(L2) one has that

$$X_t = \sum_{k=1}^{n} \xi_{k,n}(t) \quad \text{and} \quad \xi_{k,n}(t) = (X_{\frac{kt}{n}} - X_{\frac{(k-1)t}{n}}) \;\text{ are iid} \tag{1.1}$$

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_1

for every $n \in \mathbb{N}$. Letting $n \to \infty$ shows that X_t arises as a suitable limit of (a triangular array of) iid random variables which transforms the problem into a question of limit theorems and **infinite divisibility**.

This was first observed in 1929 by de Finetti [15] who introduces (without naming it, the name is due to Bawly [6] and Khintchine [29]) the concept of infinite divisibility of a random variable X

$$\forall n \quad \exists \text{ iid random variables } \xi_{i,n} \;:\; X \sim \sum_{i=1}^{n} \xi_{i,n} \tag{1.2}$$

and asks for the general structure of infinitely divisible random variables. His paper contains two remarkable results on the characteristic function $\chi(\xi) = \mathbb{E}\, \mathrm{e}^{\mathrm{i}\,\xi\cdot X}$ of an infinite divisible random variable (taken from [39]):

De Finetti's first theorem. *A random variable X is infinitely divisible if, and only if, its characteristic function is of the form*

$$\chi(\xi) = \lim_{n\to\infty} \exp\big[-p_n(1-\phi_n(\xi))\big]$$

where $p_n \geqslant 0$ and ϕ_n is a characteristic function.

De Finetti's second theorem. *The characteristic function of an infinitely divisible random variable X is the limit of finite products of Poissonian characteristic functions*

$$\chi_n(\xi) = \exp\big[-p_n(1-\mathrm{e}^{\mathrm{i}\,h_n\xi})\big],$$

and the converse is also true. In particular, all infinitely divisible laws are limits of convolutions of Poisson distributions.

Because of (1.1), X_t is infinitely divisible and as such one can construct, *in principle*, all independent-increment processes X_t as limits of sums of Poisson random variables. The contributions of Kolmogorov [31], Lévy [37] and Khintchine [28] show the exact form of the characteristic function of an infinitely divisible random variable

$$-\log \mathbb{E}\, \mathrm{e}^{\mathrm{i}\,\xi\cdot X} = -\mathrm{i}\, l\cdot\xi + \frac{1}{2}\xi\cdot Q\xi + \int_{y\neq 0}\big(1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi} + \mathrm{i}\,\xi\cdot y\mathbb{1}_{(0,1)}(|y|)\big)\,\nu(dy) \tag{1.3}$$

where $l \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d\times d}$ is a positive semidefinite symmetric matrix, and ν is a measure on $\mathbb{R}^d \setminus \{0\}$ such that $\int_{y\neq 0} \min\{1, |y|^2\}\,\nu(dy) < \infty$. This is the famous Lévy–Khintchine formula. The exact knowledge of (1.3) makes it possible to find the approximating Poisson variables in de Finetti's theorem explicitly, thus leading to a construction of X_t.

A little later, and without knowledge of de Finetti's results, Lévy came up in his seminal paper [37] (see also [38, Chap. VII]) with a decomposition of X_t in four independent components: a deterministic drift, a Gaussian part, the compensated small jumps and the large jumps $\Delta X_s := X_s - X_{s-}$. This is now known as Lévy–Itô decomposition:

$$X_t = tl + \sqrt{Q}W_t + \lim_{\epsilon\to 0}\Bigg(\sum_{\substack{0<s\leqslant t\\ \epsilon\leqslant|\Delta X_s|<1}} \Delta X_s - t\int_{\epsilon\leqslant|y|<1} y\,\nu(dy)\Bigg) + \sum_{\substack{0<s\leqslant t\\ |\Delta X_s|\geqslant 1}} \Delta X_s \tag{1.4}$$

$$= tl + \sqrt{Q}W_t + \iint_{(0,t]\times B_1(0)} y\,(N(ds,dy) - dt\nu(dy)) + \iint_{(0,t]\times B_1(0)^c} y\,N(ds,dy). \tag{1.5}$$

Lévy uses results from the convergence of random series, notably Kolmogorov's three series theorem, in order to explain the convergence of the series appearing in (1.4). A rigorous proof based on the representation (1.5) are due to Itô [23] who completed Lévy's programme to construct X_t. The coefficients l, Q, ν are the same as in (1.3), W is a d-dimensional standard Brownian motion, and $N_\omega((0,t]\times B)$ is the random measure $\#\{s \in (0,t] : X_s(\omega) - X_{s-}(\omega) \in B\}$ counting the jumps of X; it is a Poisson random variable with intensity $\mathbb{E}N((0,t]\times B) = t\nu(B)$ for all Borel sets $B \subset \mathbb{R}^d \setminus \{0\}$ such that $0 \notin \overline{B}$.

Nowadays there are at least six possible approaches to constructing processes with (stationary and) independent increments $X = (X_t)_{t\geqslant 0}$.

The de Finetti–Lévy(–Kolmogorov–Khintchine) construction. The starting point is the observation that each X_t satisfies (1.1) and is, therefore, infinitely divisible. Thus, the characteristic exponent $\log \mathbb{E}\, e^{i\,\xi\cdot X_t}$ is given by the Lévy–Khintchine formula (1.3), and using the triplet (l, Q, ν) one can construct a drift lt, a Brownian motion $\sqrt{Q}W_t$ and compound Poisson processes, i.e., Poisson processes whose intensities $y \in \mathbb{R}^d$ are mixed with respect to the finite measure

$$\nu_\epsilon(dy) := \mathbb{1}_{[\epsilon,\infty)}(|y|)\nu(dy).$$

Using a suitable compensation (in the spirit of Kolmogorov's three series theorem) of the small jumps, it is possible to show that the limit $\epsilon \to 0$ exists locally uniformly in t. A very clear presentation of this approach can be found in Breiman [10, Chapter 14.7–8], see also Chapter 7.

The Lévy–Itô construction. This is currently the most popular approach to independent-increment processes, see, e.g., Applebaum [2, Chapter 2.3–4] or Kyprianou [36, Chapter 2]. Originally the idea is due to Lévy [37], but Itô [23] gave the first rigorous construction. It is based on the observation that the jumps of a process with stationary and independent increments define a Poisson random measure $N_\omega([0,t]\times B)$ and this can be used to obtain the Lévy–Itô decomposition (1.5). The Lévy–Khintchine formula is then a corollary of the pathwise decomposition. Some of the best presentations can be found in Gikhman–Skorokhod [18, Chapter VI], Itô [24, Chapter 4.3] and Bretagnolle [11]. A proof based on additive functionals and martingale stochastic integrals is due to Kunita & Watanabe [35, Section 7]. We follow this approach in Chapter 9.

Variants of the Lévy–Itô construction. The Lévy–Itô decomposition (1.5) is, in fact, the semimartingale decomposition of a process with stationary and independent increments. Using the general theory of semimartingales – which heavily relies on general random measures – we can identify processes with independent increments as those semimartingales whose semimartingale characteristics are deterministic, cf. Jacod & Shiryaev [27, Chapter II.4c]. A further interesting derivation of the Lévy–Itô decomposition is based on stochastic integrals driven by martingales. The key is Itô's formula and, again, the fact that the jumps of a process with stationary and independent increments defines a Poisson point process which can be used as a good stochastic integrator; this unique approach[1] can be found in Kunita [34, Chapter 2].

Kolmogorov's construction. This is the classic construction of stochastic processes starting from the finite-dimensional distributions. For a process with stationary and independent increments these are given as iterated convolutions of the form

$$\begin{aligned}&\mathbb{E}f(X_{t_0},\dots,X_{t_n})\\&\quad=\int\!\cdots\!\int f(y_0,y_0+y_1,\dots,y_0+\dots+y_n)\,p_{t_0}(dy_0)p_{t_1-t_0}(dy_1)\dots p_{t_n-t_{n-1}}(dy_n)\end{aligned}$$

with $p_t(dy)=\mathbb{P}(X_t\in dy)$ or $\int e^{i\,\xi\cdot y}p_t(dy)=\exp\left[-t\psi(\xi)\right]$ where ψ is the characteristic exponent (1.3). Particularly nice presentations are those of Sato [51, Chapter 2.10–11] and Bauer [5, Chapter 37].

The invariance principle. Just as for a Brownian motion, it is possible to construct Lévy processes as limits of (suitably interpolated) random walks. For finite-dimensional distributions this is done in Gikhman & Skorokhod [18, Chapter IX.6]; for the whole trajectory, i.e., in the space of càdlàg[2] functions $D[0,1]$ equipped with the Skorokhod topology, the proper references are Prokhorov [43] and Grimvall [19].

Random series constructions. A series representation of an independent-increment process $(X_t)_{t\in[0,1]}$ is an expression of the form

$$X_t=\lim_{n\to\infty}\sum_{k=1}^{n}\left(J_k\mathbb{1}_{[0,t]}(U_k)-tc_k\right)\quad\text{a.s.}$$

The random variables J_k represent the jumps, U_k are iid uniform random variables and c_k are suitable deterministic centering terms. Compared with the Lévy–Itô decomposition (1.4), the main difference is the fact that the jumps are summed over a deterministic index set $\{1,2,\dots,n\}$ while the summation in (1.4) extends

[1]It reminds of the elegant use of Itô's formula in Kunita-and-Watanabe's proof of Lévy's characterization of Brownian motion, see, e.g., Schilling & Partzsch [56, Chapter 18.2].

[2]A french acronym meaning 'right-continuous and finite limits from the left'.

over the random set $\{s : |\Delta X_s| > 1/n\}$. In order to construct a process with characteristic exponent (1.3) where $l = 0$ and $Q = 0$, one considers a disintegration

$$\nu(dy) = \int_0^\infty \sigma(r, dy)\, dr.$$

It is possible, cf. Rosiński [47], to choose $\sigma(r, dy) = \mathbb{P}(H(r, V_k) \in dy)$ where $V = (V_k)_{k \in \mathbb{N}}$ is any sequence of d-dimensional iid random variables and $H : (0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is measurable. Now let $\Gamma = (\Gamma_k)_{k \in \mathbb{N}}$ be a sequence of partial sums of iid standard exponential random variables and $U = (U_k)_{k \in \mathbb{N}}$ iid uniform random variables on $[0, 1]$ such that U, V, Γ are independent. Then

$$J_k := H(\Gamma_k, V_k) \quad \text{and} \quad c_k = \int_{k-1}^{k} \int_{|y|<1} y\, \sigma(r, dy)\, dr$$

is the sought-for series representation, cf. Rosiński [47] and [46]. This approach is important if one wants to simulate independent-increment processes. Moreover, it still holds for Banach space valued random variables.

Chapter 2

Lévy Processes

Throughout this chapter, $(\Omega, \mathscr{A}, \mathbb{P})$ is a fixed probability space, $t_0 = 0 \leqslant t_1 \leqslant \cdots \leqslant t_n$ and $0 \leqslant s < t$ are positive real numbers, and ξ_k, η_k, $k = 1, \dots, n$, denote vectors from $\mathbb{R}^d$; we write $\xi \cdot \eta$ for the Euclidean scalar product.

Definition 2.1. A **Lévy process** $X = (X_t)_{t \geqslant 0}$ is a random process $X_t : \Omega \to \mathbb{R}^d$ satisfying (L0)–(L3); this is to say that X starts at zero, has stationary and independent increments and is continuous in probability.

One should understand Lévy processes as continuous-time versions of sums of iid random variables. This can easily be seen from the telescopic sum

$$X_t - X_s = \sum\nolimits_{k=1}^{n} \big(X_{t_k} - X_{t_{k-1}}\big), \quad s < t,\ n \in \mathbb{N}, \tag{2.1}$$

where $t_k = s + \frac{k}{n}(t-s)$. Since the increments $X_{t_k} - X_{t_{k-1}}$ are iid random variables, we see that all X_t of a Lévy process are **infinitely divisible**, i.e., (1.2) holds. Many properties of a Lévy process will, therefore, resemble those of sums of iid random variables.

Let us briefly discuss the conditions (L0)–(L3).

Remark 2.2. We have used in (L2) the **canonical filtration** $\mathscr{F}_t^X := \sigma(X_r,\ r \leqslant t)$ of the process X. Often this condition is written in the following way

$$\begin{aligned} X_{t_n} - X_{t_{n-1}}, \dots, X_{t_1} - X_{t_0} \quad &\text{are independent random variables} \\ \text{for all}\ \ n \in \mathbb{N},\ t_0 = 0 < t_1 < \cdots &< t_n. \end{aligned} \tag{L2$'$}$$

It is easy to see that this is actually equivalent to (L2): From

$$(X_{t_1}, \dots, X_{t_n}) \xleftrightarrow{\text{bi-measurable}} (X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}})$$

it follows that

$$\begin{aligned} \mathscr{F}_t^X &= \sigma\big((X_{t_1}, \dots, X_{t_n}),\ 0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t\big) \\ &= \sigma\big((X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}),\ 0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t\big) \\ &= \sigma\big(X_u - X_v,\ 0 \leqslant v \leqslant u \leqslant t\big), \end{aligned} \tag{2.2}$$

and we conclude that (L2) and (L2$'$) are indeed equivalent.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_2

The condition (L3) is equivalent to either of the following

- '$t \mapsto X_t$ is continuous in probability';
- '$t \mapsto X_t$ is a.s. càdlàg'[1] (up to a modification of the process).

The equivalence with the first claim, and the direction '$\Leftarrow$' of the second claim are easy:

$$\lim_{u\to t} \mathbb{P}(|X_u - X_t| > \epsilon) = \lim_{|t-u|\to 0} \mathbb{P}(|X_{|t-u|}| > \epsilon) \leqslant \lim_{h\to 0} \frac{1}{\epsilon}\mathbb{E}(|X_h| \wedge \epsilon), \tag{2.3}$$

but it takes more effort to show that continuity in probatility (L3) guarantees that almost all paths are càdlàg.[2] Usually, this is proved by controlling the oscillations of the paths of a Lévy process, cf. Sato [51, Theorem 11.1], or by the fundamental regularization theorem for submartingales, see Revuz & Yor [45, Theorem II.(2.5)] and Remark 11.2; in contrast to the general martingale setting [45, Theorem II.(2.9)], we do not need to augment the natural filtration because of (L1) and (L3). Since our construction of Lévy processes gives directly a càdlàg version, we do not go into further detail.

The condition (L3) has another consequence. Recall that the **Cauchy–Abel functional equations** have unique solutions if, say, ϕ, ψ **and** θ **are (right-) continuous**:

$$\begin{aligned} \phi(s+t) &= \phi(s)\cdot\phi(t) & & \phi(t) = \phi(1)^t, \\ \psi(s+t) &= \psi(s)+\psi(t) \quad (s,t\geqslant 0) \Longrightarrow & & \psi(t) = \psi(1)\cdot t \\ \theta(st) &= \theta(s)\cdot\theta(t) & & \theta(t) = t^c,\ c\geqslant 0. \end{aligned} \tag{2.4}$$

The first equation is treated in Theorem A.1 in the appendix. For a thorough discussion on conditions ensuring uniqueness we refer to Aczel [1, Chapter 2.1].

Proposition 2.3. *Let* $(X_t)_{t\geqslant 0}$ *be a Lévy process in* $\mathbb{R}^d$. *Then*

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t} = \left[\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_1}\right]^t, \quad t\geqslant 0,\ \xi\in\mathbb{R}^d. \tag{2.5}$$

Proof. Fix $s,t \geqslant 0$. We get

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_{t+s}-X_s)+\mathrm{i}\,\xi\cdot X_s} \overset{\text{(L2)}}{=} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_{t+s}-X_s)}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_s} \overset{\text{(L1)}}{=} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_s},$$

or $\phi(t+s) = \phi(t)\cdot\phi(s)$, if we write $\phi(t) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}$. Since $x \mapsto \mathrm{e}^{\mathrm{i}\,\xi\cdot x}$ is continuous, there is for every $\epsilon > 0$ some $\delta > 0$ such that

$$|\phi(t)-\phi(s)| \leqslant \mathbb{E}\left|\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)} - 1\right| \leqslant \epsilon + 2\mathbb{P}(|X_t - X_s| \geqslant \delta) = \epsilon + 2\mathbb{P}(|X_{|t-s|}| \geqslant \delta).$$

Thus, (L3) guarantees that $t \mapsto \phi(t)$ is continuous, and the claim follows from (2.4). □

[1] 'Right-continuous and finite limits from the left'

[2] More precisely: that there exists a modification of X which has almost surely càdlàg paths.

Notice that any solution $f(t)$ of (2.4) also satisfies (L0)–(L2); by Proposition 2.3 $X_t + f(t)$ is a Lévy process if, and only if, $f(t)$ is continuous. On the other hand, Hamel, cf. [1, p. 35], constructed discontinuous (non-measurable and locally unbounded) solutions to (2.4). Thus, (L3) means that $t \mapsto X_t$ has no fixed discontinuities, i.e., all jumps occur at random times.

Corollary 2.4. *The finite-dimensional distributions* $\mathbb{P}(X_{t_1} \in dx_1, \dots, X_{t_n} \in dx_n)$ *of a Lévy process are uniquely determined by*

$$\mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^{n} \xi_k \cdot X_{t_k}\right) = \prod_{k=1}^{n}\Big[\mathbb{E}\exp\left(\mathrm{i}(\xi_k + \dots + \xi_n)\cdot X_1\right)\Big]^{t_k - t_{k-1}} \tag{2.6}$$

for all $\xi_1, \dots, \xi_n \in \mathbb{R}^d$, $n \in \mathbb{N}$ *and* $0 = t_0 \leqslant t_1 \leqslant \dots \leqslant t_n$.

Proof. The left-hand side of (2.6) is the characteristic function of $(X_{t_1}, \dots, X_{t_n})$. Consequently, the assertion follows from (2.6). Using Proposition 2.3, we have

$$\begin{aligned}
&\mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^{n} \xi_k \cdot X_{t_k}\right)\\
&\quad= \mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^{n-2} \xi_k \cdot X_{t_k} + \mathrm{i}(\xi_n + \xi_{n-1})\cdot X_{t_{n-1}} + \mathrm{i}\,\xi_n\cdot(X_{t_n} - X_{t_{n-1}})\right)\\
&\quad\overset{(L2)}{\underset{(L1)}{=}} \mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^{n-2} \xi_k \cdot X_{t_k} + \mathrm{i}(\xi_n + \xi_{n-1})\cdot X_{t_{n-1}}\right)\left(\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi_n\cdot X_1}\right)^{t_n - t_{n-1}}.
\end{aligned}$$

Since the first half of the right-hand side has the same structure as the original expression, we can iterate this calculation and obtain (2.6). □

It is not hard to invert the Fourier transform in (2.6). Writing $p_t(dx) := \mathbb{P}(X_t \in dx)$ we get

$$\mathbb{P}(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = \int\cdots\int \prod_{k=1}^{n} \mathbb{1}_{B_k}(x_1 + \dots + x_k)\, p_{t_k - t_{k-1}}(dx_k) \tag{2.7}$$

$$= \int\cdots\int \prod_{k=1}^{n} \mathbb{1}_{B_k}(y_k)\, p_{t_k - t_{k-1}}(dy_k - y_{k-1}). \tag{2.8}$$

Let us discuss the structure of the characteristic function $\chi(\xi) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_1}$ of X_1. From (2.1) we see that each random variable X_t of a Lévy process is infinitely divisible. Clearly, $|\chi(\xi)|^2$ is the (real-valued) characteristic function of the symmetrization $\widetilde{X}_1 = X_1 - X_1'$ (X_1' is an independent copy of X_1) and $\widetilde{X}_1$ is again infinitely divisible:

$$\widetilde{X}_1 = \sum_{k=1}^{n}(\widetilde{X}_{\frac{k}{n}} - \widetilde{X}_{\frac{k-1}{n}}) = \sum_{k=1}^{n}\Big[(X_{\frac{k}{n}} - X_{\frac{k-1}{n}}) - (X'_{\frac{k}{n}} - X'_{\frac{k-1}{n}})\Big].$$

In particular, $|\chi|^2 = |\chi_{1/n}|^{2n}$ where $|\chi_{1/n}|^2$ is the characteristic function of $\widetilde{X}_{1/n}$. Since everything is real and $|\chi(\xi)| \leqslant 1$, we get

$$\theta(\xi) := \lim_{n\to\infty} |\chi_{1/n}(\xi)|^2 = \lim_{n\to\infty} |\chi(\xi)|^{2/n}, \quad \xi \in \mathbb{R}^d,$$

which is 0 or 1 depending on $|\chi(\xi)| = 0$ or $|\chi(\xi)| > 0$, respectively. As $\chi(\xi)$ is continuous at $\xi = 0$ with $\chi(0) = 1$, we have $\theta \equiv 1$ in a neighbourhood $B_r(0)$ of 0. Now we can use Lévy's continuity theorem (Theorem A.5) and conclude that the limiting function $\theta(\xi)$ is continuous everywhere, hence $\theta \equiv 1$. In particular, $\chi(\xi)$ has no zeroes.

Corollary 2.5. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process in $\mathbb{R}^d$. There exists a unique continuous function $\psi : \mathbb{R}^d \to \mathbb{C}$ such that*

$$\mathbb{E}\exp(\mathrm{i}\,\xi\cdot X_t) = \mathrm{e}^{-t\psi(\xi)}, \quad t \geqslant 0,\ \xi \in \mathbb{R}^d.$$

The function ψ is called the **characteristic exponent**.

Proof. In view of Proposition 2.3 it is enough to consider $t = 1$. Set $\chi(\xi) := \mathbb{E}\exp(\mathrm{i}\,\xi\cdot X_1)$. An obvious candidate for the exponent is $\psi(\xi) = -\log\chi(\xi)$, but with complex logarithms there is always the trouble which branch of the logarithm one should take. Let us begin with the uniqueness:

$$\mathrm{e}^{-\psi} = \mathrm{e}^{-\phi} \implies \mathrm{e}^{-(\psi-\phi)} = 1 \implies \psi(\xi) - \phi(\xi) = 2\pi\,\mathrm{i}\,k_\xi$$

for some integer $k_\xi \in \mathbb{Z}$. Since ϕ, ψ are continuous and $\phi(0) = \psi(0) = 1$, we get $k_\xi \equiv 0$.

To prove the existence of the logarithm, it is not sufficient to take the principal branch of the logarithm. As we have seen above, $\chi(\xi)$ is continuous and has no zeroes, i.e., $\inf_{|\xi|\leqslant r} |\chi(\xi)| > 0$ for any $r > 0$; therefore, there is a 'distinguished', continuous[3] version of the argument $\arg_\circ \chi(\xi)$ such that $\arg_\circ \chi(0) = 0$. This allows us to take a continuous version $\log\chi(\xi) = \log|\chi(\xi)| + \arg_\circ \chi(\xi)$. □

Corollary 2.6. *Let Y be an infinitely divisible random variable. Then there exists at most one[4] Lévy process $(X_t)_{t\geqslant 0}$ such that $X_1 \sim Y$.*

Proof. Since $X_1 \sim Y$, infinite divisibility is a necessary requirement for Y. On the other hand, Proposition 2.3 and Corollary 2.4 show how to construct the finite-dimensional distributions of a Lévy process, hence the process, from X_1. □

So far, we have seen the following one-to-one correspondences

$$(X_t)_{t\geqslant 0}\ \text{Lévy process} \overset{1:1}{\longleftrightarrow} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_1} \overset{1:1}{\longleftrightarrow} \psi(\xi) = -\log\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_1}$$

and the next step is to find **all possible characteristic exponents**. This will lead us to the Lévy–Khintchine formula.

[3] A very detailed argument is given in Sato [51, Lemma 7.6], a completely different proof can be found in Dieudonné [16, Chapter IX, Appendix 2].

[4] We will see in Chapter 7 how to construct this process. It is unique in the sense that its finite-dimensional distributions are uniquely determined by Y.

Chapter 3

Examples

We begin with a useful alternative characterisation of Lévy processes.

Theorem 3.1. *Let $X = (X_t)_{t\geqslant 0}$ be a stochastic process with values in $\mathbb{R}^d$, $X_0 = 1$ a.s., and $\mathscr{F}_t = \mathscr{F}_t^X = \sigma(X_r,\ r \leqslant t)$. The process X is a Lévy process if, and only if, there exists an exponent $\psi : \mathbb{R}^d \to \mathbb{C}$ such that*

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\,\middle|\,\mathscr{F}_s\right) = \mathrm{e}^{-(t-s)\psi(\xi)} \quad \textit{for all}\ \ s<t,\ \xi\in\mathbb{R}^d. \tag{3.1}$$

Proof. If X is a Lévy process, we get

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\,\middle|\,\mathscr{F}_s\right) \overset{\text{(L2)}}{=} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)} \overset{\text{(L1)}}{=} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_{t-s}} \overset{\text{Cor. 2.5}}{=} \mathrm{e}^{-(t-s)\psi(\xi)}.$$

Conversely, assume that $X_0 = 0$ a.s. and (3.1) holds. Then

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)} = \mathrm{e}^{-(t-s)\psi(\xi)} = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_{t-s}-X_0)}$$

which shows $X_t - X_s \sim X_{t-s} - X_0 = X_{t-s}$, i.e., (L1).

For any $F\in\mathscr{F}_s$ we find from the tower property of conditional expectation

$$\mathbb{E}\left(\mathbb{1}_F\cdot\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\right) = \mathbb{E}\left(\mathbb{1}_F\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\,\middle|\,\mathscr{F}_s\right]\right) = \mathbb{E}\mathbb{1}_F\cdot\mathrm{e}^{-(t-s)\psi(\xi)}. \tag{3.2}$$

Observe that $\mathrm{e}^{\mathrm{i}\,u\mathbb{1}_F} = \mathbb{1}_{F^c} + \mathrm{e}^{\mathrm{i}\,u}\mathbb{1}_F$ for any $u\in\mathbb{R}$; since both F and F^c are in $\mathscr{F}_s$, we get

$$\begin{aligned}
\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,u\mathbb{1}_F}\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\right) &= \mathbb{E}\left(\mathbb{1}_{F^c}\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\right) + \mathbb{E}\left(\mathbb{1}_F\mathrm{e}^{\mathrm{i}\,u}\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\right)\\
&\overset{(3.2)}{=} \mathbb{E}\left(\mathbb{1}_{F^c} + \mathrm{e}^{\mathrm{i}\,u}\mathbb{1}_F\right)\mathrm{e}^{-(t-s)\psi(\xi)}\\
&\overset{(3.2)}{=} \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,u\mathbb{1}_F}\,\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}.
\end{aligned}$$

Thus, $\mathbb{1}_F \perp\!(X_t - X_s)$ for any $F\in\mathscr{F}_s$, and (L2) follows.

Finally, $\lim_{t\to 0}\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t} = \lim_{t\to 0}\mathrm{e}^{-t\psi(\xi)} = 1$ proves that $X_t\to 0$ in distribution, hence in probability. This gives (L3). □

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_3

Theorem 3.1 allows us to give concrete examples of Lévy processes.

Example 3.2. The following processes are Lévy processes.

a) **Drift** in direction $l/|l|$, $l \in \mathbb{R}^d$, with speed $|l|$: $X_t = tl$ and $\psi(\xi) = -\mathrm{i}\, l \cdot \xi$.

b) **Brownian motion** with (positive semi-definite) covariance matrix $Q \in \mathbb{R}^{d\times d}$: Let $(W_t)_{t\geqslant 0}$ be a standard Wiener process on $\mathbb{R}^d$ and set $X_t := \sqrt{Q}W_t$. Then $\psi(\xi) = \frac{1}{2}\xi \cdot Q\xi$ and $\mathbb{P}(X_t \in dy) = (2\pi t)^{-d/2}(\det Q)^{-1/2} \exp(-y \cdot Q^{-1}y/2t)\, dy$.

c) **Poisson process** in $\mathbb{R}$ with jump height 1 and intensity λ. This is an integer-valued counting process $(N_t)_{t\geqslant 0}$ which increases by 1 after an independent exponential waiting time with mean λ. Thus,

$$N_t = \sum_{k=1}^{\infty} \mathbb{1}_{[0,t]}(\tau_k), \quad \tau_k = \sigma_1 + \cdots + \sigma_k, \quad \sigma_k \sim \mathsf{Exp}(\lambda) \ \text{ iid.}$$

Using this definition, it is a bit messy to show that N is indeed a Lévy process (see, e.g., Çinlar [12, Chapter 4]). We will give a different proof in Theorem 3.4 below. Usually, the first step is to show that its law is a Poisson distribution

$$\mathbb{P}(N_t = k) = \mathrm{e}^{-t\lambda} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

(thus the name!) and from this one can calculate the characteristic exponent

$$\mathbb{E}\, \mathrm{e}^{\mathrm{i}\, uN_t} = \sum_{k=0}^{\infty} \mathrm{e}^{\mathrm{i}\, uk} \mathrm{e}^{-t\lambda} \frac{(\lambda t)^k}{k!} = \mathrm{e}^{-t\lambda} \exp\left[\lambda t \mathrm{e}^{\mathrm{i}\, u}\right] = \exp\left[-t\lambda(1 - \mathrm{e}^{\mathrm{i}\, u})\right],$$

i.e., $\psi(u) = \lambda(1 - \mathrm{e}^{\mathrm{i}\, u})$. Mind that this is strictly weaker than (3.1) and does not prove that N is a Lévy process.

d) **Compound Poisson process** in $\mathbb{R}^d$ with jump distribution μ and intensity λ. Let $N = (N_t)_{t\geqslant 0}$ be a Poisson process with intensity λ and replace the jumps of size 1 by independent iid jumps of random height $H_1, H_2, \dots$ with values in $\mathbb{R}^d$ and $H_1 \sim \mu$. This is a compound Poisson process:

$$C_t = \sum_{k=1}^{N_t} H_k, \quad H_k \sim \mu \ \text{ iid and independent of } (N_t)_{t\geqslant 0}.$$

We will see in Theorem 3.4 that compound Poisson processes are Lévy processes.

Let us show that the Poisson and compound Poisson processes are Lévy processes. For this we need the following auxiliary result. Since $t \mapsto C_t$ is a step function, the Riemann–Stieltjes integral $\int f(u)\, dC_u$ is well-defined.

Lemma 3.3 (Campbell's formula). *Let $C_t = H_1 + \cdots + H_{N_t}$ be a compound Poisson process as in Example* 3.2.d) *with iid jumps $H_k \sim \mu$ and an independent Poisson process $(N_t)_{t\geqslant 0}$ with intensity λ. Then*

$$\mathbb{E}\exp\left(\mathrm{i}\int_0^\infty f(t+s)\,dC_t\right) = \exp\left(\lambda\int_0^\infty\int_{y\neq 0}(\mathrm{e}^{\mathrm{i}\,yf(s+t)}-1)\,\mu(dy)\,dt\right) \tag{3.3}$$

holds for all $s \geqslant 0$ and bounded measurable functions $f : [0,\infty) \to \mathbb{R}^d$ with compact support.

Proof. Set $\tau_k = \sigma_1 + \cdots + \sigma_k$ where $\sigma_k \sim \mathsf{Exp}(\lambda)$ are iid. Then

$$\begin{aligned}
\phi(s) &:= \mathbb{E}\exp\left(\mathrm{i}\int_0^\infty f(s+t)\,dC_t\right)\\
&= \mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^\infty f(s+\sigma_1+\cdots+\sigma_k)H_k\right)\\
&\overset{\text{iid}}{=} \int_0^\infty \underbrace{\mathbb{E}\exp\left(\mathrm{i}\sum_{k=2}^\infty f(s+x+\sigma_2+\cdots+\sigma_k)H_k\right)}_{=\phi(s+x)}\underbrace{\mathbb{E}\exp(\mathrm{i}\,f(s+x)H_1)}_{=:\gamma(s+x)}\underbrace{\mathbb{P}(\sigma_1\in dx)}_{=\lambda\mathrm{e}^{-\lambda x}\,dx}\\
&= \lambda\int_0^\infty \phi(s+x)\gamma(s+x)\mathrm{e}^{-\lambda x}\,dx\\
&= \lambda\mathrm{e}^{\lambda s}\int_s^\infty \gamma(t)\phi(t)\mathrm{e}^{-\lambda t}\,dt.
\end{aligned}$$

This is equivalent to

$$\mathrm{e}^{-\lambda s}\phi(s) = \lambda\int_s^\infty (\phi(t)\mathrm{e}^{-\lambda t})\gamma(t)\,dt$$

and $\phi(\infty) = 1$ since f has compact support. This integral equation has a unique solution; it is now a routine exercise to verify that the right-hand side of (3.3) is indeed a solution. □

Theorem 3.4. *Let $C_t = H_1 + \cdots + H_{N_t}$ be a compound Poisson process as in Example 3.2.*d) *with iid jumps $H_k \sim \mu$ and an independent Poisson process $(N_t)_{t\geqslant 0}$ with intensity λ. Then $(C_t)_{t\geqslant 0}$ (and also $(N_t)_{t\geqslant 0}$) is a d-dimensional Lévy process with characteristic exponent*

$$\psi(\xi) = \lambda\int_{y\neq 0}(1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi})\,\mu(dy). \tag{3.4}$$

Proof. Since the trajectories of $t \mapsto C_t$ are càdlàg step functions with $C_0 = 0$, the properties (L0) and (L3), see (2.3), are satisfied. We will show (L1) and (L2). Let

$\xi_k \in \mathbb{R}^d$, $a < b$ and $0 = t_0 \leqslant \ldots \leqslant t_n$. Then the Riemann–Stieltjes integral

$$\int_0^\infty \mathbb{1}_{(a,b]}(t)\,dC_t = \sum_{k=1}^\infty \mathbb{1}_{(a,b]}(\tau_k)H_k = C_b - C_a$$

exists. We apply the Campbell formula (3.3) to the function

$$f(t) := \sum_{k=1}^n \xi_k \mathbb{1}_{(t_{k-1},t_k]}(t)$$

and with $s = 0$. Then the left-hand side of (3.3) becomes the characteristic function of the increments

$$\mathbb{E}\exp\left(\mathrm{i}\sum_{k=1}^n \xi_k\cdot(C_{t_k} - C_{t_{k-1}})\right),$$

while the right-hand side is equal to

$$\begin{aligned}
&\exp\left[\lambda\int_{y\neq 0}\sum_{k=1}^n\int_{t_{k-1}}^{t_k}(\mathrm{e}^{\mathrm{i}\,\xi_k\cdot y} - 1)\,dt\,\mu(dy)\right]\\
&= \prod_{k=1}^n \exp\left[\lambda(t_k - t_{k-1})\int_{y\neq 0}(\mathrm{e}^{\mathrm{i}\,\xi_k\cdot y} - 1)\,\mu(dy)\right]\\
&= \prod_{k=1}^n \mathbb{E}\exp\left[\mathrm{i}\,\xi_k\cdot C_{t_k - t_{k-1}}\right]
\end{aligned}$$

(use Campbell's formula with $n = 1$ for the last equality). This shows that the increments are independent, i.e., (L2$'$) holds, as well as (L1): $C_{t_k} - C_{t_{k-1}} \sim C_{t_k - t_{k-1}}$.

If $d = 1$ and $H_k \sim \delta_1$, C_t is a Poisson process. □

Denote by μ^{*k} the k-fold convolution of the measure μ; as usual, $\mu^{*0} := \delta_0$.

Corollary 3.5. *Let $(N_t)_{t\geqslant 0}$ be a Poisson process with intensity λ and denote by $C_t = H_1 + \cdots + H_{N_t}$ a compound Poisson process with iid jumps $H_k \sim \mu$. Then, for all $t \geqslant 0$,*

$$\mathbb{P}(N_t = k) = \mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \ldots \tag{3.5}$$

$$\mathbb{P}(C_t \in B) = \mathrm{e}^{-\lambda t}\sum_{k=0}^\infty \frac{(\lambda t)^k}{k!}\mu^{*k}(B), \quad B \subset \mathbb{R}^d \text{ Borel.} \tag{3.6}$$

Proof. If we use Theorem 3.4 for $d = 1$ and $\mu = \delta_1$, we see that the characteristic function of N_t is $\chi_t(u) = \exp[-\lambda t(1 - \mathrm{e}^{\mathrm{i}\,u})]$. Since this is also the characteristic function of the Poisson distribution (i.e., the r.h.s. of (3.5)), we get $N_t \sim \mathsf{Poi}(\lambda t)$.

Since $(H_k)_{k\in\mathbb{N}} \perp (N_t)_{t\geqslant 0}$, we have for any Borel set B

$$\begin{aligned}\mathbb{P}(C_t \in B) &= \sum_{k=0}^{\infty} \mathbb{P}(C_t \in B,\ N_t = k)\\ &= \delta_0(B)\mathbb{P}(N_t = 0) + \sum_{k=1}^{\infty} \mathbb{P}(H_1 + \cdots + H_k \in B)\mathbb{P}(N_t = k)\\ &= \mathrm{e}^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \mu^{*k}(B). \qquad \square\end{aligned}$$

Example 3.2 contains the basic Lévy processes which will also be the building blocks for all Lévy processes. In order to define more specialized Lévy processes, we need further assumptions on the distributions of the random variables X_t.

Definition 3.6. Let $(X_t)_{t\geqslant 0}$ be a stochastically continuous process in $\mathbb{R}^d$. It is called **self-similar**, if

$$\forall a \geqslant 0 \quad \exists b = b(a) \;:\; (X_{at})_{t\geqslant 0} \sim (bX_t)_{t\geqslant 0} \tag{3.7}$$

in the sense that both sides have the same finite-dimensional distributions.

Lemma 3.7 (Lamperti). *If $(X_t)_{t\geqslant 0}$ is self-similar and non-degenerate, then there exists a unique* **index of self-similarity** *$H \geqslant 0$ such that $b(a) = a^H$. If $(X_t)_{t\geqslant 0}$ is a self-similar Lévy process, then $H \geqslant \frac{1}{2}$.*

Proof. Since $(X_t)_{t\geqslant 0}$ is self-similar, we find for $a, a' \geqslant 0$ and each $t > 0$

$$b(aa')X_t \sim X_{aa't} \sim b(a)X_{a't} \sim b(a)b(a')X_t,$$

and so $b(aa') = b(a)b(a')$ as X_t is non-degenerate.[1] By the convergence of types theorem (Theorem A.6) and the continuity in probability of $t \mapsto X_t$ we see that $a \mapsto b(a)$ is continuous. Thus, the Cauchy functional equation $b(aa') = b(a)b(a')$ has the unique continuous solution $b(a) = a^H$ for some $H \geqslant 0$.

Assume now that $(X_t)_{t\geqslant 0}$ is a Lévy process. We are going to show that $H \geqslant \frac{1}{2}$. Using self-similarity and the properties (L1), (L2) we get (primes always denote iid copies of the respective random variables)

$$(n+m)^H X_1 \sim X_{n+m} = (X_{n+m} - X_m) + X_m \sim X''_n + X'_m \sim n^H X''_1 + m^H X'_1. \tag{3.8}$$

Any standard normal random variable X_1 satisfies (3.8) with $H = \frac{1}{2}$.

[1] We use here that $bX \sim cX \implies b = c$ if X is non-degenerate. To see this, set $\chi(\xi) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X}$ and notice

$$|\chi(\xi)| = \left|\chi(\tfrac{b}{c}\xi)\right| = \cdots = \left|\chi((\tfrac{b}{c})^n \xi)\right|.$$

If $b < c$, the right-hand side converges for $n \to \infty$ to $\chi(0) = 1$, hence $|\chi| \equiv 1$, contradicting the fact that X is non-degenerate. Since b, c play symmetric roles, we conclude that $b = c$.

On the other hand, if X_1 has a second moment, we get

$$(n+m)\mathbb{V}X_1 = \mathbb{V}X_{n+m} = \mathbb{V}X_n'' + \mathbb{V}X_m' = n\mathbb{V}X_1'' + m\mathbb{V}X_1'$$

by Bienaymés identity for variances, i.e., (3.8) can only hold with $H = \frac{1}{2}$. Thus, any self-similar X_1 with finite second moment has to satisfy (3.8) with $H = \frac{1}{2}$. If we can show that $H < \frac{1}{2}$ implies the existence of a second moment, we have reached a contradiction.

If X_n is symmetric and $H < \frac{1}{2}$, we find because of $X_n \sim n^H X_1$ some $u > 0$ such that

$$\mathbb{P}(|X_n| > un^H) = \mathbb{P}(|X_1| > u) < \frac{1}{4}.$$

By the symmetrization inequality (Theorem A.7),

$$\frac{1}{2}\left(1 - \exp\{-n\mathbb{P}(|X_1| > un^H)\}\right) \leqslant \mathbb{P}(|X_n| > un^H) < \frac{1}{4}$$

which means that $n\mathbb{P}(|X_1| > un^H) \leqslant c$ for all $n \in \mathbb{N}$. Therefore wee see that $\mathbb{P}(|X_1| > x) \leqslant c' x^{-1/H}$ for all $x > u+1$, and so

$$\mathbb{E}|X_1|^2 = 2\int_0^\infty x\,\mathbb{P}(|X_1| > x)\,dx \leqslant 2(u+1) + 2c'\int_{u+1}^\infty x^{1-1/H}\,dx < \infty$$

as $H < \frac{1}{2}$. If X_n is not symmetric, we use its symmetrization $X_n - X_n'$ where X_n' are iid copies of X_n. □

Definition 3.8. A random variable X is called **stable** if

$$\forall n \in \mathbb{N} \quad \exists b_n \geqslant 0,\ c_n \in \mathbb{R}^d \ :\ X_1' + \cdots + X_n' \sim b_n X + c_n \tag{3.9}$$

where $X_1', \dots, X_n'$ are iid copies of X. If (3.9) holds with $c_n = 0$, the random variable is called **strictly stable**. A Lévy process $(X_t)_{t\geqslant 0}$ is (strictly) stable if X_1 is a (strictly) stable random variable.

Note that the symmetrization $X - X'$ of a stable random variable is strictly stable. Setting $\chi(\xi) = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X}$ it is easy to see that (3.9) is equivalent to

$$\forall n \in \mathbb{N} \quad \exists b_n \geqslant 0,\ c_n \in \mathbb{R}^d \ :\ \chi(\xi)^n = \chi(b_n\xi)\mathrm{e}^{\mathrm{i}\,c_n\cdot\xi}. \tag{3.9$'$}$$

Example 3.9. a) **Stable processes.** By definition, any stable random variable is infinitely divisible, and for every stable X there is a unique Lévy process on $\mathbb{R}^d$ such that $X_1 \sim X$, cf. Corollary 2.6.

A Lévy process $(X_t)_{t\geqslant 0}$ is stable if, and only if, all random variables X_t are stable. This follows at once from (3.9′) if we use $\chi_t(\xi) := \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}$:

$$\chi_t(\xi)^n \overset{(2.5)}{=} \left(\chi_1(\xi)^n\right)^t \overset{(3.9')}{=} \chi_1(b_n\xi)^t\mathrm{e}^{\mathrm{i}(tc_n)\cdot\xi} \overset{(2.5)}{=} \chi_t(b_n\xi)\mathrm{e}^{\mathrm{i}(tc_n)\cdot\xi}.$$

It is possible to determine the characteristic exponent of a stable process, cf. Sato [51, Theorem 14.10] and (3.10) further down.

b) **Self-similar processes.** Assume that $(X_t)_{t\geqslant 0}$ is a self-similar Lévy process. Then

$$\forall n \in \mathbb{N} \;:\; b(n)X_1 \sim X_n = \sum_{k=1}^{n}(X_k - X_{k-1}) \sim X'_{1,n} + \cdots + X'_{n,n}$$

where the $X'_{k,n}$ are iid copies of X_1. This shows that X_1, hence $(X_t)_{t\geqslant 0}$, is strictly stable. In fact, the converse is also true:

c) **A strictly stable Lévy process is self-similar**. We have already seen in b) that self-similar Lévy processes are strictly stable. Assume now that $(X_t)_{t\geqslant 0}$ is strictly stable. Since $X_{nt} \sim b_n X_t$ we get

$$\mathrm{e}^{-nt\psi(\xi)} = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_{nt}} = \mathbb{E}\,\mathrm{e}^{\mathrm{i}\,b_n\xi\cdot X_t} = \mathrm{e}^{-t\psi(b_n\xi)}.$$

Taking $n = m$, $t \rightsquigarrow t/m$ and $\xi \rightsquigarrow b_m^{-1}\xi$ we see

$$\mathrm{e}^{-\frac{t}{m}\psi(\xi)} = \mathrm{e}^{-t\psi(b_m^{-1}\xi)}.$$

From these equalities we obtain for $q = n/m \in \mathbb{Q}^+$ and $b(q) := b_n/b_m$

$$\mathrm{e}^{-qt\psi(\xi)} = \mathrm{e}^{-t\psi(b(q)\xi)} \implies X_{qt} \sim b(q)X_t \implies X_{at} \sim b(a)X_t$$

for all $t \geqslant 0$ because of the continuity in probability of $(X_t)_{t\geqslant 0}$. Since, by Corollary 2.4, the finite-dimensional distributions are determined by the one-dimensional distributions, we conclude that (3.7) holds.

This means, in particular, that strictly stable Lévy processes have an index of self-similarity $H \geqslant \frac{1}{2}$. It is common to call $\alpha = 1/H \in (0, 2]$ the **index of stability** of $(X_t)_{t\geqslant 0}$, and we have $X_{nt} \sim n^{1/\alpha}X_t$.

If X is 'only' stable, its symmetrization is strictly stable and, thus, every stable Lévy process has an index $\alpha \in (0, 2]$. It plays an important role for the characteristic exponent. For a general stable process the characteristic exponent is of the form

$$\psi(\xi) = \begin{cases} \displaystyle\int_{\mathbb{S}^d} |z\cdot\xi|^\alpha \big(1 - \mathrm{i}\,\mathrm{sgn}(z\cdot\xi)\tan\tfrac{\alpha\pi}{2}\big)\,\sigma(dz) - \mathrm{i}\,\mu\cdot\xi, & (\alpha \neq 1), \\ \displaystyle\int_{\mathbb{S}^d} |z\cdot\xi|\big(1 + \tfrac{2}{\pi}\,\mathrm{i}\,\mathrm{sgn}(z\cdot\xi)\log|z\cdot\xi|\big)\,\sigma(dz) - \mathrm{i}\,\mu\cdot\xi, & (\alpha = 1), \end{cases} \tag{3.10}$$

where σ is a finite measure on $\mathbb{S}^d$ and $\mu \in \mathbb{R}^d$. The strictly stable exponents have $\mu = 0$ (if $\alpha \neq 1$) and $\int_{\mathbb{S}^d} z_k\,\sigma(dz) = 0$, $k = 1, \ldots, d$ (if $\alpha = 1$). These formulae can be derived from the general Lévy–Khintchine formula; a good reference is the monograph by Samorodnitsky & Taqqu [48, Chapters 2.3–4].

If X is strictly stable such that the distribution of X_t is rotationally invariant, it is clear that $\psi(\xi) = c|\xi|^\alpha$. If X_t is symmetric, i.e., $X_t \sim -X_t$, then the exponent is $\psi(\xi) = \int_{\mathbb{S}^d} |z\cdot\xi|^\alpha\,\sigma(dz)$ for some finite, symmetric measure σ on the unit sphere $\mathbb{S}^d \subset \mathbb{R}^d$.

Let us finally show Kolmogorov's proof of the Lévy–Khintchine formula for one-dimensional Lévy processes admitting second moments. We need the following auxiliary result.

Lemma 3.10. *Let* $(X_t)_{t\geqslant 0}$ *be a Lévy process on* $\mathbb{R}$. *If* $\mathbb{V}X_1 < \infty$, *then* $\mathbb{V}X_t < \infty$ *for all* $t > 0$ *and*

$$\mathbb{E}X_t = t\mathbb{E}X_1 =: t\mu \quad \textit{and} \quad \mathbb{V}X_t = t\mathbb{V}X_1 =: t\sigma^2.$$

Proof. If $\mathbb{V}X_1 < \infty$, then $\mathbb{E}|X_1| < \infty$. With Bienaymé's identity, we get

$$\mathbb{V}X_m = \sum_{k=1}^m \mathbb{V}(X_k - X_{k-1}) = m\mathbb{V}X_1 \quad \text{and} \quad \mathbb{V}X_1 = n\mathbb{V}X_{1/n}.$$

In particular, $\mathbb{V}X_m, \mathbb{V}X_{1/n} < \infty$. This, and a similar argument for the expectation, show

$$\mathbb{V}X_q = q\mathbb{V}X_1 \quad \text{and} \quad \mathbb{E}X_q = q\mathbb{E}X_1 \quad \text{for all} \;\; q \in \mathbb{Q}^+.$$

Moreover, $\mathbb{V}(X_q - X_r) = \mathbb{V}X_{q-r} = (q-r)\mathbb{V}X_1$ for all rational $r \leqslant q$, and this shows that $X_q - \mathbb{E}X_q = X_q - q\mu$ converges in L^2 as $q \to t$. Since $t \mapsto X_t$ is continuous in probability, we can identify the limit and find $X_q - q\mu \to X_t - t\mu$. Consequenctly, $\mathbb{V}X_t = t\sigma^2$ and $\mathbb{E}X_t = t\mu$. □

We have seen in Proposition 2.3 that the characteristic function of a Lévy process is of the form

$$\chi_t(\xi) = \mathbb{E}\,e^{i\,\xi X_t} = \left[\mathbb{E}\,e^{i\,\xi X_1}\right]^t = \chi_1(\xi)^t.$$

Let us assume that X is real-valued and has finite (first and) second moments $\mathbb{V}X_1 = \sigma^2$ and $\mathbb{E}X_1 = \mu$. By Taylor's formula

$$\begin{aligned}\mathbb{E}\,e^{i\,\xi(X_t - t\mu)} &= \mathbb{E}\left[1 + i\,\xi(X_t - t\mu) - \int_0^1 \xi^2(X_t - t\mu)^2\,(1-\theta)e^{i\,\theta\xi(X_t - t\mu)}\,d\theta\right]\\ &= 1 - \mathbb{E}\left[\xi^2(X_t - t\mu)^2 \int_0^1 (1-\theta)e^{i\,\theta\xi(X_t - t\mu)}\,d\theta\right].\end{aligned}$$

Since

$$\left|\int_0^1 (1-\theta)e^{i\,\theta\xi(X_t - t\mu)}\,d\theta\right| \leqslant \int_0^1 (1-\theta)\,d\theta = \frac{1}{2},$$

we get

$$\left|\mathbb{E}\,e^{i\,\xi X_t}\right| = \left|\mathbb{E}\,e^{i\,\xi(X_t - t\mu)}\right| \geqslant 1 - \frac{\xi^2}{2}\,t\sigma^2.$$

Thus, $\chi_{1/n}(\xi) \neq 0$ if $n \geqslant N(\xi) \in \mathbb{N}$ is large, hence $\chi_1(\xi) = \chi_{1/n}(\xi)^n \neq 0$. For $\xi \in \mathbb{R}$ we find (using a suitable branch of the complex logarithm)

$$\begin{aligned}\psi(\xi) := -\log \chi_1(\xi) &= -\frac{\partial}{\partial t}\big[\chi_1(\xi)\big]^t\Big|_{t=0}\\ &= \lim_{t\to 0}\frac{1-\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi X_t}}{t}\\ &= \lim_{t\to 0}\frac{1}{t}\int_{-\infty}^{\infty}\big(1-\mathrm{e}^{\mathrm{i}\,y\xi}+\mathrm{i}\,y\xi\big)\,p_t(dy) - \mathrm{i}\,\xi\mu\\ &= \lim_{t\to 0}\int_{-\infty}^{\infty}\frac{1-\mathrm{e}^{\mathrm{i}\,y\xi}+\mathrm{i}\,y\xi}{y^2}\,\pi_t(dy) - \mathrm{i}\,\xi\mu \qquad (3.11)\end{aligned}$$

where $p_t(dy) = \mathbb{P}(X_t \in dy)$ and $\pi_t(dy) := y^2\,t^{-1}p_t(dy)$. Yet another application of Taylor's theorem shows that the integrand in the above integral is bounded, vanishes at infinity, and admits a continuous extension onto the whole real line if we choose the value $\frac{1}{2}\xi^2$ at $y=0$. The family $(\pi_t)_{t\in(0,1]}$ is uniformly bounded,

$$\frac{1}{t}\int y^2\,p_t(dy) = \frac{1}{t}\,\mathbb{E}(X_t^2) = \frac{1}{t}\Big(\mathbb{V}X_t + \big[\mathbb{E}X_t\big]^2\Big) = \sigma^2 + t\mu^2 \xrightarrow{t\to 0} \sigma^2,$$

hence sequentially vaguely relatively compact (see Theorem A.3). We conclude that every sequence $(\pi_{t(n)})_{n\in\mathbb{N}} \subset (\pi_t)_{t\in(0,1]}$ with $t(n)\to 0$ as $n\to\infty$ has a vaguely convergent subsequence. But since the limit (3.11) exists, all subsequential limits coincide which means[2] that π_t converges vaguely to a finite measure π on $\mathbb{R}$. This proves that

$$\psi(\xi) = -\log\chi_1(\xi) = \int_{-\infty}^{\infty}\frac{1-\mathrm{e}^{\mathrm{i}\,y\xi}+\mathrm{i}\,y\xi}{y^2}\,\pi(dy) - \mathrm{i}\,\xi\mu$$

for some finite measure π on $(-\infty,\infty)$ with total mass $\pi(\mathbb{R}) = \sigma^2$. This is sometimes called the **de Finetti–Kolmogorov formula**.

If we set $\nu(dy) := y^{-2}\mathbb{1}_{\{y\neq 0\}}\,\pi(dy)$ and $\sigma_0^2 := \pi\{0\}$, we obtain the **Lévy–Khintchine formula**

$$\psi(\xi) = -\,\mathrm{i}\,\mu\xi + \frac{1}{2}\,\sigma_0^2\,\xi^2 + \int_{y\neq 0}\big(1-\mathrm{e}^{\mathrm{i}\,y\xi}+\mathrm{i}\,y\xi\big)\,\nu(dy)$$

where $\sigma^2 = \sigma_0^2 + \int_{y\neq 0} y^2\,\nu(dy)$.

[2]Note that $\mathrm{e}^{\mathrm{i}\,y\xi} = \partial_\xi^2(1-\mathrm{e}^{\mathrm{i}\,y\xi}+\mathrm{i}\,y\xi)/y^2$, i.e., the kernel appearing in (3.11) is indeed measure-determining.

Chapter 4

On the Markov Property

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space with some filtration $(\mathscr{F}_t)_{t\geqslant 0}$ and a d-dimensional **adapted** stochastic process $X = (X_t)_{t\geqslant 0}$, i.e., each X_t is $\mathscr{F}_t$ measurable. We write $\mathscr{B}(\mathbb{R}^d)$ for the Borel sets and set $\mathscr{F}_\infty := \sigma(\bigcup_{t\geqslant 0} \mathscr{F}_t)$.

The process X is said to be a **simple Markov process**, if

$$\mathbb{P}(X_t \in B \mid \mathscr{F}_s) = \mathbb{P}(X_t \in B \mid X_s), \quad s \leqslant t,\ B \in \mathscr{B}(\mathbb{R}^d), \tag{4.1}$$

holds true. This is pretty much the most general definition of a Markov process, but it is usually too general to work with. It is more convenient to consider Markov families.

Definition 4.1. A (temporally homogeneous) **Markov transition function** is a measure kernel $p_t(x, B)$, $t \geqslant 0$, $x \in \mathbb{R}^d$, $B \in \mathscr{B}(\mathbb{R}^d)$ such that

a) $B \mapsto p_s(x, B)$ is a probability measure for every $s \geqslant 0$ and $x \in \mathbb{R}^d$;

b) $(s, x) \mapsto p_s(x, B)$ is a Borel measurable function for every $B \in \mathscr{B}(\mathbb{R}^d)$;

c) the **Chapman–Kolmogorov equations** hold

$$p_{s+t}(x, B) = \int p_t(y, B)\, p_s(x, dy) \quad \text{for all} \quad s, t \geqslant 0,\ x \in \mathbb{R}^d,\ B \in \mathscr{B}(\mathbb{R}^d). \tag{4.2}$$

Definition 4.2. A stochastic process $(X_t)_{t\geqslant 0}$ is called a (temporally homogeneous) **Markov process with transition function** if there exists a Markov transition function $p_t(x, B)$ such that

$$\mathbb{P}(X_t \in B \mid \mathscr{F}_s) = p_{t-s}(X_s, B) \quad \text{a.s. for all} \quad s \leqslant t,\ B \in \mathscr{B}(\mathbb{R}^d). \tag{4.3}$$

Conditioning w.r.t. $\sigma(X_s)$ and using the tower property of conditional expectation shows that (4.3) implies the simple Markov property (4.1). Nowadays the following definition of a Markov process is commonly used.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_4

Definition 4.3. A **(universal) Markov process** is a tuple $(\Omega, \mathscr{A}, \mathscr{F}_t, X_t, t \geqslant 0, \mathbb{P}^x, x \in \mathbb{R}^d)$ such that $p_t(x,B) = \mathbb{P}^x(X_t \in B)$ is a Markov transition function and $(X_t)_{t\geqslant 0}$ is for each $\mathbb{P}^x$ a Markov process in the sense of Definition 4.2 such that $\mathbb{P}^x(X_0 = x) = 1$. In particular,

$$\mathbb{P}^x(X_t \in B \mid \mathscr{F}_s) = \mathbb{P}^{X_s}(X_{t-s} \in B) \quad \mathbb{P}^x\text{-a.s. for all } \ s \leqslant t, \ B \in \mathscr{B}(\mathbb{R}^d). \tag{4.4}$$

We are going to show that a Lévy process is a (universal) Markov process. Assume that $(X_t)_{t\geqslant 0}$ is a Lévy process and set $\mathscr{F}_t := \mathscr{F}_t^X = \sigma(X_r, \, r \leqslant t)$. Define probability measures

$$\mathbb{P}^x(X_\bullet \in \Gamma) := \mathbb{P}(X_\bullet + x \in \Gamma), \quad x \in \mathbb{R}^d,$$

where Γ is a Borel set of the path space $(\mathbb{R}^d)^{[0,\infty)} = \{w \mid w : [0,\infty) \to \mathbb{R}^d\}$.[1] We set $\mathbb{E}^x := \int \dots d\mathbb{P}^x$. By construction, $\mathbb{P} = \mathbb{P}^0$ and $\mathbb{E} = \mathbb{E}^0$.

Note that $X_t^x := X_t + x$ satisfies the conditions (L1)–(L3), and it is common to call $(X_t^x)_{t\geqslant 0}$ a **Lévy process starting from** x.

Lemma 4.4. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process on $\mathbb{R}^d$. Then*

$$p_t(x,B) := \mathbb{P}^x(X_t \in B) := \mathbb{P}(X_t + x \in B), \quad t \geqslant 0, \ x \in \mathbb{R}^d, \ B \in \mathscr{B}(\mathbb{R}^d),$$

is a Markov transition function.

Proof. Since $p_t(x,B) = \mathbb{E}\mathbb{1}_B(X_t + x)$ (the proof of) Fubini's theorem shows that $x \mapsto p_t(x,B)$ is a measurable function and $B \mapsto p_t(x,B)$ is a probability measure. The Chapman–Kolmogorov equations follow from

$$\begin{aligned} p_{s+t}(x,B) = \mathbb{P}(X_{s+t} + x \in B) &= \mathbb{P}((X_{s+t} - X_t) + x + X_t \in B) \\ &\overset{\text{(L2)}}{=} \int_{\mathbb{R}^d} \mathbb{P}(y + X_t \in B)\, \mathbb{P}((X_{s+t} - X_t) + x \in dy) \\ &\overset{\text{(L1)}}{=} \int_{\mathbb{R}^d} \mathbb{P}(y + X_t \in B)\, \mathbb{P}(X_s + x \in dy) \\ &= \int_{\mathbb{R}^d} p_t(y,B)\, p_s(x,dy). \end{aligned}$$

□

Remark 4.5. The proof of Lemma 4.4 shows a bit more: From

$$p_t(x,B) = \int \mathbb{1}_B(x+y)\, \mathbb{P}(X_t \in dy) = \int \mathbb{1}_{B-x}(y)\, \mathbb{P}(X_t \in dy) = p_t(0, B-x)$$

we see that the kernels $p_t(x,B)$ are invariant under shifts in $\mathbb{R}^d$ (**translation invariant**). In slight abuse of notation we write $p_t(x,B) = p_t(B-x)$. From this it

[1] Recall that $\mathscr{B}((\mathbb{R}^d)^{[0,\infty)})$ is the smallest σ-algebra containing the cylinder sets $Z = \times_{t\geqslant 0} B_t$ where $B_t \in \mathscr{B}(\mathbb{R}^d)$ and only finitely many $B_t \neq \mathbb{R}^d$.

becomes clear that the Chapman–Kolmogorov equations are convolution identities $p_{t+s}(B) = p_t * p_s(B)$, and $(p_t)_{t\geqslant 0}$ is a **convolution semigroup** of probability measures; because of (L3), this semigroup is weakly continuous at $t = 0$, i.e., $p_t \to \delta_0$ as $t \to 0$, cf. Theorem A.3 *et seq.* for the weak convergence of measures.

Lévy processes enjoy an even stronger version of the above Markov property.

Theorem 4.6 (Markov property for Lévy processes)**.** *Let X be a d-dimensional Lévy process and set $Y := (X_{t+a} - X_a)_{t\geqslant 0}$ for some fixed $a \geqslant 0$. Then Y is again a Lévy process satisfying*

a) *$Y \perp (X_r)_{r\leqslant a}$, i.e., $\mathscr{F}^Y_\infty \perp \mathscr{F}^X_a$.*
b) *$Y \sim X$, i.e., X and Y have the same finite-dimensional distributions.*

Proof. Observe that $\mathscr{F}^Y_s = \sigma(X_{r+a} - X_a,\ r \leqslant s) \subset \mathscr{F}^X_{s+a}$. Using Theorem 3.1 and the tower property of conditional expectation yields for all $s \leqslant t$

$$\mathbb{E}\left(e^{i\,\xi\cdot(Y_t - Y_s)} \;\middle|\; \mathscr{F}^Y_s\right) = \mathbb{E}\left[\mathbb{E}\left(e^{i\,\xi\cdot(X_{t+a}-X_{s+a})} \;\middle|\; \mathscr{F}^X_{s+a}\right) \;\middle|\; \mathscr{F}^Y_s\right] = e^{-(t-s)\psi(\xi)}.$$

Thus, $(Y_t)_{t\geqslant 0}$ is a Lévy process with the same characteristic function as $(X_t)_{t\geqslant 0}$. The property (L2$'$) for X gives

$$X_{t_n+a} - X_{t_{n-1}+a},\ X_{t_{n-1}+a} - X_{t_{n-2}+a},\ \dots,\ X_{t_1+a} - X_a \perp \mathscr{F}^X_a.$$

As $\sigma(Y_{t_1}, \dots, Y_{t_n}) = \sigma(Y_{t_n} - Y_{t_{n-1}}, \dots, Y_{t_1} - Y_{t_0}) \perp \mathscr{F}^X_a$ for all $t_0 = 0 < t_1 < \dots < t_n$, we get

$$\mathscr{F}^Y_\infty = \sigma\left(\bigcup_{t_1\leqslant\dots\leqslant t_n} \sigma(Y_{t_1}, \dots, Y_{t_n})\right) \perp \mathscr{F}^X_a. \qquad \square$$

Using the Markov transition function $p_t(x, B)$ we can define a linear operator on the bounded Borel measurable functions $f : \mathbb{R}^d \to \mathbb{R}$:

$$P_t f(x) := \int f(y)\, p_t(x, dy) = \mathbb{E}^x f(X_t), \quad f \in \mathcal{B}_b(\mathbb{R}^d),\ t \geqslant 0,\ x \in \mathbb{R}^d. \tag{4.5}$$

For a Lévy process, cf. Remark 4.5, we have $p_t(x, B) = p_t(B - x)$ and the operators P_t are actually **convolution operators**:

$$P_t f(x) = \mathbb{E} f(X_t + x) = \int f(y + x)\, p_t(dy) = f * \widetilde{p}_t(x) \quad \text{where} \quad \widetilde{p}_t(B) := p_t(-B). \tag{4.6}$$

Definition 4.7. Let P_t, $t \geqslant 0$, be defined by (4.5). The operators are said to be

a) **acting on** $\mathcal{B}_b(\mathbb{R}^d)$, if $P_t : \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d)$.
b) an operator **semigroup**, if $P_{t+s} = P_t \circ P_s$ for all $s, t \geqslant 0$ and $P_0 = \mathrm{id}$.

c) **sub-Markovian** if $0 \leqslant f \leqslant 1 \implies 0 \leqslant P_t f \leqslant 1$.
d) **contractive** if $\|P_t f\|_\infty \leqslant \|f\|_\infty$ for all $f \in \mathcal{B}_b(\mathbb{R}^d)$.
e) **conservative** if $P_t 1 = 1$.
f) **Feller operators**, if $P_t : \mathcal{C}_\infty(\mathbb{R}^d) \to \mathcal{C}_\infty(\mathbb{R}^d)$.[2]
g) **strongly continuous** on $\mathcal{C}_\infty(\mathbb{R}^d)$, if $\lim_{t\to 0} \|P_t f - f\|_\infty = 0$ for all $f \in \mathcal{C}_\infty(\mathbb{R}^d)$.

h) **strong Feller operators**, if $P_t : \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$.

Lemma 4.8. *Let $(P_t)_{t\geqslant 0}$ be defined by* (4.5). *The properties* 4.7.a)–e) *hold for any Markov process,* 4.7.a)–g) *hold for any Lévy process, and* 4.7.a)–h) *hold for any Lévy process such that all $p_t(dy) = \mathbb{P}(X_t \in dy)$, $t > 0$, are absolutely continuous w.r.t. Lebesgue measure.*

Proof. We only show the assertions about Lévy processes $(X_t)_{t\geqslant 0}$.

a) Since $P_t f(x) = \mathbb{E} f(X_t + x)$, the boundedness of $P_t f$ is obvious, and the measurability in x follows from (the proof of) Fubini's theorem.
b) By the tower property of conditional expectation, we get for $s, t \geqslant 0$

$$
\begin{aligned}
P_{t+s} f(x) = \mathbb{E}^x f(X_{t+s}) &= \mathbb{E}^x\big(\mathbb{E}^x[f(X_{t+s}) \mid \mathscr{F}_s]\big) \\
&\overset{(4.4)}{=} \mathbb{E}^x\big(\mathbb{E}^{X_s} f(X_t)\big) = P_s \circ P_t f(x).
\end{aligned}
$$

For the Markov transition functions this is the Chapman–Kolmogorov identity (4.2).
c) and d), e) follow directly from the fact that $B \mapsto p_t(x, B)$ is a probability measure.
f) Let $f \in \mathcal{C}_\infty(\mathbb{R}^d)$. Since $x \mapsto f(x + X_t)$ is continuous and bounded, the claim follows from dominated convergence as $P_t f(x) = \mathbb{E} f(x + X_t)$.
g) $f \in \mathcal{C}_\infty$ is uniformly continuous, i.e., for every $\epsilon > 0$ there is some $\delta > 0$ such that $|x - y| \leqslant \delta$ implies $|f(x) - f(y)| \leqslant \epsilon$. Hence,

$$
\begin{aligned}
\|P_t f - f\|_\infty &\leqslant \sup_{x\in\mathbb{R}^d} \int |f(X_t) - f(x)|\, d\mathbb{P}^x \\
&= \sup_{x\in\mathbb{R}^d} \left(\int_{|X_t - x| \leqslant \delta} |f(X_t) - f(x)|\, d\mathbb{P}^x + \int_{|X_t - x| > \delta} |f(X_t) - f(x)|\, d\mathbb{P}^x \right) \\
&\leqslant \epsilon + 2\|f\|_\infty \sup_{x\in\mathbb{R}^d} \mathbb{P}^x(|X_t - x| > \delta) \\
&= \epsilon + 2\|f\|_\infty\, \mathbb{P}(|X_t| > \delta) \xrightarrow[t\to 0]{\text{(L3)}} \epsilon.
\end{aligned}
$$

[2] $\mathcal{C}_\infty(\mathbb{R}^d)$ denotes the space of continuous functions vanishing at infinity. It is a Banach space when equipped with the uniform norm $\|f\|_\infty = \sup_{x\in\mathbb{R}^d} |f(x)|$.

Since $\epsilon > 0$ is arbitrary, the claim follows. Note that this proof shows that **uniform continuity in probability** is responsible for the strong continuity of the semigroup.

h) See Lemma 4.9. □

Lemma 4.9 (Hawkes)**.** *Let $X = (X_t)_{t\geqslant 0}$ be a Lévy process on $\mathbb{R}^d$. Then the operators P_t defined by (4.5) are strong Feller if, and only if, $X_t \sim p_t(y)\,dy$ for all $t > 0$.*

Proof. '$\Leftarrow$': Let $X_t \sim p_t(y)\,dy$. Since $p_t \in L^1$ and since convolutions have a smoothing property (e.g., [54, Theorem 14.8] or [55, Satz 18.9]), we get with $\widetilde{p}_t(y) = p_t(-y)$

$$P_t f = f * \widetilde{p}_t \in L^\infty * L^1 \subset \mathcal{C}_b(\mathbb{R}^d).$$

'$\Rightarrow$': We show that $p_t(dy) \ll dy$. Let $N \in \mathscr{B}(\mathbb{R}^d)$ be a Lebesgue null set $\lambda^d(N) = 0$ and $g \in \mathcal{B}_b(\mathbb{R}^d)$. Then, by the Fubini–Tonelli theorem

$$\begin{aligned}\int g(x) P_t \mathbb{1}_N(x)\,dx &= \iint g(x)\mathbb{1}_N(x+y)\,p_t(dy)\,dx \\ &= \int \underbrace{\int g(x)\mathbb{1}_N(x+y)\,dx}_{=0}\; p_t(dy) = 0.\end{aligned}$$

Take $g = P_t\mathbb{1}_N$, then the above calculation shows

$$\int (P_t\mathbb{1}_N(x))^2\,dx = 0.$$

Hence, $P_t\mathbb{1}_N = 0$ Lebesgue-a.e. By the strong Feller property, $P_t\mathbb{1}_N$ is continuous, and so $P_t\mathbb{1}_N \equiv 0$, hence

$$p_t(N) = P_t\mathbb{1}_N(0) = 0.$$ □

Remark 4.10. The existence and smoothness of densities for a Lévy process are time-dependent properties, cf. Sato [51, Chapter V.23]. The typical example is the **Gamma process**. This is a (one-dimensional) Lévy process with characteristic exponent

$$\psi(\xi) = \frac{1}{2}\log(1 + |\xi|^2) - \mathrm{i}\arctan\xi, \quad \xi \in \mathbb{R},$$

and this process has the transition density

$$p_t(x) = \frac{1}{\Gamma(t)}\, x^{t-1}\mathrm{e}^{-x}\,\mathbb{1}_{(0,\infty)}(x), \quad t > 0.$$

The factor x^{t-1} gives a time-dependent condition for the property $p_t \in L^p(dx)$. One can show, cf. [30], that

$$\lim_{|\xi|\to\infty} \frac{\operatorname{Re}\psi(\xi)}{\log(1+|\xi|^2)} = \infty \implies \forall t > 0 \quad \exists p_t \in \mathcal{C}^\infty(\mathbb{R}^d).$$

The converse direction remains true if $\psi(\xi)$ is rotationally invariant or if it is replaced by its symmetric rearrangement.

Remark 4.11. If $(P_t)_{t\geqslant 0}$ is a **Feller semigroup**, i.e., a semigroup satisfying the conditions 4.7.a)-g), then there exists a unique stochastic process (a **Feller process**) with $(P_t)_{t\geqslant 0}$ as transition semigroup. The idea is to use Kolmogorov's consistency theorem for the following family of finite-dimensional distributions

$$p^x_{t_1,\dots,t_n}(B_1\times\dots\times B_n) = P_{t_1}\Big(\mathbb{1}_{B_1}P_{t_2-t_1}\big(\mathbb{1}_{B_2}P_{t_3-t_2}(\dots P_{t_n-t_{n-1}}(\mathbb{1}_{B_n}))\big)\Big)(x)$$

Here $X_{t_0}=X_0=x$ a.s. Note: It is **not enough** to have a semigroup on L^p as we need pointwise evaluations.

If the operators P_t are not *a priori* given on $\mathcal{B}_b(\mathbb{R}^d)$ but only on $\mathcal{C}_\infty(\mathbb{R}^d)$, one still can use the Riesz representation theorem to construct Markov kernels $p_t(x,B)$ representing and extending P_t onto $\mathcal{B}_b(\mathbb{R}^d)$, cf. Lemma 5.2.

Recall that a **stopping time** is a random time $\tau:\Omega\to[0,\infty]$ such that $\{\tau\leqslant t\}\in\mathscr{F}_t$ for all $t\geqslant 0$. It is not hard to see that $\tau_n := (\lfloor 2^n\tau\rfloor+1)2^{-n}$, $n\in\mathbb{N}$, is a sequence of stopping times with values $k2^{-n}$, $k=1,2,\dots$, such that

$$\tau_1\geqslant\tau_2\geqslant\ldots\geqslant\tau_n\downarrow\tau=\inf_{n\in\mathbb{N}}\tau_n.$$

This approximation is the key ingredient to extend the Markov property (Theorem 4.6) to random times.

Theorem 4.12 (Strong Markov property for Lévy processes)**.** *Let X be a Lévy process on $\mathbb{R}^d$ and set $Y:=(X_{t+\tau}-X_\tau)_{t\geqslant 0}$ for some a.s. finite stopping time τ. Then Y is again a Lévy process satisfying*

a) *$Y\perp(X_r)_{r\leqslant\tau}$, i.e., $\mathscr{F}^Y_\infty\perp\mathscr{F}^X_{\tau+}:=\big\{F\in\mathscr{F}^X_\infty : F\cap\{\tau<t\}\in\mathscr{F}^X_t\ \ \forall t\geqslant 0\big\}$.*
b) *$Y\sim X$, i.e., X and Y have the same finite-dimensional distributions.*

Proof. Let $\tau_n:=(\lfloor 2^n\tau\rfloor+1)2^{-n}$. For all $0\leqslant s<t$, $\xi\in\mathbb{R}^d$ and $F\in\mathscr{F}^X_{\tau+}$ we find by the right-continuity of the sample paths (or by the continuity in probability (L3))

$$\begin{aligned}
\mathbb{E}\Big[e^{i\,\xi\cdot(X_{t+\tau}-X_{s+\tau})}\,\mathbb{1}_F\Big] &= \lim_{n\to\infty}\mathbb{E}\Big[e^{i\,\xi\cdot(X_{t+\tau_n}-X_{s+\tau_n})}\,\mathbb{1}_F\Big]\\
&= \lim_{n\to\infty}\sum_{k=1}^\infty\mathbb{E}\Big[e^{i\,\xi\cdot(X_{t+k2^{-n}}-X_{s+k2^{-n}})}\,\mathbb{1}_{\{\tau_n=k2^{-n}\}}\cdot\mathbb{1}_F\Big]\\
&= \lim_{n\to\infty}\sum_{k=1}^\infty\mathbb{E}\Big[\underbrace{e^{i\,\xi\cdot(X_{t+k2^{-n}}-X_{s+k2^{-n}})}}_{\perp\,\mathscr{F}^X_{k2^{-n}}\text{ by (L2)}}\ \underbrace{\mathbb{1}_{\{(k-1)2^{-n}\leqslant\tau<k2^{-n}\}}\,\mathbb{1}_F}_{\in\mathscr{F}^X_{k2^{-n}}\text{ as }F\in\mathscr{F}^X_{\tau+}}\Big]\\
&= \lim_{n\to\infty}\sum_{k=1}^\infty\mathbb{E}\big[e^{i\,\xi\cdot X_{t-s}}\big]\,\mathbb{P}\Big(\big\{(k-1)2^{-n}\leqslant\tau<k2^{-n}\big\}\cap F\Big)\\
&= \mathbb{E}\big[e^{i\,\xi\cdot X_{t-s}}\big]\,\mathbb{P}(F).
\end{aligned}$$

In the last equality we use $\bigcupdot_{k=1}^{\infty}\{(k-1)2^{-n} \leqslant \tau < k2^{-n}\} = \{\tau < \infty\}$ for all $n \geqslant 1$.

The same calculation applies to finitely many increments. Let $F \in \mathscr{F}_{\tau+}^X$ and $t_0 = 0 < t_1 < \cdots < t_n$, $\xi_1, \ldots, \xi_n \in \mathbb{R}^d$. Then

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\sum_{k=1}^{n}\xi_k\cdot(X_{t_k+\tau}-X_{t_{k-1}+\tau})}\,\mathbb{1}_F\right] = \prod_{k=1}^{n}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\,\xi_k\cdot X_{t_k-t_{k-1}}}\right]\mathbb{P}(F).$$

This shows that the increments $X_{t_k+\tau} - X_{t_{k-1}+\tau}$ are independent and distributed like $X_{t_k-t_{k-1}}$. Moreover, all increments are independent of $F \in \mathscr{F}_{\tau+}^X$. Therefore, all random vectors of the form $(X_{t_1+\tau}-X_\tau, \ldots, X_{t_n+\tau}-X_{t_{n-1}+\tau})$ are independent of $\mathscr{F}_{\tau+}^X$, and we conclude that $\mathscr{F}_\infty^Y = \sigma(X_{t+\tau} - X_\tau,\, t \geqslant 0) \perp \mathscr{F}_{\tau+}^X$. □

Chapter 5

A Digression: Semigroups

We have seen that the Markov kernel $p_t(x, B)$ of a Lévy or Markov process induces a semigroup of linear operators $(P_t)_{t\geqslant 0}$. In this chapter we collect a few tools from functional analysis for the study of operator semigroups. By $\mathcal{B}_b(\mathbb{R}^d)$ we denote the bounded Borel functions $f : \mathbb{R}^d \to \mathbb{R}$, and $\mathcal{C}_\infty(\mathbb{R}^d)$ are the continuous functions vanishing at infinity, i.e., $\lim_{|x|\to\infty} f(x) = 0$; when equipped with the uniform norm $\|\cdot\|_\infty$ both sets become Banach spaces.

Definition 5.1. A **Feller semigroup** is a family of linear operators

$$P_t : \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d)$$

satisfying the properties a)–g) of Definition 4.7: $(P_t)_{t\geqslant 0}$ is a semigroup of conservative, sub-Markovian operators which enjoy the **Feller property** $P_t(\mathcal{C}_\infty(\mathbb{R}^d)) \subset \mathcal{C}_\infty(\mathbb{R}^d)$ and which are strongly continuous on $\mathcal{C}_\infty(\mathbb{R}^d)$.

Notice that $(t, x) \mapsto P_t f(x)$ is for every $f \in \mathcal{C}_\infty(\mathbb{R}^d)$ continuous. This follows from

$$\begin{aligned}|P_t f(x) - P_s f(y)| &\leqslant |P_t f(x) - P_t f(y)| + |P_t f(y) - P_s f(y)|\\ &\leqslant |P_t f(x) - P_t f(y)| + \|P_{|t-s|} f - f\|_\infty,\end{aligned}$$

the Feller property 4.7.f) and the strong continuity 4.7.g).

Lemma 5.2. *If $(P_t)_{t\geqslant 0}$ is a Feller semigroup, then there exists a Markov transition function $p_t(x, dy)$ (Definition 4.1) such that $P_t f(x) = \int f(y)\, p_t(x, dy)$.*

Proof. By the Riesz representation theorem we see that the operators P_t are of the form $P_t f(x) = \int f(y)\, p_t(x, dy)$ where $p_t(x, dy)$ is a Markov kernel. The tricky part is to show the joint measurability $(t, x) \mapsto p_t(x, B)$ and the Chapman–Kolmogorov identities (4.2).

For every compact set $K \subset \mathbb{R}^d$ the functions defined by

$$f_n(x) := \frac{d(x, U_n^c)}{d(x, K) + d(x, U_n^c)}, \quad d(x, A) := \inf_{a\in A} |x - a|, \quad U_n := \{y : d(y, K) < 1/n\},$$

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_5

are in $\mathcal{C}_\infty(\mathbb{R}^d)$ and $f_n \downarrow \mathbb{1}_K$. By monotone convergence, $p_t(x,K) = \inf_{n\in\mathbb{N}} P_t f_n(x)$ which proves the joint measurability in (t,x) for all compact sets.

By the same, the semigroup property $P_{t+s}f_n = P_s P_t f_n$ entails the Chapman–Kolmogorov identities for compact sets: $p_{t+s}(x,K) = \int p_t(y,K)\,p_s(x,dy)$. Since

$$\mathscr{D} := \left\{ B \in \mathscr{B}(\mathbb{R}^d) \;\middle|\; \begin{array}{l} (t,x) \mapsto p_t(x,B) \;\text{ is measurable \&} \\ p_{t+s}(x,B) = \displaystyle\int p_t(y,B)\,p_s(x,dy) \end{array} \right\}$$

is a Dynkin system containing the compact sets, we have $\mathscr{D} = \mathscr{B}(\mathbb{R}^d)$. □

To get an intuition for semigroups it is a good idea to view the semigroup property

$$P_{t+s} = P_s \circ P_t \quad \text{and} \quad P_0 = \mathrm{id}$$

as an operator-valued Cauchy functional equation. If $t \mapsto P_t$ is – in a suitable sense – continuous, the unique solution will be of the form $P_t = e^{tA}$ for some **operator** A. This can be easily made rigorous for matrices $A, P_t \in \mathbb{R}^{n\times n}$ since the matrix exponential is well-defined by the uniformly convergent series

$$P_t = \exp(tA) := \sum_{k=0}^\infty \frac{t^k A^k}{k!} \quad \text{and} \quad A = \frac{d}{dt}P_t\Big|_{t=0}$$

with $A^0 := \mathrm{id}$ and $A^k = A \circ A \circ \cdots \circ A$ (k times). With a bit more care, this can be made to work also in general settings.

Definition 5.3. Let $(P_t)_{t\geqslant 0}$ be a Feller semigroup. The (infinitesimal) **generator** is a linear operator defined by

$$\mathcal{D}(A) := \left\{ f \in \mathcal{C}_\infty(\mathbb{R}^d) \;\middle|\; \exists g \in \mathcal{C}_\infty(\mathbb{R}^d) \,:\, \lim_{t\to 0} \left\| \frac{P_t f - f}{t} - g \right\|_\infty = 0 \right\} \tag{5.1}$$

$$Af := \lim_{t\to 0} \frac{P_t f - f}{t}, \quad f \in \mathcal{D}(A). \tag{5.2}$$

The next lemma is the rigorous version for the symbolic notation '$P_t = e^{tA}$'.

Lemma 5.4. *Let $(P_t)_{t\geqslant 0}$ be a Feller semigroup with infinitesimal generator $(A, \mathcal{D}(A))$. Then $P_t(\mathcal{D}(A)) \subset \mathcal{D}(A)$ and*

$$\frac{d}{dt}P_t f = AP_t f = P_t A f \quad \text{for all } \; f \in \mathcal{D}(A),\; t \geqslant 0. \tag{5.3}$$

Moreover, $\int_0^t P_s f\,ds \in \mathcal{D}(A)$ for any $f \in \mathcal{C}_\infty(\mathbb{R}^d)$, and

$$P_t f - f = A\int_0^t P_s f\,ds, \quad f \in \mathcal{C}_\infty(\mathbb{R}^d),\; t > 0 \tag{5.4}$$

$$= \int_0^t P_s A f\,ds, \quad f \in \mathcal{D}(A),\; t > 0. \tag{5.5}$$

Proof. Let $0 < \epsilon < t$ and $f \in \mathcal{D}(A)$. The semigroup and contraction properties give

$$\begin{aligned}\left\|\frac{P_t f - P_{t-\epsilon} f}{\epsilon} - P_t A f\right\|_\infty &\leqslant \left\|P_{t-\epsilon}\frac{P_\epsilon f - f}{\epsilon} - P_{t-\epsilon} A f\right\|_\infty + \left\|P_{t-\epsilon} A f - P_{t-\epsilon} P_\epsilon A f\right\|_\infty \\ &\leqslant \left\|\frac{P_\epsilon f - f}{\epsilon} - A f\right\|_\infty + \left\|A f - P_\epsilon A f\right\|_\infty \xrightarrow[\epsilon\to 0]{} 0\end{aligned}$$

where we use the strong continuity in the last step. This shows $\frac{d^-}{dt} P_t f = A P_t f = P_t A f$; a similar (but simpler) calculation proves this also for $\frac{d^+}{dt} P_t f$.

Let $f \in \mathcal{C}_\infty(\mathbb{R}^d)$ and $t, \epsilon > 0$. By Fubini's theorem and the representation of P_t with a Markov transition function (Lemma 5.2) we get

$$P_\epsilon \int_0^t P_s f(x)\, ds = \int_0^t P_\epsilon P_s f(x)\, ds,$$

and so,

$$\begin{aligned}\frac{P_\epsilon - \mathrm{id}}{\epsilon} \int_0^t P_s f(x)\, ds &= \frac{1}{\epsilon} \int_0^t \big(P_{s+\epsilon} f(x) - P_s f(x)\big)\, ds \\ &= \frac{1}{\epsilon} \int_t^{t+\epsilon} P_s f(x)\, ds - \frac{1}{\epsilon} \int_0^\epsilon P_s f(x)\, ds.\end{aligned}$$

Since $t \mapsto P_t f(x)$ is continuous, the fundamental theorem of calculus applies, and we get $\lim_{\epsilon\to 0} \frac{1}{\epsilon} \int_r^{r+\epsilon} P_s f(x)\, ds = P_r f(x)$ for $r \geqslant 0$. This shows that $\int_0^t P_s f\, ds \in \mathcal{D}(A)$ as well as (5.4). If $f \in \mathcal{D}(A)$, then we deduce (5.5) from

$$\int_0^t P_s A f(x)\, ds \overset{(5.3)}{=} \int_0^t \frac{d}{ds} P_s f(x)\, ds = P_t f(x) - f(x) \overset{(5.4)}{=} A \int_0^t P_s f(x)\, ds. \quad \square$$

Remark 5.5 (Consequences of Lemma 5.4)**.** Write $\mathcal{C}_\infty := \mathcal{C}_\infty(\mathbb{R}^d)$.

a) (5.4) shows that $\mathcal{D}(A)$ is dense in $\mathcal{C}_\infty$, since $\mathcal{D}(A) \ni t^{-1} \int_0^t P_s f\, ds \xrightarrow[t\to 0]{} f$ for any $f \in \mathcal{C}_\infty$.

b) (5.5) shows that A is a **closed operator**, i.e.,

$$f_n \in \mathcal{D}(A),\ (f_n, A f_n) \xrightarrow[n\to\infty]{\text{uniformly}} (f, g) \in \mathcal{C}_\infty \times \mathcal{C}_\infty \implies f \in \mathcal{D}(A)\ \&\ Af = g.$$

c) (5.3) means that A determines $(P_t)_{t\geqslant 0}$ uniquely.

Let us now consider the Laplace transform of $(P_t)_{t\geqslant 0}$.

Definition 5.6. Let $(P_t)_{t\geqslant 0}$ be a Feller semigroup. The **resolvent** is a linear operator on $\mathcal{B}_b(\mathbb{R}^d)$ given by

$$R_\lambda f(x) := \int_0^\infty \mathrm{e}^{-\lambda t} P_t f(x)\, dt, \quad f \in \mathcal{B}_b(\mathbb{R}^d),\ x \in \mathbb{R}^d,\ \lambda > 0. \tag{5.6}$$

The following formal calculation can easily be made rigorous. Let $f \in \mathcal{D}(A)$ and $(\lambda - A) := (\lambda \operatorname{id} - A)$ for $\lambda > 0$. Then

$$\begin{aligned}
(\lambda - A)R_\lambda f &= (\lambda - A)\int_0^\infty e^{-\lambda t} P_t f\,dt \\
&\overset{(5.4),(5.5)}{=} \int_0^\infty e^{-\lambda t}(\lambda - A)P_t f\,dt \\
&= \lambda\int_0^\infty e^{-\lambda t} P_t f\,dt - \int_0^\infty e^{-\lambda t}\left(\frac{d}{dt}P_t f\right)dt \\
&\overset{\text{parts}}{=} \lambda\int_0^\infty e^{-\lambda t} P_t f\,dt - \lambda\int_0^\infty e^{-\lambda t} P_t f\,dt - [e^{-\lambda t}P_t f]_{t=0}^\infty = f.
\end{aligned}$$

A similar calculation for $R_\lambda(\lambda - A)$ gives

Theorem 5.7. *Let $(A, \mathcal{D}(A))$ and $(R_\lambda)_{\lambda>0}$ be the generator and the resolvent of a Feller semigroup. Then*

$$R_\lambda = (\lambda - A)^{-1} \quad \textit{for all} \quad \lambda > 0.$$

Since R_λ is the Laplace transform of $(P_t)_{t\geqslant 0}$, the properties of $(R_\lambda)_{\lambda>0}$ can be found from $(P_t)_{t\geqslant 0}$ and vice versa. With some effort one can even invert the (operator-valued) Laplace transform which leads to the familiar expression for e^x:

$$\left(\frac{n}{t}R_{\frac{n}{t}}\right)^n = \left(\operatorname{id} - \frac{t}{n}A\right)^{-n} \xrightarrow[n\to\infty]{\text{strongly}} e^{tA} = P_t \tag{5.7}$$

(the notation $e^{tA} = P_t$ is, for unbounded operators A, formal), see Pazy [42, Chapter 1.8].

Lemma 5.8. *Let $(R_\lambda)_{\lambda>0}$ be the resolvent of a Feller[1] semigroup $(P_t)_{t\geqslant 0}$. Then*

$$\frac{d^n}{d\lambda^n}R_\lambda = n!(-1)^n R_\lambda^{n+1} \quad n \in \mathbb{N}_0. \tag{5.8}$$

Proof. Using a symmetry argument we see

$$t^n = \int_0^t \dots \int_0^t dt_1 \dots dt_n = n! \int_0^t \int_0^{t_n} \dots \int_0^{t_2} dt_1 \dots dt_n.$$

Let $f \in \mathcal{C}_\infty(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$. Then

$$(-1)^n \frac{d^n}{d\lambda^n}R_\lambda f(x) = \int_0^\infty (-1)^n \frac{d^n}{d\lambda^n} e^{-\lambda t} P_t f(x)\,dt = \int_0^\infty t^n e^{-\lambda t} P_t f(x)\,dt$$

[1] This lemma only needs that the operators P_t are strongly continuous and contractive, Definition 4.7.g), d).

$$\begin{aligned}
&= n! \int_0^\infty \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} e^{-\lambda t} P_t f(x)\, dt_1 \ldots dt_n\, dt \\
&= n! \int_0^\infty \int_{t_n}^\infty \cdots \int_{t_1}^\infty e^{-\lambda t} P_t f(x)\, dt\, dt_1 \ldots dt_n \\
&= n! \int_0^\infty \int_0^\infty \cdots \int_0^\infty e^{-\lambda(t+t_1+\cdots+t_n)} P_{t+t_1+\cdots+t_n} f(x)\, dt\, dt_1 \ldots dt_n \\
&= n!\, R_\lambda^{n+1} f(x). \qquad \square
\end{aligned}$$

The key result identifying the generators of Feller semigroups is the following theorem due to Hille, Yosida and Ray, a proof can be found in Pazy [42, Chapter 1.4] or Ethier & Kurtz [17, Chapter 4.2]; a probabilistic approach is due to Itô [25].

Theorem 5.9 (Hille–Yosida–Ray)**.** *A linear operator $(A, \mathcal{D}(A))$ on $\mathcal{C}_\infty(\mathbb{R}^d)$ generates a Feller semigroup $(P_t)_{t\geqslant 0}$ if, and only if,*

a) $\mathcal{D}(A) \subset \mathcal{C}_\infty(\mathbb{R}^d)$ *dense.*
b) A *is* **dissipative***, i.e.,* $\|\lambda f - Af\|_\infty \geqslant \lambda \|f\|_\infty$ *for some (or all)* $\lambda > 0$.
c) $(\lambda - A)(\mathcal{D}(A)) = \mathcal{C}_\infty(\mathbb{R}^d)$ *for some (or all)* $\lambda > 0$.
d) A *satisfies the* **positive maximum principle***:*

$$f \in \mathcal{D}(A),\ f(x_0) = \sup_{x\in\mathbb{R}^d} f(x) \geqslant 0 \implies Af(x_0) \leqslant 0. \tag{PMP}$$

This variant of the Hille–Yosida theorem is not the standard version from functional analysis since we are interested in positivity preserving (sub-Markov) semigroups. Let us briefly discuss the role of the positive maximum principle.

Remark 5.10. Let $(P_t)_{t\geqslant 0}$ be a strongly continuous contraction semigroup on $\mathcal{C}_\infty(\mathbb{R}^d)$, i.e., $\|P_t f\|_\infty \leqslant \|f\|_\infty$ and $\lim_{t\to 0} \|P_t f - f\|_\infty = 0$, cf. Definition 4.7.d),g).[2]

1° Sub-Markov ⇒ (PMP). Assume that $f \in \mathcal{D}(A)$ is such that $f(x_0) = \sup f \geqslant 0$. Then

$$P_t f(x_0) - f(x_0) \overset{f\leqslant f^+}{\leqslant} P_t^+ f(x_0) - f^+(x_0) \leqslant \|f^+\|_\infty - f^+(x_0) = 0.$$

$$\implies Af(x_0) = \lim_{t\to 0} \frac{P_t f(x_0) - f(x_0)}{t} \leqslant 0.$$

Thus, (PMP) holds.

2° (PMP) ⇒ dissipativity. Assume that (PMP) holds and let $f \in \mathcal{D}(A)$. Since $f \in \mathcal{C}_\infty(\mathbb{R}^d)$, we may assume that $f(x_0) = |f(x_0)| = \sup|f|$ (otherwise $f \rightsquigarrow -f$). Then

$$\|\lambda f - Af\|_\infty \geqslant \lambda f(x_0) - \underbrace{Af(x_0)}_{\leqslant 0} \geqslant \lambda f(x_0) = \lambda \|f\|_\infty.$$

[2] These properties are essential for the existence of a generator and the resolvent on $\mathcal{C}_\infty(\mathbb{R}^d)$.

3° (PMP) ⇒ sub-Markov. Since P_t is contractive, we have $P_t f(x) \leqslant \|P_t f\|_\infty \leqslant \|f\|_\infty \leqslant 1$ for all $f \in \mathcal{C}_\infty(\mathbb{R}^d)$ such that $|f| \leqslant 1$. In order to see positivity, let $f \in \mathcal{C}_\infty(\mathbb{R}^d)$ be non-negative. We distinguish between two cases:

a) $R_\lambda f$ does not attain its infimum.] Since $R_\lambda f \in \mathcal{C}_\infty(\mathbb{R}^d)$ vanishes at infinity, we have necessarily $R_\lambda f \geqslant 0$.

b) $\exists x_0 \ :\ R_\lambda f(x_0) = \inf R_\lambda f$. Because of the (PMP) we find

$$\lambda R_\lambda f(x_0) - f(x_0) = A R_\lambda f(x_0) \geqslant 0$$
$$\implies \lambda R_\lambda f(x) \geqslant \inf \lambda R_\lambda f = \lambda R_\lambda f(x_0) \geqslant f(x_0) \geqslant 0.$$

This proves that $f \geqslant 0 \implies \lambda R_\lambda f \geqslant 0$. From (5.8) we see that $\lambda \mapsto R_\lambda f(x)$ is completely monotone, hence it is the Laplace transform of a positive measure. Since $R_\lambda f(x)$ has the integral representation (5.6), we see that $P_t f(x) \geqslant 0$ (for all $t \geqslant 0$ as $t \mapsto P_t f$ is continuous).

Using the Riesz representation theorem (as in Lemma 5.2) we can extend P_t as a sub-Markov operator onto $\mathcal{B}_b(\mathbb{R}^d)$.

In order to determine the domain $\mathcal{D}(A)$ of the generator the following 'maximal dissipativity' result is handy.

Lemma 5.11 (Dynkin, Reuter)**.** *Assume that $(A, \mathcal{D}(A))$ generates a Feller semigroup and that $(\mathfrak{A}, \mathcal{D}(\mathfrak{A}))$ extends A, i.e., $\mathcal{D}(A) \subset \mathcal{D}(\mathfrak{A})$ and $\mathfrak{A}|_{\mathcal{D}(A)} = A$. If*

$$u \in \mathcal{D}(\mathfrak{A}),\ u - \mathfrak{A}u = 0 \implies u = 0, \tag{5.9}$$

then $(A, \mathcal{D}(A)) = (\mathfrak{A}, \mathcal{D}(\mathfrak{A}))$.

Proof. Since A is a generator, $(\mathrm{id} - A) : \mathcal{D}(A) \to \mathcal{C}_\infty(\mathbb{R}^d)$ is bijective. On the other hand, the relation (5.9) means that $(\mathrm{id} - \mathfrak{A})$ is injective, but $(\mathrm{id} - A)$ cannot have a proper injective extension. □

Theorem 5.12. *Let $(P_t)_{t\geqslant 0}$ be a Feller semigroup with generator $(A, \mathcal{D}(A))$. Then*

$$\mathcal{D}(A) = \left\{ f \in \mathcal{C}_\infty(\mathbb{R}^d) \;\middle|\; \exists g \in \mathcal{C}_\infty(\mathbb{R}^d)\ \forall x \ :\ \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} = g(x) \right\}. \tag{5.10}$$

Proof. Denote by $\mathcal{D}(\mathfrak{A})$ the right-hand side of (5.10) and define

$$\mathfrak{A} f(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} \quad \text{for all} \quad f \in \mathcal{D}(\mathfrak{A}),\ x \in \mathbb{R}^d.$$

Obviously, $(\mathfrak{A}, \mathcal{D}(\mathfrak{A}))$ is a linear operator which extends $(A, \mathcal{D}(A))$. Since (PMP) is, essentially, a pointwise assertion (see Remark 5.10, 1°), $\mathfrak{A}$ inherits (PMP); in particular, $\mathfrak{A}$ is dissipative (see Remark 5.10, 2°):

$$\|\mathfrak{A} f - \lambda f\|_\infty \geqslant \lambda \|f\|_\infty.$$

This implies (5.9), and the claim follows from Lemma 5.11. □

Chapter 6

The Generator of a Lévy Process

We want to study the structure of the generator of (the semigroup corresponding to) a Lévy process $X = (X_t)_{t\geqslant 0}$. This will also lead to a proof of the Lévy–Khintchine formula.

Our approach uses some Fourier analysis. We denote by $\mathcal{C}_c^\infty(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d)$ the smooth, compactly supported functions and the smooth, rapidly decreasing 'Schwartz functions'.[1] The **Fourier transform** is denoted by

$$\widehat{f}(\xi) = \mathcal{F}f(\xi) := (2\pi)^{-d}\int_{\mathbb{R}^d} f(x)\,\mathrm{e}^{-\mathrm{i}\,\xi\cdot x}\,dx, \quad f \in L^1(dx).$$

Observe that $\mathcal{F}f$ is chosen in such a way that the **characteristic function becomes the inverse Fourier transform**.

We have seen in Proposition 2.3 and its Corollaries 2.4 and 2.5 that X is completely characterized by the characteristic exponent $\psi : \mathbb{R}^d \to \mathbb{C}$

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t} = \left[\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_1}\right]^t = \mathrm{e}^{-t\psi(\xi)}, \quad t \geqslant 0,\ \xi \in \mathbb{R}^d.$$

We need a few more properties of ψ which result from the fact that $\chi(\xi) = \mathrm{e}^{-\psi(\xi)}$ is a characteristic function.

Lemma 6.1. *Let $\chi(\xi)$ be the characteristic function of a probability measure μ. Then*

$$|\chi(\xi+\eta) - \chi(\xi)\chi(\eta)|^2 \leqslant (1-|\chi(\xi)|^2)(1-|\chi(\eta)|^2), \quad \xi,\eta \in \mathbb{R}^d. \tag{6.1}$$

Proof. Since μ is a probability measure, we find from the definition of χ

$$\begin{aligned}\chi(\xi+\eta) - \chi(\xi)\chi(\eta) &= \iint \left(\mathrm{e}^{\mathrm{i}\,x\cdot\xi}\mathrm{e}^{\mathrm{i}\,x\cdot\eta} - \mathrm{e}^{\mathrm{i}\,x\cdot\xi}\mathrm{e}^{\mathrm{i}\,y\cdot\eta}\right)\mu(dx)\,\mu(dy)\\ &= \frac{1}{2}\iint \left(\mathrm{e}^{\mathrm{i}\,x\cdot\xi} - \mathrm{e}^{\mathrm{i}\,y\cdot\xi}\right)\left(\mathrm{e}^{\mathrm{i}\,x\cdot\eta} - \mathrm{e}^{\mathrm{i}\,y\cdot\eta}\right)\mu(dx)\,\mu(dy).\end{aligned}$$

[1] To be precise, $f \in \mathcal{S}(\mathbb{R}^d)$, if $f \in \mathcal{C}^\infty(\mathbb{R}^d)$ and if $\sup_{x\in\mathbb{R}^d}(1+|x|^N)|\partial^\alpha f(x)| \leqslant c_{N,\alpha}$ for any $N \in \mathbb{N}_0$ and any multiindex $\alpha \in \mathbb{N}_0^d$.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_6

In the last equality we use that the integrand is symmetric in x and y, which allows us to interchange the variables. Using the elementary formula $|e^{i a} - e^{i b}|^2 = 2 - 2\cos(b-a)$ and the Cauchy–Schwarz inequality yield

$$\begin{aligned}
&|\chi(\xi+\eta) - \chi(\xi)\chi(\eta)| \\
&\quad\leqslant \frac{1}{2}\iint \left|e^{i x\cdot\xi} - e^{i y\cdot\xi}\right| \cdot \left|e^{i x\cdot\eta} - e^{i y\cdot\eta}\right| \mu(dx)\,\mu(dy) \\
&\quad= \iint \sqrt{1-\cos(y-x)\cdot\xi}\sqrt{1-\cos(y-x)\cdot\eta}\,\mu(dx)\,\mu(dy) \\
&\quad\leqslant \sqrt{\iint \big(1-\cos(y-x)\cdot\xi\big)\,\mu(dx)\,\mu(dy)}\sqrt{\iint \big(1-\cos(y-x)\cdot\eta\big)\,\mu(dx)\,\mu(dy)}.
\end{aligned}$$

This finishes the proof as

$$\iint \cos(y-x)\cdot\xi\,\mu(dx)\,\mu(dy) = \mathrm{Re}\left[\int e^{i y\cdot\xi}\,\mu(dy)\int e^{-i x\cdot\xi}\,\mu(dx)\right] = |\chi(\xi)|^2. \qquad \square$$

Theorem 6.2. *Let $\psi : \mathbb{R}^d \to \mathbb{C}$ be the characteristic exponent of a Lévy process. Then $\xi \mapsto \sqrt{|\psi(\xi)|}$ is subadditive and*

$$|\psi(\xi)| \leqslant c_\psi(1+|\xi|^2), \quad \xi \in \mathbb{R}^d. \tag{6.2}$$

Proof. We use (6.1) with $\chi = e^{-t\psi}$, divide by $t>0$ and let $t\to 0$. Since we have $|\chi| = e^{-t\,\mathrm{Re}\,\psi}$, this gives

$$|\psi(\xi+\eta) - \psi(\xi) - \psi(\eta)|^2 \leqslant 4\,\mathrm{Re}\,\psi(\xi)\,\mathrm{Re}\,\psi(\eta) \leqslant 4|\psi(\xi)|\cdot|\psi(\eta)|.$$

By the lower triangle inequality,

$$|\psi(\xi+\eta)| - |\psi(\xi)| - |\psi(\eta)| \leqslant 2\sqrt{|\psi(\xi)|}\sqrt{|\psi(\eta)|}$$

and this is the same as subadditivity: $\sqrt{|\psi(\xi+\eta)|} \leqslant \sqrt{|\psi(\xi)|} + \sqrt{|\psi(\eta)|}$.

In particular, $|\psi(2\xi)| \leqslant 4|\psi(\xi)|$. For any $\xi \neq 0$ there is some $n = n(\xi) \in \mathbb{Z}$ such that $2^{n-1} \leqslant |\xi| \leqslant 2^n$, so

$$|\psi(\xi)| = |\psi(2^n 2^{-n}\xi)| \leqslant \max\{1, 2^{2n}\} \sup_{|\eta|\leqslant 1}|\psi(\eta)| \leqslant 2\sup_{|\eta|\leqslant 1}|\psi(\eta)|(1+|\xi|^2). \qquad \square$$

Lemma 6.3. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process and denote by $(A, \mathcal{D}(A))$ its generator. Then $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$.*

Proof. Let $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. By definition, $P_t f(x) = \mathbb{E} f(X_t + x)$. Using the differentiation lemma for parameter-dependent integrals (e.g., [54, Theorem 11.5] or [55, 12.2]) it is not hard to see that $P_t : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$. Obviously,

$$e^{-t\psi(\xi)} = \mathbb{E}\, e^{i\,\xi\cdot X_t} = \mathbb{E}^x e^{i\,\xi\cdot(X_t - x)} = e_{-\xi}(x) P_t e_\xi(x) \tag{6.3}$$

for $\mathrm{e}_\xi(x) := \mathrm{e}^{\mathrm{i}\,\xi\cdot x}$. Recall that the Fourier transform of $f \in \mathcal{S}(\mathbb{R}^d)$ is again in $\mathcal{S}(\mathbb{R}^d)$. From

$$P_t f \;=\; P_t \int \widehat{f}(\xi)\mathrm{e}_\xi(\cdot)\,d\xi \;=\; \int \widehat{f}(\xi)P_t\mathrm{e}_\xi(\cdot)\,d\xi \stackrel{(6.3)}{=} \int \widehat{f}(\xi)\mathrm{e}_\xi(\cdot)\mathrm{e}^{-t\psi(\xi)}\,d\xi \tag{6.4}$$

we conclude that $\widehat{P_t f} = \widehat{f}\mathrm{e}^{-t\psi}$. Hence,

$$P_t f = \mathcal{F}^{-1}(\widehat{f}\mathrm{e}^{-t\psi}). \tag{6.5}$$

Consequently,

$$\frac{\widehat{P_t f - f}}{t} = \frac{\mathrm{e}^{-t\psi}\widehat{f} - \widehat{f}}{t} \xrightarrow[t\to 0]{} -\psi\widehat{f}$$

$$\xRightarrow{\widehat{f}\in\mathcal{S}(\mathbb{R}^d)} \quad \frac{P_t f(x) - f(x)}{t} \xrightarrow[t\to 0]{} g(x) := \mathcal{F}^{-1}(-\psi\widehat{f})(x).$$

Since ψ grows at most polynomially (Lemma 6.2) and $\widehat{f} \in \mathcal{S}(\mathbb{R}^d)$, we know that $\psi\widehat{f} \in L^1(dx)$ and, by the Riemann–Lebesgue lemma, $g \in \mathcal{C}_\infty(\mathbb{R}^d)$. Using Theorem 5.12 it follows that $f \in \mathcal{D}(A)$. □

Definition 6.4. Let $L : \mathcal{C}_b^2(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$ be a linear operator. Then

$$L(x,\xi) := \mathrm{e}_{-\xi}(x)L_x\mathrm{e}_\xi(x) \tag{6.6}$$

is the **symbol** of the operator $L = L_x$, where $\mathrm{e}_\xi(x) := \mathrm{e}^{\mathrm{i}\,\xi\cdot x}$.

The proof of Lemma 6.3 actually shows that we can recover an operator L from its symbol $L(x,\xi)$ if, say, $L : \mathcal{C}_b^2(\mathbb{R}^d) \to \mathcal{C}_b(\mathbb{R}^d)$ is continuous:[2] Indeed, for all $u \in \mathcal{C}_c^\infty(\mathbb{R}^d)$

$$\begin{aligned} Lu(x) &= L \int \widehat{u}(\xi)\mathrm{e}_\xi(x)\,d\xi \\ &= \int \widehat{u}(\xi)L_x\mathrm{e}_\xi(x)\,d\xi \\ &= \int \widehat{u}(\xi)L(x,\xi)\mathrm{e}_\xi(x)d\xi = \mathcal{F}^{-1}(L(x,\cdot)\mathcal{F}u(\cdot))(x). \end{aligned}$$

Example 6.5. A typical example would be the Laplace operator (i.e., the generator of a Brownian motion)

$$\tfrac{1}{2}\Delta f(x) = -\tfrac{1}{2}(\tfrac{1}{\mathrm{i}}\partial_x)^2 f(x) = \int \widehat{f}(\xi)\big(-\tfrac{1}{2}|\xi|^2\big)\mathrm{e}^{\mathrm{i}\,\xi\cdot x}\,d\xi, \quad \text{i.e.,} \quad L(x,\xi) = -\frac{1}{2}|\xi|^2,$$

[2] As usual, $\mathcal{C}_b^2(\mathbb{R}^d)$ is endowed with the norm $\|u\|_{(2)} = \sum_{0\leqslant|\alpha|\leqslant 2}\|\partial^\alpha u\|_\infty$.

or the fractional Laplacian of order $\frac{1}{2}\alpha \in (0,1)$ which generates a rotationally symmetric α-stable Lévy process

$$-(-\Delta)^{\alpha/2} f(x) = \int \widehat{f}(\xi)\big(-|\xi|^{\alpha}\big)\mathrm{e}^{\mathrm{i}\,\xi\cdot x}\,d\xi, \quad \text{i.e.,} \quad L(x,\xi) = -|\xi|^{\alpha}.$$

More generally, if $P(x,\xi)$ is a polynomial in ξ, then the corresponding operator is obtained by replacing ξ by $\frac{1}{\mathrm{i}}\nabla_x$ and formally expanding the powers.

Definition 6.6. An operator of the form

$$L(x,D)f(x) = \int \widehat{f}(\xi)L(x,\xi)\mathrm{e}^{\mathrm{i}\,x\cdot\xi}\,d\xi, \quad f \in \mathcal{S}(\mathbb{R}^d), \tag{6.7}$$

is called (if defined) a **pseudo-differential operator** with (non-classical) symbol $L(x,\xi)$.

Remark 6.7. The **symbol of a Lévy process** does not depend on x, i.e., $L(x,\xi) = L(\xi)$. This is a consequence of the spatial homogeneity of the process which is encoded in the translation invariance of the semigroup (cf. (4.6) and Lemma 4.4):

$$P_t f(x) = \mathbb{E} f(X_t + x) \implies P_t f(x) = \vartheta_x(P_t f)(0) = P_t(\vartheta_x f)(0)$$

where $\vartheta_x u(y) = u(y+x)$ is the shift operator. This property is obviously inherited by the generator, i.e.,

$$Af(x) = \vartheta_x(Af)(0) = A(\vartheta_x f)(0), \quad f \in \mathcal{D}(A).$$

In fact, the converse is also true: If $L : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d)$ is a linear operator satisfying $\vartheta_x(Lf) = L(\vartheta_x f)$, then $Lf = f * \lambda$ where λ is a distribution, i.e., a continuous linear functional $\lambda : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathbb{R}$, cf. Theorem A.10.

Theorem 6.8. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process with generator A. Then*

$$Af(x) = l\cdot\nabla f(x) + \frac{1}{2}\nabla\cdot Q\nabla f(x) + \int\limits_{y\neq 0} \big[f(x+y) - f(x) - \nabla f(x)\cdot y\mathbb{1}_{(0,1)}(|y|)\big]\,\nu(dy) \tag{6.8}$$

for any $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, where $l \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d\times d}$ is a positive semidefinite matrix, and ν is a measure on $\mathbb{R}^d \setminus \{0\}$ such that $\int_{y\neq 0} \min\{1, |y|^2\}\,\nu(dy) < \infty$.

Equivalently, A is a pseudo-differential operator

$$Au(x) = -\psi(D)u(x) = -\int \widehat{u}(\xi)\psi(\xi)\mathrm{e}^{\mathrm{i}\,x\cdot\xi}\,d\xi, \quad u \in \mathcal{C}_c^\infty(\mathbb{R}^d), \tag{6.9}$$

whose symbol is the characteristic exponent $-\psi$ of the Lévy process. It is given by the Lévy–Khintchine formula

$$\psi(\xi) = -\mathrm{i}\,l\cdot\xi + \frac{1}{2}\xi\cdot Q\xi + \int_{y\neq 0}\big(1 - \mathrm{e}^{\mathrm{i}\,y\cdot\xi} + \mathrm{i}\,\xi\cdot y\mathbb{1}_{(0,1)}(|y|)\big)\,\nu(dy) \tag{6.10}$$

where the triplet (l, Q, ν) as above.

I learned the following proof from Francis Hirsch; it is based on arguments by Courrège [13] and Herz [20]. The presentation below follows the version in Böttcher, Schilling & Wang [9, Section 2.3].

Proof. The proof is divided into several steps.

1° We have seen in Lemma 6.3 that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$.

2° Set $A_0 f := (Af)(0)$ for $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. This is a linear functional on $\mathcal{C}_c^\infty$. Observe that

$$f \in \mathcal{C}_c^\infty(\mathbb{R}^d), \quad f \geqslant 0, \quad f(0) = 0 \quad \xRightarrow{\text{(PMP)}} \quad A_0 f \geqslant 0.$$

3° By 2°, $f \mapsto A_{00} f := A_0(|\cdot|^2 \cdot f)$ is a positive linear functional on $\mathcal{C}_c^\infty(\mathbb{R}^d)$. Therefore it is bounded. Indeed, let $f \in \mathcal{C}_c^\infty(K)$ for a compact set $K \subset \mathbb{R}^d$ and let $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ be a cut-off function such that $\mathbb{1}_K \leqslant \phi \leqslant 1$. Then

$$\|f\|_\infty \phi \pm f \geqslant 0.$$

By linearity and positivity $\|f\|_\infty A_{00}\phi \pm A_{00} f \geqslant 0$ which shows $|A_{00} f| \leqslant C_K \|f\|_\infty$ with $C_K = A_{00}\phi$.

By Riesz' representation theorem, there exists a Radon measure[3] μ such that

$$A_0(|\cdot|^2 f) = \int f(y)\,\mu(dy) = \int |y|^2 f(y) \underbrace{\frac{\mu(dy)}{|y|^2}}_{=:\nu(dy)} = \int |y|^2 f(y)\,\nu(dy).$$

This implies that

$$A_0 f_0 = \int_{y\neq 0} f_0(y)\,\nu(dy) \quad \text{for all} \quad f_0 \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\});$$

since any compact subset of $\mathbb{R}^d \setminus \{0\}$ is contained in an annulus $B_R(0) \setminus B_\epsilon(0)$, we have $\operatorname{supp} f_0 \cap B_\epsilon(0) = \emptyset$ for some sufficiently small $\epsilon > 0$. The measure ν is a Radon measure on $\mathbb{R}^d \setminus \{0\}$.

4° Let $f, g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, $0 \leqslant f, g \leqslant 1$, $\operatorname{supp} f \subset \overline{B_1(0)}$, $\operatorname{supp} g \subset \overline{B_1(0)}^c$ and $f(0) = 1$. From

$$\sup_{y\in\mathbb{R}^d} \big(\|g\|_\infty f(y) + g(y)\big) = \|g\|_\infty = \|g\|_\infty f(0) + g(0)$$

and (PMP), it follows that $A_0(\|g\|_\infty f + g) \leqslant 0$. Consequently,

$$A_0 g \leqslant -\|g\|_\infty A_0 f.$$

[3] A Radon measure on a topological space E is a Borel measure which is finite on compact subsets of E and regular: for all open sets $\mu(U) = \sup_{K\subset U} \mu(K)$ and for all Borel sets $\mu(B) = \inf_{U\supset B} \mu(U)$ (K, U are generic compact and open sets, respectively).

If $g \uparrow 1 - \mathbb{1}_{\overline{B_1(0)}}$, then this shows

$$\int_{|y|>1} \nu(dy) \leqslant -A_0 f < \infty.$$

Hence, $\int_{y\neq 0}(|y|^2 \wedge 1)\,\nu(dy) < \infty$.

5° Let $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $\phi(y) = \mathbb{1}_{(0,1)}(|y|)$. Define

$$S_0 f := \int_{y\neq 0} \big[f(y) - f(0) - y\cdot\nabla f(0)\phi(y)\big]\,\nu(dy). \tag{6.11}$$

By Taylor's formula, there is some $\theta \in (0,1)$ such that

$$f(y) - f(0) - y\cdot\nabla f(0)\phi(y) = \frac{1}{2}\sum_{k,l=1}^{d} \frac{\partial^2 f(\theta y)}{\partial x_k \partial x_l} y_k y_l.$$

Using the elementary inequality $2y_k y_l \leqslant y_k^2 + y_l^2 \leqslant |y|^2$, we obtain

$$\begin{aligned}|f(y) - f(0) - y\cdot\nabla f(0)\phi(y)| &\leqslant \begin{cases} \frac14 \sum_{k,l=1}^d \big\|\frac{\partial^2}{\partial x_k \partial x_l} f\big\|_\infty |y|^2, & |y| < 1\\ 2\|f\|_\infty, & |y| \geqslant 1\end{cases}\\ &\leqslant 2\|f\|_{(2)}(|y|^2 \wedge 1).\end{aligned}$$

This means that S_0 defines a distribution (generalized function) of order 2.

6° Set $L_0 := A_0 - S_0$. The steps 2° and 5° show that A_0 is a distribution of order 2. Moreover,

$$L_0 f_0 = \int_{y\neq 0} \big[f_0(0) - y\cdot\nabla f_0(0)\phi(y)\big]\,\nu(dy) = 0$$

for any $f_0 \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ with $f_0|_{B_\epsilon(0)} = 0$ for some $\epsilon > 0$. This implies that $\operatorname{supp}(L_0) \subset \{0\}$.

Let us show that L_0 is **almost positive** (also: 'fast positiv', 'prèsque positif'):

$$f_0 \in \mathcal{C}_c^\infty(\mathbb{R}^d),\quad f_0(0) = 0,\quad f_0 \geqslant 0 \quad\Longrightarrow\quad L_0 f_0 \geqslant 0. \tag{PP}$$

Indeed: Pick $0 \leqslant \phi_n \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\})$, $\phi_n \uparrow \mathbb{1}_{\mathbb{R}^d\setminus\{0\}}$ and let f_0 be as in (PP). Then

$$\begin{aligned} L_0 f_0 &\overset{\operatorname{supp} L_0 \subset\{0\}}{=} L_0[(1-\phi_n)f_0)]\\ &= A_0[(1-\phi_n)f_0] - S_0[(1-\phi_n)f_0]\\ &\underset{\nabla f_0(0)=0}{\overset{f_0(0)=0}{=}} A_0[(1-\phi_n)f_0] - \int_{y\neq 0}(1-\phi_n(y))f_0(y)\,\nu(dy)\\ &\underset{\text{(PMP)}}{\overset{2^\circ}{\geqslant}} -\int (1-\phi_n(y))f_0(y)\,\nu(dy) \xrightarrow[n\to\infty]{} 0\end{aligned}$$

by the monotone convergence theorem.

7° As in 5° we find with Taylor's formula for $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, $\operatorname{supp} f \subset K$ and $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying $\mathbb{1}_K \leqslant \phi \leqslant 1$

$$(f(y) - f(0) - \nabla f(0) \cdot y)\, \phi(y) \leqslant 2\|f\|_{(2)}|y|^2\phi(y).$$

(As usual, $\|f\|_{(2)} = \sum_{0\leqslant|\alpha|\leqslant 2} \|\partial^\alpha f\|_\infty$.) Therefore,

$$2\|f\|_{(2)}|y|^2\phi(y) + f(0)\phi(y) + \nabla f(0) \cdot y\phi(y) - f(y) \geqslant 0,$$

and (PP) implies

$$L_0 f \leqslant f(0)L_0\phi + |\nabla f(0)|L_0(|\cdot|\phi) + 2\|f\|_{(2)}L_0(|\cdot|^2\phi) \leqslant C_K\|f\|_{(2)}.$$

8° We have seen in 6° that L_0 is of order 2 and $\operatorname{supp} L_0 \subset \{0\}$. Therefore,

$$L_0 f = \frac{1}{2}\sum_{k,l=1}^d q_{kl}\frac{\partial^2 f(0)}{\partial x_k \partial x_l} + \sum_{k=1}^d l_k \frac{\partial f(0)}{\partial x_k} - cf(0). \tag{6.12}$$

We will show that $(q_{kl})_{k,l}$ is positive semidefinite. Set $g(y) := (y \cdot \xi)^2 f(y)$ where $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ is such that $\mathbb{1}_{B_1(0)} \leqslant f \leqslant 1$. By (PP), $L_0 g \geqslant 0$. It is not difficult to see that this implies

$$\sum_{k,l=1}^d q_{kl}\xi_k\xi_l \geqslant 0, \quad \text{for all } \xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d.$$

9° Since Lévy processes and their semigroups are invariant under translations, cf. Remark 6.7, we get $Af(x) = A_0[f(x + \cdot)]$. If we replace f by $f(x + \cdot)$, we get

$$\begin{aligned} Af(x) &= cf(x) + l \cdot \nabla f(x) + \frac{1}{2}\nabla \cdot Q\nabla f(x) \\ &\quad + \int_{y\neq 0} \big[f(x+y) - f(x) - y \cdot \nabla f(x)\mathbb{1}_{(0,1)}(|y|)\big]\, \nu(dy). \end{aligned} \tag{6.8$'$}$$

We will show in the next step that $c = 0$.

10° So far, we have seen in 5°, 7° and 9° that

$$\|Af\|_\infty \leqslant C\|f\|_{(2)} = C\sum_{|\alpha|\leqslant 2}\|\partial^\alpha f\|_\infty, \quad f \in \mathcal{C}_c^\infty(\mathbb{R}^d),$$

which means that A (has an extension which) is continuous as an operator from $\mathcal{C}_b^2(\mathbb{R}^d)$ to $\mathcal{C}_b(\mathbb{R}^d)$. Therefore, $(A, \mathcal{C}_c^\infty(\mathbb{R}^d))$ is a pseudo-differential operator with symbol

$$-\psi(\xi) = \mathrm{e}_{-\xi}(x)A_x\mathrm{e}_\xi(x), \quad \mathrm{e}_\xi(x) = \mathrm{e}^{\mathrm{i}\,\xi\cdot x}.$$

Inserting e_ξ into (6.8′) proves (6.10) and, as $\psi(0) = 0$, $c = 0$. □

Remark 6.9. In step 8° of the proof of Theorem 6.8 one can use the (PMP) to show that the coefficient c appearing in (6.8′) is positive. For this, let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{C}_c^\infty(\mathbb{R}^d)$, $f_n \uparrow 1$ and $f_n|_{B_1(0)} = 1$. By (PMP), $A_0 f_n \leqslant 0$. Moreover,

$\nabla f_n(0) = 0$ and, therefore,

$$S_0 f_n = -\int (1 - f_n(y))\,\nu(dy) \xrightarrow[n\to\infty]{} 0.$$

Consequently,

$$\limsup_{n\to\infty} L_0 f_n = \limsup_{n\to\infty} (A_0 f_n - S_0 f_n) \leqslant 0 \implies c \geqslant 0.$$

For Lévy processes we have $c = \psi(0) = 0$ and this is a consequence of the **infinite life-time** of the process:

$$\mathbb{P}(X_t \in \mathbb{R}^d) = P_t 1 = 1 \quad \text{for all} \quad t \geqslant 0,$$

and we can use the formula $P_t f - f = A \int_0^t P_s f\,ds$, cf. Lemma 5.4, for $f \equiv 1$ to show that $c = A1 = 0 \iff P_t 1 = 1$.

Definition 6.10. A **Lévy measure** is a Radon measure ν on $\mathbb{R}^d \setminus \{0\}$ such that $\int_{y\neq 0} (|y|^2 \wedge 1)\,\nu(dy) < \infty$. A **Lévy triplet** is a triplet (l, Q, ν) consisting of a vector $l \in \mathbb{R}^d$, positive semi-definite matrix $Q \in \mathbb{R}^{d\times d}$ and a Lévy measure ν.

The proof of Theorem 6.8 incidentally shows that the Lévy triplet defining the exponent (6.10) or the generator (6.8) is unique. The following corollary can easily be checked using the representation (6.8).

Corollary 6.11. *Let A be the generator and $(P_t)_{t\geqslant 0}$ the semigroup of a Lévy process. Then the Lévy triplet is given by*

$$\begin{aligned}
\int f_0\,d\nu &= A f_0(0) = \lim_{t\to 0} \frac{P_t f_0(0)}{t} \quad \forall f_0 \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \{0\}),\\
l_k &= A\phi_k(0) - \int y_k \left[\phi(y) - \mathbb{1}_{(0,1)}(|y|)\right] \nu(dy), \quad k = 1, \dots, d, \qquad (6.13)\\
q_{kl} &= A(\phi_k \phi_l)(0) - \int_{y\neq 0} \phi_k(y)\phi_l(y)\,\nu(dy), \quad k, l = 1, \dots d,
\end{aligned}$$

where $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfies $\mathbb{1}_{B_1(0)} \leqslant \phi \leqslant 1$ and $\phi_k(y) := y_k \phi(y)$. In particular, (l, Q, ν) is uniquely determined by A or the characteristic exponent ψ.

We will see an alternative uniqueness proof in the next Chapter 7.

Remark 6.12. Setting $p_t(dy) = \mathbb{P}(X_t \in dy)$, we can recast the formula for the Lévy measure as

$$\nu(dy) = \lim_{t\to 0} \frac{p_t(dy)}{t} \qquad \text{(vague limit of measures on the set } \mathbb{R}^d \setminus \{0\}\text{)}.$$

Moreover, a direct calculation using the Lévy–Khintchine formula (6.13) gives the following alternative representation for the q_{kl}:

$$\frac{1}{2} \xi \cdot Q\xi = \lim_{n\to\infty} \frac{\psi(n\xi)}{n^2}, \quad \xi \in \mathbb{R}^d.$$

Chapter 7

Construction of Lévy Processes

Our starting point is now the Lévy–Khintchine formula for the characteristic exponent ψ of a Lévy process

$$\psi(\xi) = -\,\mathrm{i}\, l \cdot \xi + \frac{1}{2}\xi \cdot Q\xi + \int_{y\neq 0} \left[1 - \mathrm{e}^{\mathrm{i}\, y\cdot\xi} + \mathrm{i}\,\xi \cdot y \mathbb{1}_{(0,1)}(|y|)\right] \nu(dy) \tag{7.1}$$

where (l, Q, ν) is a Lévy triplet in the sense of Definition 6.10; a proof of (7.1) is contained in Theorem 6.8, but the exposition below is independent of this proof, see however Remark 7.7 at the end of this chapter.

What will be needed is that a compound Poisson process is a Lévy process with càdlàg paths and characteristic exponent of the form

$$\phi(\xi) = \int_{y\neq 0} \left[1 - \mathrm{e}^{\mathrm{i}\, y\cdot\xi}\right] \rho(dy) \tag{7.2}$$

(ρ is any finite measure), see Example 3.2.d), where $\rho(dy) = \lambda \cdot \mu(dy)$.

Let ν be a Lévy measure and denote by $A_{a,b} = \{y : a \leqslant |y| < b\}$ an annulus. Since $\int_{y\neq 0} |y|^2 \wedge 1\, \nu(dy) < \infty$, the measure $\rho(B) := \nu(B \cap A_{a,b})$ is a finite measure, and there is a corresponding compound Poisson process. Adding a drift with $l = -\int y\,\rho(dy)$ shows that for every exponent

$$\psi^{a,b}(\xi) = \int_{a\leqslant|y|<b} \left[1 - \mathrm{e}^{\mathrm{i}\, y\cdot\xi} + \mathrm{i}\, y \cdot \xi\right] \nu(dy) \tag{7.3}$$

there is some Lévy process $X^{a,b} = (X_t^{a,b})_{t\geqslant 0}$. In fact,

Lemma 7.1. *Let $0 < a < b \leqslant \infty$ and $\psi^{a,b}$ given by (7.3). Then the corresponding Lévy process $X^{a,b}$ is an $L^2(\mathbb{P})$-martingale with càdlàg paths such that*

$$\mathbb{E}\left[X_t^{a,b}\right] = 0 \quad \text{and} \quad \mathbb{E}\left[X_t^{a,b} \cdot (X_t^{a,b})^\top\right] = \left(t \int_{a\leqslant|y|<b} y_k y_l\, \nu(dy)\right)_{k,l}.$$

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_7

Proof. Set $X_t := X_t^{a,b}$, $\psi := \psi^{a,b}$ and $\mathscr{F}_t := \sigma(X_r,\, r \leqslant t)$. Using the differentiation lemma for parameter-dependent integrals we see that ψ is twice continuously differentiable and

$$\frac{\partial\psi(0)}{\partial\xi_k} = 0 \quad \text{and} \quad \frac{\partial^2\psi(0)}{\partial\xi_k\partial\xi_l} = \int_{a\leqslant|y|<b} y_k y_l\,\nu(dy).$$

Since the characteristic function $\mathrm{e}^{-t\psi(\xi)}$ is twice continuously differentiable, X has first and second moments, cf. Theorem A.2, and these can be obtained by differentiation:

$$\mathbb{E}X_t^{(k)} = \frac{1}{\mathrm{i}}\frac{\partial}{\partial\xi_k}\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}\bigg|_{\xi=0} = \mathrm{i}\,t\frac{\partial\psi(0)}{\partial\xi_k} = 0$$

and

$$\begin{aligned}\mathbb{E}(X_t^{(k)}X_t^{(l)}) = -\frac{\partial^2}{\partial\xi_k\partial\xi_l}\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}\bigg|_{\xi=0} &= \frac{t\partial^2\psi(0)}{\partial\xi_k\partial\xi_l} - \frac{t\partial\psi(0)}{\partial\xi_k}\frac{t\partial\psi(0)}{\partial\xi_l}\\ &= t\int_{a\leqslant|y|<b} y_k y_l\,\nu(dy).\end{aligned}$$

The martingale property now follows from the independence of the increments: Let $s \leqslant t$, then

$$\begin{aligned}\mathbb{E}(X_t \mid \mathscr{F}_s) &= \mathbb{E}(X_t - X_s + X_s \mid \mathscr{F}_s)\\ &= \mathbb{E}(X_t - X_s \mid \mathscr{F}_s) + X_s \overset{\text{(L2)}}{\underset{\text{(L1)}}{=}} \mathbb{E}(X_{t-s}) + X_s = X_s. \qquad \square\end{aligned}$$

We will use the processes from Lemma 7.1 as main building blocks for the Lévy process. For this we need some preparations.

Lemma 7.2. *Let $(X_t^k)_{t\geqslant 0}$ be Lévy processes with characteristic exponents ψ_k. If $X^1 \perp X^2$, then $X := X^1 + X^2$ is a Lévy process with characteristic exponent $\psi = \psi_1 + \psi_2$.*

Proof. Set $\mathscr{F}_t^k := \sigma(X_s^k,\ s \leqslant t)$ and $\mathscr{F}_t = \sigma(\mathscr{F}_t^1, \mathscr{F}_t^2)$. Since $X^1 \perp X^2$, we get for $F = F_1 \cap F_2$, $F_k \in \mathscr{F}_s^k$,

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_s)}\mathbb{1}_F\right) &= \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^1-X_s^1)}\mathbb{1}_{F_1}\cdot\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^2-X_s^2)}\mathbb{1}_{F_2}\right)\\ &= \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^1-X_s^1)}\mathbb{1}_{F_1}\right)\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^2-X_s^2)}\mathbb{1}_{F_2}\right)\\ &\overset{\text{(L2)}}{=} \mathrm{e}^{-(t-s)\psi_1(\xi)}\mathbb{P}(F_1)\cdot\mathrm{e}^{-(t-s)\psi_2(\xi)}\mathbb{P}(F_2)\\ &= \mathrm{e}^{-(t-s)(\psi_1(\xi)+\psi_2(\xi))}\mathbb{P}(F).\end{aligned}$$

As $\{F_1 \cap F_2 : F_k \in \mathscr{F}_s^k\}$ is a $\cap$-stable generator of $\mathscr{F}_s$, we find

$$\mathbb{E}\left(e^{i\,\xi\cdot(X_t - X_s)} \,\middle|\, \mathscr{F}_s\right) = e^{-(t-s)\psi(\xi)}.$$

Observe that $\mathscr{F}_s$ could be larger than the canonical filtration $\mathscr{F}_s^X$. Therefore, we first condition w.r.t. $\mathbb{E}(\cdots \mid \mathscr{F}_s^X)$ and then use Theorem 3.1, to see that X is a Lévy process with exponent $\psi = \psi_1 + \psi_2$. ☐

Lemma 7.3. *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of Lévy processes with characteristic exponents ψ_n. Assume that $X_t^n \to X_t$ converges in probability for every $t \geqslant 0$. If*

either: the convergence is uniform in probability, i.e.,

$$\forall\epsilon > 0 \ \ \forall t \geqslant 0 : \lim_{n\to\infty} \mathbb{P}\left(\sup_{s\leqslant t} |X_s^n - X_s| > \epsilon\right) = 0,$$

or: the limiting process X has càdlàg paths,

then X is a Lévy process with characteristic exponent $\psi := \lim_{n\to\infty} \psi_n$.

Proof. Let $0 = t_0 < t_1 < \cdots < t_m$ and $\xi_1, \ldots, \xi_m \in \mathbb{R}^d$. Since the X^n are Lévy processes,

$$\mathbb{E}\exp\left[i\sum_{k=1}^m \xi_k \cdot (X_{t_k}^n - X_{t_{k-1}}^n)\right] \overset{\text{(L2),(L1)}}{=} \prod_{k=1}^m \mathbb{E}\exp\left[i\,\xi_k \cdot X_{t_k - t_{k-1}}^n\right]$$

and, because of convergence in probability, this equality is inherited by the limiting process X. This proves that X has independent (L2′) and stationary (L1) increments.

The condition (L3) follows either from the uniformity of the convergence in probability or the càdlàg property. Thus, X is a Lévy process. From

$$\lim_{n\to\infty} \mathbb{E}\, e^{i\,\xi\cdot X_1^n} = \mathbb{E}\, e^{i\,\xi\cdot X_1}$$

we get that the limit $\lim_{n\to\infty}\psi_n = \psi$ exists. ☐

Lemma 7.4 (Notation of Lemma 7.1)**.** *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence $a_1 > a_2 > \cdots$ decreasing to zero and assume that the processes $(X^{a_{n+1},a_n})_{n\in\mathbb{N}}$ are independent Lévy processes with characteristic exponents ψ_{a_{n+1},a_n}. Then $X := \sum_{n=1}^\infty X^{a_{n+1},a_n}$ is a Lévy process with characteristic exponent $\psi := \sum_{n=1}^\infty \psi_{a_{n+1},a_n}$ and càdlàg paths. Moreover, X is an $L^2(\mathbb{P})$-martingale.*

Proof. Lemmas 7.1, 7.2 show that $X^{a_{n+m},a_n} = \sum_{k=1}^m X^{a_{n+k},a_{n+k-1}}$ is a Lévy process with characteristic exponent $\psi_{a_{n+m},a_n} = \sum_{k=1}^m \psi_{a_{n+k},a_{n+k-1}}$, and X^{a_{n+m},a_n}

is an $L^2(\mathbb{P})$-martingale. By Doob's inequality and Lemma 7.1

$$\begin{aligned}\mathbb{E}\left(\sup_{s\leqslant t}|X_s^{a_{n+m},a_n}|^2\right) &\leqslant 4\,\mathbb{E}\big(|X_t^{a_{n+m},a_n}|^2\big)\\ &= 4t\int_{a_{n+m}\leqslant|y|<a_n} y^2\,\nu(dy) \xrightarrow[m,n\to\infty]{\text{dom. convergence}} 0.\end{aligned}$$

Hence, the limit $X = \lim_{n\to\infty} X^{a_n,a_1}$ exists (uniformly in t) in L^2, i.e., X is an $L^2(\mathbb{P})$-martingale; since the convergence is also uniform in probability, Lemma 7.3 shows that X is a Lévy process with exponent $\psi = \sum_{n=1}^\infty \psi_{a_{n+1},a_n}$. Taking a uniformly convergent subsequence, we also see that the limit inherits the càdlàg property from the approximating Lévy processes X^{a_n,a_1}. □

We can now prove the main result of this chapter.

Theorem 7.5. *Let (l, Q, ν) be a Lévy triplet and ψ be given by (7.1). Then there exists a Lévy process X with càdlàg paths and characteristic exponent ψ.*

Proof. Because of Lemma 7.1, 7.2 and 7.4 we can construct X piece by piece.

1° Let $(W_t)_{t\geqslant 0}$ be a Brownian motion and set

$$X_t^c := tl + \sqrt{Q}W_t \quad\text{and}\quad \psi_c(\xi) := -\,\mathrm{i}\,l\cdot\xi + \frac{1}{2}\xi\cdot Q\xi.$$

2° Write $\mathbb{R}^d\setminus\{0\} = \dot{\bigcup}_{n=0}^\infty A_n$ with $A_0 := \{|y|\geqslant 1\}$ and $A_n := \left\{\frac{1}{n+1}\leqslant |y| < \frac{1}{n}\right\}$ and set $\mu_n := \nu(\cdot\cap A_n)/\nu(A_n)$, $\lambda_n := \nu(A_n)$.

3° Construct, as in Example 3.2.d), a compound Poisson process comprising the large jumps

$$X_t^0 := X_t^{1,\infty} \quad\text{and}\quad \psi_0(\xi) := \int_{1\leqslant|y|<\infty}\left[1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi}\right]\nu(dy)$$

and compensated compound Poisson processes taking account of all small jumps

$$X_t^n := X_t^{a_{n+1},a_n},\quad a_n := \frac{1}{n}\quad\text{and}\quad \psi_n(\xi) := \int_{A_n}\left[1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi}+\mathrm{i}\,y\cdot\xi\right]\nu(dy).$$

We can construct the processes X^n to be stochastically independent (just choose independent jump time processes and independent iid jump heights when constructing the compound Poisson processes) and independent of the Wiener process W.

4° Setting $\psi = \psi_0+\psi_c+\sum_{n=1}^\infty \psi_n$, Lemma 7.2 and 7.4 prove the theorem. Since all approximating processes have càdlàg paths, this property is inherited by the sums and the limit (Lemma 7.4). □

The proof of Theorem 7.5 also implies the following pathwise decomposition of a Lévy process. We write $\Delta X_t := X_t - X_{t-}$ for the jump at time t. From the construction we know that

(large jumps) $$J_t^{[1,\infty)} = \sum_{s\leqslant t} \Delta X_s \mathbb{1}_{[1,\infty)}(|\Delta X_s|) \tag{7.4}$$

(small jumps) $$J_t^{[1/n,1)} = \sum_{s\leqslant t} \Delta X_s \mathbb{1}_{[1/n,1)}(|\Delta X_s|) \tag{7.5}$$

(compensated small jumps) $$\begin{aligned} \widetilde{J}_t^{[1/n,1)} &= J_t^{[1/n,1)} - \mathbb{E} J_t^{[1/n,1)} \\ &= \sum_{s\leqslant t} \Delta X_s \mathbb{1}_{[1/n,1)}(|\Delta X_s|) - t \int\limits_{\frac{1}{n}\leqslant|y|<1} y\,\nu(dy). \end{aligned} \tag{7.6}$$

are Lévy processes and $J^{[1,\infty)} \perp \widetilde{J}^{[1/n,1)}$.

Corollary 7.6. *Let ψ be a characteristic exponent given by* (7.1) *and let X be the Lévy process constructed in Theorem* 7.5*. Then*

$$\begin{aligned} X_t = \underbrace{\sqrt{Q}W_t}_{} + \lim_{n\to\infty} \Bigg(\sum_{s\leqslant t} \Delta X_s \mathbb{1}_{[1/n,1)}(|\Delta X_s|) - t \int\limits_{[1/n,1)} y\,\nu(dy) \Bigg) &\Bigg] =: M_t, \ L^2\text{-martingale} \\ \underbrace{tl}_{\text{continuous Gaussian}} \quad \underbrace{+ \sum_{s\leqslant t} \Delta X_s \mathbb{1}_{[1,\infty)}(|\Delta X_s|).}_{\text{pure jump part}} &\Bigg] =: A_t, \ \text{bdd. variation,} \end{aligned}$$

where all appearing processes are independent.

Proof. The decomposition follows directly from the construction in Theorem 7.5. By Lemma 7.4, $\lim_{n\to\infty} X^{1/n,1}$ is an $L^2(\mathbb{P})$-martingale, and since the (independent!) Wiener process W is also an $L^2(\mathbb{P})$-martingale, so is their sum M.

The paths $t \mapsto A_t(\omega)$ are a.s. of bounded variation since, by construction, on any time-interval $[0,t]$ there are $N_t^0(\omega)$ jumps of size $\geqslant 1$. Since $N_t^0(\omega) < \infty$ a.s., the total variation of $A_t(\omega)$ is less or equal than $|l|t + \sum_{s\leqslant t} |\Delta X_s| \mathbb{1}_{[1,\infty)}(|\Delta X_s|)$ which is a.s. finite. □

Remark 7.7. A word of **caution**: Theorem 7.5 associates with any ψ given by the Lévy–Khintchine formula (7.1) a Lévy process. Unless we know that all characteristic exponents are of this form (this was proved in Theorem 6.8), it does not follow that we have constructed **all** Lévy processes.

On the other hand, Theorem 7.5 shows that the Lévy triplet determining ψ is **unique**. Indeed, assume that (l, Q, ν) and (l', Q', ν') are two Lévy triplets which yield the same exponent ψ. Now we can associate, using Theorem 7.5, with each triplet a Lévy process X and X' such that

$$\mathbb{E}\,e^{i\,\xi\cdot X_t} = e^{-t\psi(\xi)} = \mathbb{E}\,e^{i\,\xi\cdot X'_t}.$$

Thus, $X \sim X'$ and so these processes have (in law) the same pathwise decomposition, i.e., the same drift, diffusion and jump behaviour. This, however, means that $(l, Q, \nu) = (l', Q', \nu')$.

Chapter 8

Two Special Lévy Processes

We will now study the structure of the paths of a Lévy process. We begin with two extreme cases: Lévy processes which only grow by jumps of size 1 and Lévy processes with continuous paths.

Throughout this chapter we assume that all paths $[0,\infty) \ni t \mapsto X_t(\omega)$ are right-continuous with finite left-hand limits (càdlàg). This is a bit stronger than (L3), but it is always possible to construct a càdlàg version of a Lévy process (see the discussion on page 14). This allows us to consider the **jumps** of the process X

$$\Delta X_t := X_t - X_{t-} = X_t - \lim_{s \uparrow t} X_s.$$

Theorem 8.1. *Let X be a one-dimensional Lévy process which moves only by jumps of size* 1*. Then X is a Poisson process.*

Proof. Set $\mathscr{F}_t^X := \sigma(X_s,\, s \leqslant t)$ and let $T_1 = \inf\{t > 0 \,:\, \Delta X_t = 1\}$ be the time of the first jump. Since $\{T_1 > t\} = \{X_t = 0\} \in \mathscr{F}_t^X$, T_1 is a stopping time.

Let $T_0 = 0$ and $T_k = \inf\{t > T_{k-1} \,:\, \Delta X_t = 1\}$, be the time of the kth jump; this is also a stopping time. By the Markov property ((4.4) and Lemma 4.4),

$$\begin{aligned}
\mathbb{P}(T_1 > s+t) &= \mathbb{P}(T_1 > s,\ T_1 > s+t) \\
&= \mathbb{E}\big[\mathbb{1}_{\{T_1 > s\}} \mathbb{P}^{X_s}(T_1 > t)\big] \\
&= \mathbb{E}\big[\mathbb{1}_{\{T_1 > s\}} \mathbb{P}^{0}(T_1 > t)\big] \\
&= \mathbb{P}(T_1 > s)\mathbb{P}(T_1 > t)
\end{aligned}$$

where we use that $X_s = 0$ if $T_1 > s$ (the process hasn't yet moved!) and $\mathbb{P} = \mathbb{P}^0$.

Since $t \mapsto \mathbb{P}(T_1 > t)$ is right-continuous, this functional equation has the unique solution $\mathbb{P}(T_1 > t) = \exp[t \log \mathbb{P}(T_1 > 1)]$ (Theorem A.1). Thus, the sequence of inter-jump times $\sigma_k := T_k - T_{k-1}$, $k \in \mathbb{N}$, is an iid sequence of exponential times. This follows immediately from the strong Markov property (Theorem 4.12) for Lévy processes and the observation that

$$T_{k+1} - T_k = T_1^Y \quad \text{where} \quad Y = (Y_{t+T_k} - Y_{T_k})_{t \geqslant 0}$$

and T_1^Y is the first jump time of the process Y.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_8

Obviously, $X_t = \sum_{k=1}^\infty \mathbb{1}_{[0,t]}(T_k)$, $T_k = \sigma_1 + \dots + \sigma_k$, and Example 3.2.c) (and Theorem 3.4) show that X is a Poisson process. □

A Lévy process with uniformly bounded jumps admits moments of all orders.

Lemma 8.2. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process such that $|\Delta X_t(\omega)| \leqslant c$ for all $t \geqslant 0$ and some constant $c > 0$. Then $\mathbb{E}(|X_t|^p) < \infty$ for all $p \geqslant 0$.*

Proof. Let $\mathscr{F}_t^X := \sigma(X_s,\, s \leqslant t)$ and define the stopping times

$$\tau_0 := 0, \quad \tau_n := \inf\left\{t > \tau_{n-1} \,:\, |X_t - X_{\tau_{n-1}}| \geqslant c\right\}.$$

Since X has càdlàg paths, $\tau_0 < \tau_1 < \tau_2 < \cdots$. Let us show that $\tau_1 < \infty$ a.s. For fixed $t > 0$ and $n \in \mathbb{N}$ we have

$$\begin{aligned}\mathbb{P}(\tau_1 = \infty) \leqslant \mathbb{P}(\tau_1 \geqslant nt) &\leqslant \mathbb{P}(|X_{kt} - X_{(k-1)t}| \leqslant 2c, \;\; \forall k = 1, \dots, n)\\ &\underset{(\mathrm{L}2')}{\overset{(\mathrm{L}2)}{=}} \prod_{k=1}^n \mathbb{P}(|X_{kt} - X_{(k-1)t}| \leqslant 2c) \overset{(\mathrm{L}1)}{=} \mathbb{P}(|X_t| \leqslant 2c)^n.\end{aligned}$$

Letting $n \to \infty$ we see that $\mathbb{P}(\tau_1 = \infty) = 0$ if $\mathbb{P}(|X_t| \leqslant 2c) < 1$ for some $t > 0$. (In the alternative case, we have $\mathbb{P}(|X_t| \leqslant 2c) = 1$ for all $t > 0$ which makes the lemma trivial.)

By the strong Markov property (Theorem 4.12)

$$\tau_n - \tau_{n-1} \sim \tau_1 \quad \text{and} \quad \tau_n - \tau_{n-1} \perp \mathscr{F}^X_{\tau_{n-1}},$$

i.e., $(\tau_n - \tau_{n-1})_{n\in\mathbb{N}}$ is an iid sequence. Therefore,

$$\mathbb{E}\, \mathrm{e}^{-\tau_n} = \left(\mathbb{E}\, \mathrm{e}^{-\tau_1}\right)^n = q^n$$

for some $q \in [0, 1)$. From the very definition of the stoppping times we infer

$$|X_{t\wedge\tau_n}| \leqslant \sum_{k=1}^n |X_{\tau_k} - X_{\tau_{k-1}}| \leqslant \sum_{k=1}^n \Big(\underbrace{|\Delta X_{\tau_k}|}_{\leqslant c} + \underbrace{|X_{\tau_k -} - X_{\tau_{k-1}}|}_{\leqslant c}\Big) \leqslant 2nc.$$

Thus, $|X_t| > 2nc$ implies that $\tau_n < t$, and by Markov's inequality

$$\mathbb{P}(|X_t| > 2nc) \leqslant \mathbb{P}(\tau_n < t) \leqslant \mathrm{e}^t\, \mathbb{E}\, \mathrm{e}^{-\tau_n} = \mathrm{e}^t\, q^n.$$

Finally,

$$\begin{aligned}\mathbb{E}\big(|X_t|^p\big) &= \sum_{n=0}^\infty \mathbb{E}\big(|X_t|^p \mathbb{1}_{\{2nc < |X_t| \leqslant 2(n+1)c\}}\big)\\ &\leqslant (2c)^p \sum_{n=0}^\infty (n+1)^p\, \mathbb{P}(|X_t| > 2nc) \leqslant (2c)^p \mathrm{e}^t \sum_{n=0}^\infty (n+1)^p q^n < \infty.\end{aligned}$$ □

Recall that a Brownian motion $(W_t)_{t\geqslant 0}$ on $\mathbb{R}^d$ is a Lévy process such that W_t is a normal random variable with mean 0 and covariance matrix $t\,\mathrm{id}$. We will need Paul Lévy's characterization of Brownian motion which we state without proof. An elementary proof can be found in [56, Chapter 9.4].

Theorem 8.3 (Lévy). *Let $M = (M_t, \mathscr{F}_t)$, $M_0 = 0$, be a one-dimensional martingale with continuous sample paths such that $(M_t^2 - t, \mathscr{F}_t)_{t\geqslant 0}$ is also a martingale. Then M is a one-dimensional standard Brownian motion.*

Theorem 8.4. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process in $\mathbb{R}^d$ whose sample paths are a.s. continuous. Then $X_t \sim tl + \sqrt{Q}W_t$ where $l \in \mathbb{R}^d$, Q is a positive semidefinite symmetric matrix, and W is a standard Brownian motion in $\mathbb{R}^d$.*

We will give two proofs of this result.

Proof (*using Theorem* 8.3). By Lemma 8.2, $(X_t)_{t\geqslant 0}$ has moments of all orders. Therefore, $M_t := M_t^\xi := \xi \cdot (X_t - \mathbb{E}X_t)$ exists for any $\xi \in \mathbb{R}^d$ and is a martingale for the canonical filtration $\mathscr{F}_t := \sigma(X_s,\, s \leqslant t)$. Indeed, for all $s \leqslant t$

$$\mathbb{E}(M_t \mid \mathscr{F}_s) = \mathbb{E}(M_t - M_s \mid \mathscr{F}_s) - M_s \underset{\text{(L1)}}{\overset{\text{(L2)}}{=}} \mathbb{E}M_{t-s} + M_s = M_s.$$

Moreover

$$\begin{aligned}\mathbb{E}(M_t^2 - M_s^2 \mid \mathscr{F}_s) \quad &= \quad \mathbb{E}((M_t - M_s)^2 + 2M_s(M_t - M_s) \mid \mathscr{F}_s)\\ &\overset{\text{(L2)}}{=} \quad \mathbb{E}((M_t - M_s)^2) + 2M_s\mathbb{E}(M_t - M_s)\\ &\overset{\text{Lemma 3.10}}{=} (t-s)\mathbb{E}M_1^2 = (t-s)\mathbb{V}M_1,\end{aligned}$$

and so $(M_t^2 - t\mathbb{V}M_1)_{t\geqslant 0}$ and $(M_t)_{t\geqslant 0}$ are martingales with continuous paths.

Now we can use Theorem 8.3 and deduce that $\xi \cdot (X_t - \mathbb{E}X_t)$ is a one-dimensional Brownian motion with variance $\xi \cdot Q\xi$ where tQ is the covariance matrix of the random variable X_t (cf. the proof of Lemma 7.1 or Lemma 3.10). Thus, $X_t - \mathbb{E}X_t = \sqrt{Q}W_t$ where W_t is a d-dimensional standard Brownian motion. Finally, $\mathbb{E}X_t = t\mathbb{E}X_1 =: tl$. $\square$

Standard proof (*using the CLT*). Fix $\xi \in \mathbb{R}^d$ and set $M(t) := \xi \cdot (X_t - \mathbb{E}X_t)$. Since X has moments of all orders, M is well-defined and it is again a Lévy process. Moreover,

$$\mathbb{E}M(t) = 0 \quad \text{and} \quad t\sigma^2 = \mathbb{V}M(t) = \mathbb{E}[(\xi \cdot (X_t - \mathbb{E}X_t))^2] = t\xi \cdot Q\xi$$

where Q is the covariance matrix of X, cf. the proof of Lemma 7.1. We proceed as in the proof of the CLT: Using a Taylor expansion we get

$$\begin{aligned}\mathbb{E}\,e^{i\,M(t)} &= \mathbb{E}\,e^{i\sum_{k=1}^n [M(tk/n) - M(t(k-1)/n)]}\\ &\underset{\text{(L1)}}{\overset{\text{(L2')}}{=}} \left(\mathbb{E}\,e^{i\,M(t/n)}\right)^n = \left(1 - \frac{1}{2}\,\mathbb{E}M^2(t/n) + R_n\right)^n.\end{aligned}$$

The remainder term R_n is estimated by $\frac{1}{6}\mathbb{E}|M^3(\frac{t}{n})|$. If we can show that $|R_n| \leqslant \epsilon\frac{t}{n}$ for large $n = n(\epsilon)$ and any $\epsilon > 0$, we get because of $\mathbb{E}M^2(\frac{t}{n}) = \frac{t}{n}\sigma^2$

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,M(t)} = \lim_{n\to\infty}\left(1 - \frac{1}{2}(\sigma^2 + 2\epsilon)\frac{t}{n}\right)^n = \mathrm{e}^{-\frac{1}{2}(\sigma^2+2\epsilon)t} \xrightarrow[\epsilon\to 0]{} \mathrm{e}^{-\frac{1}{2}\sigma^2 t}.$$

This shows that $\xi \cdot (X_t - \mathbb{E}X_t)$ is a centered Gaussian random variable with variance σ^2. Since $\mathbb{E}X_t = t\mathbb{E}X_1$ we conclude that X_t is Gaussian with mean tl and covariance tQ.

We will now estimate $\mathbb{E}|M^3(\frac{t}{n})|$. For every $\epsilon > 0$ we can use the uniform continuity of $s \mapsto M(s)$ on $[0, t]$ to get

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}|M(\tfrac{k}{n}t) - M(\tfrac{k-1}{n}t)| = 0.$$

Thus, we have for all $\epsilon > 0$

$$\begin{aligned}
1 &= \lim_{n\to\infty}\mathbb{P}\Big(\max_{1\leqslant k\leqslant n}|M(\tfrac{k}{n}t) - M(\tfrac{k-1}{n}t)| \leqslant \epsilon\Big) \\
&= \lim_{n\to\infty}\mathbb{P}\Big(\bigcap_{k=1}^{n}\big\{|M(\tfrac{k}{n}t) - M(\tfrac{k-1}{n}t)| \leqslant \epsilon\big\}\Big) \\
&\overset{\text{(L2)}}{\underset{\text{(L1)}}{=}} \lim_{n\to\infty}\prod_{k=1}^{n}\mathbb{P}(|M(\tfrac{t}{n})| \leqslant \epsilon) \\
&= \lim_{n\to\infty}\big[1 - \mathbb{P}(|M(\tfrac{t}{n})| > \epsilon)\big]^n \\
&\leqslant \lim_{n\to\infty}\mathrm{e}^{-n\,\mathbb{P}(|M(t/n)|>\epsilon)} \leqslant 1
\end{aligned}$$

where we use the inequality $1+x \leqslant \mathrm{e}^x$. This proves $\lim_{n\to\infty} n\,\mathbb{P}(|M(t/n)| > \epsilon) = 0$. Therefore,

$$\begin{aligned}
\mathbb{E}|M^3(\tfrac{t}{n})| &\leqslant \epsilon\mathbb{E}M^2(\tfrac{t}{n}) + \int_{|M(t/n)|>\epsilon}|M^3(\tfrac{t}{n})|\,d\mathbb{P} \\
&\leqslant \epsilon\frac{t}{n}\,\sigma^2 + \sqrt{\mathbb{P}(|M(\tfrac{t}{n})| > \epsilon)}\sqrt{\mathbb{E}M^6(\tfrac{t}{n})}.
\end{aligned}$$

It is not hard to see that $\mathbb{E}M^6(s) = a_1 s + \cdots + a_6 s^6$ (differentiate the characteristic function $\mathbb{E}\mathrm{e}^{\mathrm{i}\,uM(s)} = \mathrm{e}^{-s\psi(u)}$ six times at $u = 0$), and so

$$\mathbb{E}|M^3(\tfrac{t}{n})| \leqslant \epsilon\frac{t}{n}\,\sigma^2 + c\frac{t}{n}\sqrt{\frac{\mathbb{P}(|M(t/n)| > \epsilon)}{t/n}} = \frac{t}{n}\big(\epsilon\sigma^2 + o(1)\big) \qquad \square$$

We close this chapter with Paul Lévy's construction of a standard Brownian motion $(W_t)_{t\geqslant 0}$. Since W is a Lévy process which has the Markov property, it is enough to construct a Brownian motion $W(t)$ only for $t \in [0,1]$, then produce independent copies $(W^n(t))_{t\in[0,1]}$, $n = 0,1,2,\dots$, and join them continuously:

$$W_t := \begin{cases} W^0(t), & t \in [0,1), \\ W^0(1) + \dots + W^{n-1}(1) + W^n(t-n), & t \in [n, n+1). \end{cases}$$

Since each W_t is normally distributed with mean 0 and variance t, we will get a Lévy process with characteristic exponent $\frac{1}{2}\xi^2$, $\xi \in \mathbb{R}$. In the same vein we get a d-dimensional Brownian motion by making a vector $(W_t^{(1)}, \dots, W_t^{(d)})_{t\geqslant 0}$ of d independent copies of $(W_t)_{t\geqslant 0}$. This yields a Lévy process with exponent $\frac{1}{2}(\xi_1^2 + \dots + \xi_d^2)$, $\xi_1, \dots, \xi_d \in \mathbb{R}$.

Denote a one-dimensional normal distribution with mean m and variance σ^2 as $\mathsf{N}(m, \sigma^2)$. The motivation for the construction is the observation that a Brownian motion satisfies the following mid-point law (cf. [56, Chapter 3.4]):

$$\mathbb{P}(W_{(s+t)/2} \in \bullet \mid W_s = x, W_t = y) = \mathsf{N}\big(\tfrac{1}{2}(x+y),\ \tfrac{1}{4}(t-s)\big), \quad s \leqslant t,\ x, y \in \mathbb{R}.$$

This can be turned into the following construction method:

Algorithm. Set $W(0) = 0$ and let $W(1) \sim \mathsf{N}(0,1)$. Let $n \geqslant 1$ and assume that the random variables $W(k2^{-n})$, $k = 1, \dots, 2^n - 1$ have already been constructed. Then

$$W(l2^{-n-1}) := \begin{cases} W(k2^{-n}), & l = 2k, \\ \frac{1}{2}\big(W(k2^{-n}) + W((k+1)2^{-n})\big) + \Gamma_{2^n+k}, & l = 2k+1, \end{cases}$$

where Γ_{2^n+k} is an independent (of everything else) $\mathsf{N}(0, 2^{-n}/4)$ Gaussian random variable, cf. Figure 8.1. In-between the nodes we use piecewise linear interpolation:

$$W_{2^n}(t,\omega) := \text{Linear interpolation of } \big(W(k2^{-n}, \omega),\ k = 0,1,\dots,2^n\big), \quad n \geqslant 1.$$

At the dyadic points $t = k2^{-j}$ we get the 'true' value of $W(t,\omega)$, while the linear interpolation is an approximation, see Figure 8.1.

Theorem 8.5 (Lévy 1940). *The series*

$$W(t,\omega) := \sum_{n=0}^{\infty} \big(W_{2^{n+1}}(t,\omega) - W_{2^n}(t,\omega)\big) + W_1(t,\omega), \quad t \in [0,1],$$

converges a.s. uniformly. In particular $(W(t))_{t\in[0,1]}$ is a one-dimensional Brownian motion.

Proof. Set $\Delta_n(t,\omega) := W_{2^{n+1}}(t,\omega) - W_{2^n}(t,\omega)$. By construction,

$$\Delta_n\big((2k-1)2^{-n-1}, \omega\big) = \Gamma_{2^n+(k-1)}(\omega), \quad k = 1,2,\dots,2^n,$$

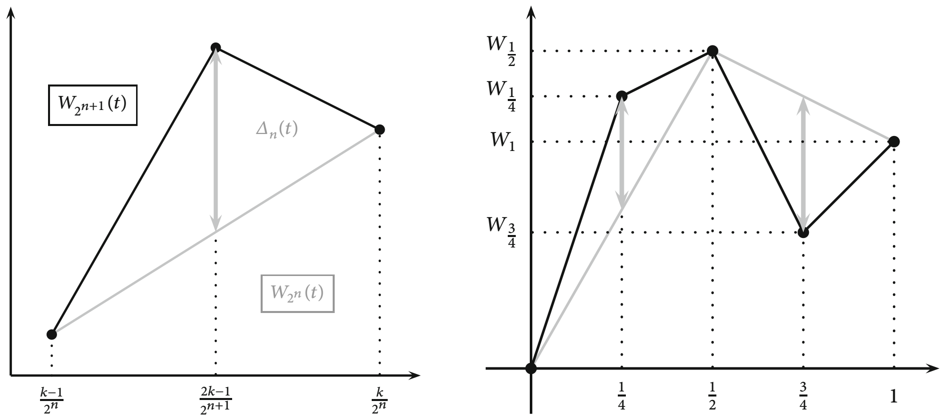


Figure 8.1: Interpolation of order four in Lévy's construction of Brownian motion.

are iid $\mathsf{N}(0, 2^{-(n+2)})$ distributed random variables. Therefore,

$$\mathbb{P}\left(\max_{1\leqslant k\leqslant 2^n}\left|\Delta_n\big((2k-1)2^{-n-1}\big)\right| > \frac{x_n}{\sqrt{2^{n+2}}}\right) \leqslant 2^n\,\mathbb{P}\left(\left|\sqrt{2^{n+2}}\,\Delta_n\big(2^{-n-1}\big)\right| > x_n\right),$$

and the right-hand side equals

$$\frac{2\cdot 2^n}{\sqrt{2\pi}}\int_{x_n}^{\infty} \mathrm{e}^{-r^2/2}\,dr \leqslant \frac{2^{n+1}}{\sqrt{2\pi}}\int_{x_n}^{\infty}\frac{r}{x_n}\,\mathrm{e}^{-r^2/2}\,dr = \frac{2^{n+1}}{x_n\sqrt{2\pi}}\,\mathrm{e}^{-x_n^2/2}.$$

Choose $c > 1$ and $x_n := c\sqrt{2n\log 2}$. Then

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{P}\left(\max_{1\leqslant k\leqslant 2^n}\left|\Delta_n\big((2k-1)2^{-n-1}\big)\right| > \frac{x_n}{\sqrt{2^{n+2}}}\right) &\leqslant \sum_{n=1}^{\infty}\frac{2^{n+1}}{c\sqrt{2\pi}}\,\mathrm{e}^{-c^2\log 2^n}\\ &= \frac{2}{c\sqrt{2\pi}}\sum_{n=1}^{\infty}2^{-(c^2-1)n} < \infty.\end{aligned}$$

Using the Borel–Cantelli lemma we find a set $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that for every $\omega \in \Omega_0$ there is some $N(\omega) \geqslant 1$ with

$$\max_{1\leqslant k\leqslant 2^n}\left|\Delta_n\big((2k-1)2^{-n-1}\big)\right| \leqslant c\sqrt{\frac{n\log 2}{2^{n+1}}} \quad \text{for all} \quad n \geqslant N(\omega).$$

$\Delta_n(t)$ is the distance between the polygonal arcs $W_{2^{n+1}}(t)$ and $W_{2^n}(t)$; the maximum is attained at one of the midpoints of the intervals $[(k-1)2^{-n}, k2^{-n}]$, $k = 1, \dots, 2^n$, see Figure 8.1. Thus,

$$\sup_{0\leqslant t\leqslant 1}\left|W_{2^{n+1}}(t,\omega) - W_{2^n}(t,\omega)\right| \leqslant \max_{1\leqslant k\leqslant 2^n}\left|\Delta_n\big((2k-1)2^{-n-1},\omega\big)\right| \leqslant c\sqrt{\frac{n\log 2}{2^{n+1}}},$$

for all $n \geqslant N(\omega)$ which means that the limit

$$W(t,\omega) := \lim_{N\to\infty} W_{2^N}(t,\omega) = \sum_{n=0}^{\infty} \big(W_{2^{n+1}}(t,\omega) - W_{2^n}(t,\omega)\big) + W_1(t,\omega)$$

exists for all $\omega \in \Omega_0$ uniformly in $t \in [0,1]$. Therefore, $t \mapsto W(t,\omega)$, $\omega \in \Omega_0$, inherits the continuity of the polygonal arcs $t \mapsto W_{2^n}(t,\omega)$. Set

$$\widetilde{W}(t,\omega) := W(t,\omega)\mathbb{1}_{\Omega_0}(\omega).$$

By construction, we find for all $0 \leqslant k \leqslant l \leqslant 2^n$

$$\begin{aligned}
\widetilde{W}(l2^{-n}) - \widetilde{W}(k2^{-n}) &= W_{2^n}(l2^{-n}) - W_{2^n}(k2^{-n}) \\
&= \sum_{l=k+1}^{l} \big(W_{2^n}(l2^{-n}) - W_{2^n}((l-1)2^{-n})\big) \\
&\overset{\text{iid}}{\sim} \mathsf{N}(0,(l-k)2^{-n}).
\end{aligned}$$

Since $t \mapsto \widetilde{W}(t)$ is continuous and the dyadic numbers are dense in $[0,t]$, we conclude that the increments $\widetilde{W}(t_k) - \widetilde{W}(t_{k-1})$, $0 = t_0 < t_1 < \cdots < t_N \leqslant 1$ are independent $\mathsf{N}(0, t_k - t_{k-1})$ distributed random variables. This shows that $(\widetilde{W}(t))_{t\in[0,1]}$ is a Brownian motion. □

Chapter 9

Random Measures

We continue our investigations of the paths of càdlàg Lévy processes. Independently of Chapters 5 and 6 we will show in Theorem 9.12 that the processes constructed in Theorem 7.5 are indeed **all** Lévy processes; this gives also a new proof of the Lévy–Khintchine formula, cf. Corollary 9.13. As before, we denote the jumps of $(X_t)_{t\geqslant 0}$ by

$$\Delta X_t := X_t - X_{t-} = X_t - \lim_{s\uparrow t} X_s.$$

Definition 9.1. Let X be a Lévy process. The counting measure

$$N_t(B,\omega) := \#\{s \in (0,t] : \Delta X_s(\omega) \in B\}, \quad B \in \mathscr{B}(\mathbb{R}^d \setminus \{0\}) \tag{9.1}$$

is called the **jump measure** of the process X.

Since a càdlàg function $x : [0,\infty) \to \mathbb{R}^d$ has on any compact interval $[a,b]$ at most finitely many jumps $|\Delta x_t| > \epsilon$ exceeding a fixed size,[1] we see that

$$N_t(B,\omega) < \infty \quad \forall t > 0,\ B \in \mathscr{B}(\mathbb{R}^d) \text{ such that } 0 \notin \overline{B}.$$

Notice that $0 \notin \overline{B}$ is equivalent to $B_\epsilon(0) \cap B = \emptyset$ for some $\epsilon > 0$. Thus, $B \mapsto N_t(B,\omega)$ is for every ω a locally finite Borel measure on $\mathbb{R}^d \setminus \{0\}$.

Definition 9.2. Let $N_t(B)$ be the jump measure of the Lévy process X. For every Borel function $f : \mathbb{R}^d \to \mathbb{R}$ with $0 \notin \operatorname{supp} f$ we define

$$N_t(f,\omega) := \int f(y)\, N_t(dy,\omega). \tag{9.2}$$

Since $0 \notin \operatorname{supp} f$, it is clear that $N_t(\operatorname{supp} f, \omega) < \infty$, and for every ω

$$N_t(f,\omega) = \sum_{0<s\leqslant t} f(\Delta X_s(\omega)) = \sum_{n=1}^{\infty} f(\Delta X_{\tau_n}(\omega))\mathbb{1}_{(0,t]}(\tau_n(\omega)). \tag{9.2$'$}$$

[1] Otherwise we would get an accumulation point of jumps within $[a,b]$, and x would be unbounded on $[a,b]$.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_9

Both sums are finite sums, extending only over those s where $\Delta X_s(\omega) \neq 0$. This is obvious in the second sum where $\tau_1(\omega), \tau_2(\omega), \tau_3(\omega) \dots$ are the jump times of X.

Lemma 9.3. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X, take $s < t$, $t_{k,n} = s + \frac{k}{n}(t-s)$ and $f \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\}, \mathbb{R}^m)$, i.e., f takes values in $\mathbb{R}^m$. Then*

$$N_t(f,\omega) - N_s(f,\omega) = \sum_{s<u\leqslant t} f\big(\Delta X_u(\omega)\big) = \lim_{n\to\infty} \sum_{k=0}^{n-1} f\big(X_{t_{k+1,n}}(\omega) - X_{t_{k,n}}(\omega)\big). \tag{9.3}$$

Proof. Throughout the proof ω is fixed and we will omit it in our notation. Since $0 \notin \operatorname{supp} f$, there is some $\epsilon > 0$ such that $B_\epsilon(0) \cap \operatorname{supp} f = \emptyset$; therefore, we need only consider jumps of size $|\Delta X_t| \geqslant \epsilon$. Denote by $J = \{\tau_1, \dots, \tau_N\}$ those jumps. For sufficiently large n we can achieve that

- $\#\big(J \cap (t_{k,n}, t_{k+1,n}]\big) \leqslant 1$ for all $k = 0, \dots, n-1$;
- $|X_{t_{\kappa+1,n}} - X_{t_{\kappa,n}}| < \epsilon$ if κ is such that $J \cap (t_{\kappa,n}, t_{\kappa+1,n}] = \emptyset$.

  Indeed: Assume this is not the case, then we could find sequences $s < s_k < t_k \leqslant t$ such that $t_k - s_k \to 0$, $J \cap (s_k, t_k] = \emptyset$ and $|X_{t_k} - X_{s_k}| \geqslant \epsilon$. Without loss of generality we may assume that $s_k \uparrow u$ and $t_k \downarrow u$ for some $u \in (s,t]$; $u = s$ can be ruled out because of right-continuity. By the càdlàg property of the paths, $|\Delta X_u| \geqslant \epsilon$, i.e., $u \in J$, which is a contradiction.

Since we have $f(X_{t_{\kappa+1,n}} - X_{t_{\kappa,n}}) = 0$ for intervals of the 'second kind', only the intervals containing some jump contribute to the (finite!) sum (9.3), and the claim follows. □

Lemma 9.4. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X.*

a) *$(N_t(f))_{t\geqslant 0}$ is a Lévy process on $\mathbb{R}^m$ for all $f \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\}, \mathbb{R}^m)$.*

b) *$(N_t(B))_{t\geqslant 0}$ is a Poisson process for all $B \in \mathscr{B}(\mathbb{R}^d)$ such that $0 \notin \overline{B}$.*

c) *$\nu(B) := \mathbb{E} N_1(B)$ is a locally finite measure on $\mathbb{R}^d \setminus \{0\}$.*

Proof. Set $\mathscr{F}_t := \sigma(X_s,\, s \leqslant t)$.

a) Let $f \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\}, \mathbb{R}^m)$. From Lemma 9.3 and (L2′) we see that $N_t(f)$ is $\mathscr{F}_t$ measurable and $N_t(f) - N_s(f) \perp \mathscr{F}_s$, $s \leqslant t$. Moreover, if $N_t^Y(\cdot)$ denotes the jump measure of the Lévy process $Y = (X_{t+s} - X_s)_{t\geqslant 0}$, we see that $N_{t-s}^Y(f) = N_t(f) - N_s(f)$. By the Markov property (Theorem 4.6), $X \sim Y$, and we get $N_{t-s}^Y(f) \sim N_{t-s}(f)$. Since $t \mapsto N_t(f)$ is càdlàg, $(N_t(f))_{t\geqslant 0}$ is a Lévy process.

b) By definition, $N_0(B) = 0$ and $t \mapsto N_t(B)$ is càdlàg. Since X is a Lévy process, we see as in the proof of Theorem 8.1 that the jump times

$$\tau_0 := 0, \quad \tau_1 := \inf\big\{t > 0 : \Delta X_t \in B\big\}, \quad \tau_k := \inf\big\{t > \tau_{k-1} : \Delta X_t \in B\big\}$$

satisfy $\tau_1 \sim \mathsf{Exp}(\nu(B))$, and the inter-jump times $(\tau_k - \tau_{k-1})_{k\in\mathbb{N}}$ are an iid sequence. The condition $0 \notin \overline{B}$ ensures that $N_t(B) < \infty$ a.s., which means that the intensity $\nu(B)$ is finite. Indeed, we have

$$1 - \mathrm{e}^{-t\nu(B)} = \mathbb{P}(\tau_1 \leqslant t) = \mathbb{P}(N_t(B) > 0) \xrightarrow[t\to 0]{} 0;$$

this shows that $\nu(B) < \infty$. Thus,

$$N_t(B) = \sum_{k=1}^{\infty} \mathbb{1}_{(0,t]}(\tau_k)$$

is a Poisson process (Example 3.2) and, in particular, a Lévy process (Theorem 3.4).

c) The intensity of $(N_t(B))_{t\geqslant 0}$ is $\nu(B) = \mathbb{E}N_1(B)$. By Fubini's theorem it is clear that ν is a measure. □

Definition 9.5. Let $N_t(\cdot)$ be the jump measure of a Lévy process X. The **intensity measure** is the measure $\nu(B) := \mathbb{E}N_1(B)$ from Lemma 9.4.

We will see in Corollary 9.13 that ν is the Lévy measure of $(X_t)_{t\geqslant 0}$ appearing in the Lévy–Khintchine formula.

Lemma 9.6. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X and ν the intensity measure. For every $f \in L^1(\nu)$, $f : \mathbb{R}^d \to \mathbb{R}^m$, the random variable $N_t(f) := \int f(y)\, N_t(dy)$ exists as L^1-limit of integrals of simple functions and satisfies*

$$\mathbb{E}N_t(f) = \mathbb{E}\int f(y)\, N_t(dy) = t\int_{y\neq 0} f(y)\, \nu(dy). \tag{9.4}$$

Proof. For any step function f of the form $f(y) = \sum_{k=1}^m f_k \mathbb{1}_{B_k}(y)$ with $0 \notin \overline{B}_k$ the formula (9.4) follows from $\mathbb{E}N_t(B_k) = t\nu(B_k)$ and the linearity of the integral.

Since ν is defined on $\mathbb{R}^d \setminus \{0\}$, any $f \in L^1(\nu)$ can be approximated by a sequence of step functions $(f_n)_{n\in\mathbb{N}}$ in $L^1(\nu)$-sense, and we get

$$\mathbb{E}|N_t(f_n) - N_t(f_m)| \leqslant t\int |f_n - f_m|\, d\nu \xrightarrow[m,n\to\infty]{} 0.$$

Because of the completeness of $L^1(\mathbb{P})$, the limit $\lim_{n\to\infty} N_t(f_n)$ exists, and with a routine argument we see that it is independent of the approximating sequence $f_n \to f \in L^1(\nu)$. This allows us to define $N_t(f)$ for $f \in L^1(\nu)$ as $L^1(\mathbb{P})$-limit of stochastic integrals of simple functions; obviously, (9.4) is preserved under this limiting procedure. □

Theorem 9.7. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X and ν the intensity measure.*

a) *$N_t(f) := \int f(y)\, N_t(dy)$ is a Lévy process for every $f \in L^1(\nu)$, $f : \mathbb{R}^d \to \mathbb{R}^m$.*
b) *$X_t^B := N_t(y\mathbb{1}_B(y))$ and $X_t - X_t^B$ are for every $B \in \mathscr{B}(\mathbb{R}^d)$, $0 \notin \overline{B}$, Lévy processes.*

Proof. a) Note that ν is a locally finite measure on $\mathbb{R}^d \setminus \{0\}$. This means that, by standard density results from integration theory, the family $\mathcal{C}_c(\mathbb{R}^d \setminus \{0\})$ is dense in $L^1(\nu)$. Fix $f \in L^1(\nu)$ and choose $f_n \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\})$ such that $f_n \to f$ in $L^1(\nu)$. Then, as in Lemma 9.6,

$$\mathbb{E}|N_t(f) - N_t(f_n)| \leqslant t \int |f - f_n| \, d\nu \xrightarrow[n\to\infty]{} 0.$$

Since $\mathbb{P}(|N_t(f)| > \epsilon) \leqslant \frac{t}{\epsilon} \int |f| \, d\nu \to 0$ for every $\epsilon > 0$ as $t \to 0$, the process $N_t(f)$ is continuous in probability. Moreover, it is the limit (in L^1, hence in probability) of the Lévy processes $N_t(f_n)$ (Lemma 9.4); therefore it is itself a Lévy process, see Lemma 7.3.

b) Set $f(y) := y\mathbb{1}_B(y)$ and $B_n := B \cap B_n(0)$. Then $f_n(y) = y\mathbb{1}_{B_n}(y)$ is bounded and $0 \notin \operatorname{supp} f_n$, hence $f_n \in L^1(\nu)$. This means that $N_t(f_n)$ is for every $n \in \mathbb{N}$ a Lévy process. Moreover,

$$N_t(f_n) = \int_{B_n} y \, N_t(dy) \xrightarrow[n\to\infty]{} \int_B y \, N_t(dy) = N_t(f) \quad \text{a.s.}$$

Since $N_t(f)$ changes its value only by jumps,

$$\begin{aligned} \mathbb{P}(|N_t(f)| > \epsilon) &\leqslant \mathbb{P}(X \text{ has at least one jump of size } B \text{ in } [0,t]) \\ &= \mathbb{P}(N_t(B) > 0) = 1 - \mathrm{e}^{-t\nu(B)}, \end{aligned}$$

which proves that the process $N_t(f)$ is continuous in probability. Lemma 7.3 shows that $N_t(f)$ is a Lévy process.

Finally, approximate $f(y) := y\mathbb{1}_B(y)$ by a sequence $\phi_l \in \mathcal{C}_c(\mathbb{R}^d \setminus \{0\}, \mathbb{R}^d)$. Now we can use Lemma 9.3 to get

$$X_t - N_t(\phi_l) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \left[(X_{t_{k+1,n}} - X_{t_{k,n}}) - \phi_l(X_{t_{k+1,n}} - X_{t_{k,n}}) \right],$$

The increments of X are stationary and independent, and so we conclude from the above formula that $X - N(\phi_l)$ has also stationary and independent increments. Since both X and $N(\phi_l)$ are continuous in probability, so is their difference, i.e., $X - N(\phi_l)$ is a Lévy process. Finally,

$$N_t(\phi_l) \xrightarrow[l\to\infty]{} N_t(f) \quad \text{and} \quad X_t - N_t(\phi_l) \xrightarrow[l\to\infty]{} X_t - N_t(f),$$

and since X and $N(f)$ are continuous in probability, Lemma 7.3 tells us that $X - N(f)$ is a Lévy process. □

We will now show that Lévy processes with 'disjoint jump heights' are independent. For this we need the following immediate consequence of Theorem 3.1:

Lemma 9.8 (Exponential martingale)**.** *Let $(X_t)_{t\geqslant 0}$ be a Lévy process. Then*

$$M_t := \frac{e^{i\,\xi\cdot X_t}}{\mathbb{E}\, e^{i\,\xi\cdot X_t}} = e^{i\,\xi\cdot X_t} e^{t\psi(\xi)}, \quad t \geqslant 0,$$

is a martingale for the canonical filtration $\mathscr{F}_t^X = \sigma(X_s,\, s \leqslant t)$ satisfying the inequality $\sup_{s\leqslant t} |M_s| \leqslant e^{t\,\mathrm{Re}\,\psi(\xi)}$.

Theorem 9.9. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X and $U, V \in \mathscr{B}(\mathbb{R}^d)$, $0 \notin \overline{U}, 0 \notin \overline{V}$ and $U \cap V = \emptyset$. Then the processes*

$$X_t^U := N_t(y\mathbb{1}_U(y)), \quad X_t^V := N_t(y\mathbb{1}_V(y)), \quad X_t - X_t^{U\cup V}$$

are independent Lévy processes in $\mathbb{R}^d$.

Proof. Set $W := U \cup V$. By Theorem 9.7, X^U, X^V and $X - X^W$ are Lévy processes. In fact, a slight variation of that argument even shows that $(X^U, X^V, X - X^W)$ is a Lévy process in $\mathbb{R}^{3d}$.

To see their independence, fix $s > 0$ and define for $t > s$ and $\xi, \eta, \theta \in \mathbb{R}^d$ the processes

$$C_t := \frac{e^{i\,\xi\cdot(X_t^U - X_s^U)}}{\mathbb{E}\big[e^{i\,\xi\cdot(X_t^U - X_s^U)}\big]} - 1, \qquad D_t := \frac{e^{i\,\eta\cdot(X_t^V - X_s^V)}}{\mathbb{E}\big[e^{i\,\eta\cdot(X_t^V - X_s^V)}\big]} - 1,$$
$$E_t := \frac{e^{i\,\theta\cdot(X_t - X_t^W - X_s + X_s^W)}}{\mathbb{E}\big[e^{i\,\theta\cdot(X_t - X_t^W - X_s + X_s^W)}\big]} - 1.$$

By Lemma 9.8, these processes are bounded martingales and $\mathbb{E}C_t = \mathbb{E}D_t = \mathbb{E}E_t = 0$. Set $t_{k,n} = s + \frac{k}{n}(t-s)$. Observe that

$$\begin{aligned}\mathbb{E}(C_t D_t E_t) &= \mathbb{E}\left(\sum_{k,l,m=0}^{n-1} (C_{t_{k+1,n}} - C_{t_{k,n}})(D_{t_{l+1,n}} - D_{t_{l,n}})(E_{t_{m+1,n}} - E_{t_{m,n}})\right)\\ &= \mathbb{E}\left(\sum_{k=0}^{n-1} (C_{t_{k+1,n}} - C_{t_{k,n}})(D_{t_{k+1,n}} - D_{t_{k,n}})(E_{t_{k+1,n}} - E_{t_{k,n}})\right).\end{aligned}$$

In the second equality we use that martingale increments $C_t - C_s, D_t - D_s, E_t - E_s$ are independent of $\mathscr{F}_s^X$, and by the tower property

$$\mathbb{E}\big[(C_{t_{k+1,n}} - C_{t_{k,n}})(D_{t_{l+1,n}} - D_{t_{l,n}})(E_{t_{m+1,n}} - E_{t_{m,n}})\big] = 0 \quad \text{unless } k = l = m.$$

An argument along the lines of Lemma 9.3 gives

$$\mathbb{E}(C_t D_t E_t) = \mathbb{E}\Bigg(\sum_{s<u\leqslant t} \underbrace{\Delta C_u\, \Delta D_u\, \Delta E_u}_{=0}\Bigg) = 0$$

as X_t^U, X_t^V and $Y_t := X_t - X_t^W$ cannot jump simultaneously since U, V and $\mathbb{R}^d \setminus W$ are mutually disjoint. Thus,

$$\begin{aligned}&\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^U - X_s^U)}\mathrm{e}^{\mathrm{i}\,\eta\cdot(X_t^V - X_s^V)}\mathrm{e}^{\mathrm{i}\,\theta\cdot(Y_t - Y_s)}\Big]\\&\quad = \mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t^U - X_s^U)}\Big]\cdot\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\,\eta\cdot(X_t^V - X_s^V)}\Big]\cdot\mathbb{E}\Big[\mathrm{e}^{\mathrm{i}\,\theta\cdot(Y_t - Y_s)}\Big].\end{aligned} \tag{9.5}$$

Since all processes are Lévy processes, (9.5) already proves the independence of X^U, X^V and $Y = X - X^W$. Indeed, we find for $0 = t_0 < t_1 < \cdots < t_m = t$ and $\xi_k, \eta_k, \theta_k \in \mathbb{R}^d$

$$\begin{aligned}&\mathbb{E}\Big(\mathrm{e}^{\mathrm{i}\sum_k \xi_k\cdot(X^U_{t_{k+1}} - X^U_{t_k})}\mathrm{e}^{\mathrm{i}\sum_k \eta_k\cdot(X^V_{t_{k+1}} - X^V_{t_k})}\mathrm{e}^{\mathrm{i}\sum_k \theta_k\cdot(Y_{t_{k+1}} - Y_{t_k})}\Big)\\
&\qquad = \mathbb{E}\Big(\prod_k \mathrm{e}^{\mathrm{i}\,\xi_k\cdot(X^U_{t_{k+1}} - X^U_{t_k})}\mathrm{e}^{\mathrm{i}\,\eta_k\cdot(X^V_{t_{k+1}} - X^V_{t_k})}\mathrm{e}^{\mathrm{i}\,\theta_k\cdot(Y_{t_{k+1}} - Y_{t_k})}\Big)\\
&\qquad \overset{(\mathrm{L2}')}{=} \prod_k \mathbb{E}\Big(\mathrm{e}^{\mathrm{i}\,\xi_k\cdot(X^U_{t_{k+1}} - X^U_{t_k})}\mathrm{e}^{\mathrm{i}\,\eta_k\cdot(X^V_{t_{k+1}} - X^V_{t_k})}\mathrm{e}^{\mathrm{i}\,\theta_k\cdot(Y_{t_{k+1}} - Y_{t_k})}\Big)\\
&\qquad \overset{(9.5)}{=} \prod_k \mathbb{E}\Big(\mathrm{e}^{\mathrm{i}\,\xi_k\cdot(X^U_{t_{k+1}} - X^U_{t_k})}\Big)\mathbb{E}\Big(\mathrm{e}^{\mathrm{i}\,\eta_k\cdot(X^V_{t_{k+1}} - X^V_{t_k})}\Big)\mathbb{E}\Big(\mathrm{e}^{\mathrm{i}\,\theta_k\cdot(Y_{t_{k+1}} - Y_{t_k})}\Big).\end{aligned}$$

The last equality follows from (9.5); the second equality uses (L2′) for the Lévy process $(X_t^U, X_t^V, X_t - X_t^W)$.

This shows that the families

$$(X^U_{t_{k+1}} - X^U_{t_k})_k, \quad (X^V_{t_{k+1}} - X^V_{t_k})_k \quad \text{and} \quad (Y_{t_{k+1}} - Y_{t_k})_k$$

are independent, hence the canonical σ-algebras $\sigma(X_t^U,\, t \geqslant 0)$, $\sigma(X_t^V,\, t \geqslant 0)$ and $\sigma(X_t - X_t^W,\, t \geqslant 0)$ are independent. □

Corollary 9.10. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X and ν the intensity measure.*

a) $(N_t(U))_{t\geqslant 0} \perp\!\!\!\perp (N_t(V))_{t\geqslant 0}$ *for* $U, V \in \mathscr{B}(\mathbb{R}^d)$, $0 \notin \overline{U}$, $0 \notin \overline{V}$, $U \cap V = \emptyset$.

b) *For all measurable $f : \mathbb{R}^d \to \mathbb{R}^m$ satisfying $f(0) = 0$ and $f \in L^1(\nu)$*

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\,\xi\cdot N_t(f)}\right) = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\int \xi\cdot f(y)\, N_t(dy)}\right) = \mathbb{E}\left(\mathrm{e}^{-t\int_{y\neq 0}\left[1 - \mathrm{e}^{\mathrm{i}\,\xi\cdot f(y)}\right]\nu(dy)}\right). \tag{9.6}$$

c) $\displaystyle\int_{y\neq 0} \left(|y|^2 \wedge 1\right)\, \nu(dy) < \infty.$

Proof. a) Since $(N_t(U))_{t\geqslant 0}$ and $(N_t(V))_{t\geqslant 0}$ are completely determined by the independent processes $(X_t^U)_{t\geqslant 0}$ and $(X_t^V)_{t\geqslant 0}$, cf. Theorem 9.9, the independence is clear.

b) Let us first prove (9.6) for step functions $f(x) = \sum_{k=1}^n f_k \mathbb{1}_{U_k}(x)$ with $f_k \in \mathbb{R}^m$ and disjoint sets $U_1, \dots, U_n \in \mathscr{B}(\mathbb{R}^d)$ such that $0 \notin \overline{U}_k$. Then

$$\begin{aligned}
\mathbb{E}\exp\left[\mathrm{i}\int \xi\cdot f(y)\,N_t(dy)\right] &= \mathbb{E}\exp\left[\mathrm{i}\sum_{k=1}^n \int \xi\cdot f_k\mathbb{1}_{U_k}(y)\,N_t(dy)\right]\\
&\overset{\text{a)}}{=} \prod_{k=1}^n \mathbb{E}\exp\left[\mathrm{i}\,\xi\cdot f_k\,N_t(U_k)\right]\\
&\overset{9.4.\text{b)}}{=} \prod_{k=1}^n \exp\left[t\nu(U_k)\left[\mathrm{e}^{\mathrm{i}\,\xi\cdot f_k}-1\right]\right]\\
&= \exp\left[t\sum_{k=1}^n\left[\mathrm{e}^{\mathrm{i}\,\xi\cdot f_k}-1\right]\nu(U_k)\right]\\
&= \exp\left[-t\int\left[1-\mathrm{e}^{\mathrm{i}\,\xi\cdot f(y)}\right]\nu(dy)\right].
\end{aligned}$$

For any $f \in L^1(\nu)$ the integral on the right-hand side of (9.6) exists. Indeed, the elementary inequality $|1-\mathrm{e}^{\mathrm{i}\,u}| \leqslant |u|\wedge 2$ and $\nu\{|y|\geqslant 1\} < \infty$ (Corollary 9.4.c)) yield

$$\left|\int_{y\neq 0}\left[1-\mathrm{e}^{\mathrm{i}\,\xi\cdot f(y)}\right]\nu(dy)\right| \leqslant |\xi|\int_{0<|y|<1}|f(y)|\,\nu(dy) + 2\int_{|y|\geqslant 1}\nu(dy) < \infty.$$

Therefore, (9.6) follows with a standard approximation argument and dominated convergence.

c) We have already seen in Lemma 9.4.c) that $\nu\{|y|\geqslant 1\} < \infty$.

Let us show that $\int_{0<|y|<1}|y|^2\,\nu(dy) < \infty$. For this we take $U = \{\delta < |y| < 1\}$. Again by Theorem 9.9, the processes X_t^U and $X_t - X_t^U$ are independent, and we get

$$0 < \left|\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t}\right| = \left|\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t^U}\right|\cdot\left|\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t-X_t^U)}\right| \leqslant \left|\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t^U}\right|.$$

Since X_t^U is a compound Poisson process – use part b) with $f(y) = y\mathbb{1}_U(y)$ – we get for all $|\xi|\leqslant 1$

$$0 < \left|\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot X_t^U}\right| = \mathrm{e}^{-t\int_U(1-\cos\xi\cdot y)\,\nu(dy)} \leqslant \mathrm{e}^{-t\int_{\delta<|y|\leqslant 1}\frac14(\xi\cdot y)^2\,\nu(dy)}.$$

For the equality we use $|\mathrm{e}^z| = \mathrm{e}^{\mathrm{Re}\,z}$, the inequality follows from the elementary estimate $\frac14 u^2 \leqslant 1-\cos u$ if $|u|\leqslant 1$. Letting $\delta\to 0$ we see that

$$\int_{0<|y|<1}|y|^2\,\nu(dy) < \infty. \qquad \square$$

Corollary 9.11. *Let $N_t(\cdot)$ be the jump measure of a Lévy process X and ν the intensity measure. For all $f:\mathbb{R}^d\to\mathbb{R}^m$ satisfying $f(0)=0$ and $f\in L^2(\nu)$ we* *have*[2]

$$\mathbb{E}\left(\left|\int f(y)\left[N_t(dy) - t\nu(dy)\right]\right|^2\right) = t\int_{y\neq 0} |f(y)|^2\,\nu(dy). \tag{9.7}$$

Proof. It is clearly enough to show (9.7) for step functions of the form

$$f(x) = \sum_{k=1}^{n} f_k \mathbb{1}_{B_k}(x), \quad B_k \text{ disjoint, } 0 \notin \overline{B}_k,\ f_k \in \mathbb{R}^m,$$

and then use an approximation argument.

Since the processes $N_t(B_k, \cdot)$ are independent Poisson processes with mean $\mathbb{E}N_t(B_k) = t\nu(B_k)$ and variance $\mathbb{V}N_t(B_k) = t\nu(B_k)$, we find

$$\begin{aligned}
&\mathbb{E}\left[(N_t(B_k) - t\nu(B_k))(N_t(B_l) - t\nu(B_l))\right]\\
&= \left\{\begin{array}{ll} 0, & \text{if } B_k \cap B_l = \emptyset, \text{ i.e., } k \neq l, \\ \mathbb{V}N_t(B_k) = t\nu(B_k), & \text{if } k = l, \end{array}\right\}\\
&= t\nu(B_k \cap B_l).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\left(\left|\int f(y)\left(N_t(dy) - t\nu(dy)\right)\right|^2\right)\\
&= \mathbb{E}\left(\iint f(y)f(z)\left(N_t(dy) - t\nu(dy)\right)\left(N_t(dz) - t\nu(dz)\right)\right)\\
&= \sum_{k,l=1}^{n} f_k f_l \underbrace{\mathbb{E}\Big(\big(N_t(B_k) - t\nu(B_k)\big)\big(N_t(B_l) - t\nu(B_l)\big)\Big)}_{= t\nu(B_k \cap B_l)}\\
&= t\sum_{k=1}^{n} |f_k|^2\,\nu(B_k) = t\int |f(y)|^2\,\nu(dy).
\end{aligned}$$

□

In contrast to Corollary 7.6 the following theorem does not need (but constructs) the Lévy triplet (l, Q, ν).

Theorem 9.12 (Lévy–Itô decomposition)**.** *Let X be a Lévy process and denote by $N_t(\cdot)$ and ν the jump and intensity measures. Then*

$$\begin{aligned}
X_t = \underbrace{\sqrt{Q}W_t}_{} &+ \int_{0<|y|<1} y\big(N_t(dy) - t\nu(dy)\big) \quad \Big] =: M_t,\ \ L^2\text{-martingale}\\
\underbrace{tl}_{\text{continuous Gaussian}} &+ \underbrace{\int_{|y|\geqslant 1} y\,N_t(dy).}_{\text{pure jump part}} \quad \Big] =: A_t,\ \ \text{bdd. variation}
\end{aligned} \tag{9.8}$$

[2] This is a special case of an Itô isometry, cf. (10.9) in the following chapter.

where $l \in \mathbb{R}^d$ and $Q \in \mathbb{R}^{d\times d}$ is a positive semidefinite symmetric matrix and W is a standard Brownian motion in $\mathbb{R}^d$. The processes on the right-hand side of (9.8) *are independent Lévy processes.*

Proof. 1° Set $U_n := \{\frac{1}{n} < |y| < 1\}$, $V = \{|y| \geqslant 1\}$, $W_n := U_n \cup V$ and define

$$X_t^V := \int_V y\, N_t(dy) \quad\text{and}\quad \widetilde{X}_t^{U_n} := \int_{U_n} y\, N_t(dy) - t\int_{U_n} y\,\nu(dy).$$

By Theorem 9.9 $(X_t^V)_{t\geqslant 0}$, $(\widetilde{X}_t^{U_n})_{t\geqslant 0}$ and $\big(X_t - X_t^{W_n} + t\int_{U_n} y\,\nu(dy)\big)_{t\geqslant 0}$ are independent Lévy processes. Since

$$X = (X - \widetilde{X}^{U_n} - X^V) + \widetilde{X}^{U_n} + X^V,$$

the theorem follows if we can show that the three terms on the right-hand side converge separately as $n \to \infty$.

2° Lemma 9.6 shows $\mathbb{E}\widetilde{X}_t^{U_n} = 0$; since $\widetilde{X}^{U_n}$ is a Lévy process, it is a martingale: for $s \leqslant t$

$$\begin{aligned}\mathbb{E}\left(\widetilde{X}_t^{U_n} \,\middle|\, \mathscr{F}_s\right) &= \mathbb{E}\left(\widetilde{X}_t^{U_n} - \widetilde{X}_s^{U_n} \,\middle|\, \mathscr{F}_s\right) + \widetilde{X}_s^{U_n}\\ &\overset{\text{(L2)}}{\underset{\text{(L1)}}{=}} \mathbb{E}\left(\widetilde{X}_{t-s}^{U_n}\right) + \widetilde{X}_s^{U_n} = \widetilde{X}_s^{U_n}.\end{aligned}$$

($\mathscr{F}_s$ can be taken as the natural filtration of X^{U_n} or X). By Doob's L^2 martingale inequality we find for any $t > 0$ and $m < n$

$$\begin{aligned}\mathbb{E}\left(\sup_{s\leqslant t}\big|\widetilde{X}_s^{U_n} - \widetilde{X}_s^{U_m}\big|^2\right) &\leqslant 4\mathbb{E}\left(\big|\widetilde{X}_t^{U_n} - \widetilde{X}_t^{U_m}\big|^2\right)\\ &= 4t\int_{\frac{1}{n}<|y|\leqslant\frac{1}{m}} |y|^2\,\nu(dy) \xrightarrow[m,n\to\infty]{} 0.\end{aligned}$$

Therefore, the limit $\int_{0<|y|<1} y\,\big(N_t(dy) - t\nu(dy)\big) = L^2\text{-}\lim_{n\to\infty}\widetilde{X}_t^{U_n}$ exists locally uniformly (in t). The limit is still an L^2 martingale with càdlàg paths (take a locally uniformly a.s. convergent subsequence) and, by Lemma 7.3, also a Lévy process.

3° Observe that

$$(X - \widetilde{X}^{U_n} - X^V) - (X - \widetilde{X}^{U_m} - X^V) = \widetilde{X}^{U_m} - \widetilde{X}^{U_n},$$

and so $X_t^c := L^2\text{-}\lim_{n\to\infty}(X_t - \widetilde{X}_t^{U_n} - X_t^V)$ exists locally uniformly (in t) Since, by construction $|\Delta(X_t - \widetilde{X}_t^{U_n} - X_t^V)| \leqslant \frac{1}{n}$, it is clear that X^c has a.s. continuous sample paths. By Lemma 7.3 it is a Lévy process. From Theorem 8.4 we know that all Lévy processes with continuous sample paths are of the form $tl + \sqrt{Q}W_t$ where

W is a Brownian motion, $Q \in \mathbb{R}^{d\times d}$ a symmetric positive semidefinite matrix and $l \in \mathbb{R}^d$.

4° Since independence is preserved under L^2-limits, the decomposition (9.8) follows. Finally,

$$\int_{|y|\geqslant 1} y\, N_t(dy,\omega) = \sum_{0<s\leqslant t} \Delta X_s(\omega)\mathbb{1}_{\{\Delta X_s(\omega)\geqslant 1\}} - \sum_{0<s\leqslant t} |\Delta X_s(\omega)|\mathbb{1}_{\{\Delta X_s(\omega)\leqslant -1\}}$$

is the difference of two increasing processes, i.e., it is of bounded variation. □

Corollary 9.13 (Lévy–Khintchine formula)**.** *Let X be a Lévy process. Then the characteristic exponent ψ is given by*

$$\psi(\xi) = -\,\mathrm{i}\, l\cdot\xi + \frac{1}{2}\xi\cdot Q\xi + \int_{y\neq 0}\left[1 - \mathrm{e}^{\mathrm{i}\, y\cdot\xi} + \mathrm{i}\,\xi\cdot y\mathbb{1}_{(0,1)}(|y|)\right]\nu(dy) \tag{9.9}$$

where ν is the intensity measure, $l \in \mathbb{R}^d$ and $Q \in \mathbb{R}^{d\times d}$ is symmetric and positive semidefinite.

Proof. Since the processes appearing in the Lévy–Itô decomposition (9.8) are independent, we see

$$\mathrm{e}^{-\psi(\xi)} = \mathbb{E}\mathrm{e}^{\mathrm{i}\xi\cdot X_1} = \mathbb{E}\mathrm{e}^{\mathrm{i}\xi\cdot(-l+\sqrt{Q}W_1)}\cdot\mathbb{E}\mathrm{e}^{\mathrm{i}\int_{0<|y|<1}\xi\cdot y(N_1(dy)-\nu(dy))}\cdot\mathbb{E}\mathrm{e}^{\mathrm{i}\int_{|y|\geqslant 1}\xi\cdot y\, N_1(dy)}.$$

Since W is a standard Brownian motion,

$$\mathbb{E}\,\mathrm{e}^{\mathrm{i}\,\xi\cdot(l+\sqrt{Q}W_1)} = \mathrm{e}^{\mathrm{i}\, l\cdot\xi - \frac{1}{2}\xi\cdot Q\xi}.$$

Using (9.6) with $f(y) = y\mathbb{1}_{U_n}(y)$, $U_n = \{\frac{1}{n} < |y| < 1\}$, subtracting $\int_{U_n} y\,\nu(dy)$ and letting $n\to\infty$ we get

$$\begin{aligned}\mathbb{E}\exp&\left[\mathrm{i}\int_{0<|y|<1}\xi\cdot y(N_1(dy)-\nu(dy))\right]\\ &= \exp\left[-\int_{0<|y|<1}\left[1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi}+\mathrm{i}\,\xi\cdot y\right]\nu(dy)\right];\end{aligned}$$

finally, (9.6) with $f(y) = y\mathbb{1}_V(y)$, $V = \{|y|\geqslant 1\}$, once again yields

$$\mathbb{E}\exp\left[\mathrm{i}\int_{|y|\geqslant 1}\xi\cdot y\, N_1(dy)\right] = \exp\left[-\int_{|y|\geqslant 1}\left[1-\mathrm{e}^{\mathrm{i}\,y\cdot\xi}\right]\nu(dy)\right]$$

finishing the proof. □

Chapter 10

A Digression: Stochastic Integrals

In this chapter we explain how one can integrate with respect to (a certain class of) random measures. Our approach is based on the notion of **random orthogonal measures** and it will include the classical Itô integral with respect to square-integrable martingales. Throughout this chapter, $(\Omega, \mathscr{A}, \mathbb{P})$ is a probability space, $(\mathscr{F}_t)_{t\geqslant 0}$ some filtration, $(E, \mathscr{E})$ is a measurable space and μ is a (positive) measure on $(E, \mathscr{E})$. Moreover, $\mathscr{R} \subset \mathscr{E}$ is a semiring, i.e., a family of sets such that $\emptyset \in \mathscr{R}$, for all $R, S \in \mathscr{R}$ we have $R \cap S \in \mathscr{R}$, and $R \setminus S$ can be represented as a finite union of disjoint sets from $\mathscr{R}$, cf. [54, Chapter 6] or [55, Definition 5.1]. It is not difficult to check that $\mathscr{R}_0 := \{R \in \mathscr{R} : \mu(R) < \infty\}$ is again a semiring.

Definition 10.1. Let $\mathscr{R}$ be a semiring on the measure space $(E, \mathscr{E}, \mu)$. A **random orthogonal measure** with **control measure** μ is a family of random variables $N(\omega, R) \in \mathbb{R}$, $R \in \mathscr{R}_0$, such that

$$\mathbb{E}\left[|N(\cdot, R)|^2\right] < \infty \quad \forall R \in \mathscr{R}_0 \tag{10.1}$$

$$\mathbb{E}\left[N(\cdot, R)N(\cdot, S)\right] = \mu(R \cap S) \quad \forall R, S \in \mathscr{R}_0. \tag{10.2}$$

The following Lemma explains why $N(R) = N(\omega, R)$ is called a (random) measure.

Lemma 10.2. *The random set function $R \mapsto N(R) := N(\omega, R)$, $R \in \mathscr{R}_0$, is countably additive in L^2, i.e.,*

$$N\left(\dot{\bigcup_{n=1}^{\infty}} R_n\right) = L^2\text{-}\lim_{n\to\infty} \sum_{k=1}^{n} N(R_k) \quad a.s. \tag{10.3}$$

for every sequence $(R_n)_{n\in\mathbb{N}} \subset \mathscr{R}_0$ of mutually disjoint sets such that

$$R := \dot{\bigcup_{n=1}^{\infty}} R_n \in \mathscr{R}_0.$$

In particular, $N(R \cup S) = N(R) + N(S)$ a.s. for disjoint $R, S \in \mathscr{R}_0$ such that $R \cup S \in \mathscr{R}_0$ and $N(\emptyset) = 0$ a.s. (notice that the exceptional set may depend on the sets R, S).

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_10

Proof. From $R = S = \emptyset$ and $\mathbb{E}[N(\emptyset)^2] = \mu(\emptyset) = 0$ we get $N(\emptyset) = 0$ a.s. It is enough to prove (10.3) as finite additivity follows if we take $(R_1, R_2, R_3, R_4 \dots) = (R, S, \emptyset, \emptyset, \dots)$. If $R_n \in \mathscr{R}_0$ are mutually disjoint sets such that $R := \bigcupdot_{n=1}^\infty R_n$ is again contained in $\mathscr{R}_0$, then

$$\begin{aligned}
&\mathbb{E}\left[\left(N(R) - \sum_{k=1}^n N(R_k)\right)^2\right] \\
&= \mathbb{E}N^2(R) + \sum_{k=1}^n \mathbb{E}N^2(R_k) - 2\sum_{k=1}^n \mathbb{E}[N(R)N(R_k)] + \sum_{j\neq k, j,k=1}^n \mathbb{E}[N(R_j)N(R_k)] \\
&\overset{(10.2)}{=} \mu(R) - \sum_{k=1}^n \mu(R_k) \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

where we use the σ-additivity of the measure μ. □

Example 10.3. a) **(White noise)** Let $\mathscr{R} = \{(s,t] : 0 \leqslant s < t < \infty\}$ and $\mu = \lambda$ be Lebesgue measure on $(0, \infty)$. Clearly, $\mathscr{R} = \mathscr{R}_0$ is a semiring. Let $W = (W_t)_{t \geqslant 0}$ be a one-dimensional standard Brownian motion. The random set function

$$N(\omega, (s,t]) := W_t(\omega) - W_s(\omega), \quad 0 \leqslant s < t < \infty$$

is a random orthogonal measure with control measure λ. This follows at once from

$$\mathbb{E}\big[(W_t - W_s)(W_v - W_u)\big] = t \wedge v - s \vee u = \lambda\big[(s,t] \cap (u,v]\big]$$

for all $0 \leqslant s < t < \infty$ and $0 \leqslant u < v < \infty$.

Mind, however, that N is **not σ-additive**. In order to see this, take $R_n := (1/(n+1), 1/n]$, where $n \in \mathbb{N}$, and observe that $\bigcupdot_n R_n = (0,1]$. Since W has stationary and independent increments, and scales like $W_t \sim \sqrt{t} W_1$, we have

$$\begin{aligned}
\mathbb{E}\exp\left[-\sum\nolimits_{n=1}^\infty |N(R_n)|\right] &= \mathbb{E}\exp\left[-\sum\nolimits_{n=1}^\infty |W_{1/(n+1)} - W_{1/n}|\right] \\
&= \prod\nolimits_{n=1}^\infty \mathbb{E}\exp\left[-(n(n+1))^{-1/2}|W_1|\right] \\
&\overset{\text{Jensen's}}{\underset{\text{ineq.}}{\leqslant}} \prod\nolimits_{n=1}^\infty \alpha^{(n(n+1))^{-1/2}}, \quad \alpha := \mathbb{E}e^{-|W_1|} \in (0,1).
\end{aligned}$$

As the series $\sum_{n=1}^\infty (n(n+1))^{-1/2}$ diverges, we get $\mathbb{E}\exp[-\sum_{n=1}^\infty |N(R_n)|] = 0$ which means that $\sum_{n=1}^\infty |N(\omega, R_n)| = \infty$ for almost all ω. This shows that $N(\cdot)$ cannot be countably additive. Indeed, countable additivity implies that the series

$$N\left(\omega, \bigcup_{n=1}^\infty R_n\right) = \sum_{n=1}^\infty N(\omega, R_n)$$

converges. The left-hand side, hence the summation, is independent under rearrangements.

This, however, entails absolute convergence of the series $\sum_{n=1}^{\infty} |N(\omega, R_n)|$ which does not hold as we have seen above.

b) **(2nd order orthogonal noise)** Let $X = (X_t)_{t\in T}$ be a complex-valued stochastic process defined on a bounded or unbounded interval $T \subset \mathbb{R}$. We assume that X has a.s. càdlàg paths. If $\mathbb{E}(|X_t|^2) < \infty$, we call X a **second-order process**; many properties of X are characterized by the **correlation function** $K(s,t) = \mathbb{E}\big(X_s \overline{X}_t\big)$, $s, t \in T$.

If $\mathbb{E}\big[(X_t - X_s)(\overline{X}_v - \overline{X}_u)\big] = 0$ for all $s \leqslant t \leqslant u \leqslant v$, $s,t,u,v \in T$, then X is said to have **orthogonal increments**. Fix $t_0 \in T$ and define for all $t \in T$

$$F(t) := \begin{cases} \mathbb{E}(|X_t - X_{t_0}|^2), & \text{if } t \geqslant t_0, \\ -\mathbb{E}(|X_{t_0} - X_t|^2), & \text{if } t \leqslant t_0. \end{cases}$$

Clearly, F is increasing and, since $t \mapsto X_t$ is a.s. right-continuous, it is also right-continuous. Moreover,

$$F(t) - F(s) = \mathbb{E}(|X_t - X_s|^2) \qquad \text{for all } s \leqslant t,\ s,t \in T. \tag{10.4}$$

To see this, we assume without loss of generality that $s \leqslant t_0 \leqslant t$. We have

$$\begin{aligned} F(t) - F(s) &= \mathbb{E}(|X_t - X_{t_0}|^2) - \mathbb{E}(|X_s - X_{t_0}|^2) \\ &= \mathbb{E}(|(X_t - X_s) + (X_s - X_{t_0})|^2) - \mathbb{E}(|X_s - X_{t_0}|^2) \\ &\overset{\text{orth.}}{\underset{\text{incr.}}{=}} \mathbb{E}(|X_t - X_s|^2). \end{aligned}$$

This shows that $\mu(s,t] := F(t) - F(s)$ defines a measure on the family $\mathscr{R} = \mathscr{R}_0 = \{(s,t] : -\infty < s < t < \infty,\ s,t \in T\}$, which is the control measure of $N(\omega, (s,t]) := X_t(\omega) - X_s(\omega)$. In fact, for $s < t, u < v$, $s,t,u,v \in T$, we have

$$\begin{aligned} X_t - X_s &= \big(X_t - X_{t\wedge v}\big) + \big(X_{t\wedge v} - X_{s\vee u}\big) + \big(X_{s\vee u} - X_s\big) \\ \overline{X}_v - \overline{X}_u &= \big(\overline{X}_u - \overline{X}_{t\wedge v}\big) + \big(\overline{X}_{t\wedge v} - \overline{X}_{s\vee u}\big) + \big(\overline{X}_{s\vee u} - \overline{X}_u\big). \end{aligned}$$

Using the orthogonality of the increments we get

$$\begin{aligned} \mathbb{E}\big[(X_t - X_s)(\overline{X}_v - \overline{X}_u)\big] &= \mathbb{E}\big[\big(X_{t\wedge v} - X_{s\vee u}\big)\big(\overline{X}_{t\wedge v} - \overline{X}_{s\vee u}\big)\big] \\ &= F(t\wedge v) - F(s\vee u) = \mu\big((s,t] \cap (u,v]\big), \end{aligned}$$

i.e., $N(\omega, \bullet)$ is a random orthogonal measure.

c) **(Martingale noise)** Let $M = (M_t)_{t\geqslant 0}$ be a square-integrable martingale with respect to the filtration $(\mathscr{F}_t)_{t\geqslant 0}$, $M_0 = 0$, and with càdlàg paths. Denote by $\langle M\rangle$ the predictable quadratic variation, i.e., the unique ($\langle M\rangle_0 := 0$) increasing predictable process such that $M^2 - \langle M\rangle$ is a martingale. The random set function

$$N(\omega, (s,t]) := M_t(\omega) - M_s(\omega), \quad s \leqslant t,$$

is a random orthogonal measure on $\mathscr{R} = \{(s,t] \,:\, 0 \leqslant s < t < \infty\}$ with control measure $\mu(s,t] = \mathbb{E}(\langle M\rangle_t - \langle M\rangle_s)$. This follows immediately from the tower property of conditional expectation

$$\mathbb{E}[M_t M_v] \stackrel{\text{tower}}{=} \mathbb{E}[M_t \mathbb{E}(M_v \mid \mathscr{F}_t)] = \mathbb{E}[M_t^2] = \mathbb{E}\langle M\rangle_t \quad \text{if} \quad t \leqslant v$$

which, in turn, gives for all $0 \leqslant s < t$ and $0 \leqslant u < v$

$$\begin{aligned}\mathbb{E}[(M_t - M_s)(M_v - M_u)] &= \mathbb{E}\langle M\rangle_{t\wedge v} - \mathbb{E}\langle M\rangle_{s\wedge v} - \mathbb{E}\langle M\rangle_{t\wedge u} + \mathbb{E}\langle M\rangle_{s\wedge u} \\ &= \mu\big((s,t] \cap (0,v]\big) - \mu\big((s,t] \cap (0,u]\big) \\ &= \mu\big((s,t] \cap (u,v]\big).\end{aligned}$$

d) **(Poisson random measure)** Let X be a d-dimensional Lévy process,

$$\mathscr{S} := \big\{B \in \mathscr{B}(\mathbb{R}^d) \,:\, 0 \notin \overline{B}\big\}, \quad \mathscr{R} := \big\{(s,t] \times B \,:\, 0 \leqslant s < t < \infty,\ B \in \mathscr{S}\big\},$$

and $N_t(B)$ the jump measure (Definition 9.1). The random set function

$$\widetilde{N}(\omega, (s,t] \times B) := [N_t(\omega, B) - t\nu(B)] - [N_s(\omega, B) - s\nu(B)], \quad R = (s,t] \times B \in \mathscr{R},$$

is a random orthogonal measure with control measure $\lambda \times \nu$ where λ is Lebesgue measure on $(0,\infty)$ and ν is the Lévy measure of X. Indeed, by definition $\mathscr{R} = \mathscr{R}_0$, and it is not hard to see that $\mathscr{R}$ is a semiring[1].

Set $\widetilde{N}_t(B) := \widetilde{N}((0,t] \times B)$ and let $B, C \in \mathscr{S}$, $t, v \geqslant 0$. As in the proof of Corollary 9.11 we have

$$\mathbb{E}\big[\widetilde{N}_t(B)\widetilde{N}_t(C)\big] = t\nu(B \cap C).$$

Since $\mathscr{S}$ is a semiring, we get $B = (B \cap C) \cup (B \setminus C) = (B \cap C) \cup B_1 \cup \cdots \cup B_n$ with finitely many mutually disjoint $B_k \in \mathscr{S}$ such that $B_k \subset B \setminus C$.

The processes $\widetilde{N}(B_k)$ and $\widetilde{N}(C)$ are independent (Corollary 9.10) and centered. Therefore we have for $t \leqslant v$

$$\begin{aligned}\mathbb{E}\big[\widetilde{N}_t(B)\widetilde{N}_v(C)\big] &= \mathbb{E}\big[\widetilde{N}_t(B \cap C)\widetilde{N}_v(C)\big] + \sum_{k=1}^n \mathbb{E}\big[\widetilde{N}_t(B_k)\widetilde{N}_v(C)\big] \\ &= \mathbb{E}\big[\widetilde{N}_t(B \cap C)\widetilde{N}_v(C)\big] + \sum_{k=1}^n \mathbb{E}\widetilde{N}_t(B_k) \cdot \mathbb{E}\widetilde{N}_v(C) \\ &= \mathbb{E}\big[\widetilde{N}_t(B \cap C)\widetilde{N}_v(C)\big].\end{aligned}$$

[1] Both $\mathscr{S}$ and $\mathscr{I} := \{(s,t] \,:\, 0 \leqslant s < t < \infty\}$ are semirings, and so is their cartesian product $\mathscr{R} = \mathscr{I} \times \mathscr{S}$, see [54, Lemma 13.1] or [55, Lemma 15.1] for the straightforward proof.

Use the same argument over again, as well as the fact that $\widetilde{N}_t(B\cap C)$ has independent and centered increments (Lemma 9.4), to get

$$\begin{aligned}
&\mathbb{E}\big[\widetilde{N}_t(B)\widetilde{N}_v(C)\big]\\
&\quad= \mathbb{E}\big[\widetilde{N}_t(B\cap C)\widetilde{N}_v(B\cap C)\big]\\
&\quad= \mathbb{E}\big[\widetilde{N}_t(B\cap C)\widetilde{N}_t(B\cap C)\big] + \overbrace{\mathbb{E}\big[\widetilde{N}_t(B\cap C)\big\{\widetilde{N}_v(B\cap C)-\widetilde{N}_t(B\cap C)\big\}\big]}^{=\mathbb{E}\widetilde{N}_t(B\cap C)\,\mathbb{E}[\widetilde{N}_v(B\cap C)-\widetilde{N}_t(B\cap C)]=0}\\
&\quad= \mathbb{E}\big[\widetilde{N}_t(B\cap C)\widetilde{N}_t(B\cap C)\big]\\
&\quad= t\nu(B\cap C)\\
&\quad= \lambda((0,t]\cap(0,v])\nu(B\cap C).
\end{aligned} \tag{10.5}$$

For $s\leqslant t$, $u\leqslant v$ and $B,C\in\mathscr{S}$ a lengthy, but otherwise completely elementary, calculation based on (10.5) shows

$$\mathbb{E}\big[\widetilde{N}((s,t]\times B)\widetilde{N}((u,v]\times C)\big] = \lambda((s,t]\cap(u,v])\nu(B\cap C).$$

e) **(Space-time white noise)** Let $\mathscr{R} := \big\{(0,t]\times B \;:\; t>0,\ B\in\mathscr{B}(\mathbb{R}^d)\big\}$ and $\mu=\lambda$ Lebesgue measure on the half-space $\mathbb{H}^+ := [0,\infty)\times\mathbb{R}^d$.

Consider the mean-zero, real-valued Gaussian process $(\mathbb{W}(R))_{R\in\mathscr{B}(\mathbb{H}^+)}$ whose covariance function is given by $\mathbb{C}\mathrm{ov}(\mathbb{W}(R)\mathbb{W}(S)) = \lambda(R\cap S)$.[2] By its very definition $\mathbb{W}(R)$ is a random orthogonal measure on $\mathscr{R}_0$ with control measure λ.

We will now define a stochastic integral in the spirit of Itô's original construction.

Definition 10.4. Let $\mathscr{R}$ be a semiring and $\mathscr{R}_0 = \{R\in\mathscr{R} \;:\; \mu(R)<\infty\}$. A **simple function** is a deterministic function of the form

$$f(x) = \sum_{k=1}^{n} c_k \mathbb{1}_{R_k}(x), \quad n\in\mathbb{N},\ c_k\in\mathbb{R},\ R_k\in\mathscr{R}_0. \tag{10.6}$$

Intuitively, $I_N(f) = \sum_{k=1}^n c_k N(R_k)$ should be the stochastic integral of a simple function f. The only problem is the well-definedness. Since a random orthogonal measure is a.s. finitely additive, the following lemma has exactly the same proof as the usual well-definedness result for the Lebesgue integral of a step function, see, e.g., Schilling [54, Lemma 9.1] or [55, Lemma 8.1]; note that finite unions of null sets are again null sets.

[2] The map $(R,S)\mapsto\lambda(R\cap S)$ is positive semidefinite, i.e., for $R_1,\dots,R_n\in\mathscr{B}(\mathbb{H}^+)$ and $\xi_1,\dots,\xi_n\in\mathbb{R}$

$$\sum_{j,k=1}^n \xi_j\xi_k\lambda(R_j\cap R_k) = \sum_{j,k=1}^n \int \xi_j\mathbb{1}_{R_j}(x)\xi_k\mathbb{1}_{R_k}(x)\,\lambda(dx) = \int\Big(\sum_{k=1}^n \xi_k\mathbb{1}_{R_k}(x)\Big)^2\lambda(dx)\geqslant 0.$$

Lemma 10.5. *Let f be a simple function and assume that $f = \sum_{k=1}^n c_k \mathbb{1}_{R_k} = \sum_{j=1}^m b_j \mathbb{1}_{S_j}$ has two representations as step-function. Then*

$$\sum_{k=1}^n c_k N(R_k) = \sum_{j=1}^m b_j N(S_j) \quad a.s.$$

Definition 10.6. Let $N(R)$, $R \in \mathscr{R}_0$, be a random orthogonal measure with control measure μ. The **stochastic integral** of a simple function f given by (10.6) is the random variable

$$I_N(\omega, f) := \sum_{k=1}^n c_k N(\omega, R_k). \tag{10.7}$$

The following properties of the stochastic integral are more or less immediate from the definition.

Lemma 10.7. *Let $N(R)$, $R \in \mathscr{R}_0$, be a random orthogonal measure with control measure μ, f, g simple functions, and $\alpha, \beta \in \mathbb{R}$.*

a) $I_N(\mathbb{1}_R) = N(R)$ *for all* $R \in \mathscr{R}_0$;
b) $S \mapsto I_N(\mathbb{1}_S)$ *extends N uniquely to $S \in \rho(\mathscr{R}_0)$, the ring generated by $\mathscr{R}_0$;*[3]
c) $I_N(\alpha f + \beta g) = \alpha I_N(f) + \beta I_N(g)$; *(linearity)*
d) $\mathbb{E}\left[I_N(f)^2\right] = \int f^2 \, d\mu$; *(Itô's isometry)*

Proof. The properties a) and c) are clear. For b) we note that $\rho(\mathscr{R}_0)$ can be constructed from $\mathscr{R}_0$ by adding all possible finite unions of (disjoint) sets (see, e.g., [54, Proof of Theorem 6.1, Step 2]). In order to see d), we use (10.6) and the orthogonality relation $\mathbb{E}\left[N(R_j)N(R_k)\right] = \mu(R_j \cap R_k)$ to get

$$\begin{aligned}
\mathbb{E}\big[I_N(f)^2\big] &= \sum_{j,k=1}^n c_j c_k \mathbb{E}\left[N(R_j)N(R_k)\right] \\
&= \sum_{j,k=1}^n c_j c_k \mu(R_j \cap R_k) \\
&= \int \sum_{j,k=1}^n c_j \mathbb{1}_{R_j}(x)\, c_k \mathbb{1}_{R_k}(x)\, \mu(dx) \\
&= \int f^2(x)\, \mu(dx).
\end{aligned}$$ □

Itô's isometry now allows us to extend the stochastic integral to the $L^2(\mu)$-closure of the simple functions: $L^2(E, \sigma(\mathscr{R}), \mu)$. For this take $f \in L^2(E, \sigma(\mathscr{R}), \mu)$

[3] A ring is a family of sets which contains $\emptyset$ and which is stable under unions and differences of finitely many sets. Since $R \cap S = R \setminus (R \setminus S)$, it is automatically stable under finite intersections. The ring generated by $\mathscr{R}_0$ is the smallest ring containing $\mathscr{R}_0$.

and any approximating sequence $(f_n)_{n\in\mathbb{N}}$ of simple functions, i.e.,

$$\lim_{n\to\infty}\int |f-f_n|^2\,d\mu = 0.$$

In particular, $(f_n)_{n\in\mathbb{N}}$ is an $L^2(\mu)$ Cauchy sequence, and Itô's isometry shows that the random variables $(I_N(f_n))_{n\in\mathbb{N}}$ are a Cauchy sequence in $L^2(\mathbb{P})$:

$$\mathbb{E}\left[(I_N(f_n)-I_N(f_m))^2\right] = \mathbb{E}\left[I_N(f_n-f_m)^2\right] = \int (f_n-f_m)^2\,d\mu \xrightarrow[m,n\to\infty]{} 0.$$

Because of the completeness of $L^2(\mathbb{P})$, the limit $\lim_{n\to\infty} I_N(f_n)$ exists and, by a standard argument, it does not depend on the approximating sequence.

Definition 10.8. Let $N(R)$, $R\in\mathscr{R}_0$, be a random orthogonal measure with control measure μ. The **stochastic integral** of a function $f\in L^2(E,\sigma(\mathscr{R}),\mu)$ is the random variable

$$\int f(x)\,N(\omega,dx) := L^2(\mathbb{P})\text{-}\lim_{n\to\infty} I_N(\omega,f_n) \tag{10.8}$$

where $(f_n)_{n\in\mathbb{N}}$ is any sequence of simple functions which approximate f in $L^2(\mu)$.

It is immediate from the definition of the stochastic integral, that $f\mapsto\int f\,dN$ is linear and enjoys **Itô's isometry**

$$\mathbb{E}\left[\left(\int f(x)\,N(dx)\right)^2\right] = \int f^2(x)\,\mu(dx). \tag{10.9}$$

Remark 10.9. Assume that the random orthogonal measure N is of **space-time type**, i.e., $E=(0,\infty)\times X$ where $(X,\mathscr{X})$ is some measurable space, and $\mathscr{R}=\{(0,t]\times B \;:\; B\in\mathscr{S}\}$ where $\mathscr{S}$ is a semiring in $\mathscr{X}$. If for $B\in\mathscr{S}$ the stochastic process $N_t(B) := N((0,t]\times B)$ is a martingale with respact to the filtration $\mathscr{F}_t := \sigma(N((0,s]\times B),\ s\leqslant t, B\in\mathscr{S})$, then

$$N_t(f) := \iint \mathbb{1}_{(0,t]}(s)f(x)\,N(ds,dx),\quad \mathbb{1}_{(0,t]}\otimes f\in L^2(\mu),\ t\geqslant 0,$$

is again a(n L^2-)martingale. For simple functions f this follows immediately from the fact that sums and differences of finitely many martingales (with a common filtration) are again a martingale. Since $L^2(\mathbb{P})$-limits preserve the martingale property, the claim follows. □

At first sight, the stochastic integral defined in 10.8 looks rather restrictive since we can only integrate deterministic functions f. As all randomness can be put into the random orthogonal measure, we have considerable flexibility, and the following construction shows that Definition 10.8 covers pretty much the most general stochastic integrals.

From now on we assume that

- the random measure $N(dt,dx)$ on $(E,\mathscr{E}) = ((0,\infty)\times X,\ \mathscr{B}(\mathbb{R}^d)\otimes\mathscr{X})$ is of space-time type, cf. Remark 10.9, with control measure $\mu(dt,dx)$;

- $(\mathscr{F}_t)_{t\geqslant 0}$ is some filtration in $(\Omega, \mathscr{A}, \mathbb{P})$.

Let τ be a stopping time; the set $]\!]0,\tau]\!] := \{(\omega,t) : 0 < t \leqslant \tau(\omega)\}$ is called **stochastic interval**. We define

- $E^\circ := \Omega \times (0,\infty) \times X$;
- $\mathscr{E}^\circ := \mathscr{P} \otimes \mathscr{X}$ where $\mathscr{P}$ is the predictable σ-algebra in $\Omega \times (0,\infty)$, see Definition A.8 in the appendix;
- $\mathscr{R}^\circ := \{]\!]0,\tau]\!] \times B : \tau \text{ bounded stopping time, } B \in \mathscr{S}\}$;
- $\mu^\circ(d\omega, dt, dx) := \mathbb{P}(d\omega)\mu(dt,dx)$ as control measure;
- $N^\circ\big(\omega,]\!]0,\tau]\!] \times B\big) := N\big(\omega, (0,\tau(\omega)] \times B\big)$ as random orthogonal measure[4].

Lemma 10.10. *Let N°, $\mathscr{R}^\circ_0$ and μ° be as above. The $\mathscr{R}^\circ_0$-simple processes*

$$f(\omega,t,x) := \sum_{k=1}^n c_k \mathbb{1}_{]\!]0,\tau_k]\!]}(\omega,t)\mathbb{1}_{B_k}(x), \qquad c_k \in \mathbb{R},\ \ ]\!]0,\tau_k]\!] \times B_k \in \mathscr{R}^\circ_0$$

are $L^2(\mu^\circ)$-dense in $L^2(E^\circ, \mathscr{P} \otimes \sigma(\mathscr{S}), \mu^\circ)$.

Proof. This follows from standard arguments from measure and integration; notice that the predictable σ-algebra $\mathscr{P} \otimes \sigma(\mathscr{S})$ is generated by sets $]\!]0,\tau]\!] \times B$ where τ is a bounded stopping time and $B \in \mathscr{S}$, cf. Theorem A.9 in the appendix. □

Observe that for simple processes appearing in Lemma 10.10

$$\int f(\omega,t,x) N(\omega,dt,dx) := \sum_{k=1}^n c_k N^\circ(\omega,]\!]0,\tau_k]\!] \times B_k) = \sum_{k=1}^n c_k N(\omega,(0,\tau_k(\omega)] \times B_k)$$

is a stochastic integral which satisfies

$$\mathbb{E}\left[\Big(\int f(\cdot,t,x)\, N(\cdot,dt,dx)\Big)^2\right] = \sum_{k=1}^n c_k^2\, \mu^\circ(]\!]0,\tau_k]\!] \times B_k) = \sum_{k=1}^n c_k^2\, \mathbb{E}\mu((0,\tau_k] \times B_k).$$

Just as above we can now extend the stochastic integral to $L^2(E^\circ, \mathscr{P} \otimes \sigma(\mathscr{S}), \mu^\circ)$.

Corollary 10.11. *Let $N(\omega, dt, dx)$ be a random orthogonal measure on E of space-time type with control measure $\mu(dt,dx)$ and $f : \Omega \times (0,\infty) \times X \to \mathbb{R}$ be an element of $L^2(E^\circ, \mathscr{P} \otimes \sigma(\mathscr{S}), \mu^\circ)$. Then the stochastic integral*

$$\iint f(\omega,t,x)\, N(\omega,dt,dx)$$

exists and satisfies the following **Itô isometry**

$$\mathbb{E}\left[\Big(\iint f(\cdot,t,x)\, N(\cdot,dt,dx)\Big)^2\right] = \iint \mathbb{E} f^2(\cdot,t,x)\, \mu(dt,dx). \tag{10.10}$$

[4] To see that it is indeed a random orthogonal measure, use a discrete approximation of τ.

Let us show that the stochastic integral w.r.t. a space-time random orthogonal measure extends the usual Itô integral. To do so we need the following auxiliary result.

Lemma 10.12. *Let $N(\omega, dt, dx)$ be a random orthogonal measure on E of space-time type (cf. Remark 10.9) with control measure $\mu(dt, dx)$ and τ a stopping time. Then*

$$\begin{aligned}\iint \phi(\omega)\mathbb{1}_{]\!]\tau,\infty[\![}(\omega,t)f(\omega,t,x)\,N(\omega,dt,dx)&\\ = \phi(\omega)\iint \mathbb{1}_{]\!]\tau,\infty[\![}(\omega,t)f(\omega,t,x)\,N(\omega,dt,dx)&\end{aligned} \tag{10.11}$$

for all $\phi \in L^\infty(\mathscr{F}_\tau)$ and $f \in L^2(E^\circ, \mathscr{P}\otimes\sigma(\mathscr{S}), \mu^\circ)$.

Proof. Since $t \mapsto \mathbb{1}_{]\!]\tau,\infty[\![}(\omega,t)$ is adapted to the filtration $(\mathscr{F}_t)_{t\geqslant 0}$ and left-continuous, the integrands appearing in (10.11) are predictable, hence all stochastic integrals are well defined.

1° Assume that $\phi(\omega) = \mathbb{1}_F(\omega)$ for some $F \in \mathscr{F}_\tau$ and $f(\omega,t,x) = \mathbb{1}_{]\!]0,\sigma]\!]}(\omega,t)\mathbb{1}_B(x)$ for some bounded stopping time σ and $]\!]0,\sigma]\!] \times B \in \mathscr{R}_0^\circ$. Define

$$N_\sigma(\omega, B) := N(\omega, (0,\sigma(\omega)] \times B) = N^\circ(\omega,]\!]0,\sigma]\!] \times B)$$

for any $B \in \mathscr{S}_0$. The random time $\tau_F := \tau\mathbb{1}_F + \infty\mathbb{1}_{F^c}$ is a stopping time[5], and we have

$$\phi\mathbb{1}_{]\!]\tau,\infty[\![}f = \mathbb{1}_F\mathbb{1}_{]\!]\tau,\infty[\![}\mathbb{1}_{]\!]0,\sigma]\!]}\mathbb{1}_B = \mathbb{1}_{]\!]\tau_F,\infty[\![}\mathbb{1}_{]\!]0,\sigma]\!]}\mathbb{1}_B = \mathbb{1}_{]\!]\tau_F\wedge\sigma,\sigma]\!]}\mathbb{1}_B.$$

From this we get (10.11) for our choice of ϕ and f:

$$\begin{aligned}\iint \phi\mathbb{1}_{]\!]\tau,\infty[\![}(t)f(t,x)\,N(dt,dx) &= \iint \mathbb{1}_{]\!]\tau_F\wedge\sigma,\sigma]\!]}(t)\mathbb{1}_B(x)\,N(dt,dx)\\ &= N_\sigma(B) - N_{\tau_F\wedge\sigma}(B)\\ &= \mathbb{1}_F\cdot(N_\sigma(B) - N_{\tau\wedge\sigma}(B))\\ &= \mathbb{1}_F\iint \mathbb{1}_{]\!]\tau\wedge\sigma,\sigma]\!]}(t)\mathbb{1}_B(x)\,N(dt,dx)\\ &= \mathbb{1}_F\iint \mathbb{1}_{]\!]\tau,\infty[\![}\mathbb{1}_{]\!]0,\sigma]\!]}(t)\mathbb{1}_B(x)\,N(dt,dx)\\ &= \phi\iint \mathbb{1}_{]\!]\tau,\infty[\![}(t)f(t,x)\,N(dt,dx).\end{aligned}$$

2° If $\phi = \mathbb{1}_F$ for some $F \in \mathscr{F}_\tau$ and f is a simple process, then (10.11) follows from 1° because of the linearity of the stochastic integral.

[5] Indeed, $\{\tau_F \leqslant t\} = \{\tau \leqslant t\} \cap F = \begin{Bmatrix} \emptyset, & \tau > t \\ F, & \tau \leqslant t\end{Bmatrix} \in \mathscr{F}_t$ for all $t \geqslant 0$.

3° If $\phi = \mathbb{1}_F$ for some $F \in \mathscr{F}_\tau$ and $f \in L^2(\mu^\circ)$, then (10.11) follows from 2° and Itô's isometry: Let f_n be a sequence of simple processes which approximate f. Then

$$\begin{aligned}\mathbb{E}\left[\Big(\iint \big(\phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f_n(t,x) - \phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x)\big)\,N(dt,dx)\Big)^2\right] \\ = \iint \mathbb{E}\left[\big(\phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f_n(t,x) - \phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x)\big)^2\right]\mu(dt,dx) \\ \leqslant \iint \mathbb{E}\left[\big(f_n(t,x) - f(t,x)\big)^2\right]\mu(dt,dx) \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

4° If ϕ is an $\mathscr{F}_\tau$ measurable step-function and $f \in L^2(\mu^\circ)$, then (10.11) follows from 3° because of the linearity of the stochastic integral.

5° Since we can approximate $\phi \in L^\infty(\mathscr{F}_\tau)$ uniformly by $\mathscr{F}_\tau$ measurable step functions ϕ_n, (10.11) follows from 4° and Itô's isometry because of the following inequality:

$$\begin{aligned}\mathbb{E}\left[\Big(\iint \big[\phi_n\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x) - \phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x)\big]\;N(dt,dx)\Big)^2\right] \\ = \iint \mathbb{E}\left(\big[\phi_n\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x) - \phi\mathbb{1}_{\rrbracket\tau,\infty\llbracket}(t)f(t,x)\big]^2\right)\mu(dt,dx) \\ \leqslant \|\phi_n - \phi\|^2_{L^\infty(\mathbb{P})} \iint \mathbb{E}\big[f^2(t,x)\big]\,\mu(dt,dx).\end{aligned}$$

□

We will now consider 'martingale noise' random orthogonal measures, see Example 10.3.c), which are given by (the predictable quadratic variation of) a square-integrable martingale M. For these random measures our definition of the stochastic integral coincides with Itô's definition. Recall that the Itô integral driven by M is first defined for simple, left-continuous processes of the form

$$f(\omega,t) := \sum_{k=1}^{n} \phi_k(\omega)\mathbb{1}_{\rrbracket\tau_k,\tau_{k+1}\rrbracket}(\omega,t), \quad t \geqslant 0, \tag{10.12}$$

where $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \cdots \leqslant \tau_{n+1}$ are bounded stopping times and ϕ_k bounded $\mathscr{F}_{\tau_k}$ measurable random variables. The Itô integral for such simple processes is

$$\int f(\omega,t)\,dM_t(\omega) := \sum_{k=1}^{n} \phi_k(\omega)\big(M_{\tau_{k+1}}(\omega) - M_{\tau_k}(\omega)\big)$$

and it is extended by Itô's isometry to all integrands from $L^2(\Omega\times(0,\infty), \mathscr{P}, d\mathbb{P}\otimes d\langle M\rangle_t)$. For details we refer to any standard text on Itô integration, e.g., Protter [44, Chapter II] or Revuz & Yor [45, Chapter IV].

We will now use Lemma 10.12 in the particular situation where the space component dx is not present.

Theorem 10.13. *Let $N(dt)$ be a 'martingale noise' random orthogonal measure induced by the square-integrable martingale M (Example 10.3). The stochastic integral w.r.t. the random orthogonal measure $N(dt)$ and Itô's stochastic integral w.r.t. M coincide.*

Proof. Let $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \cdots \leqslant \tau_{n+1}$ be bounded stopping times, $\phi_k \in L^\infty(\mathscr{F}_{\tau_k})$ bounded random variables and $f(\omega, t)$ be a simple stochastic process of the form (10.12). From Lemma 10.12 we get

$$\begin{aligned}
\int f(t)\,N(dt) &= \int \sum_{k=1}^{n} \phi_k \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}(t)\,N(dt)\\
&= \sum_{k=1}^{n} \int \phi_k \mathbb{1}_{]\!]\tau_k,\infty[\![}(t)\mathbb{1}_{]\!]0,\tau_{k+1}]\!]}(t)\,N(dt)\\
&= \sum_{k=1}^{n} \phi_k \int \mathbb{1}_{]\!]\tau_k,\infty[\![}(t)\mathbb{1}_{]\!]0,\tau_{k+1}]\!]}(t)\,N(dt)\\
&= \sum_{k=1}^{n} \phi_k (M_{\tau_{k+1}} - M_{\tau_k}).
\end{aligned}$$

This means that both stochastic integrals coincide on the simple stochastic processes. Since both integrals are extended by Itô's isometry, the assertion follows. □

Example 10.14. Using random orthogonal measures we can re-state the Lévy–Itô decomposition appearing in Theorem 9.12. For this, let $\widetilde{N}(dt, dx)$ be the Poisson random orthogonal measure (Example 10.3.d) on $E = (0,\infty) \times (\mathbb{R}^d \setminus \{0\})$ with control measure $dt \times \nu(dx)$ (ν is a Lévy measure). Additionally, we **define** for all deterministic functions $h : (0,\infty) \times \mathbb{R}^d \to \mathbb{R}$

$$\iint h(s,x)\,N(\omega, ds, dx) := \sum_{0<s<\infty} h(s, \Delta X_s(\omega)) \quad \forall \omega \in \Omega$$

provided that the sum $\sum_{0<s<\infty} |h(s, \Delta X_s(\omega))| < \infty$ for each ω.[6] If X is a Lévy process with characteristic exponent ψ and Lévy triplet (l, Q, ν), then

$$\begin{aligned}
X_t = \sqrt{Q}W_t \;&+ \iint \mathbb{1}_{(0,t]}(s) y \mathbb{1}_{(0,1)}(|y|)\,\widetilde{N}(ds,dy) \quad \Big] =: M_t,\ \ L^2\text{-martingale}\\
tl \;&+ \iint \mathbb{1}_{(0,t]}(s) y \mathbb{1}_{\{|y|\geqslant 1\}}\,N(ds,dy). \quad \Big] =: A_t,\ \ \text{bdd. variation}
\end{aligned}$$

continuous Gaussian
pure jump part

[6] This is essentially an ω-wise Riemann–Stieltjes integral. A sufficient condition for the absolute convergence is, e.g., that h is continuous and $h(t,\cdot)$ vanishes uniformly in t in some neighbourhood of $x = 0$. The reason for this is the fact that $N_t(\omega, B_\epsilon^c(0)) = N(\omega, (0,t] \times B_\epsilon^c(0)) < \infty$, i.e., there are at most finitely many jumps of size exceeding $\epsilon > 0$.

Example 10.14 is quite particular in the sense that $N(\cdot, dt, dx)$ is a bona fide positive measure, and the control measure $\mu(dt, dx)$ is also the compensator, i.e., a measure such that $\widetilde{N}((0,t]\times B) = N((0,t]\times B) - \mu((0,t]\times B)$ is a square-integrable martingale.

Following Ikeda & Watanabe [22, Chapter II.4] we can generalize the set-up of Example 10.14 in the following way: Let $N(\omega, dt, dx)$ be for each ω a positive measure of space-time type. Since $t \mapsto N(\omega, (0,t] \times B)$ is increasing, there is a unique **compensator** $\widehat{N}(\omega, dt, dx)$ such that for all B with $\mathbb{E}\widehat{N}((0,t] \times B) < \infty$

$$\widetilde{N}(\omega, (0,t] \times B) := N(\omega, (0,t] \times B) - \widehat{N}(\omega, (0,t] \times B), \quad t \geqslant 0,$$

is a square-integrable martingale. If $t \mapsto \widehat{N}((0,t] \times B)$ is continuous and $B \mapsto \widehat{N}((0,t] \times B)$ a σ-finite measure, then one can show that the angle bracket satisfies

$$\left\langle \widetilde{N}((0,\cdot] \times B),\ \widetilde{N}((0,\cdot] \times C) \right\rangle_t = \widehat{N}\big((0,t] \times (B \cap C)\big).$$

This means, in particular, that $\widetilde{N}(\omega, dt, dx)$ is a random orthogonal measure with control measure $\mu((0,t] \times B) = \mathbb{E}\widehat{N}((0,t] \times B)$, and we are back in the theory which we have developed in the first part of this chapter.

It is possible to develop a fully-fledged stochastic calculus for this kind of random measures.

Definition 10.15. Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space with a filtration $(\mathscr{F}_t)_{t\geqslant 0}$. A **semimartingale** is a stochastic process X of the form

$$\begin{aligned} X_t = X_0 + A_t + M_t &+ \int_0^t\!\!\int f(\cdot, s, x)\, \widetilde{N}(ds, dx) \\ &+ \int_0^t\!\!\int g(\cdot, s, x)\, N(ds, dx) \end{aligned}$$

($\int_0^t := \int_{(0,t]}$) where

- X_0 is an $\mathscr{F}_0$ measurable random variable,
- M is a continuous square-integrable local martingale (w.r.t. $\mathscr{F}_t$),
- A is a continuous $\mathscr{F}_t$ adapted process of bounded variation,
- $N(ds, dx)$, $\widetilde{N}(ds, dx)$ and $\widehat{N}(ds, dx)$ are as described above,
- $f\mathbb{1}_{]\!]0,\tau_n]\!]} \in L^2(\Omega \times (0,\infty) \times X, \mathscr{P} \otimes \mathscr{X}, \mu^\circ)$ for some increasing sequence $\tau_n \uparrow \infty$ of bounded stopping times,
- g is such that $\int_0^t\!\int g(\omega, s, x)\, N(\omega, ds, dx)$ exists as an ω-wise integral,
- $f(\cdot, s, x) g(\cdot, s, x) \equiv 0$.

In this case, we even have **Itô's formula**, see [22, Chapter II.5], for any twice continuously differentiable $F \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$:

$$
\begin{aligned}
F(X_t) - F(X_0) = &\int_0^t F'(X_{s-})\, dA_s + \int_0^t F'(X_{s-})\, dM_s + \frac{1}{2}\int_0^t F''(X_{s-})\, d\langle M\rangle_s \\
&+ \int_0^t \int \big[F(X_{s-} + f(s,x)) - F(X_{s-})\big]\, \widetilde{N}(ds, dx) \\
&+ \int_0^t \int \big[F(X_{s-} + g(s,x)) - F(X_{s-})\big]\, N(ds, dx) \\
&+ \int_0^t \int \big[F(X_s + f(s,x)) - F(X_s) - f(s,x)F'(X_s)\big]\, \widehat{N}(ds, dx)
\end{aligned}
$$

where we use again the convention that $\int_0^t := \int_{(0,t]}$.

Chapter 11

From Lévy to Feller Processes

We have seen in Lemma 4.8 that the semigroup $P_t f(x) := \mathbb{E}^x f(X_t) = \mathbb{E} f(X_t + x)$ of a Lévy process $(X_t)_{t\geqslant 0}$ is a Feller semigroup. Moreover, the convolution structure of the semigroup $\mathbb{E} f(X_t + x) = \int f(x+y)\mathbb{P}(X_t \in dy)$ is a consequence of the spatial homogeneity (translation invariance) of the Lévy process, see Remark 4.5 and the characterization of translation invariant linear functionals (Theorem A.10). Lemma 4.4 shows that the translation invariance of a Lévy process is due to the assumptions (L1) and (L2).

It is, therefore, a natural question to ask what we get if we consider stochastic processes whose semigroups are Feller semigroups which are **not** translation invariant. Since every Feller semigroup admits a Markov transition kernel (Lemma 5.2), we can use Kolmogorov's construction to obtain a Markov process. Thus, the following definition makes sense.

Definition 11.1. A **Feller process** is a càdlàg Markov process $(X_t)_{t\geqslant 0}$, $X_t : \Omega \to \mathbb{R}^d$, $t \geqslant 0$, whose transition semigroup $P_t f(x) = \mathbb{E}^x f(X_t)$ is a Feller semigroup.

Remark 11.2. It is no restriction to require that a Feller process has càdlàg paths. By a fundamental result in the theory of stochastic processes we can construct such modifications. Usually, one argues like this: It is enough to study the coordinate processes, i.e., $d = 1$. Rather than looking at $t \mapsto X_t$ we consider a (countable, point-separating) family of functions $u : \mathbb{R} \to \mathbb{R}$ and show that each $t \mapsto u(X_t)$ has a càdlàg modification. One way of achieving this is to use martingale regularization techniques (e.g., Revuz & Yor [45, Chapter II.2]) which means that we should pick u in such a way that $u(X_t)$ is a supermartingale. The usual candidate for this is the resolvent $e^{-\lambda t} R_\lambda f(X_t)$ for some $f \in \mathcal{C}_\infty^+(\mathbb{R})$. Indeed, if $\mathscr{F}_t = \sigma(X_s,\, s \leqslant t)$ is the natural filtration, $f \geqslant 0$ and $s \leqslant t$, then

$$\begin{aligned}\mathbb{E}^x\big[R_\lambda f(X_t) \mid \mathscr{F}_s\big] &= \mathbb{E}^{X_s} \int_0^\infty e^{-\lambda r} P_r f(X_{t-s})\,dr = \int_0^\infty e^{-\lambda r} P_r P_{t-s} f(X_s)\,dr \\ &= e^{\lambda(t-s)} \int_{t-s}^\infty e^{-\lambda u} P_u f(X_s)\,du \leqslant e^{\lambda(t-s)} \int_0^\infty e^{-\lambda u} P_u f(X_s)\,du \\ &= e^{\lambda(t-s)} R_\lambda f(X_s).\end{aligned}$$

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_11

Let $\mathscr{F}_t = \mathscr{F}_t^X := \sigma(X_s,\, s \leqslant t)$ be the canonical filtration.

Lemma 11.3. *Every Feller process* $(X_t)_{t\geqslant 0}$ *is a* **strong Markov process**, *i.e.,*

$$\mathbb{E}^x\big[f(X_{t+\tau}) \mid \mathscr{F}_\tau\big] = \mathbb{E}^{X_\tau} f(X_t), \quad \mathbb{P}^x\text{-a.s. on } \{\tau < \infty\},\ t \geqslant 0, \tag{11.1}$$

holds for any stopping time τ, $\mathscr{F}_\tau := \{F \in \mathscr{F}_\infty \,:\, F \cap \{\tau \leqslant t\} \in \mathscr{F}_t\ \forall t \geqslant 0\}$ *and* $f \in \mathcal{C}_\infty(\mathbb{R}^d)$.

A routine approximation argument shows that (11.1) extends to $f(y) = \mathbb{1}_K(y)$ (where K is a compact set) and then, by a Dynkin-class argument, to any $f(y) = \mathbb{1}_B(y)$ where $B \in \mathscr{B}(\mathbb{R}^d)$.

Proof. In order to prove (11.1), we approximate τ from above by discrete stopping times $\tau_n = \big(\lfloor 2^n\tau\rfloor + 1\big)2^{-n}$ and observe that for $F \in \mathscr{F}_\tau \cap \{\tau < \infty\}$

$$\begin{aligned}
\mathbb{E}^x[\mathbb{1}_F f(X_{t+\tau})] &\overset{\text{(i)}}{=} \lim_{n\to\infty} \mathbb{E}^x[\mathbb{1}_F f(X_{t+\tau_n})] \\
&\overset{\text{(ii)}}{=} \lim_{n\to\infty} \mathbb{E}^x\big[\mathbb{1}_F \mathbb{E}^{X_{\tau_n}} f(X_t)\big] \\
&\overset{\text{(iii)}}{=} \mathbb{E}^x\big[\mathbb{1}_F \mathbb{E}^{X_\tau} f(X_t)\big].
\end{aligned}$$

Here we use that $t \mapsto X_t$ is right-continuous, plus (i) dominated convergence and (iii) the Feller continuity 4.7.f); (ii) is the strong Markov property for discrete stopping times which follows directly from the (ordinary) Markov property: Since $\{\tau_n < \infty\} = \{\tau < \infty\}$, we get

$$\begin{aligned}
\mathbb{E}^x[\mathbb{1}_F f(X_{t+\tau_n})] &= \sum_{k=1}^\infty \mathbb{E}^x\big[\mathbb{1}_{F\cap\{\tau_n = k2^{-n}\}} f(X_{t+k2^{-n}})\big] \\
&= \sum_{k=1}^\infty \mathbb{E}^x\big[\mathbb{1}_{F\cap\{\tau_n = k2^{-n}\}} \mathbb{E}^{X_{k2^{-n}}} f(X_t)\big] \\
&= \mathbb{E}^x\big[\mathbb{1}_F \mathbb{E}^{X_{\tau_n}} f(X_t)\big].
\end{aligned}$$

In the last calculation we use that $F \cap \{\tau_n = k2^{-n}\} \in \mathscr{F}_{k2^{-n}}$ for all $F \in \mathscr{F}_\tau$. □

Once we know the generator of a Feller process, we can construct many important martingales with respect to the canonical filtration of the process.

Corollary 11.4. *Let* $(X_t)_{t\geqslant 0}$ *be a Feller process with generator* $(A, \mathcal{D}(A))$ *and semi-group* $(P_t)_{t\geqslant 0}$. *For every* $f \in \mathcal{D}(A)$ *the process*

$$M_t^{[f]} := f(X_t) - \int_0^t Af(X_r)\,dr, \quad t \geqslant 0, \tag{11.2}$$

is a martingale for the canonical filtration $\mathcal{F}_t^X := \sigma(X_s,\ s \leqslant t)$ *and any* $\mathbb{P}^x$, $x \in \mathbb{R}^d$.

Proof. Let $s \leqslant t$, $f \in \mathcal{D}(A)$ and write, for short, $\mathcal{F}_s := \mathcal{F}_s^X$ and $M_t := M_t^{[f]}$. By the Markov property

$$\begin{aligned}\mathbb{E}^x\left[M_t - M_s \mid \mathcal{F}_s\right] &= \mathbb{E}^x\left[f(X_t) - f(X_s) - \int_s^t Af(X_r)\,dr \;\middle|\; \mathcal{F}_s\right] \\ &= \mathbb{E}^{X_s} f(X_{t-s}) - f(X_s) - \int_0^{t-s} \mathbb{E}^{X_s} Af(X_u)\,du.\end{aligned}$$

On the other hand, we get from the semigroup identity (5.5)

$$\begin{aligned}\int_0^{t-s} \mathbb{E}^{X_s} Af(X_u)\,du &= \int_0^{t-s} P_u Af(X_s)\,du \\ &= P_{t-s}f(X_s) - f(X_s) \\ &= \mathbb{E}^{X_s} f(X_{t-s}) - f(X_s)\end{aligned}$$

which shows that $\mathbb{E}^x\left[M_t - M_s \mid \mathcal{F}_s\right] = 0$. □

Our approach from Chapter 6 to prove the structure of a Lévy generator 'only' uses the positive maximum principle. Therefore, it can be adapted to Feller processes provided that the domain $\mathcal{D}(A)$ is rich in the sense that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$. All we have to do is to take into account that Feller processes are not any longer invariant under translations. The following theorem is due to Courrège [14] and von Waldenfels [61, 62].

Theorem 11.5 (von Waldenfels, Courrège). *Let $(A, \mathcal{D}(A))$ be the generator of a Feller process such that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$. Then $A|_{\mathcal{C}_c^\infty(\mathbb{R}^d)}$ is a* **pseudo-differential operator**

$$Au(x) = -q(x, D)u(x) := -\int q(x,\xi)\widehat{u}(\xi)\mathrm{e}^{\mathrm{i}\,x\cdot\xi}\,d\xi \tag{11.3}$$

whose **symbol** $q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ *is a measurable function of the form*

$$q(x,\xi) = \underbrace{q(x,0)}_{\geqslant 0} - \mathrm{i}\,l(x)\cdot\xi + \frac{1}{2}\xi\cdot Q(x)\xi + \int_{y\neq 0}\left[1 - \mathrm{e}^{\mathrm{i}\,y\cdot\xi} + \mathrm{i}\,y\cdot\xi\mathbb{1}_{(0,1)}(|y|)\right]\nu(x,dy) \tag{11.4}$$

and $(l(x), Q(x), \nu(x, dy))$ is a Lévy triplet[1] *for every fixed $x \in \mathbb{R}^d$.*

If we insert (11.4) into (11.3) and invert the Fourier transform we obtain the following **integro-differential representation** of the Feller generator A:

$$\begin{aligned}Af(x) &= l(x)\cdot\nabla f(x) + \frac{1}{2}\nabla\cdot Q(x)\nabla f(x) \\ &\quad + \int_{y\neq 0}\left[f(x+y) - f(x) - \nabla f(x)\cdot y\mathbb{1}_{(0,1)}(|y|)\right]\nu(x,dy).\end{aligned} \tag{11.5}$$

[1] Cf. Definition 6.10

This formula obviously extends to all functions $f \in \mathcal{C}_b^2(\mathbb{R}^d)$. In particular, we may use $f(x) = e_\xi(x) = e^{i\,x\cdot\xi}$, and get[2]

$$e_{-\xi}(x) A e_\xi(x) = -q(x,\xi). \tag{11.6}$$

Proof of Theorem 11.5 (*sketch*). *For a worked-out version cf.* [9, Chapter 2.3]. In the proof of Theorem 6.8 use, instead of A_0 and A_{00}

$$A_0 f \rightsquigarrow A_x f := (Af)(x) \quad \text{and} \quad A_{00} f \rightsquigarrow A_{xx} f := A_x(|\cdot - x|^2 f)$$

for every $x \in \mathbb{R}^d$. This is needed since P_t and A are not any longer translation invariant, i.e., we cannot shift $A_0 f$ to get $A_x f$. Then follow the steps 1°–4° to get $\nu(dy) \rightsquigarrow \nu(x,dy)$ and 6°–9° for $(l(x), Q(x))$. Remark 6.9 shows that the term $q(x,0)$ is non-negative.

The key observation is, as in the proof of Theorem 6.8, that we can use in steps 3° and 7° the positive maximum principle[3] to make sure that $A_x f$ is a distribution of order 2, i.e.,

$$|A_x f| = |L_x f + S_x f| \leqslant C_K \|f\|_{(2)} \text{ for all } f \in \mathcal{C}_c^\infty(K) \text{ and all compact sets } K \subset \mathbb{R}^d.$$

Here L_x is the local part with support in $\{x\}$ accounting for $(q(x,0), l(x), Q(x))$, and S_x is the non-local part supported in $\mathbb{R}^d \setminus \{x\}$ giving $\nu(x,dy)$. □

With some abstract functional analysis we can show some (local) boundedness properties of $x \mapsto Af(x)$ and $(x,\xi) \mapsto q(x,\xi)$.

Corollary 11.6. *In the situation of Theorem* 11.5*, the condition* (PP) *shows that*

$$\sup_{|x|\leqslant r} |Af(x)| \leqslant C_r \|f\|_{(2)} \quad \textit{for all} \;\; f \in \mathcal{C}_c^\infty(\overline{B_r(0)}) \tag{11.7}$$

and the positive maximum principle (PMP) *gives*

$$\sup_{|x|\leqslant r} |Af(x)| \leqslant C_{r,A} \|f\|_{(2)} \quad \textit{for all} \;\; f \in \mathcal{C}_c^\infty(\mathbb{R}^d),\; r > 0. \tag{11.8}$$

Proof. In the above sketched analogue of the proof of Theorem 6.8 we have seen that the family of linear functionals

$$\Big\{ \mathcal{C}_c^\infty(\overline{B_r(0)}) \ni f \mapsto A_x f \;:\; x \in \overline{B_r(0)} \Big\}$$

where $A_x f := (Af)(x)$ satisfies

$$|A_x f| \leqslant c_{r,x} \|f\|_{(2)}, \quad f \in \mathcal{C}_c^\infty(\overline{B_r(0)}),$$

[2] This should be compared with Definition 6.4 and the subsequent comments.
[3] To be precise: its weakened form (PP), cf. page 46.

i.e., $A_x : (\mathcal{C}_b^2(\overline{B_r(0)}), \|\cdot\|_{(2)}) \to (\mathbb{R}, |\cdot|)$ is bounded. By the Banach–Steinhaus theorem (uniform boundedness principle)

$$\sup_{|x|\leqslant r} |A_x f| \leqslant C_r \|f\|_{(2)}.$$

Since A also satisfies the positive maximum principle (PMP), we know from step 4° of the (suitably adapted) proof of Theorem 6.8 that

$$\int_{|y|>1} \nu(x, dy) \leqslant A\phi_0(x) \quad \text{for some} \quad \phi_0 \in \mathcal{C}_c(\overline{B_1(0)}).$$

Let $r > 1$, pick $\chi = \chi_r \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ such that $\mathbb{1}_{B_{2r}(0)} \leqslant \chi \leqslant \mathbb{1}_{B_{3r}(0)}$. We get for $|x| \leqslant r$

$$\begin{aligned} Af(x) &= A[\chi f](x) + A[(1-\chi)f](x) \\ &= A[\chi f](x) + \int_{|y|\geqslant r} \big[(1-\chi(x+y))f(x+y) - \underbrace{(1-\chi(x))f(x)}_{=0}\big]\, \nu(x, dy), \end{aligned}$$

and so

$$\sup_{|x|\leqslant r} |Af(x)| \leqslant C_r \|\chi f\|_{(2)} + \|f\|_\infty \|A\phi_0\|_\infty \leqslant C_{r,A} \|f\|_{(2)}. \qquad \square$$

Corollary 11.7. *In the situation of Theorem* 11.5 *there exists a locally bounded function* $\gamma : \mathbb{R}^d \to [0, \infty)$ *such that*

$$|q(x, \xi)| \leqslant \gamma(x)(1 + |\xi|^2), \quad x, \xi \in \mathbb{R}^d. \tag{11.9}$$

Proof. Using (11.8) we can extend A by continuity to $\mathcal{C}_b^2(\mathbb{R}^d)$ and, therefore,

$$-q(x, \xi) = \mathrm{e}_{-\xi}(x) A \mathrm{e}_\xi(x), \quad \mathrm{e}_\xi(x) = \mathrm{e}^{\mathrm{i}\, x\cdot\xi}$$

makes sense. Moreover, we have $\sup_{|x|\leqslant r} |A\mathrm{e}_\xi(x)| \leqslant C_{r,A} \|\mathrm{e}_\xi\|_{(2)}$ for any $r \geqslant 1$; since $\|\mathrm{e}_\xi\|_{(2)}$ is a polynomial of order 2 in the variable ξ, the claim follows. $\square$

For a Lévy process we have $\psi(0) = 0$ since $\mathbb{P}(X_t \in \mathbb{R}^d) = 1$ for all $t \geqslant 0$, i.e., the process X_t does not explode in finite time. For Feller processes the situation is more complicated. We need the following technical lemmas.

Lemma 11.8. *Let* $q(x, \xi)$ *be the symbol of* (*the generator of*) *a Feller process as in Theorem* 11.5 *and* $F \subset \mathbb{R}^d$ *be a closed set. Then the following assertions are equivalent.*

a) $|q(x, \xi)| \leqslant C(1 + |\xi|^2)$ *for all* $x, \xi \in \mathbb{R}^d$ *where* $C = 2 \sup_{|\xi|\leqslant 1} \sup_{x\in F} |q(x, \xi)|$.

b) $\sup_{x\in F} q(x, 0) + \sup_{x\in F} |l(x)| + \sup_{x\in F} \|Q(x)\| + \sup_{x\in F} \int_{y\neq 0} \frac{|y|^2}{1 + |y|^2}\, \nu(x, dy) < \infty.$

If $F = \mathbb{R}^d$, then the equivalent properties of Lemma 11.8 are often referred to as 'the symbol has **bounded coefficients**'.

Outline of the proof (*see* [57, *Appendix*] *for a complete proof*).
The direction b)⇒a) can be proved as Theorem 6.2. Observe that $\xi \mapsto q(x,\xi)$ is for fixed x the characteristic exponent of a Lévy process. The finiteness of the constant C follows from the assumption b) and the Lévy–Khintchine formula (11.4).

For the converse a)⇒b) we note that the integrand appearing in (11.4) can be estimated by $c\frac{|y|^2}{1+|y|^2}$ which is itself a Lévy exponent:

$$\frac{|y|^2}{1+|y|^2} = \int [1-\cos(y\cdot\xi)]\, g(\xi)\, d\xi, \quad g(\xi) = \frac{1}{2}\int_0^\infty (2\pi\lambda)^{-d/2} \mathrm{e}^{-|\xi|^2/2\lambda} \mathrm{e}^{-\lambda/2}\, d\lambda.$$

Therefore, by Tonelli's theorem,

$$\begin{aligned}\int_{y\neq 0} \frac{|y|^2}{1+|y|^2}\,\nu(x,dy) &= \iint_{y\neq 0} [1-\cos(y\cdot\xi)]\ \nu(x,dy)\, g(\xi)\, d\xi \\ &= \int g(\xi)\left(\operatorname{Re} q(x,\xi) - \frac{1}{2}\xi\cdot Q(x)\xi - q(x,0)\right) d\xi. \qquad \square\end{aligned}$$

Lemma 11.9. *Let A be the infinitesimal generator of a Feller process, assume that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ and denote by $q(x,\xi)$ the symbol of A. For any cut-off function $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying $\mathbb{1}_{B_1(0)} \leqslant \chi \leqslant \mathbb{1}_{B_2(0)}$ and $\chi_r(x) := \chi(x/r)$ one has*

$$|q(x,D)(\chi_r \mathrm{e}_\xi)(x)| \leqslant 4 \sup_{|\eta|\leqslant 1} |q(x,\eta)| \int_{\mathbb{R}^d} \left[1 + r^{-2}|\rho|^2 + |\xi|^2\right] \left|\widehat{\chi}(\rho)\right| d\rho, \tag{11.10}$$

$$\lim_{r\to\infty} A(\chi_r \mathrm{e}_\xi)(x) = -\mathrm{e}_\xi(x) q(x,\xi) \quad \textit{for all} \ \ x,\xi\in\mathbb{R}^d. \tag{11.11}$$

Proof. Observe that $\widehat{\chi_r \mathrm{e}_\xi}(\eta) = r^d \widehat{\chi}(r(\eta-\xi))$ and

$$\begin{aligned} q(x,D)(\chi_r\mathrm{e}_\xi)(x) &= \int q(x,\eta)\, \mathrm{e}^{\mathrm{i}\,x\cdot\eta}\, \widehat{\chi_r \mathrm{e}_\xi}(\eta)\, d\eta \\ &= \int q(x,\eta)\, \mathrm{e}^{\mathrm{i}\,x\cdot\eta}\, r^d\, \widehat{\chi}(r(\eta-\xi))\, d\eta \\ &= \int q(x,\xi+r^{-1}\rho)\, \mathrm{e}^{\mathrm{i}\,x\cdot(\xi+\rho/r)}\, \widehat{\chi}(\rho)\, d\rho. \end{aligned} \tag{11.12}$$

Therefore we can use the estimate (11.9) with the optimal constant $\gamma(x) = 2\sup_{|\eta|\leqslant 1}|q(x,\eta)|$ and the elementary estimate $(a+b)^2 \leqslant 2(a^2+b^2)$ to obtain

$$\begin{aligned} |q(x,D)(\chi_r\mathrm{e}_\xi)(x)| &\leqslant \int \left|q(x,\xi+r^{-1}\rho)\right|\left|\widehat{\chi}(\rho)\right| d\rho \\ &\leqslant 4 \sup_{|\eta|\leqslant 1} |q(x,\eta)| \int \left[1+r^{-2}|\rho|^2 + |\xi|^2\right] \left|\widehat{\chi}(\rho)\right| d\rho. \end{aligned}$$

This proves (11.10); it also allows us to use dominated convergence in (11.12) to get (11.11). Just observe that $\widehat{\chi} \in \mathcal{S}(\mathbb{R}^d)$ and $\int \widehat{\chi}(\rho)\,d\rho = \chi(0) = 1$. □

Lemma 11.10. *Let $q(x,\xi)$ be the symbol of (the generator of) a Feller process. Then the following assertions are equivalent:*

a) *$x \mapsto q(x,\xi)$ is continuous for all ξ.*

b) *$x \mapsto q(x,0)$ is continuous.*

c) *Tightness: $\lim\limits_{r\to\infty} \sup\limits_{x\in K} \nu(x, \mathbb{R}^d \setminus B_r(0)) = 0$ for all compact sets $K \subset \mathbb{R}^d$.*

d) *Uniform continuity at the origin: $\lim\limits_{|\xi|\to 0} \sup\limits_{x\in K} |q(x,\xi) - q(x,0)| = 0$ for all compact sets $K \subset \mathbb{R}^d$.*

Proof. Let $\chi : [0,\infty) \to [0,1]$ be a decreasing $\mathcal{C}^\infty$-function satisfying $\mathbb{1}_{[0,1)} \leqslant \chi \leqslant \mathbb{1}_{[0,4)}$. The functions $\chi_n(x) := \chi(|x|^2/n^2)$, $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, are radially symmetric, smooth functions with $\mathbb{1}_{B_n(0)} \leqslant \chi_n \leqslant \mathbb{1}_{B_{2n}(0)}$. Fix any compact set $K \subset \mathbb{R}^d$ and some sufficiently large n_0 such that $K \subset B_{n_0}(0)$.

a)⇒b) is obvious.

b)⇒c) For $m \geqslant n \geqslant 2n_0$ the positive maximum principle implies

$$\mathbb{1}_K(x) A(\chi_n - \chi_m)(x) \geqslant 0.$$

Therefore, $-\mathbb{1}_K(x) q(x,0) = \lim_{n\to\infty} \mathbb{1}_K(x) A\chi_{n+n_0}(x)$ is a decreasing limit of continuous functions. Since the limit function $q(x,0)$ is continuous, Dini's theorem implies that the limit is uniform on the set K. From the integro-differential representation (11.5) of the generator we get

$$\mathbb{1}_K(x)|A\chi_m(x) - A\chi_n(x)| = \mathbb{1}_K(x) \int\limits_{n-n_0 \leqslant |y| \leqslant 2m+n_0} \big(\chi_m(x+y) - \chi_n(x+y)\big)\,\nu(x,dy).$$

Letting $m \to \infty$ yields

$$\begin{aligned} \mathbb{1}_K(x)|q(x,0) + A\chi_n(x)| &= \mathbb{1}_K(x) \int_{|y|\geqslant n-n_0} \big(1 - \chi_n(x+y)\big)\,\nu(x,dy) \\ &\geqslant \mathbb{1}_K(x) \int_{|y|\geqslant n-n_0} \big(1 - \mathbb{1}_{B_{2n+n_0}(0)}(y)\big)\,\nu(x,dy) \\ &\geqslant \mathbb{1}_K(x) \nu\left(x, B^c_{2n+n_0}(0)\right). \end{aligned}$$

where we use that $K \subset B_{n_0}(0)$ and

$$\chi_n(x+y) \leqslant \mathbb{1}_{B_{2n}(0)}(x+y) = \mathbb{1}_{B_{2n}(0)-x}(y) \leqslant \mathbb{1}_{B_{2n+n_0}(0)}(y).$$

Since the left-hand side converges uniformly to 0 as $n \to \infty$, c) follows.

c)⇒d) Since the function $x \mapsto q(x,\xi)$ is locally bounded, we conclude from Lemma 11.8 that $\sup_{x\in K}|l(x)| + \sup_{x\in K}\|Q(x)\| < \infty$. Thus,

$$\lim_{|\xi|\to 0}\sup_{x\in K}\big(|l(x)\cdot\xi| + |\xi\cdot Q(x)\xi|\big) = 0,$$

and we may safely assume that $l \equiv 0$ and $Q \equiv 0$. If $|\xi| \leqslant 1$ we find, using (11.4) and Taylor's formula for the integrand,

$$\begin{aligned}
&|q(x,\xi) - q(x,0)|\\
&= \left|\int_{y\neq 0}\left[1 - \mathrm{e}^{\mathrm{i}\,y\cdot\xi} + \mathrm{i}\,y\cdot\xi\mathbb{1}_{(0,1)}(y)\right]\,\nu(x,dy)\right|\\
&\leqslant \int_{0<|y|^2<1}\frac{1}{2}|y|^2|\xi|^2\,\nu(x,dy) + \int_{1\leqslant|y|^2<1/|\xi|}|y||\xi|\,\nu(x,dy) + \int_{|y|^2\geqslant 1/|\xi|}2\,\nu(x,dy)\\
&\leqslant \int_{0<|y|^2<1/|\xi|}\frac{|y|^2}{1+|y|^2}\,\nu(x,dy)\big(1+|\xi|^{-1/2}\big)|\xi| + 2\nu\big(x,\{y\,:\,|y|^2\geqslant 1/|\xi|\}\big).
\end{aligned}$$

Since this estimate is uniform for $x \in K$, we get d) as $|\xi| \to 0$.

d)⇒c) As before, we may assume that $l \equiv 0$ and $Q \equiv 0$. For every $r > 0$

$$\begin{aligned}
\frac{1}{2}\nu(x, B_r^c(0)) &\leqslant \int_{|y|\geqslant r}\frac{|y/r|^2}{1+|y/r|^2}\,\nu(x,dy)\\
&= \int_{|y|\geqslant r}\int_{\mathbb{R}^d}\left(1-\cos\frac{\eta\cdot y}{r}\right)g(\eta)\,d\eta\,\nu(x,dy)\\
&\leqslant \int_{\mathbb{R}^d}\left[\operatorname{Re}q\big(x,\tfrac{\eta}{r}\big) - q(x,0)\right]\,g(\eta)\,d\eta,
\end{aligned}$$

where $g(\eta)$ is as in the proof of Lemma 11.8. Since $\int(1+|\eta|^2)g(\eta)\,d\eta < \infty$, we can use (11.9) and find

$$\nu(x, B_r^c(0)) \leqslant c_g \sup_{|\eta|\leqslant 1/r}\Big|\operatorname{Re}q(x,\eta) - q(x,0)\Big|.$$

Taking the supremum over all $x \in K$ and letting $r \to \infty$ proves c).

c)⇒a) From Lemma 11.9 we know that $\lim_{n\to\infty}\mathrm{e}_{-\xi}(x)A[\chi_n\mathrm{e}_\xi](x) = -q(x,\xi)$. Let us show that this convergence is uniform for $x \in K$. Let $m \geqslant n \geqslant 2n_0$. For $x \in K$

$$\begin{aligned}
&\big|\mathrm{e}_{-\xi}(x)A[\mathrm{e}_\xi\chi_n](x) - \mathrm{e}_{-\xi}(x)A[\mathrm{e}_\xi\chi_m](x)\big|\\
&\qquad= \left|\int_{y\neq 0}\left[\mathrm{e}_\xi(y)\chi_n(x+y) - \mathrm{e}_\xi(y)\chi_m(x+y)\right]\nu(x,dy)\right|\\
&\qquad\leqslant \int_{y\neq 0}\left[\chi_m(x+y) - \chi_n(x+y)\right]\nu(x,dy)\\
&\qquad\leqslant \int_{n-n_0\leqslant|y|\leqslant 2m+n_0}\left[\chi_m(x+y) - \chi_n(x+y)\right]\nu(x,dy)\\
&\qquad\leqslant \nu(x, B_{n-n_0}^c(0)).
\end{aligned}$$

In the penultimate step we use that, because of the definition of the functions χ_n,

$$\operatorname{supp}\big(\chi_m(x+\cdot) - \chi_n(x+\cdot)\big) \subset B_{2m}(x) \setminus B_n(x) \subset B_{2m+n_0}(0) \setminus B_{n-n_0}(0)$$

for all $x \in K \subset B_{n_0}(0)$. The right-hand side tends to 0 uniformly for $x \in K$ as $n \to \infty$, hence $m \to \infty$. $\square$

Remark 11.11. The argument used in the first three lines of the step b)$\Rightarrow$c) in the proof of Lemma 11.10 shows, incidentally, that

$$x \mapsto q(x,\xi) \quad \text{is always upper semicontinuous}$$

since it is (locally) a decreasing limit of continuous functions.

Remark 11.12. Let A be the infinitesimal generator of a Feller process, and assume that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$; although A maps $\mathcal{D}(A)$ into $\mathcal{C}_\infty(\mathbb{R}^d)$, this is not enough to guarantee that the symbol $q(x,\xi)$ is continuous in the variable x. On the other hand, if the Feller process X has only bounded jumps, i.e., if the support of the Lévy measure $\nu(x,\cdot)$ is uniformly bounded, then $q(\cdot,\xi)$ is continuous. This is, in particular, true for diffusions.

This follows immediately from Lemma 11.10.c) which holds if $\nu(x, B_r^c(0)) = 0$ for some $r > 0$ and all $x \in \mathbb{R}^d$.

We can also give a direct argument: pick $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying $\mathbb{1}_{B_{3r}(0)} \leqslant \chi \leqslant 1$. From the representation (11.5) it is not hard to see that

$$Af(x) = A[\chi f](x) \quad \text{for all} \quad f \in \mathcal{C}_b^2(\mathbb{R}^d) \quad \text{and} \quad x \in B_r(0);$$

in particular, $A[\chi f]$ is continuous.

If we take $f(x) := \mathrm{e}_\xi(x)$, then, by (11.6),

$$-q(x,\xi) = \mathrm{e}_{-\xi}(x) A\mathrm{e}_\xi(x) = \mathrm{e}_{-\xi}(x) A[\chi \mathrm{e}_\xi](x)$$

which proves that $x \mapsto q(x,\xi)$ is continuous on every ball $B_r(0)$, hence everywhere.

We can now discuss the role of $q(x,\xi)$ for the conservativeness of a Feller process.

Theorem 11.13. *Let $(X_t)_{t\geqslant 0}$ be a Feller process with an infinitesimal generator $(A, \mathcal{D}(A))$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$, symbol $q(x,\xi)$ and semigroup $(P_t)_{t\geqslant 0}$.*

a) *If $x \mapsto q(x,\xi)$ is continuous for all $\xi \in \mathbb{R}^d$ and $P_t 1 = 1$, then $q(x,0) = 0$.*
b) *If $q(x,\xi)$ has bounded coefficients and $q(x,0) = 0$, then $x \mapsto q(x,\xi)$ is continuous and $P_t 1 = 1$.*

Proof. Let χ and χ_r, $r \in \mathbb{N}$, be as in Lemma 11.9. Then $\mathrm{e}_\xi \chi_r \in \mathcal{D}(A)$ and, by Corollary 11.4,

$$M_t := \mathrm{e}_\xi \chi_r(X_t) - \mathrm{e}_\xi \chi_r(x) - \int_{[0,t)} A(\mathrm{e}_\xi \chi_r)(X_s)\,ds, \quad t \geqslant 0,$$

is a martingale. Using optional stopping for the stopping time

$$\tau := \tau_R^x := \inf\{s \geqslant 0 \,:\, |X_s - x| \geqslant R\}, \quad x \in \mathbb{R}^d,\ R > 0,$$

the stopped process $(M_{t\wedge\tau})_{t\geqslant 0}$ is still a martingale. Since $\mathbb{E}^x M_{t\wedge\tau} = 0$, we get

$$\mathbb{E}^x(\chi_r \mathrm{e}_\xi)(X_{t\wedge\tau}) - \chi_r \mathrm{e}_\xi(x) = \mathbb{E}^x \int_{[0,t\wedge\tau)} A(\chi_r \mathrm{e}_\xi)(X_s)\,ds.$$

Observe that the integrand is evaluated only for times $s < t \wedge \tau$ where $|X_s| \leqslant R + x$. Since $A(\mathrm{e}_\xi \chi_r)(x)$ is locally bounded, we can use dominated convergence and Lemma 11.9 and we find, as $r \to \infty$,

$$\mathbb{E}^x \mathrm{e}_\xi(X_{t\wedge\tau}) - \mathrm{e}_\xi(x) = -\mathbb{E}^x \int_{[0,t\wedge\tau)} \mathrm{e}_\xi(X_s) q(X_s, \xi)\,ds.$$

a) Set $\xi = 0$ and observe that $P_t 1 = 1$ implies that $\tau = \tau_R^x \to \infty$ a.s. as $R \to \infty$. Therefore,

$$\mathbb{P}^x(X_{t\wedge\tau} \in \mathbb{R}^d) - 1 = -\mathbb{E}^x \int_{[0,\tau\wedge t)} q(X_s, 0)\,ds,$$

and with Fatou's Lemma we can let $R \to \infty$ to get

$$\begin{aligned} 0 &= \liminf_{R\to\infty} \mathbb{E}^x \int_{[0,\tau\wedge t)} q(X_s, 0)\,ds \\ &\geqslant \mathbb{E}^x \left[\liminf_{R\to\infty} \int_{[0,\tau\wedge t)} q(X_s, 0)\,ds \right] = \mathbb{E}^x \int_0^t q(X_s, 0)\,ds. \end{aligned}$$

Since $x \mapsto q(x, 0)$ is continuous and $q(x, 0)$ non-negative, we conclude with Tonelli's theorem that

$$q(x, 0) = \lim_{t\to 0} \frac{1}{t} \int_0^t \mathbb{E}^x q(X_s, 0)\,ds = 0.$$

b) Set $\xi = 0$ and observe that the boundedness of the coefficients implies that

$$\mathbb{P}^x(X_t \in \mathbb{R}^d) - 1 = -\mathbb{E}^x \int_{[0,t)} q(X_s, 0)\,ds$$

as $R \to \infty$. Since the right-hand side is 0, we get $P_t 1 = \mathbb{P}^x(X_t \in \mathbb{R}^d) = 1$. □

Remark 11.14. The boundedness of the coefficients in Theorem 11.13.b) is important. If the coefficients of $q(x, \xi)$ grow too rapidly, we may observe explosion in finite time even if $q(x, 0) = 0$. A typical example in dimension 3 is the diffusion process given by the generator

$$Lf(x) = \frac{1}{2} a(x) \Delta f(x)$$

where $a(x)$ is continuous, rotationally symmetric $a(x) = \alpha(|x|)$ for a suitable function $\alpha(r)$, and satisfies $\int_1^\infty 1/\alpha(\sqrt{r})\,dr < \infty$, see Stroock & Varadhan [60, p. 260, 10.3.3]; the corresponding symbol is $q(x,\xi) = \frac{1}{2}a(x)|\xi|^2$. This process explodes in finite time. Since this is essentially a time-changed Brownian motion (see Böttcher, Schilling & Wang [9, Chapter 4.1]), this example works only if Brownian motion is transient, i.e., in dimensions $d = 3$ and higher. A sufficient criterion for conservativeness in terms of the symbol is

$$\liminf_{r\to\infty} \sup_{|y-x|\leqslant 2r} \sup_{|\eta|\leqslant 1/r} |q(y,\eta)| < \infty \quad \text{for all} \quad x \in \mathbb{R}^d,$$

see [9, Theorem 2.34].

Chapter 12

Symbols and Semimartingales

So far, we have been treating the symbol $q(x,\xi)$ of (the generator of) a Feller process X as an analytic object. On the other hand, Theorem 11.13 indicates, that there should be some probabilistic consequences. In this chapter we want to follow this lead, show a probabilistic method to calculate the symbol and link it to the semimartingale characteristics of a Feller process. The blueprint for this is the relation of the Lévy–Itô decomposition (which is the semimartingale decomposition of a Lévy process, cf. Theorem 9.12) with the Lévy–Khintchine formula for the characteristic exponent (which coincides with the symbol of a Lévy process, cf. Corollary 9.13).

For a Lévy process X_t with semigroup $P_t f(x) = \mathbb{E}^x f(X_t) = \mathbb{E} f(X_t + x)$ the symbol can be calculated in the following way:

$$\lim_{t\to 0} \frac{\mathrm{e}_{-\xi}(x) P_t \mathrm{e}_\xi(x) - 1}{t} = \lim_{t\to 0} \frac{\mathbb{E}^x \mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t - x)} - 1}{t} = \lim_{t\to 0} \frac{\mathrm{e}^{-t\psi(\xi)} - 1}{t} = -\psi(\xi). \tag{12.1}$$

For a Feller process a similar formula is true.

Theorem 12.1. *Let $X = (X_t)_{t\geqslant 0}$ be a Feller process with transition semigroup $(P_t)_{t\geqslant 0}$ and generator $(A, \mathcal{D}(A))$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$. If $x \mapsto q(x,\xi)$ is continuous and q has bounded coefficients (Lemma 11.8 with $F = \mathbb{R}^d$), then*

$$-q(x,\xi) = \lim_{t\to 0} \frac{\mathbb{E}^x \mathrm{e}^{\mathrm{i}\,\xi\cdot(X_t - x)} - 1}{t}. \tag{12.2}$$

Proof. Pick $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, $\mathbb{1}_{B_1(0)} \leqslant \chi \leqslant \mathbb{1}_{B_2(0)}$ and set $\chi_n(x) := \chi\left(\frac{x}{n}\right)$. Obviously, $\chi_n \to 1$ as $n \to \infty$. By Lemma 5.4,

$$\begin{aligned} P_t[\chi_n \mathrm{e}_\xi](x) - \chi_n(x)\mathrm{e}_\xi(x) &= \int_0^t A P_s[\chi_n \mathrm{e}_\xi](x)\,ds \\ &= \mathbb{E}^x \int_0^t A[\chi_n \mathrm{e}_\xi](X_s)\,ds \end{aligned}$$

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_12

$$= -\int_0^t \int_{\mathbb{R}^d} \mathbb{E}^x \Big(e_\eta(X_s) q(X_s, \eta) \underbrace{\widehat{\chi_n e_\xi}(\eta)}_{=n^d \widehat{\chi}(n(\eta-\xi))} \Big) \, d\eta \, ds$$

$$\xrightarrow[n\to\infty]{} -\int_0^t P_s\big[e_\xi q(\cdot, \xi)\big](x) \, ds.$$

Since $x \mapsto q(x, \xi)$ is continuous, we can divide by t and let $t \to 0$; this yields

$$\lim_{t\to 0} \frac{1}{t} \left(P_t e_\xi(x) - e_\xi(x) \right) = -e_\xi(x) q(x, \xi). \qquad \square$$

Theorem 12.1 is a relatively simple probabilistic formula to calculate the symbol. We want to relax the boundedness and continuity assumptions. Here Dynkin's characteristic operator becomes useful.

Lemma 12.2 (Dynkin's formula). *Let $(X_t)_{t\geqslant 0}$ be a Feller process with semigroup $(P_t)_{t\geqslant 0}$ and generator $(A, \mathcal{D}(A))$. For every stopping time σ with $\mathbb{E}^x\sigma < \infty$ one has*

$$\mathbb{E}^x f(X_\sigma) - f(x) = \mathbb{E}^x \int_{[0,\sigma)} Af(X_s) \, ds, \quad f \in \mathcal{D}(A). \tag{12.3}$$

Proof. From Corollary 11.4 we know that $M_t^{[f]} := f(X_t) - f(X_0) - \int_0^t Af(X_s)\, ds$ is a martingale; thus (12.3) follows from the optional stopping theorem. $\square$

Definition 12.3. Let $(X_t)_{t\geqslant 0}$ be a Feller process. A point $a \in \mathbb{R}^d$ is an **absorbing point**, if

$$\mathbb{P}^a(X_t = a, \ \forall t \geqslant 0) = 1.$$

Denote by $\tau_r := \inf\{t > 0 \,:\, |X_t - X_0| \geqslant r\}$ the first exit time from the ball $B_r(x)$ centered at the starting position $x = X_0$.

Lemma 12.4. *Let $(X_t)_{t\geqslant 0}$ be a Feller process and assume that $b \in \mathbb{R}^d$ is not absorbing. Then there exists some $r > 0$ such that $\mathbb{E}^b \tau_r < \infty$.*

Proof. 1° If b is not absorbing, then there is some $f \in \mathcal{D}(A)$ such that $Af(b) \neq 0$. Assume the contrary, i.e.,

$$Af(b) = 0 \quad \text{for all} \quad f \in \mathcal{D}(A).$$

By Lemma 5.4, $P_s f \in \mathcal{D}(A)$ for all $s \geqslant 0$, and

$$P_t f(b) - f(b) = \int_0^t A(P_s f)(b) \, ds = 0.$$

So, $P_t f(b) = f(b)$ for all $f \in \mathcal{D}(A)$. Since $\mathcal{D}(A)$ is dense in $\mathcal{C}_\infty(\mathbb{R}^d)$ (Remark 5.5), we get $P_t f(b) = f(b)$ for all $f \in \mathcal{C}_\infty(\mathbb{R}^d)$, hence $\mathbb{P}^b(X_t = b) = 1$ for any $t \geqslant 0$. Therefore,

$$\mathbb{P}^b(X_q = b, \ \forall q \in \mathbb{Q}, \ q \geqslant 0) = 1 \quad \text{and} \quad \mathbb{P}^b(X_t = b, \ \forall t \geqslant 0) = 1,$$

because of the right-continuity of the sample paths. This means that b is an absorbing point, contradicting our assumption.

2° Pick $f \in \mathcal{D}(A)$ such that $Af(b) > 0$. Since Af is continuous, there exist $\epsilon > 0$ and $r > 0$ such that $Af|_{B_r(b)} \geqslant \epsilon > 0$. From Dynkin's formula (12.3) with $\sigma = \tau_r \wedge n$, $n \geqslant 1$, we deduce

$$\epsilon \mathbb{E}^b(\tau_r \wedge n) \leqslant \mathbb{E}^b \int_{[0,\tau_r\wedge n)} Af(X_s)\,ds = \mathbb{E}^b f(X_{\tau_r \wedge n}) - f(b) \leqslant 2\|f\|_\infty.$$

Finally, monotone convergence shows that $\mathbb{E}^b\tau_r \leqslant 2\|f\|_\infty/\epsilon < \infty$. □

Definition 12.5 (Dynkin's operator)**.** Let $(X_t)_{t\geqslant 0}$ be a Feller process. The linear operator $(\mathfrak{A}, \mathcal{D}(\mathfrak{A}))$ defined by

$$\mathfrak{A}f(x) := \begin{cases} \lim\limits_{r\to 0} \dfrac{\mathbb{E}^x f(X_{\tau_r}) - f(x)}{\mathbb{E}^x\tau_r}, & \text{if } x \text{ is not absorbing,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\mathcal{D}(\mathfrak{A}) := \left\{ f \in \mathcal{C}_\infty(\mathbb{R}^d) \;:\; \text{the above limit exists pointwise} \right\},$$

is called **Dynkin's (characteristic) operator**.

Lemma 12.6. *Let $(X_t)_{t\geqslant 0}$ be a Feller process with generator $(A, \mathcal{D}(A))$ and characteristic operator $(\mathfrak{A}, \mathcal{D}(\mathfrak{A}))$.*

a) *$\mathfrak{A}$ is an extension of A, i.e., $\mathcal{D}(A) \subset \mathcal{D}(\mathfrak{A})$ and $\mathfrak{A}|_{\mathcal{D}(A)} = A$.*
b) *$(\mathfrak{A}, \mathcal{D}) = (A, \mathcal{D}(A))$ if $\mathcal{D} = \{f \in \mathcal{D}(\mathfrak{A}) : f, \mathfrak{A}f \in \mathcal{C}_\infty(\mathbb{R}^d)\}$.*

Proof. a) Let $f \in \mathcal{D}(A)$ and assume that $x \in \mathbb{R}^d$ is not absorbing. By Lemma 12.4 there is some $r = r(x) > 0$ with $\mathbb{E}^x\tau_r < \infty$. Since we have $Af \in \mathcal{C}_\infty(\mathbb{R}^d)$, there exists for every $\epsilon > 0$ some $\delta > 0$ such that

$$|Af(y) - Af(x)| < \epsilon \quad \text{for all} \;\; y \in B_\delta(x).$$

Without loss of generality let $\delta < r$. Using Dynkin's formula (12.3) with $\sigma = \tau_\delta$, we see

$$\left|\mathbb{E}^x f(X_{\tau_\delta}) - f(x) - Af(x)\mathbb{E}^x\tau_\delta\right| \leqslant \mathbb{E}^x \int_{[0,\tau_\delta)} \underbrace{|Af(X_s) - Af(x)|}_{\leqslant\epsilon}\,ds \leqslant \epsilon\mathbb{E}^x\tau_\delta.$$

Thus, $\lim_{r\to 0}\left(\mathbb{E}^x f(X_{\tau_r}) - f(x)\right)/\mathbb{E}^x\tau_r = Af(x)$.

If x is absorbing and $f \in \mathcal{D}(A)$, then $Af(x) = 0$ and so $Af(x) = \mathfrak{A}f(x)$.

b) Since $(\mathfrak{A}, \mathcal{D})$ satisfies the positive maximum principle (PMP), the claim follows from Lemma 5.11. □

Theorem 12.7. *Let $(X_t)_{t\geqslant 0}$ be a Feller process with an infinitesimal generator $(A,\mathcal{D}(A))$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d)\subset\mathcal{D}(A)$ and $x\mapsto q(x,\xi)$ is continuous*[1]*. Then*

$$-q(x,\xi)=\lim_{r\to 0}\frac{\mathbb{E}^x e^{i(X_{\tau_r}-x)\cdot\xi}-1}{\mathbb{E}^x\tau_r} \tag{12.4}$$

for all $x\in\mathbb{R}^d$ (as usual, $\frac{1}{\infty}:=0$). In particular, $q(a,\xi)=0$ for all absorbing states $a\in\mathbb{R}^d$.

Proof. Let $\chi_n\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ such that $\mathbb{1}_{B_n(0)}\leqslant\chi_n\leqslant 1$. By Dynkin's formula (12.3)

$$e_{-\xi}(x)\mathbb{E}^x[\chi_n(X_{\tau_r\wedge t})e_\xi(X_{\tau_r\wedge t})]-\chi_n(x)=\mathbb{E}^x\int_{[0,\tau_r\wedge t)}e_{-\xi}(x)A[\chi_n e_\xi](X_s)\,ds.$$

Observe that $A[\chi_n e_\xi](X_s)$ is bounded if $s<\tau_r$, see Corollary 11.6. Using the dominated convergence theorem, we can let $n\to\infty$ to get

$$\begin{aligned} e_{-\xi}(x)\mathbb{E}^x e_\xi(X_{\tau_r\wedge t})-1 &= \mathbb{E}^x\int_{[0,\tau_r\wedge t)}e_{-\xi}(x)Ae_\xi(X_s)\,ds\\ &\overset{(11.6)}{=} -\mathbb{E}^x\int_{[0,\tau_r\wedge t)}e_\xi(X_s-x)q(X_s,\xi)\,ds. \end{aligned} \tag{12.5}$$

If x is absorbing, we have $q(x,\xi)=0$, and (12.4) holds trivially. For non-absorbing x, we pass to the limit $t\to\infty$ and get, using $\mathbb{E}^x\tau_r<\infty$ (see Lemma 12.4),

$$\frac{e_{-\xi}(x)\mathbb{E}^x e_\xi(X_{\tau_r})-1}{\mathbb{E}^x\tau_r}=-\frac{1}{\mathbb{E}^x\tau_r}\mathbb{E}^x\int_{[0,\tau_r)}e_\xi(X_s-x)q(X_s,\xi)\,ds.$$

Since $s\mapsto q(X_s,\xi)$ is right-continuous at $s=0$, the limit $r\to 0$ exists, and (12.4) follows. □

A small variation of the above proof yields

Corollary 12.8 (Schilling, Schnurr [57])**.** *Let $(X_t)_{t\geqslant 0}$ be a Feller process with generator $(A,\mathcal{D}(A))$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d)\subset\mathcal{D}(A)$ and $x\mapsto q(x,\xi)$ is continuous. Then*

$$-q(x,\xi)=\lim_{t\to 0}\frac{\mathbb{E}^x e^{i(X_{t\wedge\tau_r}-x)\cdot\xi}-1}{t} \tag{12.6}$$

for all $x\in\mathbb{R}^d$ and $r>0$.

Proof. We follow the proof of Theorem 12.7 up to (12.5). This relation can be rewritten as

$$\frac{\mathbb{E}^x e^{i(X_{t\wedge\tau_r}-x)\cdot\xi}-1}{t}=-\frac{1}{t}\mathbb{E}^x\int_0^t e_\xi(X_s-x)q(X_s,\xi)\mathbb{1}_{[0,\tau_r)}(s)\,ds.$$

Observe that X_s is bounded if $s<\tau_r$ and that $s\mapsto q(X_s,\xi)$ is right-continuous. Therefore, the limit $t\to 0$ exists and yields (12.6). □

[1] For instance, if X has bounded jumps, see Lemma 11.10 and Remark 11.12. Our proof will show that it is actually enough to assume that $s\mapsto q(X_s,\xi)$ is right-continuous.

Every Feller process $(X_t)_{t\geqslant 0}$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ is a semimartingale. Moreover, the semimartingale characteristics can be expressed in terms of the Lévy triplet $(l(x), Q(x), \nu(x, dy))$ of the symbol. Recall that a (d-dimensional) **semimartingale** is a stochastic process of the form

$$\begin{aligned} X_t = X_0 + X_t^{\mathsf{c}} + \int_0^t \int y \mathbb{1}_{(0,1)}(|y|) \big[\mu^X(\cdot, ds, dy) - \nu(\cdot, ds, dy)\big] \\ + \sum_{s\leqslant t} \mathbb{1}_{[1,\infty)}(|\Delta X_s|) \Delta X_s + B_t \end{aligned}$$

where X^{c} is the continuous martingale part, B is a previsible process with paths of finite variation (on compact time intervals) and with the jump measure

$$\mu^X(\omega, ds, dy) = \sum_{s\,:\,\Delta X_s(\omega)\neq 0} \delta_{(s, \Delta X_s(\omega))}(ds, dy)$$

whose compensator is $\nu(\omega, ds, dy)$. The triplet (B, C, ν) with the (predictable) quadratic variation $C = [X^{\mathsf{c}}, X^{\mathsf{c}}]$ of X^{c} is called the **semimartingale characteristics**.

Theorem 12.9 (Schilling [53], Schnurr [59])**.** *Let $(X_t)_{t\geqslant 0}$ be a Feller process with infinitesimal generator $(A, \mathcal{D}(A))$ such that $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ and symbol $q(x, \xi)$ given by* (11.4)*. If $q(x, 0) = 0$,*[2] *then X is a semimartingale whose semimartingale characteristics can be expressed by the Lévy triplet $(l(x), Q(x), \nu(x, dy))$*

$$B_t = \int_0^t l(X_s)\,ds, \quad C_t = \int_0^t Q(X_s)\,ds, \quad \nu(\cdot, ds, dy) = \nu(X_s(\cdot), dy)\,ds.$$

Proof. 1° Combining Corollary 11.4 with the integro-differential representation (11.5) of the generator shows that

$$\begin{aligned} M_t^{[f]} &= f(X_t) - \int_{[0,t)} Af(X_s)\,ds \\ &= f(X_t) - \int_{[0,t)} l(X_s)\cdot\nabla f(X_s)\,ds - \frac{1}{2}\int_{[0,t)} \nabla\cdot Q(X_s)\nabla f(X_s)\,ds \\ &\quad - \int_{[0,t)}\int_{y\neq 0} \big[f(X_s + y) - f(X_s) - \nabla f(X_s)\cdot y\mathbb{1}_{(0,1)}(|y|)\big]\,\nu(X_s, dy)\,ds \end{aligned}$$

is for all $f \in \mathcal{C}_b^2(\mathbb{R}^d) \cap \mathcal{D}(A)$ a martingale.

2° We claim that $\mathcal{C}_c^2(\mathbb{R}^d) \subset \mathcal{D}(A)$. Indeed, let $f \in \mathcal{C}_c^2(\mathbb{R}^d)$ with $\operatorname{supp} f \subset B_r(0)$ for some $r > 0$ and pick a sequence $f_n \in \mathcal{C}_c^\infty(B_{2r}(0))$ such that $\lim_{n\to\infty} \|f - f_n\|_{(2)} = 0$. Using (11.7) we get

$$\sup_{|x|\leqslant 3r} |Af_n(x) - Af_m(x)| \leqslant c_{3r}\|f_n - f_m\|_{(2)} \xrightarrow[m,n\to\infty]{} 0.$$

[2]A sufficient condition is, for example, that X has infinite life-time and $x \mapsto q(x, \xi)$ is continuous (either for all ξ or just for $\xi = 0$), cf. Theorem 11.13 and Lemma 11.10.

Since $\operatorname{supp} f_n \subset B_{2r}(0)$ and $f_n \to f$ uniformly, there is some $u \in \mathcal{C}_c^\infty(B_{3r}(0))$ such that $|f_n(x)| \leqslant u(x)$. Therefore, we get for $|x| \geqslant 2r$

$$\begin{aligned} |Af_n(x) - Af_m(x)| &\leqslant \int_{y\neq 0} |f_n(x+y) - f_m(x+y)|\, \nu(x,dy) \\ &\leqslant 2\int_{y\neq 0} u(x+y)\,\nu(x,dy) = 2Au(x) \xrightarrow[|x|\to\infty]{} 0 \end{aligned}$$

uniformly for all m, n. This shows that $(Af_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{C}_\infty(\mathbb{R}^d)$. By the closedness of $(A, \mathcal{D}(A))$ we gather that $f \in \mathcal{D}(A)$ and $Af = \lim_{n\to\infty} Af_n$.

3° Fix $r > 1$, pick $\chi_r \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ such that $\mathbb{1}_{B_{3r}(0)} \leqslant \chi \leqslant 1$, and set

$$\sigma = \sigma_r := \inf\{t > 0 \,:\, |X_t - X_0| \geqslant r\} \wedge \inf\{t > 0 \,:\, |\Delta X_t| \geqslant r\}.$$

Since $(X_t)_{t\geqslant 0}$ has càdlàg paths and infinite life-time, σ_r is a family of stopping times with $\sigma_r \uparrow \infty$. For any $f \in \mathcal{C}^2(\mathbb{R}^d) \cap \mathcal{C}_b(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ we set $f_r := \chi_r f$, $f^x := f(\cdot - x)$, $f_r^x(\cdot - x)$, and consider

$$\begin{aligned} M_{t\wedge\sigma}^{[f^x]} &= f^x(X_{t\wedge\sigma}) - \int_{[0,t\wedge\sigma)} l(X_s)\cdot\nabla f^x(X_s)\,ds - \frac{1}{2}\int_{[0,t\wedge\sigma)} \nabla\cdot Q(X_s)\nabla f^x(X_s)\,ds \\ &\quad - \int_{[0,t\wedge\sigma)}\int_{y\neq 0} \left[f^x(X_s+y) - f^x(X_s) - \nabla f^x(X_s)\cdot y\mathbb{1}_{(0,1)}(|y|)\right]\nu(X_s,dy)\,ds \\ &= f_r^x(X_{t\wedge\sigma}) - \int_{[0,t\wedge\sigma)} l(X_s)\cdot\nabla f_r^x(X_s)\,ds - \frac{1}{2}\int_{[0,t\wedge\sigma)} \nabla\cdot Q(X_s)\nabla f_r^x(X_s)\,ds \\ &\quad - \int_{[0,t\wedge\sigma)}\int_{0<|y|<r} \left[f_r^x(X_s+y) - f_r^x(X_s) - \nabla f_r^x(X_s)\cdot y\mathbb{1}_{(0,1)}(|y|)\right]\nu(X_s,dy)\,ds \\ &= M_{t\wedge\sigma}^{[f_r^x]}. \end{aligned}$$

Since $f_r \in \mathcal{C}_c^2(\mathbb{R}^d) \subset \mathcal{D}(A)$, we see that $M_t^{[f^x]}$ is a local martingale (with reducing sequence σ_r, $r > 0$), and by a general result of Jacod & Shiryaev [27, Theorem II.2.42], it follows that X is a semimartingale with the characteristics mentioned in the theorem. □

We close this chapter with the discussion of a very important example: Lévy driven stochastic differential equations. From now on we assume that

$$\begin{aligned} \Phi : \mathbb{R}^d \to \mathbb{R}^{d\times n} \quad &\text{is a matrix-valued, measurable function,} \\ L = (L_t)_{t\geqslant 0} \quad &\text{is an } n\text{-dimensional Lévy process with exponent } \psi(\xi), \end{aligned}$$

and consider the following Itô stochastic differential equation (SDE)

$$dX_t = \Phi(X_{t-})\,dL_t, \quad X_0 = x. \tag{12.7}$$

If Φ is globally Lipschitz continuous, then the SDE (12.7) has a unique solution which is a strong Markov process, see Protter [44, Chapter V, Theorem 32][3]. If we write X_t^x for the solution of (12.7) with initial condition $X_0 = x = X_0^x$, then the flow $x \mapsto X_t^x$ is continuous [44, Chapter V, Theorem 38].

If we use $L_t = (t, W_t, J_t)^\top$ as driving Lévy process where W is a Brownian motion and J is a pure-jump Lévy process (we assume[4] that $W \perp J$), and if Φ is a block-matrix, then we see that (12.7) covers also SDEs of the form

$$dX_t = f(X_{t-})\,dt + F(X_{t-})\,dW_t + G(X_{t-})\,dJ_t.$$

Lemma 12.10. *Let Φ be bounded and Lipschitz, X the unique solution of the SDE* (12.7)*, and denote by A the generator of X. Then $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$.*

Proof. Because of Theorem 5.12 (this theorem does not only hold for Feller semigroups, but for any strongly continuous contraction semigroup satisfying the positive maximum principle), it suffices to show

$$\lim_{t\to 0} \frac{1}{t}\big(\mathbb{E}^x f(X_t) - f(x)\big) = g(x) \quad \text{and} \quad g \in \mathcal{C}_\infty(\mathbb{R}^d)$$

for any $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. For this we use, as in the proof of the following theorem, Itô's formula to get

$$\mathbb{E}^x f(X_t) - f(x) = \mathbb{E}^x \int_0^t Af(X_s)\,ds,$$

and a calculation similar to the one in the proof of the next theorem. A complete proof can be found in Schilling & Schnurr [57, Theorem 3.5]. □

Theorem 12.11. *Let $(L_t)_{t\geqslant 0}$ be an n-dimensional Lévy process with exponent ψ and assume that Φ is Lipschitz continuous. Then the unique Markov solution X of the SDE* (12.7) *admits a generalized symbol in the sense that*

$$\lim_{t\to 0} \frac{\mathbb{E}^x e^{i(X_{t\wedge\tau_r} - x)\cdot\xi} - 1}{t} = -\psi(\Phi^\top(x)\xi), \quad r > 0.$$

If $\Phi \in \mathcal{C}_b^1(\mathbb{R}^d)$, then X is Feller, $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ and $q(x,\xi) = \psi(\Phi^\top(x)\xi)$ is the symbol of the generator.

Theorem 12.11 indicates that the formulae (12.4) and (12.6) may be used to associate symbols not only with Feller processes but with more general Markov semimartingales. This has been investigated by Schnurr [58] and [59] who shows

[3] Protter requires that L has independent coordinate processes, but this is not needed in his proof. For existence and uniqueness the local Lipschitz and linear growth conditions are enough; the strong Lipschitz condition is used for the Markov nature of the solution.

[4] This assumption can be relaxed under suitable assumptions on the (joint) filtration, see for example Ikeda & Watanabe [22, Theorem II.6.3]

that the class of Itô processes with jumps is essentially the largest class that can be described by symbols; see also [9, Chapters 2.4–5]. This opens up the way to analyze rather general semimartingales using the symbol. Let us also point out that the boundedness of Φ is only needed to ensure that X is a Feller process.

Proof. Let $\tau = \tau_r$ be the first exit time for the process X from the ball $B_r(x)$ centered at $x = X_0$. We use the Lévy–Itô decomposition (9.8) of L. From Itô's formula for jump processes (see, e.g., Protter [44, Chapter II, Theorem 33]) we get

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}^x\left(\mathrm{e}_\xi(X_{t\wedge\tau}-x)-1\right)\\
&\quad=\frac{1}{t}\mathbb{E}^x\bigg[\int_0^{t\wedge\tau}\mathrm{i}\,\mathrm{e}_\xi(X_{s-}-x)\xi\,dX_s-\frac{1}{2}\int_0^{t\wedge\tau}\mathrm{e}_\xi(X_{s-}-x)\xi\cdot d[X,X]_s^{\mathrm{c}}\xi\\
&\qquad\qquad+\mathrm{e}_{-\xi}(x)\sum_{s\leqslant\tau\wedge t}\big(\mathrm{e}_\xi(X_s)-\mathrm{e}_\xi(X_{s-})-\mathrm{i}\,\mathrm{e}_\xi(X_{s-})\xi\cdot\Delta X_s\big)\bigg]\\
&\quad=:I_1+I_2+I_3.
\end{aligned}$$

We consider the three terms separately.

$$\begin{aligned}
I_1&=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{i}\,\mathrm{e}_\xi(X_{s-}-x)\xi\,dX_s\\
&=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{i}\,\mathrm{e}_\xi(X_{s-}-x)\xi\cdot\Phi(X_{s-})\,dL_s\\
&=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{i}\,\mathrm{e}_\xi(X_{s-}-x)\xi\cdot\Phi(X_{s-})\,d_s\left[ls+\int_{|y|\geqslant 1}y\,N_s(dy)\right]\\
&=:I_{11}+I_{12}
\end{aligned}$$

where we use that the diffusion part and the compensated small jumps of a Lévy process are a martingale, cf. Theorem 9.12. Further,

$$\begin{aligned}
&I_3+I_{12}\\
&=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\int_{y\neq 0}\mathrm{e}_\xi(X_{s-}-x)\left[\mathrm{e}_\xi(\Phi(X_{s-})y)-1-\mathrm{i}\xi\cdot\Phi(X_{s-})y\mathbb{1}_{(0,1)}(|y|)\right]d_sN_s(dy)\\
&=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\int_{y\neq 0}\mathrm{e}_\xi(X_{s-}-x)\left[\mathrm{e}_\xi(\Phi(X_{s-})y)-1-\mathrm{i}\xi\Phi(X_{s-})y\mathbb{1}_{(0,1)}(|y|)\right]\nu(dy)\,ds\\
&\xrightarrow[t\to 0]{}\int_{y\neq 0}\left[\mathrm{e}^{\mathrm{i}\xi\cdot\Phi(x)y}-1-\mathrm{i}\xi\cdot\Phi(x)y\mathbb{1}_{(0,1)}(|y|)\right]\nu(dy).
\end{aligned}$$

Here we use that $\nu(dy)\,ds$ is the compensator of $d_sN_s(dy)$, see Lemma 9.4. This requires that the integrand is ν-integrable, but this is ensured by the local boundedness of $\Phi(\cdot)$ and the fact that $\int_{y\neq 0}\min\{|y|^2,1\}\,\nu(dy)<\infty$. Moreover,

$$I_{11}=\frac{1}{t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{i}\,\mathrm{e}_\xi(X_{s-}-x)\xi\cdot\Phi(X_{s-})l\,ds\xrightarrow[t\to 0]{}\mathrm{i}\,\xi\cdot\Phi(x)l,$$

and, finally, we observe that

$$\begin{aligned}[X,X]^{\mathsf{c}}&=\Big[\int\Phi(X_{s-})\,dL_s,\int\Phi(X_{s-})\,dL_s\Big]^{\mathsf{c}}\\&=\int\Phi(X_{s-})\,d[L,L]^{\mathsf{c}}_s\,\Phi^\top(X_{s-})\\&=\int\Phi(X_{s-})\,Q\,\Phi^\top(X_{s-})\,ds\;\in\;\mathbb{R}^{d\times d}\end{aligned}$$

which gives

$$\begin{aligned}I_2&=-\frac{1}{2t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{e}_\xi(X_{s-}-x)\xi\cdot d[X,X]^{\mathsf{c}}_s\xi\\&=-\frac{1}{2t}\mathbb{E}^x\int_0^{t\wedge\tau}\mathrm{e}_\xi(X_{s-}-x)\xi\cdot\Phi(X_{s-})Q\Phi^\top(X_{s-})\xi\,ds\\&\xrightarrow[t\to 0]{}-\frac{1}{2}\xi\cdot\Phi(x)Q\Phi^\top(x)\xi.\end{aligned}$$

This proves

$$\begin{aligned}q(x,\xi)&=-\mathrm{i}\,l\cdot\Phi^\top(x)\xi+\frac{1}{2}\xi\cdot\Phi(x)Q\Phi^\top(x)\xi\\&\qquad+\int_{y\neq 0}\big[1-\mathrm{e}^{\mathrm{i}\,y\cdot\Phi^\top(x)\xi}+\mathrm{i}\,y\cdot\Phi^\top(x)\xi\mathbb{1}_{(0,1)}(|y|)\big]\,\nu(dy)\\&=\psi(\Phi^\top(x)\xi).\end{aligned}$$

For the second part of the theorem we use Lemma 12.10 to see that $\mathcal{C}_c^\infty(\mathbb{R}^d)\subset\mathcal{D}(A)$. The continuity of the flow $x\mapsto X_t^x$ (Protter [44, Chapter V, Theorem 38]) – X^x is the unique solution of the SDE with initial condition $X_0^x=x$ – ensures that the semigroup $P_tf(x):=\mathbb{E}^xf(X_t)=\mathbb{E}f(X_t^x)$ maps $f\in\mathcal{C}_\infty(\mathbb{R}^d)$ to $\mathcal{C}_b(\mathbb{R}^d)$. In order to see that $P_tf\in\mathcal{C}_\infty(\mathbb{R}^d)$ we need that $\lim_{|x|\to\infty}|X_t^x|=\infty$ a.s. This requires some longer calculations, see, e.g., Schnurr [58, Theorem 2.49] or Kunita [34, Proof of Theorem 3.5, p. 353, line 13 from below] (for Brownian motion this argument is fully worked out in Schilling & Partzsch [56, Corollary 19.31]). □

Chapter 13

Dénouement

It is well known that the characteristic exponent $\psi(\xi)$ of a Lévy process $L = (L_t)_{t\geqslant 0}$ can be used to describe many probabilistic properties of the process. The key is the formula

$$\mathbb{E}^x e^{i\,\xi\cdot(L_t - x)} = \mathbb{E} e^{i\,\xi\cdot L_t} = e^{-t\psi(\xi)} \tag{13.1}$$

which gives direct access to the Fourier transform of the transition probability $\mathbb{P}^x(L_t \in dy) = \mathbb{P}(L_t + x \in dy)$. Although it is not any longer true that the symbol $q(x,\xi)$ of a Feller process $X = (X_t)_{t\geqslant 0}$ is the characteristic *exponent*, we may interpret formulae like (12.4)

$$-q(x,\xi) = \lim_{r\to 0} \frac{\mathbb{E}^x e^{i(X_{\tau_r} - x)\cdot\xi} - 1}{\mathbb{E}^x \tau_r}$$

as infinitesimal versions of the relation (13.1). What is more, both $\psi(\xi)$ and $q(x,\xi)$ are the Fourier symbols of the generators of the processes. We have already used these facts to discuss the conservativeness of Feller processes (Theorem 11.13) and the semimartingale decomposition of Feller processes (Theorem 12.9).

It is indeed possible to investigate further path properties of a Feller process using its symbol $q(x,\xi)$. Below we will, mostly without proof, give some examples which are taken from Böttcher, Schilling & Wang [9]. Let us point out the two guiding principles.

▶ *For sample path properties, the symbol $q(x,\xi)$ of a Feller process assumes same role as the characteristic exponent $\psi(\xi)$ of a Lévy process.*

▶▶ *A Feller process is 'locally Lévy', i.e., for short-time path properties the Feller process, started at the point x_0, behaves like the Lévy process $(L_t + x_0)_{t\geqslant 0}$ with exponent $\psi(\xi) := q(x_0,\xi)$.*

The latter property is the reason why such Feller processes are often called **Lévy-type processes**. The model case is the stable-like process whose symbol is given

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_13

by $q(x,\xi)=|\xi|^{\alpha(x)}$ where $\alpha:\mathbb{R}^d\to(0,2)$ is sufficiently smooth[1]. This process behaves locally, and for short times $t\ll 1$, like an $\alpha(x)$-stable process, only that the index now depends on the starting point $X_0=x$.

The key to many path properties are the following **maximal estimates** which were first proved in [53]. The present proof is taken from [9], the observation that we may use a random time τ instead of a fixed time t is due to F. Kühn.

Theorem 13.1. *Let $(X_t)_{t\geqslant 0}$ be a Feller process with generator A, $\mathcal{C}_c^\infty(\mathbb{R}^d)\subset\mathcal{D}(A)$ and symbol $q(x,\xi)$. If τ is an integrable stopping time, then*

$$\mathbb{P}^x\left(\sup_{s\leqslant\tau}|X_s-x|>r\right)\leqslant c\,\mathbb{E}^x\tau\sup_{|y-x|\leqslant r}\sup_{|\xi|\leqslant r^{-1}}|q(y,\xi)|. \tag{13.2}$$

Proof. Denote by $\sigma_r=\sigma_r^x$ the first exit time from the closed ball $\overline{B_r(x)}$. Clearly,

$$\{\sigma_r^x<\tau\}\subset\left\{\sup_{s\leqslant\tau}|X_s-x|>r\right\}\subset\{\sigma_r^x\leqslant\tau\}.$$

Pick $u\in\mathcal{C}_c^\infty(\mathbb{R}^d)$, $0\leqslant u\leqslant 1$, $u(0)=1$, $\operatorname{supp}u\subset B_1(0)$, and set

$$u_r^x(y):=u\left(\frac{y-x}{r}\right).$$

In particular, $u_r^x|_{B_r(x)^c}=0$. Hence,

$$\mathbb{1}_{\{\sigma_r^x\leqslant\tau\}}\leqslant 1-u_r^x(X_{\tau\wedge\sigma_r^x}).$$

Now we can use (5.5) or (12.3) to get

$$\begin{aligned}
\mathbb{P}^x(\sigma_r^x\leqslant\tau) &\leqslant \mathbb{E}^x\left[1-u_r^x(X_{\tau\wedge\sigma_r^x})\right]\\
&= \mathbb{E}^x\int_{[0,\tau\wedge\sigma_r^x)} q(X_s,D)u_r^x(X_s)\,ds\\
&= \mathbb{E}^x\int_{[0,\tau\wedge\sigma_r^x)}\int \mathbb{1}_{B_r(x)}(X_s)\mathrm{e}_\xi(X_s)q(X_s,\xi)\widehat{u}_r^x(\xi)\,d\xi\,ds\\
&\leqslant \mathbb{E}^x\int_{[0,\tau\wedge\sigma_r^x)}\int\sup_{|y-x|\leqslant r}|q(y,\xi)||\widehat{u}_r^x(\xi)|\,d\xi\,ds\\
&= \mathbb{E}^x\left[\tau\wedge\sigma_r^x\right]\int\sup_{|y-x|\leqslant r}|q(y,\xi)||\widehat{u}_r^x(\xi)|\,d\xi\\
&\underset{\text{L 11.8}}{\overset{(11.9)}{\leqslant}} c\,\mathbb{E}^x\tau\sup_{|y-x|\leqslant r}\sup_{|\xi|\leqslant r^{-1}}|q(y,\xi)|\int(1+|\xi|^2)|\widehat{u}(\xi)|\,d\xi.
\end{aligned}$$
□

[1] In dimension $d=1$ Lipschitz or even Dini continuity is enough (see Bass [4]), in higher dimensions we need something like $\mathcal{C}^{5d+3}$-smoothness, cf. Hoh [21]. Meanwhile, Kühn [33] established the existence for $d\geqslant 1$ with $\alpha(x)$ satisfying any Hölder condition.

There is a also a probability estimate for $\sup_{s\leqslant t}|X_s - x| \leqslant r$, but this requires some **sector condition** for the symbol, that is an estimate of the form

$$|\operatorname{Im} q(x,\xi)| \leqslant \kappa \operatorname{Re} q(x,\xi), \qquad x,\xi \in \mathbb{R}^d. \tag{13.3}$$

One consequence of (13.3) is that the drift (which is contained in the imaginary part of the symbol) is not dominant. This is a familiar assumption in the study of path properties of Lévy processes, see, e.g., Blumenthal & Getoor [8]; a typical example where this condition is violated are (Lévy) symbols of the form $\psi(\xi) = \mathrm{i}\,\xi + |\xi|^{\frac{1}{2}}$. For a Lévy process the sector condition on ψ coincides with the sector condition for the generator and the associated non-symmetric Dirichlet form, see Jacob [26, Volume 1, 4.7.32–33].

With some more effort (see [9, Theorem 5.5]), we may replace the sector condition by imposing conditions on the expression

$$\sup_{|y-x|\leqslant r} \frac{\operatorname{Re} q(x,\xi)}{|\xi|\,|\operatorname{Im} q(y,\xi)|} \quad \text{as } r \to \infty.$$

Theorem 13.2 (see [9, pp. 117–119]). *Let $(X_t)_{t\geqslant 0}$ be a Feller process with generator A, $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ and symbol $q(x,\xi)$ satisfying the sector condition* (13.3)*. Then*

$$\mathbb{P}^x\left(\sup_{s\leqslant t}|X_s - x| < r\right) \leqslant \frac{c_{\kappa,d}}{t\sup_{|\xi|\leqslant r^{-1}}\inf_{|y-x|\leqslant 2r}|q(y,\xi)|}. \tag{13.4}$$

The maximal estimates (13.2) and (13.4) are quite useful tools. With them we can estimate the mean exit time from balls (X and $q(x,\xi)$ are as in Theorem 13.2):

$$\frac{c}{\sup_{|\xi|\leqslant 1/r}\inf_{|y-x|\leqslant r}|q(y,\xi)|} \leqslant \mathbb{E}^x \sigma_r^x \leqslant \frac{c_\kappa}{\sup_{|\xi|\leqslant k^*/r}\inf_{|y-x|\leqslant r}|q(y,\xi)|}$$

for all $x \in \mathbb{R}^d$ and $r > 0$ and with $k^* = \arccos\sqrt{2/3} \approx 0.615$.

Recently, Kühn [32] studied the existence of and estimates for generalized moments; a typical result is contained in the following theorem.

Theorem 13.3. *Let $X = (X_t)_{t\geqslant 0}$ be a Feller process with infinitesimal generator $(A, \mathcal{D}(A))$ and $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$. Assume that the symbol $q(x,\xi)$, given by* (11.4)*, satisfies $q(x,0) = 0$ and has $(l(x), Q(x), \nu(x,dy))$ as x-dependent Lévy triplet. If $f : \mathbb{R}^d \to [0,\infty)$ is* (*comparable to*) *a twice continuously differentiable submultiplicative function such that*

$$\sup_{x\in K}\int f(y)\,\nu(x,dy) < \infty \quad \textit{for a compact set } K \subset \mathbb{R}^d,$$

then the generalized moment $\sup_{x\in K}\sup_{s\leqslant t}\mathbb{E}^x f(X_{s\wedge\tau_K} - x) < \infty$ exists and

$$\mathbb{E}^x f(X_{t\wedge\tau_K}) \leqslant c f(x) \mathrm{e}^{C(M_1+M_2)t};$$

here $\tau_K = \inf\{t > 0 : X_t \notin K\}$ *is the first exit time from* K, $C = C_K$ *is some absolute constant and*

$$M_1 = \sup_{x \in K} \Big(|l(x)| + |Q(x)| + \int_{y \neq 0} (|y|^2 \wedge 1)\, \nu(x, dy) \Big),$$
$$M_2 = \sup_{x \in K} \int_{|y| \geqslant 1} f(y)\, \nu(x, dy).$$

If X *has bounded coefficients* (*Lemma* 11.8), *then* $K = \mathbb{R}^d$ *is admissible.*

There are also counterparts for the Blumenthal–Getoor and Pruitt indices. Below we give two representatives, for a full discussion we refer to [9].

Definition 13.4. Let $q(x, \xi)$ be the symbol of a Feller process. Then

$$\beta_\infty^x := \inf \left\{ \lambda > 0 \,:\, \lim_{|\xi| \to \infty} \frac{\sup_{|\eta| \leqslant |\xi|} \sup_{|y - x| \leqslant 1/|\xi|} |q(y, \eta)|}{|\xi|^\lambda} = 0 \right\}, \tag{13.5}$$
$$\delta_\infty^x := \sup \left\{ \lambda > 0 \,:\, \lim_{|\xi| \to \infty} \frac{\inf_{|\eta| \leqslant |\xi|} \inf_{|y - x| \leqslant 1/|\xi|} |q(y, \eta)|}{|\xi|^\lambda} = \infty \right\}. \tag{13.6}$$

By definition, $0 \leqslant \delta_\infty^x \leqslant \beta_\infty^x \leqslant 2$. For example, if $q(x, \xi) = |\xi|^{\alpha(x)}$ with a smooth exponent function $\alpha(x)$, then $\beta_\infty^x = \delta_\infty^x = \alpha(x)$; in general, however, we cannot expect that the two indices coincide.

As for Lévy processes, these indices can be used to describe the path behaviour.

Theorem 13.5. *Let* $(X_t)_{t \geqslant 0}$ *be a* d*-dimensional Feller process with the generator* A *such that* $\mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A)$ *and symbol* $q(x, \xi)$. *For every bounded analytic set* $E \subset [0, \infty)$, *the Hausdorff dimension*

$$\dim \{X_t \,:\, t \in E\} \leqslant \min \left\{ d,\ \sup_{x \in \mathbb{R}^d} \beta_\infty^x \dim E \right\}. \tag{13.7}$$

A proof can be found in [9, Theorem 5.15]. It is instructive to observe that we have to take the supremum w.r.t. the space variable x, as we do not know how the process X moves while we observe it during $t \in E$. This shows that we can only expect to get 'exact' results if $t \to 0$. Here is such an example.

Theorem 13.6. *Let* $(X_t)_{t \geqslant 0}$ *be a* d*-dimensional Feller process with symbol* $q(x, \xi)$ *satisfying the sector condition. Then,* $\mathbb{P}^x$*-a.s.*

$$\lim_{t \to 0} \frac{\sup_{0 \leqslant s \leqslant t} |X_s - x|}{t^{1/\lambda}} = 0 \qquad \forall \lambda > \beta_\infty^x, \tag{13.8}$$
$$\lim_{t \to 0} \frac{\sup_{0 \leqslant s \leqslant t} |X_s - x|}{t^{1/\lambda}} = \infty \qquad \forall \lambda < \delta_\infty^x. \tag{13.9}$$

As one would expect, these results are proved using the maximal estimates (13.2) and (13.4) in conjunction with the Borel–Cantelli lemma, see [9, Theorem 5.16].

If we are interested in the long-term behaviour, one could introduce indices 'at zero', where we replace in Definition 13.4 the limit $|\xi| \to \infty$ by $|\xi| \to 0$, but we will always have to pay the price that we loose the influence of the starting point $X_0 = x$, i.e., we will have to take the supremum or infimum for all $x \in \mathbb{R}^d$.

With the machinery we have developed here, one can also study further path properties, such as invariant measures, ergodicity, transience and recurrence etc. For this we refer to the monograph [9] as well as recent developments by Behme & Schnurr [7] and Sandrić [49, 50].

Appendix: Some Classical Results

In this appendix we collect some classical results from (or needed for) probability theory which are not always contained in a standard course.

The Cauchy–Abel functional equation

Below we reproduce the standard proof for **continuous** functions which, however, works also for right-continuous (or monotone) functions.

Theorem A.1. *Let $\phi : [0,\infty) \to \mathbb{C}$ be a right-continuous function satisfying the functional equation $\phi(s+t) = \phi(s)\phi(t)$. Then $\phi(t) = \phi(1)^t$.*

Proof. Assume that $\phi(a) = 0$ for some $a \geqslant 0$. Then we find for all $t \geqslant 0$

$$\phi(a+t) = \phi(a)\phi(t) = 0 \implies \phi|_{[a,\infty)} \equiv 0.$$

To the left of a we find for all $n \in \mathbb{N}$

$$0 = \phi(a) = \left[\phi\left(\tfrac{a}{n}\right)\right]^n \implies \phi\left(\tfrac{a}{n}\right) = 0.$$

Since ϕ is right-continuous, we infer that $\phi|_{[0,\infty)} \equiv 0$, and $\phi(t) = \phi(1)^t$ holds.

Now assume that $\phi(1) \neq 0$. Setting $f(t) := \phi(t)\phi(1)^{-t}$ we get

$$f(s+t) = \phi(s+t)\phi(1)^{-(s+t)} = \phi(s)\phi(1)^{-s}\phi(t)\phi(1)^{-t} = f(s)f(t)$$

as well as $f(1) = 1$. Applying the functional equation k times we conclude that

$$f\left(\tfrac{k}{n}\right) = \left[f\left(\tfrac{1}{n}\right)\right]^k \quad \text{for all} \quad k,n \in \mathbb{N}.$$

The same calculation done backwards yields

$$\left[f\left(\tfrac{1}{n}\right)\right]^k = \left[f\left(\tfrac{1}{n}\right)\right]^{n\frac{k}{n}} = \left[f\left(\tfrac{n}{n}\right)\right]^{\frac{k}{n}} = \left[f(1)\right]^{\frac{k}{n}} = 1.$$

Hence, $f|_{\mathbb{Q}_+} \equiv 1$. Since ϕ, hence f, is right-continuous, we see that $f \equiv 1$ or, equivalently, $\phi(t) = [\phi(1)]^t$ for all $t > 0$. $\square$

Characteristic functions and moments

Theorem A.2 (Even moments and characteristic functions)**.** *Let* $Y=(Y^{(1)},\dots,Y^{(d)})$ *be a random variable in* $\mathbb{R}^d$ *and let* $\chi(\xi) = \mathbb{E}\, e^{i\,\xi\cdot Y}$ *be its characteristic function. Then* $\mathbb{E}(|Y|^2)$ *exists if, and only if, the second derivatives* $\frac{\partial^2}{\partial\xi_k^2}\chi(0)$, $k = 1,\dots,d$, *exist and are finite. In this case all mixed second derivatives exist and*

$$\mathbb{E}Y^{(k)} = \frac{1}{i}\frac{\partial\chi(0)}{\partial\xi_k} \quad \textit{and} \quad \mathbb{E}(Y^{(k)}Y^{(l)}) = -\frac{\partial^2\chi(0)}{\partial\xi_k\partial\xi_l}. \tag{A.1}$$

Proof. In order to keep the notation simple, we consider only $d = 1$. If $\mathbb{E}(Y^2) < \infty$, then the formulae (A.1) are routine applications of the differentiation lemma for parameter-dependent integrals, see, e.g., [54, Theorem 11.5] or [55, Satz 12.2]. Moreover, χ is twice continuously differentiable.

Let us prove that $\mathbb{E}(Y^2) \leqslant -\chi''(0)$. An application of l'Hospital's rule gives

$$\begin{aligned}
\chi''(0) &= \lim_{h\to 0}\frac{1}{2}\left(\frac{\chi'(2h)-\chi'(0)}{2h} + \frac{\chi'(0)-\chi'(-2h)}{2h}\right)\\
&= \lim_{h\to 0}\frac{\chi'(2h)-\chi'(-2h)}{4h}\\
&= \lim_{h\to 0}\frac{\chi(2h)-2\chi(0)+\chi(-2h)}{4h^2}\\
&= \lim_{h\to 0}\mathbb{E}\left[\left(\frac{e^{i\,hY}-e^{-i\,hY}}{2h}\right)^2\right]\\
&= -\lim_{h\to 0}\mathbb{E}\left[\left(\frac{\sin hY}{h}\right)^2\right].
\end{aligned}$$

From Fatou's lemma we get

$$\chi''(0) \leqslant -\mathbb{E}\left[\lim_{h\to 0}\left(\frac{\sin hY}{h}\right)^2\right] = -\mathbb{E}\big[Y^2\big].$$

In the multivariate case observe that $\mathbb{E}|Y^{(k)}Y^{(l)}| \leqslant \mathbb{E}\big[(Y^{(k)})^2\big] + \mathbb{E}\big[(Y^{(l)})^2\big]$. □

Vague and weak convergence of measures

A sequence of locally finite[2] Borel measures $(\mu_n)_{n\in\mathbb{N}}$ on $\mathbb{R}^d$ converges **vaguely** to a locally finite measure μ if

$$\lim_{n\to\infty}\int\phi\,d\mu_n = \int\phi\,d\mu \quad \text{for all} \quad \phi\in\mathcal{C}_c(\mathbb{R}^d). \tag{A.2}$$

[2] I.e., every compact set K has finite measure.

Since the compactly supported continuous functions $\mathcal{C}_c(\mathbb{R}^d)$ are dense in the space of continuous functions vanishing at infinity

$$\mathcal{C}_\infty(\mathbb{R}^d) = \{\phi \in \mathcal{C}(\mathbb{R}^d) \,:\, \lim_{|x|\to\infty} \phi(x) = 0\},$$

we can replace in (A.2) the set $\mathcal{C}_c(\mathbb{R}^d)$ with $\mathcal{C}_\infty(\mathbb{R}^d)$. The following theorem guarantees that a family of Borel measures is sequentially relatively compact[3] for the vague convergence.

Theorem A.3. *Let $(\mu_t)_{t\geqslant 0}$ be a family of measures on $\mathbb{R}^d$ which is uniformly bounded, i.e., $\sup_{t\geqslant 0} \mu_t(\mathbb{R}^d) < \infty$. Then every sequence $(\mu_{t_n})_{n\in\mathbb{N}}$ has a vaguely convergent subsequence.*

If we test in (A.2) against all bounded continuous functions $\phi \in \mathcal{C}_b(\mathbb{R}^d)$, we get **weak convergence** of the sequence $\mu_n \to \mu$. One has

Theorem A.4. *A sequence of measures $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to μ if, and only if, μ_n converges vaguely to μ and $\lim_{n\to\infty} \mu_n(\mathbb{R}^d) = \mu(\mathbb{R}^d)$ (preservation of mass). In particular, weak and vague convergence coincide for sequences of probability measures.*

Proofs and a full discussion of vague and weak convergence can be found in Malliavin [41, Chapter III.6] or Schilling [55, Chapter 25].

For any finite measure μ on $\mathbb{R}^d$ we denote by $\check{\mu}(\xi) := \int_{\mathbb{R}^d} e^{i\,\xi\cdot y}\,\mu(dy)$ its characteristic function.

Theorem A.5 (Lévy's continuity theorem)**.** *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of finite measures on $\mathbb{R}^d$. If $\mu_n \to \mu$ weakly, then the characteristic functions $\check{\mu}_n(\xi)$ converge locally uniformly to $\check{\mu}(\xi)$.*

Conversely, if the limit $\lim_{n\to\infty} \check{\mu}_n(\xi) = \chi(\xi)$ exists for all $\xi \in \mathbb{R}^d$ and defines a function χ which is continuous at $\xi = 0$, then there exists a finite measure μ such that $\check{\mu}(\xi) = \chi(\xi)$ and $\mu_n \to \mu$ weakly.

A proof in one dimension is contained in the monograph by Breiman [10], d-dimensional versions can be found in Bauer [5, Chapter 23] and Malliavin [41, Chapter IV.4].

Convergence in distribution

By $\xrightarrow{d}$ we denote convergence in distribution.

Theorem A.6 (Convergence of types)**.** *Let $(Y_n)_{n\in\mathbb{N}}$, Y and Y' be random variables and suppose that there are constants $a_n > 0$, $c_n \in \mathbb{R}$ such that*

$$Y_n \xrightarrow[n\to\infty]{d} Y \quad \textit{and} \quad a_n Y_n + c_n \xrightarrow[n\to\infty]{d} Y'.$$

If Y and Y' are non-degenerate, then the limits $a = \lim_{n\to\infty} a_n$ and $c = \lim_{n\to\infty} c_n$ exist and $Y' \sim aY + c$.

[3]Note that compactness and sequential compactness need not coincide!

Proof. Write $\chi_Z(\xi) := \mathbb{E}e^{iup\xi Z}$ for the characteristic function of the random variable Z.

1° By Lévy's continuity theorem (Theorem A.5) convergence in distribution ensures that

$$\chi_{a_nY_n+c_n}(\xi)=e^{i\,c_n\cdot\xi}\chi_{Y_n}(a_n\xi)\xrightarrow[n\to\infty]{\text{locally unif.}}\chi_{Y'}(\xi) \quad \text{and} \quad \chi_{Y_n}(\xi)\xrightarrow[n\to\infty]{\text{locally unif.}}\chi_Y(\xi).$$

Take some subsequence $(a_{n(k)})_{k\in\mathbb{N}} \subset (a_n)_{n\in\mathbb{N}}$ such that $\lim_{k\to\infty} a_{n(k)} = a$ exists in $[0,\infty]$.

2° Claim: $a > 0$. Assume, on the contrary, that $a = 0$.

$$|\chi_{a_{n(k)}Y_{n(k)}+c_{n(k)}}(\xi)| = |\chi_{Y_{n(k)}}(a_{n(k)}\xi)| \xrightarrow[k\to\infty]{} |\chi_Y(0)| = 1.$$

Thus, $|\chi_{Y'}| \equiv 1$ which means that Y' would be degenerate, contradicting our assumption.

3° Claim: $a < \infty$. If $a = \infty$, we use $Y_n = (a_n)^{-1}(Y_n - c_n)$ and the argument from step 1° to reach the contradiction

$$(a_{n(k)})^{-1} \xrightarrow[k\to\infty]{} a^{-1} > 0 \iff a < \infty.$$

4° Claim: There exists a unique $a \in [0,\infty)$ such that $\lim_{n\to\infty} a_n = a$. Assume that there were two different subsequential limits $a_{n(k)} \to a$, $a_{m(k)} \to a'$ and $a \neq a'$. Then

$$|\chi_{a_{n(k)}Y_{n(k)}+c_{n(k)}}(\xi)| = |\chi_{a_{n(k)}Y_{n(k)}}(\xi)| \xrightarrow[k\to\infty]{} \chi_Y(a\xi),$$
$$|\chi_{a_{m(k)}Y_{m(k)}+c_{m(k)}}(\xi)| = |\chi_{a_{m(k)}Y_{m(k)}}(\xi)| \xrightarrow[k\to\infty]{} \chi_Y(a'\xi).$$

On the other hand, $\chi_Y(a\xi) = \chi_Y(a'\xi) = \chi_{Y'}(\xi)$. If $a' < a$, we get by iteration

$$|\chi_Y(\xi)| = \left|\chi_Y\left(\tfrac{a'}{a}\xi\right)\right| = \cdots = \left|\chi_Y\left(\left(\tfrac{a'}{a}\right)^N\xi\right)\right| \xrightarrow[N\to\infty]{} |\chi_Y(0)| = 1.$$

Thus, $|\chi| \equiv 1$ and Y is a.s. constant. Since a, a' can be interchanged, we conclude that $a = a'$.

5° We have

$$e^{i\,c_n\cdot\xi} = \frac{\chi_{a_nY_n+c_n}(\xi)}{\chi_{a_nY_n}(\xi)} = \frac{\chi_{a_nY_n+c_n}(\xi)}{\chi_{Y_n}(a_n\xi)} \xrightarrow[n\to\infty]{} \frac{\chi_{Y'}(\xi)}{\chi_Y(a\xi)}.$$

Since χ_Y is continuous and $\chi_Y(0) = 1$, the limit $\lim_{n\to\infty} e^{i\,c_n\cdot\xi}$ exists for all $|\xi| \leqslant \delta$ and some small δ. For $\xi = t\xi_0$ with $|\xi_0| = 1$, we get

$$0 < \left|\int_0^\delta \frac{\chi_{Y'}(t\xi_0)}{\chi_Y(ta\xi_0)}\,dt\right| = \left|\lim_{n\to\infty}\int_0^\delta e^{i\,tc_n\cdot\xi_0}\,dt\right|$$
$$= \left|\lim_{n\to\infty}\frac{e^{i\,\delta c_n\cdot\xi_0}-1}{i\,c_n\cdot\xi_0}\right| \leqslant \liminf_{n\to\infty}\frac{2}{|c_n\cdot\xi_0|},$$

and so $\limsup_{n\to\infty} |c_n| < \infty$; if there were two limit points $c \neq c'$, then $e^{i\,c\cdot\xi} = e^{i\,c'\cdot\xi}$ for all $|\xi| \leqslant \delta$. This gives $c = c'$, and we finally see that $c_n \to c$, $e^{i\,c_n\cdot\xi} \to e^{i\,c\cdot\xi}$, as well as

$$\chi_{Y'}(\xi) = e^{i\,c\cdot\xi}\chi_Y(a\xi). \qquad \square$$

A random variable is called **symmetric** if $Y \sim -Y$.

Theorem A.7 (Symmetrization inequality). *Let $Y_1, \dots, Y_n$ be independent symmetric random variables. Then the partial sum $S_n = Y_1 + \dots + Y_n$ is again symmetric and*

$$\mathbb{P}(|Y_1 + \dots + Y_n| > u) \geqslant \tfrac{1}{2}\mathbb{P}\Big(\max_{1\leqslant k\leqslant n} |Y_k| > u\Big). \tag{A.3}$$

If the Y_k are iid with $Y_1 \sim \mu$, then

$$\mathbb{P}(|Y_1 + \dots + Y_n| > u) \geqslant \tfrac{1}{2}\big(1 - e^{-n\mathbb{P}(|Y_1|>u)}\big). \tag{A.4}$$

Proof. By independence, $S_n = Y_1 + \dots + Y_n \sim -Y_1 - \dots - Y_n = -S_n$.

Let $\tau = \min\{1 \leqslant k \leqslant n : |Y_k| = \max_{1\leqslant l\leqslant n} |Y_l|\}$ and set $Y_{n,\tau} = S_n - Y_\tau$. Then the four (counting all possible $\pm$ combinations) random variables $(\pm Y_\tau, \pm Y_{n,\tau})$ have the same law. Moreover,

$$\mathbb{P}(Y_\tau > u) \leqslant \mathbb{P}(Y_\tau > u,\, Y_{n,\tau} \geqslant 0) + \mathbb{P}(Y_\tau > u,\, Y_{n,\tau} \leqslant 0) = 2\mathbb{P}(Y_\tau > u,\, Y_{n,\tau} \geqslant 0),$$

and so

$$\mathbb{P}(S_n > u) = \mathbb{P}(Y_\tau + Y_{n,\tau} > u) \geqslant \mathbb{P}(Y_\tau > u,\, Y_{n,\tau} \geqslant 0) \geqslant \tfrac{1}{2}\mathbb{P}(Y_\tau > u).$$

By symmetry, this implies (A.3). In order to see (A.4), we use that the Y_k are iid, hence

$$\mathbb{P}(\max_{1\leqslant k\leqslant n} |Y_k| \leqslant u) = \mathbb{P}(|Y_1| \leqslant u)^n \leqslant e^{-n\mathbb{P}(|Y_1|>u)},$$

along with the elementary inequality $1 - p \leqslant e^{-p}$ for $0 \leqslant p \leqslant 1$. This proves (A.4). $\square$

The predictable σ-algebra

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space and $(\mathscr{F}_t)_{t\geqslant 0}$ some filtration. A stochastic process $(X_t)_{t\geqslant 0}$ is called **adapted**, if for every $t \geqslant 0$ the random variable X_t is $\mathscr{F}_t$ measurable.

Definition A.8. The **predictable σ-algebra** $\mathscr{P}$ is the smallest σ-algebra on $\Omega \times (0,\infty)$ such that all left-continuous adapted stochastic processes $(\omega, t) \mapsto X_t(\omega)$ are measurable. A $\mathscr{P}$ measurable process X is called a **predictable process**.

For a stopping time τ we denote by

$$\rrbracket 0,\tau \rrbracket := \{(\omega,t) \,:\, 0 < t \leqslant \tau(\omega)\} \quad \text{and} \quad \rrbracket \tau,\infty\llbracket := \{(\omega,t) \,:\, t > \tau(\omega)\} \tag{A.5}$$

the left-open **stochastic intervals**. The following characterization of the predictable σ-algebra is essentially from Jacod & Shiryaev [27, Theorem I.2.2].

Theorem A.9. *The predictable σ-algebra $\mathscr{P}$ is generated by any one of the following families of random sets*

a) $\rrbracket 0,\tau\rrbracket$ *where τ is any bounded stopping time;*
b) $F_s \times (s,t]$ *where $F_s \in \mathscr{F}_s$ and $0 \leqslant s < t$.*

Proof. We write $\mathscr{P}_a$ and $\mathscr{P}_b$ for the σ-algebras generated by the families listed in a) and b), respectively.

1° Pick $0 \leqslant s < t$, $F = F_s \in \mathscr{F}_s$ and let $n > t$. Observe that the random time $s_F := s\mathbb{1}_F + n\mathbb{1}_{F^c}$ is a bounded stopping time[4] and $F \times (s,t] = \rrbracket s_F, t_F\rrbracket$. Therefore, $\rrbracket s_F, t_F\rrbracket = \rrbracket 0, t_F\rrbracket \setminus \rrbracket 0, s_F\rrbracket \in \mathscr{P}_a$, and we conclude that $\mathscr{P}_b \subset \mathscr{P}_a$.

2° Let τ be a bounded stopping time. Since $t \mapsto \mathbb{1}_{\rrbracket 0,\tau\rrbracket}(\omega,t)$ is adapted and left-continuous, we have $\mathscr{P}_a \subset \mathscr{P}$.

3° Let X be an adapted and left-continuous process and define for every $n \in \mathbb{N}$

$$X_t^n := \sum_{k=0}^\infty X_{k2^{-n}} \mathbb{1}_{\rrbracket k2^{-n},\,(k+1)2^{-n}\rrbracket}(t).$$

Obviously, $X^n = (X_t^n)_{t\geqslant 0}$ is $\mathscr{P}_b$ measurable; due to the left-continuity of $t \mapsto X_t$, the limit $\lim_{n\to\infty} X_t^n = X_t$ exists, and we conclude that X is $\mathscr{P}_b$ measurable; consequently, $\mathscr{P} \subset \mathscr{P}_b$. □

The structure of translation invariant operators

Let $\vartheta_x f(y) := f(y+x)$ be the translation operator and $\widetilde{f}(x) := f(-x)$. A linear operator $L : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d)$ is called **translation invariant** if

$$\vartheta_x(Lf) = L(\vartheta_x f). \tag{A.6}$$

A **distribution** λ is an element of the topological dual $(\mathcal{C}_c^\infty(\mathbb{R}^d))'$, i.e., a continuous linear functional $\lambda : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathbb{R}$. The **convolution of a distribution** with a function $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ is defined as

$$f * \lambda(x) := \lambda(\vartheta_{-x}\widetilde{f}), \quad \lambda \in \big(\mathcal{C}_c^\infty(\mathbb{R}^d)\big)', \; f \in \mathcal{C}_c^\infty(\mathbb{R}^d), \; x \in \mathbb{R}^d.$$

[4] Indeed, $\{s_F \leqslant t\} = \{s \leqslant t\} \cap F = \begin{Bmatrix} \emptyset, & s > t \\ F, & s \leqslant t \end{Bmatrix} \in \mathscr{F}_t$ for all $t \geqslant 0$.

If $\lambda = \mu$ is a measure, this formula generalizes the 'usual' convolution

$$f * \mu(x) = \int f(x-y)\,\mu(dy) = \int \widetilde{f}(y-x)\,\mu(dy) = \mu(\vartheta_{-x}\widetilde{f}).$$

Theorem A.10. *If $\lambda \in (\mathcal{C}_c^\infty(\mathbb{R}^d))'$ is a distribution, then $Lf(x) := f * \lambda(x)$ defines a translation invariant continuous linear map $L : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathcal{C}^\infty(\mathbb{R}^d)$.*

Conversely, every translation-invariant continuous linear map $L : \mathcal{C}_c^\infty(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d)$ is of the form $Lf = f * \lambda$ for some unique distribution $\lambda \in (\mathcal{C}_c^\infty(\mathbb{R}^d))'$.

Proof. Let $Lf = f * \lambda$. From the very definition of the convolution we get

$$\begin{aligned}(\vartheta_{-x}f) * \lambda = \lambda(\vartheta_{-x}\widetilde{\vartheta_{-x}f}) &= \lambda(\vartheta_{-x}[f(x-\cdot)])\\ &= \vartheta_{-x}\lambda(\vartheta_{-x}[f(-\cdot)])\\ &= \vartheta_{-x}(f * \lambda)\end{aligned}$$

For proving the continuity of L, it is enough to show that $L : \mathcal{C}_c^\infty(K) \to \mathcal{C}^\infty(\mathbb{R}^d)$ is continuous for every compact set $K \subset \mathbb{R}^d$ (this is because of the definition of the topology in $\mathcal{C}_c^\infty(\mathbb{R}^d)$). We will use the closed graph theorem: Assume that $f_n \to f$ in $\mathcal{C}_c^\infty(\mathbb{R}^d)$ and $f_n * \lambda \to g$ in $\mathcal{C}^\infty(\mathbb{R}^d)$, then we have to show that $g = f * \lambda$.

For every $x \in \mathbb{R}^d$ we have $\vartheta_{-x}\widetilde{f_n} \to \vartheta_{-x}\widetilde{f}$ in $\mathcal{C}_c^\infty(\mathbb{R}^d)$, and so

$$g(x) = \lim_{n\to\infty}(f_n * \lambda)(x) = \lim_{n\to\infty}\lambda(\vartheta_{-x}\widetilde{f_n}) = \lambda(\vartheta_{-x}\widetilde{f}) = (f * \lambda)(x).$$

Assume now, that L is translation invariant and continuous. Define $\lambda(f) := (L\widetilde{f})(0)$. Since L is linear and continuous, and $f \mapsto \widetilde{f}$ and the evaluation at $x = 0$ are continuous operations, λ is a continuous linear map on $\mathcal{C}_c^\infty(\mathbb{R}^d)$. Because of the translation invariance of L we get

$$(Lf)(x) = (\vartheta_x Lf)(0) = L(\vartheta_x f)(0) = \lambda(\widetilde{\vartheta_x f}) = \lambda(\vartheta_x\widetilde{f}) = (f * \lambda)(x).$$

If μ is a further distribution with $Lf(0) = f * \mu(0)$, we see

$$(\mu - \lambda)(\widetilde{f}) = f * (\mu - \lambda)(0) = f * \mu(0) - f * \lambda(0) = 0 \quad \text{for all} \quad f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$$

which proves $\mu = \lambda$. □

Bibliography

[1] Aczél, J.: *Lectures on Functional Equations and Their Applications.* Academic Press, New York (NY) 1966.

[2] Applebaum, D.: *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge 2009 (2nd ed).

[3] Barndorff-Nielsen, O.E. et al.: *Lévy Processes. Theory and Applications.* Birkhäuser, Boston (MA) 2001.

[4] Bass, R.: Uniqueness in law for pure-jump Markov processes. *Probab. Theor. Relat. Fields* **79** (1988) 271–287.

[5] Bauer, H.: *Probability Theory.* De Gruyter, Berlin 1996.

[6] Bawly, G.M.: Über einige Verallgemeinerungen der Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Mat. Sbornik* (*Réc. Math. Moscou, N.S.*) **1** (1936) 917–930.

[7] Behme, A., Schnurr, A.: A criterion for invariant measures of Itô processes based on the symbol. *Bernoulli* **21** (2015) 1697–1718.

[8] Blumenthal, R.M., Getoor, R.K.: Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10** (1961) 493–516.

[9] Böttcher, B., Schilling, R.L., Wang, J.: *Lévy-type Processes: Construction, Approximation and Sample Path Properties.* Lecture Notes in Mathematics **2099** (Lévy Matters III), Springer, Cham 2014.

[10] Breiman, L.: *Probability.* Addison–Wesley, Reading (MA) 1968 (several reprints by SIAM, Philadelphia).

[11] Bretagnolle, J.L.: Processus à accroissements independants. In: Bretagnolle, J.L. et al.: *École d'Été de Probabilités: Processus Stochastiques.* Springer, Lecture Notes in Mathematics **307**, Berlin 1973, pp. 1–26.

[12] Çinlar, E.: *Introduction to Stochastic Processes.* Prentice Hall, Englewood Cliffs (NJ) 1975 (reprinted by Dover, Mineola).

[13] Courrège, P.: Générateur infinitésimal d'un semi-groupe de convolution sur $\mathbb{R}^n$, et formule de Lévy–Khintchine. *Bull. Sci. Math.* **88** (1964) 3–30.

[14] Courrège, P.: Sur la forme intégro différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum. In: *Séminaire Brelot–Choquet–Deny. Théorie du potentiel* **10** (1965/66) exposé 2, pp. 1–38.

[15] de Finetti, B.: Sulle funzioni ad incremento aleatorio. *Rend. Accad. Lincei Ser. VI* **10** (1929) 163–168.

[16] Dieudonné, J.: *Foundations of Modern Analysis*. Academic Press, New York (NY) 1969.

[17] Ethier, S.N., Kurtz, T.G.: *Markov Processes. Characterization and Convergence*. John Wiley & Sons, New York (NY) 1986.

[18] Gikhman, I.I., Skorokhod, A.V.: *Introduction to the Theory of Random Processes*. W.B. Saunders, Philadelphia (PA) 1969 (reprinted by Dover, Mineola 1996).

[19] Grimvall, A.: A theorem on convergence to a Lévy process. *Math. Scand.* **30** (1972) 339–349.

[20] Herz, C.S.: Théorie élémentaire des distributions de Beurling. *Publ. Math. Orsay* no. **5**, 2ème année 1962/63, Paris 1964.

[21] Hoh, W.: Pseudo differential operators with negative definite symbols of variable order. *Rev. Mat. Iberoamericana* **16** (2000) 219–241.

[22] Ikeda, N., Watanabe S.: *Stochastic Differential Equations and Diffusion Processes*. North-Holland Publishing Co./Kodansha, Amsterdam and Tokyo 1989 (2nd ed).

[23] Itô, K.: On stochastic processes. I. (Infinitely divisible laws of probability). *Japanese J. Math.* **XVIII** (1942) 261–302.

[24] Itô, K.: *Lectures on Stochastic Processes*. Tata Institute of Fundamental Research, Bombay 1961 (reprinted by Springer, Berlin 1984). http://www.math.tifr.res.in/~publ/ln/tifr24.pdf

[25] Itô, K.: Semigroups in probability theory. In: *Functional Analysis and Related Topics*, Proceedings in Memory of K. Yosida (Kyoto 1991). Springer, Lecture Notes in Mathematics **1540**, Berlin 1993, 69–83.

[26] Jacob, N.: *Pseudo Differential Operators and Markov Processes* (*3 volumes*). Imperial College Press, London 2001–2005.

[27] Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*. Springer, Berlin 1987 (2nd ed Springer, Berlin 2003).

[28] Khintchine, A.Ya.: A new derivation of a formula by P. Lévy. *Bull. Moscow State Univ.* **1** (1937) 1–5 (Russian; an English translation is contained in Appendix 2.1, pp. 44–49 of [40]).

[29] Khintchine, A.Ya.: Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze. *Mat. Sbornik* (*Réc. Math. Moscou, N.S.*) **2** (1937) 79–119 (German; an English translation is contained in Appendix 3.3, pp. 79–125 of [40]).

[30] Knopova, V., Schilling, R.L.: A note on the existence of transition probability densities for Lévy processes. *Forum Math.* **25** (2013) 125–149.

[31] Kolmogorov, A.N.: Sulla forma generale di un processo stocastico omogeneo (Un problema die Bruno de Finetti). *Rend. Accad. Lincei Ser. VI* **15** (1932)

805–808 and 866–869 (an English translation is contained in: Shiryayev, A.N. (ed.): *Selected Works of A.N. Kolmogorov.* Kluwer, Dordrecht 1992, vol. 2, pp. 121–127).

[32] Kühn, F.: Existence and estimates of moments for Lévy-type processes. *Stoch. Proc. Appl.* (in press). DOI: `10.1016/j.spa.2016.07.008`

[33] Kühn, F.: *Probability and Heat Kernel Estimates for Lévy(-Type) Processes.* PhD Thesis, Technische Universität Dresden 2016.

[34] Kunita, H.: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: Rao, M.M. (ed.): *Real and Stochastic Analysis.* Birkhäuser, Boston (MA) 2004, pp. 305–373.

[35] Kunita, H., Watanabe, S.: On square integrable martingales. *Nagoya Math. J.* **30** (1967) 209–245.

[36] Kyprianou, A.: *Introductory Lectures on Fluctuations and Lévy Processes with Applications.* Springer, Berlin 2006.

[37] Lévy, P.: Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Sc. Norm. Sup. Pisa* **3** (1934) 337–366.

[38] Lévy, P.: *Théorie de l'addition des variables aléatoires.* Gauthier–Villars, Paris 1937.

[39] Mainardi, F., Rogosin, S.V.: The origin of infinitely divisible distributions: from de Finetti's problem to Lévy–Khintchine formula. *Math. Methods Economics Finance* **1** (2006) 37–55.

[40] Mainardi, F., Rogosin, S.V.: *The Legacy of A.Ya. Khintchine's Work in Probability Theory.* Cambridge Scientific Publishers, Cambridge 2010.

[41] Malliavin, P.: *Integration and Probability.* Springer, New York (NY) 1995.

[42] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York (NY) 1983.

[43] Prokhorov, Yu.V.: Convergence of random processes and limit theorems in probability theory. *Theor. Probab. Appl.* **1** (1956) 157–214.

[44] Protter, P.E.: *Stochastic Integration and Differential Equations.* Springer, Berlin 2004 (2nd ed).

[45] Revuz, D., Yor, M.: *Continuous Martingales and Brownian Motion.* Springer, Berlin 2005 (3rd printing of the 3rd ed).

[46] Rosiński, J.: On series representations of infinitely divisible random vectors. *Ann. Probab.* **18** (1990) 405–430.

[47] Rosiński, J.: Series representations of Lévy processes from the perspective of point processes. In: Barndorff-Nielsen et al. [3] (2001) 401–415.

[48] Samorodnitsky, G., Taqqu, M.S.: *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York (NY) 1994.

[49] Sandrić, N.: On recurrence and transience of two-dimensional Lévy and Lévy-type processes. *Stoch. Proc. Appl.* **126** (2016) 414–438.

[50] Sandrić, N.: Long-time behavior for a class of Feller processes. *Trans. Am. Math. Soc.* **368** (2016) 1871–1910.

[51] Sato, K.: *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge 1999 (2nd ed 2013).

[52] Schilling, R.L.: Conservativeness and extensions of Feller semigroups. *Positivity* **2** (1998) 239–256.

[53] Schilling, R.L.: Growth and Hölder conditions for the sample paths of a Feller process. *Probab. Theor. Relat. Fields* **112** (1998) 565–611.

[54] Schilling, R.L.: *Measures, Integrals and Martingales.* Cambridge University Press, Cambridge 2011 (3rd printing).

[55] Schilling, R.L.: *Maß und Integral.* De Gruyter, Berlin 2015.

[56] Schilling, R.L., Partzsch, L.: *Brownian Motion. An Introduction to Stochastic Processes.* De Gruyter, Berlin 2014 (2nd ed).

[57] Schilling, R.L., Schnurr, A.: The symbol associated with the solution of a stochastic differential equation. *El. J. Probab.* **15** (2010) 1369–1393.

[58] Schnurr, A.: *The Symbol of a Markov Semimartingale.* PhD Thesis, Technische Universität Dresden 2009. Shaker-Verlag, Aachen 2009.

[59] Schnurr, A.: On the semimartingale nature of Feller processes with killing. *Stoch. Proc. Appl.* **122** (2012) 2758–2780.

[60] Stroock, D.W., Varadhan, S.R.S.: *Multidimensional Stochastic Processes.* Springer, Berlin 1997 (2nd corrected printing).

[61] von Waldenfels, W.: *Eine Klasse stationärer Markowprozesse.* Kernforschungsanlage Jülich, Institut für Plasmaphysik, Jülich 1961.

[62] von Waldenfels, W.: Fast positive Operatoren. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **4** (1965) 159–174.

Part II

Invariance and Comparison Principles for Parabolic Stochastic Partial Differential Equations

By Davar Khoshnevisan

Preface

These notes aim to introduce the reader to aspects of the theory of parabolic stochastic partial differential equations (SPDEs, for short). As an example of the type of object that we wish to study, let us consider the following boundary value problem: we aim to find a real-valued space-time function $(t\,,x) \mapsto u_t(x)$, where $t \geqslant 0$ and $x \in [0\,,1]$, such that

$$\left[\begin{array}{ll} \dot{u}_t(x) = u_t''(x) + \sigma(u_t(x))\xi_t(x) & \text{for } t > 0 \text{ and } 0 < x < 1, \\ u_0(x) = \sin(2\pi x) & \text{for } 0 < x < 1, \\ u_t(0) = u_t(1) = 0 & \text{for all } t > 0. \end{array}\right. \tag{13.1}$$

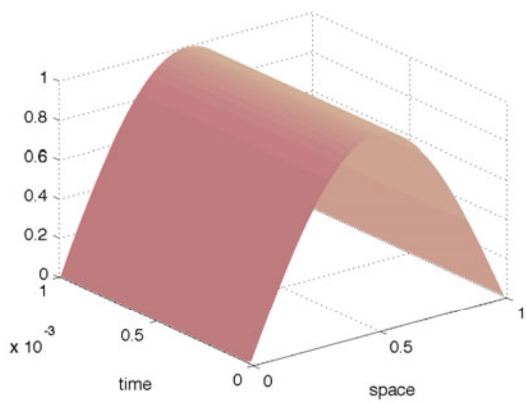
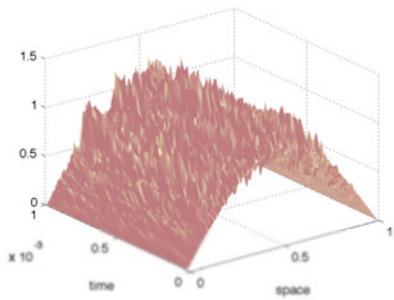


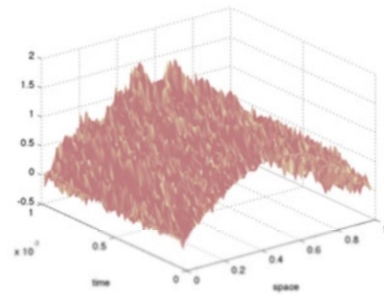
Figure 13.1: A numerical evaluation of the heat equation, where $\sigma(u) \equiv 0$.

We have written $u_t(x)$ in place of the more commonplace notation $u(t\,,x)$, as it is more natural in the probabilistic context. Thus, u_t designates the map $t \mapsto u$ and not the time derivative $\partial u/\partial t$.

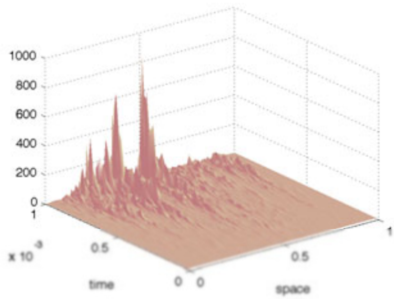
If $\sigma\colon \mathbb{R} \to \mathbb{R}$ and $\xi\colon [0\,,\infty)\times\mathbb{R} \to \mathbb{R}$ are sufficiently smooth then the preceding is a classical problem of the theory of heat flow, the solution exists, is unique, and has good regularity properties; see, for instance, Evans [17, Chapter 2, §2.3].



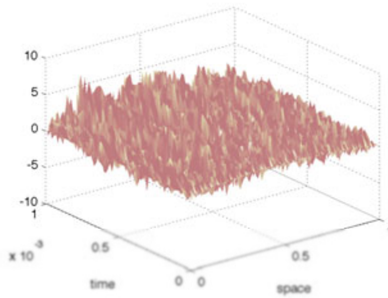
(a) A simulation of the stochastic heat equation where $\sigma(u) = u$.



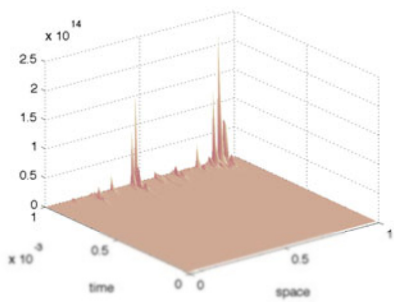
(b) A simulation of the stochastic heat equation where $\sigma(u) = 1$.



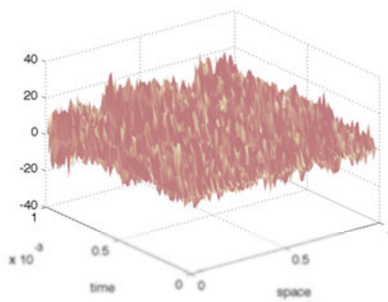
(c) A simulation of the stochastic heat equation where $\sigma(u) = 10u$.



(d) A simulation of the stochastic heat equation where $\sigma(u) = 10$.



(e) A simulation of the stochastic heat equation where $\sigma(u) = 50u$.



(f) A simulation of the stochastic heat equation where $\sigma(u) = 50$.

Figure 13.2: The left column consists of simulations of (13.1) where $\sigma(u) = \lambda u$, and the right column is for $\sigma(u) = \lambda$, as λ ranges in $\{1\,, 10\,, 50\}$.

Consider an ideal rod of length one unit, and identify the rod with the interval $[0\,,1]$. Suppose the rod is heated at time $t=0$ such that the heat density at every point $x\in[0\,,1]$ (x units along the rod) is $\sin(2\pi x)$. Then it can be argued using Fourier's law of thermal conduction that, under ideal conditions, the heat density $u_t(x)$ at place $x\in[0\,,1]$ and at time $t>0$ solves the linear heat equation $\dot{u}=\nu u''$, subject to $u_0(x)=\sin(2\pi x)$. Here, ν is a physical constant and is sometimes called "thermal conductivity", in this context.

We can always scale the problem so that $\nu=1$. Indeed, if $\dot{u}=\nu u''$ then $F_t(x):=u_{t/\nu}(x)$ solves the heat equation $\dot{F}=F''$, subject to the same initial and boundary conditions as u. In this way, we arrive at (13.1) with $\sigma:=\xi:=0$.

Suppose that the rod also feels external density $\xi_t(x)$ of heat (or cold, if $\xi_t(x)<0$) at the point $(t\,,x)$. Then, the heat density solves $\dot{u}=u''+\xi$. That is, (13.1) with $\sigma(u)\equiv 1$.

The general form of (13.1) arises when the external heating/cooling source interacts with the heat flow on the rod due to the presence of one or more feedback systems. In that case, the function σ models the nature of the feedback mechanism.

The main goal of these notes is to study the heat-flow problem (13.1) in the case whereξ denotes "space-time white noise" (a notion defined carefully below). For the time being, we can think of the $\xi_t(x)$'s as a collection of independent mean-zero normal random variables. In this sense, (13.1) describes heat flow in a random environment.

We plan to study how the solution depends on the nonlinearity σ. In order to motivate this, consider the simplest case that $\sigma\equiv 0$. In that case, we can solve u explicitly, and find that $u_t(x)=\exp(-4\pi^2 t)\sin(2\pi x)$ for all $t\geqslant 0$ and $x\in[0\,,1]$, when $\sigma(u)\equiv 0$.

Figure 13.1 shows a numerical evaluation of the solution for time values $t\in[0\,,10^{-3}]$. Figures 13.2(a) and 13.2(b) show typical simulations of the solution for $\sigma(u)=u$ and $\sigma(u)=1$, respectively. Figures 13.2(c) and 13.2(d) do the same thing for $\sigma(u)=10u$ and $\sigma(u)=10$, respectively. And Figures 13.2(e) and 13.2(f) for $\sigma(u)=50u$ and $\sigma(u)=50$. A quick inspection of these suggests that the behavior of the solution to (13.1) depends critically on the properties of the nonlinearity σ. In the last chapter of these notes, an answer on how this phenomenon can arise will be provided.

These notes are based on lectures given in the summer of 2014 at the *Second Summer School on Stochastic Analysis* held at the *Centre de Recerca Matemàtica* (CRM) in Barcelona. I would like to thank the CRM for their generous hospitality. Many hearty thanks are owed to the organizing and scientific committee, David Applebaum, Robert Dalang, Lłuis Quer-Sardanyons, Marta Sanz-Solé, Frederic Utzet, and Josep Vives for their kind invitation.

The material of these notes is based on my collaborations with Kunwoo Kim [33, 34], as well as Mathew Joseph and Carl Mueller [30]. I thank all three for many years of extremely enjoyable scientific discourse. Many thanks are due to

Pavel Bezdek, who pointed out a number of misprints and one error in an earlier draft. The remaining errors and misprints are of course mine.

My research was made possible thanks in part to a number of grants by the United States' *National Science Foundation*, in particular, DMS-1307470.

Salt Lake City (Utah), August, 2014 *Davar Khoshnevisan*

Chapter 14

White Noise

We begin our description of noisy equations with an analysis of the underlying noise. Throughout these notes, we use only space-time white noise. There is a very nice noise theory that is "white in time and colored in space", as very well described in the book of Sanz-Solé [51] on the Malliavin calculus. We will not cover such noises here, however. The material of this chapter borrows heavily from the paper [34] with Kunwoo Kim.

14.1 Some heuristics

The noise of choice in these notes in "white noise" on a nice space G such as $G = \mathbb{R}$ or $G = \mathbb{Z}$. In order to gel some of the basic ideas, consider $G = \mathbb{R}$. Then the intuitive definition of white noise on $\mathbb{R}$ is a mean-zero Gaussian process $\{\xi(x)\}_{x\in\mathbb{R}}$, indexed by $G = \mathbb{R}$, such that

$$\mathrm{Cov}[\xi(x)\,,\xi(y)] = \delta_0(x-y),$$

for all $x, y \in \mathbb{R}$. Of course, such a Gaussian process does not exist in the usual sense because $(x\,,y) \mapsto \delta_0(x-y)$ is not a function on $\mathbb{R}^2$. Regardless, if $x \mapsto \xi(x)$ *did* make sense as was prescribed (it does not!), and if all sorts of nice measurability conditions held (they do not either!), then it would follow that $f \mapsto \xi(f) := \int_{-\infty}^{\infty} f(x)\xi(x)\mathrm{d}x$ is a mean-zero Gaussian process with covariance

$$\mathrm{Cov}[\xi(f_1)\,,\xi(f_2)] = \int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\;f_1(x)f_2(y)\mathrm{Cov}[\xi(x)\,,\xi(y)],$$

which yields the following after we formally replace $\mathrm{Cov}[\xi(x)\,,\xi(y)]$ by $\delta_0(x-y)$:

$$\mathrm{Cov}[\xi(f_1)\,,\xi(f_2)] = \int_{-\infty}^{\infty} f_1(x)f_2(x)\,\mathrm{d}x. \tag{14.1}$$

Most of the preceding "analysis" is admittedly flawed. However, the end result, namely (14.1), turns out to make sense. In other words, we will be able

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_14

to construct a mean-zero Gaussian process $\varphi \mapsto \xi(\varphi)$ with covariance (14.1) for all $\varphi \in L^2(\mathbb{R})$. And if $\{\varphi_\epsilon\}_{\epsilon>0}$ is an approximation to the identity such that $\varphi_\epsilon \in L^2(\mathbb{R})$ and $\varphi_\epsilon(y) \to \delta_0(x-y)$, vaguely as $\epsilon \downarrow 0$,[1] then we may think of $\xi(\varphi_\epsilon)$ as an "approximation to $\xi(x)$", even though $\xi(x)$ itself is not a well-defined random variable.

As our first task, we plan to make rigorous the preceding discussion and thus define white noise as a mean-zero "generalized Gaussian random function" or, perhaps more appropriately, put a "random linear functional". For reasons that will become clear later on, we will benefit by studying white noise on more general objects than just the real line. Specifically, we will study white noise on "LCA groups".

14.2 LCA groups

Let G be a group with a multiplication operation, $(g_1, g_2) \mapsto g_1 g_2$, and an inversion operation, $g \mapsto g^{-1}$. We will always endow G with its *group topology*; that is, the smallest topology that makes the group operations (both multiplication and inversion) continuous. In this way, we can always view G as a *topological group*.

Let G_1 and G_2 be two topological groups. A mapping $h\colon G_1 \to G_2$ is a *homomorphism* if h is continuous and respects the group operations; namely that $h(xy) = h(x)h(y)$ and $h(x^{-1}) = (h(x))^{-1}$ for all $x, y \in G_1$. We say that $h\colon G_1 \to G_2$ is an *isomorphism* when h is a homomorphism with an inverse function $h^{-1}\colon G_2 \to G_1$ that is also a homomorphism. If we can find an isomorphism from G_1 to G_2 then we write $G_1 \cong G_2$ and might also say that we can *identify* G_1 with G_2. In words, $G_1 \cong G_2$ if and only if G_2 is a "relabeling" of G_1 in a way that is compatible with both the topology of G_1, as well as the group-theoretic properties of G_1.

We say that a topological group G is an *LCA group* if it is locally compact, Hausdorff, and abelian. Here are some examples of the sorts of LCA groups that we will be studying.

Example 14.2.1 (The trivial group)**.** Let $G := \{g\}$ be a set with one element. We can define $g^{-1} := g$ and $gg := g$ in order to see that G is an abelian group with identity g. If we endow G with the discrete topology (the only one possible here) then G becomes an LCA group that is, not surprisingly, referred to as the *trivial group*.

Example 14.2.2 (Cyclic groups)**.** Let $G := \{0, 1\}$ be endowed with the discrete topology. Then G is an LCA group in a standard way: we impose binary addition as our group multiplication on G; that is, $10 = 01 := 1$ and $00 = 11 := 0$, with

[1] More precisely, we want $\lim_{\epsilon\downarrow 0} \int_{-\infty}^{\infty} f(y)\, \varphi_\epsilon(y)\, \mathrm{d}y = f(x)$ for all $x \in \mathbb{R}$ and functions $f\colon \mathbb{R} \to \mathbb{R}$ that are continuous and have compact support.

0 being the group identity.[2] In this way, we find that $0^{-1} = 0$ and $1^{-1} = 1$, as well. The resulting LCA group can be identified with $\mathbb{Z}/2\mathbb{Z}$ and is called the *cyclic group on two elements*. The cyclic group on one element is $G := \mathbb{Z}/\mathbb{Z}$, which is defined to be the trivial group, which we just saw. In the more interesting case where $n \geqslant 2$, the cyclic group on n elements is $G := \mathbb{Z}/n\mathbb{Z}$, which can be thought of in the following way: $G = \{0, \dots, n-1\}$ and the group multiplication is given by addition mod n. The topology is, of course, the discrete topology. Moreover, $0^{-1} = n-1$, and $g^{-1} = (g-1)^{-1}$ for $g \in \{1, \dots, n-1\}$. This preceding should tell you why $\mathbb{Z}/n\mathbb{Z}$ is called a "cyclic" group. All cyclic groups are LCA groups.

Example 14.2.3 (Euclidean groups)**.** Let n be an integer greater than or equal to one. Then the additive group $\mathbb{R}^n$ is an LCA group, and the group topology is the Euclidean one. And so is the n-dimensional torus $\mathbb{T}^n := [0\,, 2\pi)^n$, once it is endowed with coordinatewise addition mod 2π. It is a good exercise to try and verify that the group topology on $\mathbb{T}^n$ is the usual Euclidean topology on $[0\,, 2\pi)^n$.

Example 14.2.4 (Positive reals)**.** We endow $G := (0\,, \infty)$ with group multiplication $ab := a \times b$ to see that G is an LCA group where the group identity is the numeral 1 and the group inverse g^{-1} is the reciprocal $1/g$ of $g \in G$. As was the case in the previous example as well, it is a good exercise to try and identify the group topology. Here too, as in the previous example, the topology is Euclidean.

Example 14.2.5 (Direct products)**.** Suppose G_a is an LCA group for every a in some index set A. The direct product $\prod_{a\in A} G_a$ is endowed with product topology, and the group operations on $\prod_{a\in G} G_a$ are defined coordinatewise in terms of the group operations of the G_a's. In this way we see that $\prod_{a\in A} G_a$ is an LCA group. If G is an LCA group then we may write G^n in place of $G \times \cdots \times G$ (n times) for every integer $n \geqslant 1$. We also might write G^0 for the trivial group and G^∞ for the *countable* direct product $\prod_{i=1}^\infty G$. A nice example of this sort of LCA group is $(\mathbb{Z}/2\mathbb{Z})^\infty$. Elements of this group are infinite sequences of 0s and 1s. Therefore, we can identify $(\mathbb{Z}/2\mathbb{Z})^\infty$ with a binary tree with a single root (or more carefully put, $(\mathbb{Z}/2\mathbb{Z})^\infty$ acts transitively on this binary tree). Somewhat more general remarks apply to $(\mathbb{Z}/n\mathbb{Z})^\infty$ (for n-ary trees).

General theory tells us that there exists a Radon Borel measure m_G on G such that $m_G(gA) = m_G(A)$ for all Borel sets $A \subseteq G$ and all elements $g \in G$. It might help to recall that: (i) "Radon" means $m_G(C) < \infty$ for all compact sets C; and (ii) $gA := \{ga \mid a \in A\}$ is the g-translate of A. The measure m_G is said to be *a Haar measure on G*. Moreover, we know also from the general theory that if m_G and m'_G are two Haar measures on G, then there exists $a \in (0\,, \infty)$ such that $m_G = am'_G$. In particular, there is a one-parameter family of possible Haar measures on G. As it turns out, however, there always is a canonical choice which will be the one that is denoted by m_G henceforth. I will not describe here the why or the how since that discussion will require topics that we have not covered; we

[2]In this example, "10" refers to the group product of the group elements 1 and 0; "10" does not refer to the numeral ten! Similar remarks apply to "00", "11," etc.

will see this later on. Instead, let me mention two examples where one can make this canonical choice more or less explicitly:

(i) if G is a discrete LCA group, then we always choose m_G to be the counting measure on G; and

(ii) when G is compact (as is, for example, a torus), then we always choose m_G to be a probability measure (also known as the "uniform distribution on G").

Rudin [49] has a more detailed account of Fourier analysis on LCA groups.

14.3 White noise on G

As in the previous section, let G denote an LCA group with a choice m_G for its Haar measure. Let $L^2(G)$ denote the usual space of all measurable functions $f\colon G \to \mathbb{R}$ such that

$$\int_G |f(x)|^2 \, m_G(\mathrm{d}x) < \infty,$$

endowed with the inner product

$$(f_1\,, f_2)_{L^2(G)} := \int_G f_1(x) f_2(x) \, m_G(\mathrm{d}x),$$

and corresponding norm

$$\|f\|_{L^2(G)} := (f\,, f)^{1/2}_{L^2(G)}.$$

We may begin with an existence theorem.

Theorem 14.3.1 (Wiener, [60, 61])**.** *On a suitable probability space, one can construct a mean-zero Gaussian process $\xi := \{\xi(f)\}_{f \in L^2(G)}$ with covariance*

$$\mathrm{Cov}[\xi(f_1)\,, \xi(f_2)] = (f_1\,, f_2)_{L^2(G)}, \tag{14.2}$$

for all $f_1, f_2 \in L^2(G)$.

The Gaussian process ξ is called *white noise on G*, and plays a central theme in these lectures. You might wish to compare the content of Theorem 14.3.1 with that of the discussion that surrounds (14.1).

Proof of Theorem 14.3.1. It is easy to see that the inner product of $L^2(G)$ is positive semidefinite bi-linear form on $L^2(G)$. This is another way to state that

$$\sum_{j=1}^n \sum_{k=1}^n z_j \overline{z_k} (f_j\,, f_k)_{L^2(G)} \geqslant 0,$$

for every $f_1, \dots, f_n \in L^2(G)$ and $z_1, \dots, z_n \in \mathbf{C}$. Indeed, a direct computation reveals that

$$\sum_{j=1}^{n}\sum_{k=1}^{n} z_j \overline{z_k} (f_j\,, f_k)_{L^2(G)} = \left\| \sum_{j=1}^{n} z_j f_j \right\|_{L^2(G)}^2 .$$

Consequently, $((f_i\,, f_j)_{L^2(G)})_{i,j=1}^n$ is a positive semidefinite n-by-n matrix for every $F := \{f_1, \dots, f_n\} \subset L^2(G)$.

Let $\Omega := [0\,,1]$ and $\mathscr{F}$:= the Borel σ-algebra on Ω. Because of the asserted positive semidefinite property, the theory of multivariate normal distributions on $\mathbb{R}^n$ tells us that for every $F := \{f_1, \dots, f_n\} \subset L^2(G)$ we can find a probability space $(\Omega^F, \mathscr{F}^F, \mathrm{P}^F)$ which supports a mean-zero Gaussian random vector $\xi := (\xi(f_1)\,, \dots, \xi(f_n))$ such that

$$\mathrm{Cov}[\xi(f_i)\,,\xi(f_j)] = (f_j\,, f_j)_{L^2(G)},$$

for all $1 \leqslant i, j \leqslant n$; see Chapter 1 of Ash–Gardner [2], for a detailed account. One checks directly that the P^F's form a consistent family of probability measures. This means the following: let $\Omega_* := \Omega^{L^2(G)}$, and endow it with the product topology and corresponding Borel σ-algebra $\mathscr{F}^{L^2(G)}$. Whenever F_1 and F_2 are finite subsets of $L^2(G)$ and $F_1 \subset F_2$,

$$\mathrm{P}^{F_2} \circ \pi_{F_2} \circ \pi_{F_1}^{-1} = \mathrm{P}^{F_1},$$

where π_F denotes the canonical projection from $\Omega^{L^2(G)}$ onto Ω^F for every finite subset $F \subset L^2(G)$. Then Kolmogorov's consistency theorem ensures the existence of a probability measure P on $(\Omega^{L^2(G)}, \mathscr{F}^{L^2(G)})$ such that $\mathrm{P} \circ \pi_F^{-1} = \mathrm{P}^F$ for every finite $F \subset L^2(G)$. This is another way to state the theorem. □

Now we can derive the central property of white noise on G.

Proposition 14.3.2 (Wiener, [60, 61])**.** *Let ξ denote white noise on G. Then for all real numbers $a_1, \dots, a_n$ and functions $f_1, \dots, f_n \in L^2(G)$,*

$$\xi\left(\sum_{j=1}^{n} a_j f_j\right) = \sum_{j=1}^{n} a_j \xi(f_j) \qquad a.s.$$

Proof. By induction it suffices to prove that, for all $a \in \mathbb{R}$ and $f, h \in L^2(G)$,

(i) $\xi(af) = a\xi(f)$ a.s.; and
(ii) $\xi(f + h) = \xi(f) + \xi(h)$ a.s.

We verify (i) by appealing to (14.2) in order to see that

$$\mathrm{E}\left(|\xi(af) - a\xi(f)|^2\right) = 0. \tag{14.3}$$

Similarly, one can prove (ii) by showing that

$$\mathrm{E}\left(|\xi(f + h) - \xi(f) - \xi(h)|^2\right) = 0. \tag{14.4}$$

This requires only (14.2). □

Exercise 14.3.3. Verify the details of formulas (14.3) and (14.4).

Theorem 14.3.1 and Proposition 14.3.2 together tell us that $\xi\colon L^2(G) \to L^2(\Omega)$ is a linear isometry. This is a way to say that $f \mapsto \xi(f)$ is a "random linear functional". The relation (14.2) is referred to as the *Wiener isometry* for similar reasons.

If $f \in L^2(G)$, then the square-integrable random variable $\xi(f)$ is called the *Wiener integral of* f. We may write this, using alternative notations, as

$$\xi(f) := \int_G f \mathrm{d}\xi := \int_G f(x)\xi(\mathrm{d}x).$$

Proposition 14.3.2 is a way to say that $\int_G f\,\mathrm{d}\xi$ is an "$L^2(\Omega)$-valued integral". We might also write the "indefinite integral"

$$\int_A f\,\mathrm{d}\xi$$

in place of

$$\int_G f\mathbf{1}_A\,\mathrm{d}\xi,$$

when A is a Borel subset of G. Define

$$\mathscr{A}(G) := \{B \subset G \mid B \text{ is Borel measurable and } m_G(B) < \infty\}.$$

It is easy to verify that $\mathscr{A}(G)$ is an algebra of Borel-measurable subsets of G. Then we abuse notation slightly and write $\xi(B)$ in place of $\xi(\mathbf{1}_B)$ for every $B \in \mathscr{A}(G)$. I have learned the following from Walsh [57].

Proposition 14.3.4. *The following is true:*

(i) *if* $B_1, B_2 \in \mathscr{A}(G)$*, then* $\xi(B_1 \cup B_2) = \xi(B_1) + \xi(B_2) - \xi(B_1 \cap B_2)$ *a.s.; and*
(ii) *if* $B_1 \supset B_2 \supset \cdots$ *are elements of* $\mathscr{A}(G)$ *such that* $\cap_{n=1}^\infty B_n = \varnothing$*, then* $\lim_{n\to\infty} \xi(B_n) = 0$ *in* $L^2(\Omega)$.

Proof. We will require the following immediate consequence of the Wiener isometry (14.2):

$$\mathrm{E}\left[\xi(B)\xi(B')\right] = m_G(B \cap B'),$$

for all $B, B' \in \mathscr{A}(G)$. From this, we can conclude readily that

$$\mathrm{E}\left(|\xi(B_1 \cup B_2) - \xi(B_1) - \xi(B_2)|^2\right) = 0,$$

whenever B_1, B_2 are disjoint elements of $\mathscr{A}(G)$. The general case of (i) reduces to this one by induction. The proof of (ii) is even quicker, since $\mathrm{E}(|\xi(B_n)|^2) = m_G(B_n)$ for every $n \geqslant 1$. □

It is easy to deduce from the preceding proposition that if $B_1, B_2, \ldots$ are disjoint (non-random) elements of $\mathscr{A}(G)$ such that $\cup_{n=1}^{\infty} B_n \in \mathscr{A}(G)$, then

$$\xi\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \xi(B_n) \qquad \text{a.s.},$$

where the sum converges in $L^2(\Omega)$. This is another way to say that ξ is "an $L^2(\Omega)$-valued measure". In other words, we have identified the "$L^2(\Omega)$-valued integral",

$$f \longmapsto \xi(f) := \int_G f \mathrm{d}\xi,$$

with the $L^2(\Omega)$-valued measure ξ in the same way that one can identify a real-valued integral with a real-valued measure. Among other things, this justifies our abuse of notation in writing ξ both for the random linear functional $f \mapsto \xi(f)$ and for the infinite-dimensional measure $A \mapsto \xi(A)$.

14.4 Space-time white noise

We continue to write G for an LCA group, but now consider the LCA group $\mathbb{R} \times G$ as well. Because of the underlying product topology, it is not hard to argue that the Haar measure on $\mathbb{R} \times G$ is the product measure $m_{\mathbb{R}\times G}(\mathrm{d}x\, \mathrm{d}y) = \mathrm{d}x \times m_G(\mathrm{d}y)$.

Let $\tilde{\xi}$ denote a white noise on $\mathbb{R} \times G$. Then we say that ξ is *space-time white noise* on $\mathbb{R}_+ \times G$ when, viewed as an $L^2(\Omega)$-valued measure, ξ is the restriction of $\tilde{\xi}$ to $\mathbb{R}_+ \times G$; that is,

$$\xi(A) := \tilde{\xi}\left(A \cap [\mathbb{R}_+ \times G]\right),$$

for all Borel sets $A \subset \mathbb{R} \times G$ of finite Haar measure. Equivalently, we can consider Wiener integrals: for all $f \in L^2(\mathbb{R}_+ \times G)$,

$$\int_{\mathbb{R}_+\times G} f\, \mathrm{d}\xi := \int_{\mathbb{R}\times G} \tilde{f}\, \mathrm{d}\tilde{\xi},$$

where $\tilde{f}(t\,, x) := f(t\,, x)$ if $t \geqslant 0$ and $\tilde{f}(t\,, x) := 0$ when $t < 0$.

The Wiener integral $\int_{\mathbb{R}_+\times G} f\, \mathrm{d}\xi$ generalizes Wiener integrals that you are likely to have seen before.

Example 14.4.1 (Group on one element)**.** Consider the case that $G := \{\mathrm{e}\}$ is the trivial group. Then $L^2(\{\mathrm{e}\}) \cong \mathbb{R}$ isometrically, as Hilbert spaces. It is easy to see from (14.2) that $B_t := \xi([0\,, t] \times \{\mathrm{e}\})$ defines a Brownian motion, $t \geqslant 0$. Also, $L^2(\mathbb{R}_+ \times \{\mathrm{e}\}) \cong L^2(\mathbb{R}_+)$. Therefore, $f \in L^2(\mathbb{R}_+ \times \{\mathrm{e}\})$ is identified with $t \mapsto f(t)$, and $\int_{\mathbb{R}_+\times\{\mathrm{e}\}} f\, \mathrm{d}\xi$ can be written more compactly as $\int_0^\infty f(t)\, \mathrm{d}B_t$; the latter is sometimes called the *Wiener integral* of the (non-random) function f against Brownian motion B.

Example 14.4.2 (Group on n elements)**.** Consider $G = \{0, \ldots, n-1\}$, as a cyclic group on $n \geqslant 1$ elements. If $f \in L^2(\mathbb{R}_+ \times G)$, then $t \mapsto f(t, \bullet)$ is a square-integrable function from $\mathbb{R}_+$ to $\mathbb{R}^n$; that is, it is an element of $L^2(\mathbb{R}_+, \mathbb{R}^n)$. And conversely, we can identify (isometrically) $L^2(\mathbb{R}_+, \mathbb{R}^n)$ with $L^2(\mathbb{R}_+ \times G)$; in symbols, $L^2(\mathbb{R}_+ \times G) \cong L^2(\mathbb{R}_+, \mathbb{R}^n)$.

Let $B_t := (B_t^{(1)}, \ldots, B_t^{(n)})$, where $B_t^{(i)} := \xi([0, t] \times \{i\})$ for $t \geqslant 0$ and $i = 0, \ldots, n-1$. According to the Wiener isometry (14.2), $B := \{B_t\}_{t \geqslant 0}$ is an n-dimensional Brownian motion. Let $f \in L^2(\mathbb{R}_+, \mathbb{R}^n)$; then people frequently write

$$\int_0^\infty f(t) \cdot \mathrm{d}B_t = \sum_{i=0}^{n-1} \int_0^\infty f(t, i)\, \mathrm{d}B_t^{(i)}$$

in place of $\int_{\mathbb{R}_+ \times G} f\, \mathrm{d}\xi$. Frequently, people refer to $\int_0^\infty f(t) \cdot \mathrm{d}B_t$ as the *Wiener integral* of $f \in L^2(\mathbb{R}_+, \mathbb{R}^n)$ against the n-dimensional Brownian motion B.

As one can check directly, (14.2) implies that, whenever $\varphi \in L^2(G)$,

$$B_t(\varphi) := \int_{[0,t] \times G} \varphi(y)\, \xi(\mathrm{d}s\, \mathrm{d}y) \qquad (t \geqslant 0)$$

defines a Brownian motion, scaled to have variance $t\|\varphi\|^2_{L^2(G)}$ at time t. Let $\{\mathscr{F}_t^\varphi\}_{t \geqslant 0}$ denote the filtration of the Brownian motion $t \mapsto B_t(\varphi)$. Define

$$\mathscr{F}_t^* := \bigvee_{\varphi \in L^2(G)} \mathscr{F}_t^\varphi \qquad (t \geqslant 0).$$

One can check that $t \mapsto B_t(\varphi)$ continues to be a Brownian motion in the filtration $\{\mathscr{F}_t^*\}_{t \geqslant 0}$. Let $\overline{\mathscr{F}}_t$ denote the P-completion of $\mathscr{F}_t^*$ for every $t \geqslant 0$. Standard methods show that $t \mapsto B_t(\varphi)$ is also a Brownian motion in the filtration $\mathscr{F} := \{\mathscr{F}_t\}_{t \geqslant 0}$, where $\mathscr{F}_t$ is the *right-continuous extension of* $\overline{\mathscr{F}}_t$ for every $t \geqslant 0$. That is,

$$\mathscr{F}_t := \bigcap_{s > t} \overline{\mathscr{F}}_s \qquad (t \geqslant 0).$$

For these reasons, the filtration $\mathscr{F}$ is called *the Brownian filtration* corresponding to the space-time white noise ξ.

14.5 The Walsh stochastic integral

Let G continue to designate an LCA group, and ξ space-time white noise on $\mathbb{R}_+ \times G$. We now extend the domain of the definition of the Wiener integral $\int_{\mathbb{R}_+ \times G} f\, \mathrm{d}\xi$ in order to include some *random functions* f as well. The resulting stochastic integral is then called the *Walsh integral.*

To be more concrete, we wish to integrate now a function $\Phi(t\,,x)(\omega)$ of three variables against $\xi(\mathrm{d}t\,\mathrm{d}x)(\omega)$, where the variables are: $t \geqslant 0$ (time variable), $x \in G$ (space variable), and $\omega \in \Omega$ (probability variable). We do this by following Walsh [57] and Itô [26] in a few simple steps.[3]

14.5.1 Simple random fields

We say that $\Phi\colon \mathbb{R}_+ \times G \times \Omega \to \mathbb{R}$ is a *simple random field* if there exists positive real numbers $a < b$, a random variables $X \in L^2(\Omega\,,\mathscr{F}_a)$, and a function $\varphi \in C_c(G)$ such that

$$\Phi(t\,,x)(\omega) = X(\omega)\varphi(x)\mathbf{1}_{(a,b]}(t), \tag{14.5}$$

for every $t > 0$, $x \in G$, and $\omega \in \Omega$.[4] We might also refer to the time interval $(a\,,b]$ as the *temporal support* of Φ.

If Φ is a simple random field with the preceding representation, then we define the stochastic integral $\int_{\mathbb{R}_+\times G} \Phi\,\mathrm{d}\xi$, ω by ω, as

$$\xi(\Phi) := \int_{\mathbb{R}_+\times G} \Phi\,\mathrm{d}\xi := X \cdot \int_{(a,b]\times G} \varphi(x)\,\xi(\mathrm{d}t\,\mathrm{d}x),$$

where the integral with respect to ξ is the Wiener integral of the preceding sections. Thanks to (14.2), the random variable $\int_{(a,b]\times G} \varphi(x)\,\xi(\mathrm{d}t\,\mathrm{d}x)$ is uncorrelated – whence independent – from $\int_{[0,a]\times G} h(x)\,\xi(\mathrm{d}t\,\mathrm{d}x)$ for all $h \in L^2(G)$; therefore, the Wiener integral $\int_{(a,b]\times G} \varphi(x)\,\xi(\mathrm{d}t\,\mathrm{d}x)$ is independent from the σ-algebra $\mathscr{F}_a$. Since X is measurable with respect to $\mathscr{F}_a$, it follows that $\int_{(a,b]\times G} \varphi(x)\,\xi(\mathrm{d}t\,\mathrm{d}x)$ and X are independent. In particular,

$$\mathrm{E}[\xi(\Phi)] = 0, \tag{14.6}$$

and

$$\mathrm{E}\left([\xi(\Phi)]^2\right) = \|\Phi\|^2_{L^2(\mathbb{R}_+\times G\times\Omega)}, \tag{14.7}$$

where we are writing $\Phi_t(x) := \Phi(t\,,x)$, as is standard in the theory of stochastic processes. (We will, in fact, write our space-time stochastic processes this way very often from now on.)

Exercise 14.5.1. Verify (14.7).

---

[3] There are other definitions of the "Itô stochastic integral in infinite dimension" than those introduced here. Other approaches can be found, for example, in Chow [6], Kotelenez [35], Krylov [36, 37], Krylov–Rozovskii [38], Mikulevicious–Rozovskii [39], da Prato–Zabczyk [11], and Prévôt–Röckner [45]. For examples of concrete families of important SPDEs see Corwin [9], Dawson–Perkins [15], Giacomin–Lebowitz–Presutti [22], and Quastel [46] and their combined bibliographies. As far as the central bulk of the theory is concerned, all of these approaches are essentially equivalent; see Dalang–Quer-Sardanyons [14] for details.

[4] As is customary, $C_c(G)$ denotes the space of all continuous functions $f\colon G \to \mathbb{R}$ that have compact support. It is well known that $C_c(G)$ is always dense in $L^2(G)$; see for example Rudin [49, **E8**, p. 268]. For this reason, we could also define our simple function with $\varphi \in L^2(G)$ without altering the end result, which is the Walsh integral.

14.5.2 Elementary random fields

We say that $\Phi\colon \mathbb{R}_+ \times G \times \Omega \to \mathbb{R}$ is an *elementary random field* if there exist simple random fields $\Phi^{(1)}, \ldots, \Phi^{(n)}$, with disjoint temporal supports, such that $\Phi = \sum_{i=1}^n \Phi^{(i)}$. In this case, we may define

$$\int_{\mathbb{R}_+ \times G} \Phi \, \mathrm{d}\xi := \sum_{i=1}^n \int_{\mathbb{R}_+ \times G} \Phi^{(i)} \, \mathrm{d}\xi.$$

We will denote by $\mathscr{E}$ the class of all elementary random fields.

The elementary properties of Wiener integrals show that our definition of the integral $\int_{\mathbb{R}_+ \times G} \Phi \, \mathrm{d}\xi$ a.s. does not depend on the choice of the $\Phi^{(i)}$'s. In other words, if we could write

$$\Phi = \sum_{i=1}^n \Phi^{(i)} = \sum_{j=1}^m \tilde{\Phi}^{(j)}$$

as a finite sum of simple random fields of disjoint temporal support in two different ways, then

$$\sum_{i=1}^n \int_{\mathbb{R}_+ \times G} \Phi^{(i)} \, \mathrm{d}\xi = \sum_{j=1}^m \int_{\mathbb{R}_+ \times G} \tilde{\Phi}^{(j)} \, \mathrm{d}\xi \qquad \text{a.s.}$$

Furthermore, one checks directly that (14.6) and (14.7) continue to hold when $\Phi \in \mathscr{E}$.

14.5.3 Walsh-integrable random fields

The identity (14.7) is called the *Walsh isometry*, and shows that the Walsh integral operator $\mathscr{E} \ni \Phi \mapsto \xi(\Phi) \in L^2(\Omega)$ is a linear isometry.

As a corollary of this development, consider a sequence $\{\Phi^{(n)}\}_{n=1}^\infty$ of elements of $\mathscr{E}$ that is a Cauchy sequence in $L^2(\mathbb{R}_+ \times G \times \Omega)$. By completeness, $\Phi := \lim_{n\to\infty} \Phi^{(n)}$ exists in $L^2(\mathbb{R}_+ \times G \times \Omega)$. According to the Walsh isometry, $\{\xi(\Phi^{(n)})\}_{n=1}^\infty$ is Cauchy in $L^2(\Omega)$. Therefore,

$$\xi(\Phi) := \lim_{n\to\infty} \xi(\Phi^{(n)})$$

exists in $L^2(\Omega)$ and satisfies (14.6) and (14.7). We have proved the following.

Theorem 14.5.2 (Walsh, [57]). *Let $\mathscr{W}$ denote the completion of $\mathscr{E}$ with respect to the $L^2(\mathbb{R}_+ \times G \times \Omega)$-norm. Then, $\xi\colon \mathscr{W} \to L^2(\Omega)$ is a linear isometry satisfying* (14.6) *and* (14.7).

Let us mention the following ready fact as well.

Proposition 14.5.3. *The class of all non-random elements of $\mathscr{W}$ agrees with all of $L^2(\mathbb{R}_+ \times G)$. Moreover, the Walsh integral of $\Phi \in L^2(\mathbb{R}_+ \times G)$ is the same as the Wiener integral of Φ a.s.*

Here is the idea of the proof: if Φ is a non-random simple function, then one sees immediately that the Walsh and Wiener integrals agree a.s. The general case follows by approximation.

From now on, we refer to the elements of $\mathscr{W}$ as *Walsh-integrable random fields*, and (14.7) as the *Walsh isometry*, valid for all $\Phi \in \mathscr{E}$. We also might write, interchangeably,

$$\xi(\Phi) := \int_{\mathbb{R}_+ \times G} \Phi \, \mathrm{d}\xi := \int_{\mathbb{R}_+ \times G} \Phi_t(x)\, \xi(\mathrm{d}t\, \mathrm{d}x) := \int_{\mathbb{R}_+ \times G} \Phi(t\,, x)\, \xi(\mathrm{d}t\, \mathrm{d}x).$$

As we did for Wiener integrals, whenever $\Lambda \subset \mathbb{R}_+ \times G$, we may also write

$$\int_\Lambda \Phi_t(x)\, \xi(\mathrm{d}t\, \mathrm{d}x) := \int_\Lambda \Phi(t\,, x)\, \xi(\mathrm{d}t\, \mathrm{d}x) := \int_\Lambda \Phi \, \mathrm{d}\xi := \xi(\Phi \mathbf{1}_\Lambda),$$

for all $\Phi \in \mathscr{W}$.

The Walsh integral is a strict generalization of the so-called *Itô integral* of stochastic calculus. The following two examples hash out some of the details for respectively 1-dimensional and n-dimensional Itô calculus of Brownian motion.

Example 14.5.4 (The trivial group)**.** Consider the case that $G = \mathbb{Z}/\mathbb{Z}$ is the trivial group. Let e denote the identity element of $\mathbb{Z}/\mathbb{Z}$, and recall that $\mathbb{R}_+ \ni t \mapsto B_t := \xi([0\,, t] \times \{\mathrm{e}\})$ defines a standard Brownian motion. In stochastic calculus terms, the class $\mathscr{W}$ of all Walsh-integrable random fields coincides with the family of all B-predictable processes X such that $\mathrm{E} \int_0^\infty X_s^2 \, \mathrm{d}s < \infty$ in the following sense: if $\Phi \in \mathscr{W}$, then $X_t := \Phi(t\,, \mathrm{e})$ is predictable and $\mathrm{E} \int_0^\infty X_s^2 \, \mathrm{d}s < \infty$, and vice versa. The integral $\int_{\mathbb{R}_+ \times \{\mathrm{e}\}} \Phi(s\,, y)\, \xi(\mathrm{d}s\, \mathrm{d}y)$ is usually written as $\int_0^\infty X_s \, \mathrm{d}B_s$, and is called the *Itô integral of* X.

Example 14.5.5 (Cyclic groups)**.** Let $G := \mathbb{Z}/n\mathbb{Z} \cong \{0\,, \ldots, n-1\}$ for some integer $n \geqslant 1$ and recall that $\mathbb{R}_+ \ni t \mapsto B_t^{(i)} := \xi([0\,, t] \times \{i\})$ defines standard n-dimensional Brownian motion, $i = 0, \ldots, n-1$. In stochastic calculus terms, the class of all $\Phi \in \mathscr{W}$ coincides with the family of all B-predictable processes taking values in $\mathbb{R}^n$ and satisfying $\mathrm{E} \int_0^\infty \|X_s\|^2 \, \mathrm{d}s < \infty$. Properties of Walsh integrals show that, since G is finite,

$$\int_{\mathbb{R}_+ \times G} \Phi(s\,, y)\, \xi(\mathrm{d}s\, \mathrm{d}y) = \sum_{i=0}^{n-1} \int_{\mathbb{R}_+ \times \{i\}} \Phi(s\,, i)\, \xi(\mathrm{d}s\, \mathrm{d}y),$$

for every $\Phi \in \mathscr{W}$. Define for all $i = 0, \ldots, n-1$ and any Borel set $A \subset \mathbb{R}_+$ of finite Lebesgue measure, $\xi^{(i)}(A) := \xi(A \times \{i\})$. It is easy to see that $\xi^{(1)}, \ldots, \xi^{(n-1)}$ are i.i.d. white noises on $\mathbb{R}_+$, the process $s \mapsto \Phi(s\,, i)$ is Walsh integrable with respect to $\xi^{(i)}$ for every i, and

$$\int_{\mathbb{R}_+ \times \{i\}} \Phi(s\,, i)\, \xi(\mathrm{d}s\, \mathrm{d}y) = \int_0^\infty \Phi(s\,, i)\, \xi^{(i)}(\mathrm{d}s) \qquad a.s.$$

for every $i = 0, \ldots, n-1$ (it suffices to prove the preceding for a simple random field Φ, in which case the identity is a tautology). Therefore,

$$t \mapsto B_t^{(i)} := \xi^{(i)}([0\,, t])$$

defines i.i.d. Brownian motions, as i ranges over $\{0, \ldots, n-1\}$ and, in terms of Itô integrals,

$$\int_{\mathbb{R}_+ \times G} \Phi(s\,, y)\, \xi(\mathrm{d}s\, \mathrm{d}y) = \sum_{i=0}^{n-1} \int_0^\infty X_s^{(i)}\, \mathrm{d}B_s^{(i)} := \int_0^\infty X_s \cdot \mathrm{d}B_s,$$

where $X_s^{(i)} := \Phi(s\,, i)$, $i = 0, \ldots, n-1$, defines a B-predictable process.

14.6 Moment inequalities

If Φ is a simple random field, then one sees directly that

$$t \longmapsto M_t(\Phi) := \int_{[0,t]\times G} \Phi\, \mathrm{d}\xi \tag{14.8}$$

defines a continuous $L^2(\Omega)$-martingale with quadratic variation process

$$t \longmapsto \langle M(\Phi)\rangle_t := \int_0^t \|\Phi_s\|^2_{L^2(G)}\, \mathrm{d}s. \tag{14.9}$$

The same fact will therefore hold when $\Phi \in \mathscr{E}$.

If $\Phi \in \mathscr{W}$, then we can find $\Phi^{(n)} \in \mathscr{E}$ such that $\Phi^{(n)}$ converges to Φ in $L^2(\mathbb{R}_+ \times G \times \Omega)$, as $n \to \infty$. And $\xi(\Phi^{(n)}) \to \xi(\Phi)$ in $L^2(\Omega)$ as $n \to \infty$, thanks to the Walsh isometry (14.7). Moreover, by Doob's maximal inequality for continuous $L^2(\Omega)$-martingales,

$$\begin{aligned} \mathrm{E}\left(\sup_{t\geqslant 0}\left|M_t(\Phi^{(n)}) - M_t(\Phi^{(m)})\right|^2\right) &\leqslant 4\sup_{t\geqslant 0}\mathrm{E}\left(\left|M_t\left(\Phi^{(n)} - \Phi^{(m)}\right)\right|^2\right) \\ &= 4\|\Phi^{(n)} - \Phi^{(m)}\|^2_{L^2(\mathbb{R}_+\times G\times\Omega)}, \end{aligned}$$

which goes to zero as $n, m \to \infty$. It follows that for all $\Phi \in \mathscr{W}$, (14.8) defines a continuous $L^2(\Omega)$-martingale. Moreover, we can see from this that

$$\int_{\mathbb{R}_+\times G} \Phi\, \mathrm{d}\xi = \lim_{t\to\infty} M_t(\Phi) \quad \text{in probability,}$$

whence

$$\mathrm{E}\left(\left|\int_{\mathbb{R}_+\times G} \Phi\, \mathrm{d}\xi\right|^k\right) \leqslant \sup_{t\geqslant 0}\mathrm{E}\left(|M_t(\Phi)|^k\right),$$

for all real numbers $k \geqslant 0$ thanks to Fatou's lemma. In this way, we obtain the following from (14.9) and the Burkholder–Davis–Gundy inequality for continuous $L^2(\Omega)$-martingales; see Revuz–Yor [47, Chapter IV], for instance.

Proposition 14.6.1. *For every real number $k \geqslant 2$ there exists a finite constant c_k such that*

$$\mathrm{E}\left(\left|\int_{\mathbb{R}_+\times G} \Phi\,\mathrm{d}\xi\right|^k\right) \leqslant c_k \mathrm{E}\left(\left|\int_0^\infty \|\Phi_s\|_{L^2(G)}^2\,\mathrm{d}s\right|^{k/2}\right),$$

for every $\Phi \in \mathscr{W}$.

Remark 14.6.2. It can be shown that $c_k \leqslant (4k)^{k/2}$ for all $k \in [2\,,\infty)$; see Carlen–Krée [4]. This bound is optimal too; in fact, if $c_{k,\mathrm{opt}}$ denotes the smallest such choice of c_k, then Carlen–Krée [4] shows that

$$\lim_{k\to\infty} \frac{c_{k,\mathrm{opt}}^{1/k}}{\sqrt{k}} = 2.$$

In these notes we will not require this bound on c_k. However, the following exercise hints at the fact that the multiplicative factor of $k^{k/2}$ can be useful in obtaining large-deviations bounds for Walsh integrals.

Exercise 14.6.3. Let W be a non-negative random variable, and suppose that there exists a finite constant a such that $\mathrm{E}[W^k] \leqslant a^k k^{k/2}$, for all integers $k \geqslant 2$. Prove that there exists a finite constant θ such that $\mathrm{P}\{W > \lambda\} \leqslant \theta \exp(-\lambda^2/\theta)$ for all $\lambda > 1$. (Hint: It suffices to prove that $\mathrm{E}\exp(\varepsilon W^2) < \infty$ for some $\varepsilon > 0$.)

14.7 Examples of Walsh-integrable random fields

One of the main themes of this chapter is that if Φ is a Walsh-integrable random field, $\Phi \in \mathscr{W}$, then we can integrate it against space-time white noise. A question that arises naturally at this point is, "what do Walsh-integrable random fields look like"? We explore some answers to this question next.

14.7.1 Integral kernels

First of all, the Walsh integral is a strict generalization of the integral of Wiener; in other words, $L^2(\mathbb{R}_+ \times G) \subset \mathscr{W}$ and $\int_{\mathbb{R}_+\times G} \Phi\,\mathrm{d}\xi$ is the Wiener integral of $\Phi \in L^2(\mathbb{R}_+ \times G)$. This assertion is obvious when $\Phi \in L^2(\mathbb{R}_+ \times G)$ is an elementary function in the sense of Lebesgue. We take limits in $L^2(\Omega)$ of such stochastic integrals for general $\Phi \in L^2(\mathbb{R}_+ \times G)$. This provides an answer to our question, but the answer is not satisfactory, since elements of $L^2(\mathbb{R}_+ \times G)$ are non-random.

Our strategy is the following: suppose we have some random field $\Phi \in \mathscr{W}$ such as $\Phi \in L^2(\mathbb{R}_+ \times G)$. Then we can try to build a new random field in $\mathscr{W}$ by choosing a good non-random "kernel" K such that:

(i) $(K\,,\Phi)_t(x) := \int_{[0,t]\times G} K_{t,s}(x\,,y)\Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y)$, ($t>0$, $x\in G$), and $(K\,,\Phi)_0(x) := 0$, is a Walsh integral for every $t \geqslant 0$ and $x \in G$; and

(ii) $(K\,,\Phi)$ itself is a Walsh-integrable random field.

As we will now see, this procedure sometimes works to produce bona fide random fields in $\mathscr{W}$.

First, suppose Φ has the following particular form:

$$\Phi_s(y) = X \cdot \mathbf{1}_{(a,b]}(s)\varphi(y), \tag{14.10}$$

where $0 < a < b$, $X \in L^2(\Omega\,,\mathscr{F}_a)$, and $\varphi \in C_c(G)$. In particular, Φ is a simple random field in the sense of (14.5). Consider also a simple kernel of the form

$$K_{t,s}(x\,,y) := \alpha\mathbf{1}_{(\beta,\gamma]}(s)\mathbf{1}_{(\beta',\gamma']}(t)\mathbf{1}_Q(x)\psi(y), \tag{14.11}$$

where $\alpha \in \mathbb{R}$, $0 < \beta < \gamma \leqslant \beta' < \gamma'$, $\psi \in L^\infty(G)$, and Q and Q' are compact subsets of G. In this particular case,

$$K_{t,s}(x\,,y)\Phi_s(y) = \alpha\mathbf{1}_{(\beta',\gamma']}(t)\mathbf{1}_Q(x)X \cdot \mathbf{1}_{(\beta,\gamma]\cap(a,b]}(s) \cdot (\varphi\psi)(y).$$

Since $\alpha\mathbf{1}_{(\beta',\gamma']}(t)\mathbf{1}_Q(x)X$ is in $L^2(\Omega\,,\mathscr{F}_a)$ and $\varphi\psi \in L^2(G)$, $(s\,,y) \mapsto K_{t,s}(x\,,y)$ is a simple random field for every $t \geqslant 0$ and $x \in G$. Consequently, $(K\,,\Phi)_t(x)$ is a bona fide Walsh integral for every $t \geqslant 0$ and $x \in G$, and

$$(K\,,\Phi)_t(x) = \alpha\mathbf{1}_{(\beta',\gamma']}(t)\mathbf{1}_Q(x) \cdot \tilde{X},$$

where $\tilde{X} := X \cdot \int_{(\beta,\gamma]\cap(a,b]} \varphi(y)\psi(y)\,\xi(\mathrm{d}s\,\mathrm{d}y)$.

Clearly, $\tilde{X} \in L^2(\Omega\,,\mathscr{F}_\gamma) \subset L^2(\Omega\,,\mathscr{F}_{\beta'})$; the last set inclusion holds since $\gamma \leqslant \beta'$. Therefore, $K\Phi$ is a simple random field in this case.

Let $\mathscr{W}_{00}$ denote the collection of all finite linear combinations of simple random fields of the form (14.10) having disjoint temporal support. To be sure, elements of $\mathscr{W}_{00}$ are elementary random fields; that is, $\mathscr{W}_{00} \subset \mathscr{E}$. Also, define $\mathscr{K}_{00}$ to be the collection of all finite linear combinations of kernels of the form (14.11) having disjoint supports. Then, clearly, $(K\,,\Phi)_t(x)$ is a well-defined Walsh integral for all $t \geqslant 0$ and $x \in G$, and $(K\,,\Phi) \in \mathscr{W}$. Furthermore, the Walsh isometry (14.7) tells us that

$$\begin{aligned}\mathrm{E}\left(|(K\,,\Phi)_t(x)|^2\right) &= \int_0^t \mathrm{d}s \int_G m_G(\mathrm{d}y)\ [K_{t,s}(x\,,y)]^2\mathrm{E}\left(|\Phi_s(y)|^2\right)\\ &\leqslant \mathcal{N}_0(\Phi) \cdot \sup_{t>0}\sup_{x\in G} \int_0^t \mathrm{d}s \int_G m_G(\mathrm{d}y)\ [K_{t,s}(x\,,y)]^2.\end{aligned}$$

The right-hand side does not depend on $(t\,,x)$. For all space-time random fields $X := \{X_t(x)\}_{t\geqslant 0, x\in G}$, let us define the norm

$$\mathcal{N}_0(X) := \sup_{t\geqslant 0}\sup_{x\in G}\left\{\mathrm{E}\left(|X_t(x)|^2\right)\right\}^{1/2}.$$

Also define, for every measurable function $F : (s\,, t\,; x\,, y) \to F_{t,s}(x\,, y)$, the norm

$$\mathcal{M}_0(F) := \sup_{t>0} \sup_{x\in G} \left[\int_0^t \mathrm{d}s \int_G m_G(\mathrm{d}y)\ [F_{t,s}(x\,, y)]^2 \right]^{1/2}.$$

In this way we find that

$$\mathcal{N}_0\left((K\,, \Phi)\right) \leqslant \mathcal{N}_0(\Phi)\mathcal{M}_0(K), \tag{14.12}$$

uniformly for all $\Phi \in \mathscr{W}_{00}$ and $K \in \mathscr{K}_{00}$. Now let $\mathscr{K}_0$ denote the completion of $\mathscr{K}_{00}$ in the norm $\mathcal{M}_0$, and $\mathscr{W}_0$ the completion of $\mathscr{W}_{00}$ in the norm $\mathcal{N}_0$. Since $(K\,, \Phi)$ is a (random) bilinear map on K and Φ, this procedure defines the random field $(t\,, x) \mapsto (K\,, \Phi)_t(x)$ for every $K \in \mathscr{K}_0$ and $\Phi \in \mathscr{W}_0$.

If $F, G \in \mathscr{K}_0$ are equal almost everywhere ($\mathrm{d}s \times \mathrm{d}M_G$) then we can identify F and G with one another. We will write $\mathscr{K}_0$ also for the family of the resulting equivalence classes (as one does in L^p-theory, for instance). In this way, we see that $(\mathscr{K}_0, \mathcal{M}_0)$ is a metric space. Similarly, if we identify two random fields that are modifications of one another (as one does in probability theory), then $(\mathscr{W}_0, \mathcal{N}_0)$ is easily seen to be a metric space.

Let Λ be a compact subset of $\mathbb{R}_+ \times G$, and define a kernel F via $F_{t,s}(x\,, y) := \mathbf{1}_\Lambda(s\,, y)\mathbf{1}_{(0,t)}(s)$ for all $s, t \geqslant 0$ and $x, y \in G$. Then it is easy to see that $F \in \mathscr{K}_0$. Let us summarize our findings.

Theorem 14.7.1. *Define $\mathscr{W}_{\mathrm{loc}}$ to be the collection of all space-time random fields Φ such that $\Phi\mathbf{1}_\Lambda \in \mathscr{W}$ for all compact non-random sets $\Lambda \subset \mathbb{R}_+ \times G$. Then, $L^\infty((0\,,\infty) \times G) \subset \mathscr{W}_0 \subset \mathscr{W}_{\mathrm{loc}}$. Moreover, $(K\,, \Phi)$ defines a bilinear form from $\mathscr{K}_0 \times \mathscr{W}_0$ to $\mathscr{W}_0$ which satisfies* (14.12)*. Finally, $(K\,, \Phi)_t(x)$ is a Walsh integral for all $K \in \mathscr{K}_0$ and $\Phi \in \mathscr{W}_0$.*

14.7.2 Stochastic convolutions

We now apply Theorem 14.7.1 in order to produce non-trivial examples of Walsh-integrable random fields that are relevant to the study of SPDEs.

Suppose $\kappa : (0\,,\infty) \times G \mapsto \mathbb{R}$ is a non-random measurable function, and write $\kappa_t(x)$ in place of $\kappa(t\,, x)$ in order to be consistent with the notation of stochastic process theory. We can define a kernel K as follows:

$$K_{t,s}(x\,, y) := \begin{cases} \kappa_{t-s}(xy^{-1}) & \text{if } s \in (0\,, t), \\ 0 & \text{otherwise,} \end{cases}$$

for all $s, t \geqslant 0$ and $x, y \in G$. According to Lebesgue's integration theory, $K \in \mathscr{K}_0$ if and only if $\kappa \in L^2((0\,,\infty) \times G)$. For a such a κ, we will always write $\kappa \circledast \Phi$ in place of $(K\,, \Phi)$, whenever $\Phi \in \mathscr{W}_0$. We might also write the Walsh integral $(\kappa \circledast \Phi)_t(x)$ as

$$(\kappa \circledast \Phi)_t(x) := \int_{[0,t]\times G} \kappa_{t-s}(xy^{-1})\Phi_s(y)\, \xi(\mathrm{d}s\, \mathrm{d}y).$$

We will refer to $\kappa \circledast \Phi$ as the *stochastic convolution* of $\kappa \in L^2((0,\infty)\times G)$ with $\Phi \in \mathscr{W}_0$. According to Theorem 14.7.1, $\kappa \circledast \Phi$ is itself in $\mathscr{W}_0$. Therefore, so are the "multifold stochastic convolutions", $\kappa \circledast_{(2)} \Phi := \kappa \circledast (\kappa \circledast \Phi)$, $\kappa \circledast_{(3)} \Phi := \kappa \circledast [\kappa \circledast (\kappa \circledast \Phi)]$, etc. And by induction,

$$\mathcal{N}_0\left(\kappa \circledast_{(n)} \Phi\right) \leqslant \mathcal{N}_0(\Phi) \cdot \left[\int_0^\infty \|\kappa_s\|^2_{L^2(G)}\,\mathrm{d}s\right]^{n/2},$$

for all $n \geqslant 1$, where $\kappa \circledast_{(1)} \Phi := \kappa \circledast \Phi$.

If $\alpha > 0$, then we can define $\mathscr{W}_\alpha$ to be the collection of all random fields Φ such that the random field $(t,x) \mapsto \exp(-\alpha t)\Phi_t(x)$ is in $\mathscr{W}_0$. Note that $\mathscr{W}_\alpha \subseteq \mathscr{W}_\beta$, whenever $0 \leqslant \alpha \leqslant \beta$. Also, if $X := \{X_t(x)\}_{t\geqslant 0, x\in G}$ is a space-time random field, then we define the norm

$$\mathcal{N}_\alpha(X) := \sup_{t\geqslant 0}\sup_{x\in G}\left\{\mathrm{e}^{-\alpha t}\mathrm{E}\left(|X_t(x)|^2\right)\right\}^{1/2}.$$

Since $\kappa_{t-s}(x-y)\Phi_s(y) = \mathrm{e}^{-\alpha(t-s)}\kappa_{t-s}(x-y)\cdot\mathrm{e}^{-\alpha s}\Phi_s(y)$, we can replace $\kappa_t(x-y)$ by $\mathrm{e}^{-\alpha t}\kappa_t(x-y)$ and $\Phi_s(y)$ and $\mathrm{e}^{-\alpha s}\Phi_s(y)$ everywhere in order to see that Theorem 14.7.1 implies the following.

Theorem 14.7.2. *Suppose $\kappa : (0,\infty)\times G \to \mathbb{R}$ is non-random and measurable, and satisfies $\int_0^\infty \exp(-\alpha s)\|\kappa_s\|^2_{L^2(G)}\,\mathrm{d}s < \infty$ for some $\alpha \in [0,\infty)$. Then, the stochastic convolution map $\Phi \mapsto \kappa \circledast \Phi$ defines a linear mapping from $\mathscr{W}_\alpha$ to $\mathscr{W}_\alpha$ such that, for all $n \geqslant 1$,*

$$\mathcal{N}_\alpha\left(\kappa \circledast_{(n)} \Phi\right) \leqslant \mathcal{N}_\alpha(\Phi) \cdot \left[\int_0^\infty \mathrm{e}^{-\alpha s}\|\kappa_s\|^2_{L^2(G)}\,\mathrm{d}s\right]^{n/2}.$$

14.7.3 Relation to Itô integrals

Since every non-random element of $L^2((0,\infty)\times G)$ is Walsh integrable, the Walsh integral generalizes the integral of Wiener. It also generalizes the integral of Itô, as we will see next. It might be a good idea to revisit Examples 14.4.1 and 14.4.2 [p. 141] before you read on.

Example 14.7.3 (The trivial group)**.** Let $G = \{\mathrm{e}\}$ denote the trivial group, and recall that $B_t := \xi([0,t]\times\{\mathrm{e}\})$ defines a Brownian motion. In this case, for every fixed $\alpha \geqslant 0$, the elements of $\mathscr{W}_\alpha$ are stochastic processes of the form $t \mapsto \Phi_t(\mathrm{e}) := \Phi_t$, since G has only one element. These processes are called *predictable* with respect to the Brownian motion $B := \{B_t\}_{t\geqslant 0}$, and satisfy $\mathrm{E}\int_0^\infty \exp(-\alpha s)\Phi_s^2\,\mathrm{d}s < \infty$. The Walsh integral $\int_{\mathbb{R}_+\times\{\mathrm{e}\}} \Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y)$ is often written as $\int_0^\infty \Phi_s\,\mathrm{d}B_s$, and is referred to as the *Itô integral of* Φ.

Example 14.7.4 (Cyclic groups)**.** Let $G = \{0,\ldots,n-1\} \cong \mathbb{Z}/n\mathbb{Z}$ for some $n \geqslant 1$. Recall that we $B_t^{(i)} := \xi([0,t]\times\{i\})$ defines an n-dimensional Brownian motion

B. If $\alpha \geqslant 0$, then elements Φ of $\mathscr{W}_\alpha$ are also called *predictable* (with respect to B). In addition, the Walsh integral

$$\int_{\mathbb{R}_+\times G} \Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y) = \sum_{i=0}^{n-1} \int_{\mathbb{R}_+\times\{i\}} \Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y)$$

is also written as

$$\int_0^\infty \Phi_s \cdot \mathrm{d}B_s := \sum_{i=0}^{n-1} \int_0^\infty \Phi_s(i)\,\mathrm{d}B_s^{(i)},$$

and called the *Itô integral* of the n-dimensional process Φ.

Example 14.7.5 (The integer group)**.** If $G = \mathbb{Z}$ denotes the additive integers, then $t \mapsto B_t^{(i)} := \xi([0\,,t]\times\{i\})$ $(i \in \mathbb{Z})$ defines an infinite sequence of Brownian motions. The sequence-valued stochastic process $B_t := (\cdots, B_t^{(-1)}, B_t^{(0)}, B_t^{(1)}, \ldots)$ is called *Brownian motion on* $\mathbb{R}^{\mathbb{Z}}$, elements Φ of $\mathscr{W}_\alpha$ are said to be *predictable processes* (with respect to B) and satisfy $\sum_{i=-\infty}^{\infty} \mathrm{E}\int_0^\infty |\Phi_s(i)|^2\,\mathrm{d}s < \infty$. Finally, the Walsh integral

$$\int_{\mathbb{R}_+\times\mathbb{Z}} \Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y) = \sum_{i=-\infty}^{\infty} \int_{\mathbb{R}_+\times\{i\}} \Phi_s(y)\,\xi(\mathrm{d}s\,\mathrm{d}y)$$

is usually written as

$$\int_0^\infty \Phi_s \cdot \mathrm{d}B_s = \sum_{i=-\infty}^{\infty} \int_0^\infty \Phi_s(i)\,\mathrm{d}B_s^{(i)},$$

and called the *Itô integral* of Φ with respect to the Brownian motion B.

Chapter 15

Lévy Processes

Before we continue our discussion of stochastic partial differential equations, we pause to recall a few facts from the theory of Lévy processes on LCA groups.

Definition 15.0.6. Let $X := \{X_t\}_{t \geqslant 0}$ be a stochastic process with values in an LCA group G. Then we say that X is a *Lévy process* if:

(i) $X_0 = \mathrm{e}$ is the group identity;

(ii) the map $t \mapsto X_t$ is right-continuous with left limits (for all $\omega \in \Omega$);

(iii) X has *stationary independent increments*, that is, for all $s, t \geqslant 0$,

  (a) (Stationarity) $X_{t+s}X_t^{-1}$ has the same distribution as X_s; and

  (b) (Independence) $X_{t+s}X_t^{-1}$ is independent of $\{X_r\}_{r\in[0,t]}$.

It is possible to prove the following.

Proposition 15.0.7. *Lévy processes satisfy the strong Markov property.*

15.1 Introduction

We will simplify the discussion greatly by concentrating only on Lévy processes that we will need later on. Therefore, from now on we assume tacitly that G is either $\mathbb{R}$, $\mathbb{T}$, $\mathbb{Z}$, or $\mathbb{Z}/n\mathbb{Z}$ for some $n \geqslant 1$. All of our remarks that refer to G continue to hold for general LCA groups, after suitable modifications are made. However, the extra gain in generality is not of direct use to us.

We will also use the additive notation for group operations; thus, we will write $x + y$ in place of xy, $-x$ in place of x^{-1}, and 0 in place of e. We will continue to write m_G for the Haar measure on G in all cases, though it might help to recall the following: (i) when $G = \mathbb{R}$, the Haar measure is the ordinary Lebesgue measure normalized to assign total mass one to $[0\,,1]$; (ii) when $G = \mathbb{T}$, the Haar measure is the Lebesgue measure normalized to assign total mass one to $\mathbb{T} = [0\,,2\pi)$; and (iii) m_G denotes the counting measure when $G = \mathbb{Z}$ or $\mathbb{Z}/n\mathbb{Z}$.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_15

We may study Lévy processes on $\mathbb{R}$, $\mathbb{T}$, $\mathbb{Z}$, and $\mathbb{Z}/n\mathbb{Z}$ separately, on a case by case basis. The general theory of Lévy processes on LCA groups requires abstract harmonic analysis, which is a topic for another lecture series. Thus, let us begin in this order with the concrete LCA groups $\mathbb{R}$, $\mathbb{T}$, $\mathbb{Z}$, and $\mathbb{Z}/n\mathbb{Z}$.

15.1.1 Lévy processes on $\mathbb{R}$

Recall, once again, that X is a Lévy process on $\mathbb{R}$ if and only if: (i) $X_0 = 0$; (ii) $t \mapsto X_t$ is right-continuous with left limits; and $X_{t+s} - X_t$ is independent of $\{X_r\}_{r\in[0,t]}$ with a distribution that does not depend on t, for every $s, t \geqslant 0$.

Definition 15.1.1. A *Lévy measure* m is a Borel measure on $\mathbb{R}$ such that $m(\{0\}) = 0$ and $\int_{-\infty}^{\infty}(1 \wedge x^2)\, m(\mathrm{d}x) < \infty$. A *characteristic exponent* is a function $\Psi\colon \mathbb{R} \to \mathbb{C}$ that has the form

$$\Psi(z) = iaz + \frac{\sigma^2 z^2}{2} + \frac{1}{2}\int_{-\infty}^{\infty} \left[1 - \mathrm{e}^{ixz} + ixz\mathbf{1}_{[-1,1]}(x)\right] m(\mathrm{d}x), \tag{15.1}$$

where m is a Lévy measure, and $a, \sigma \in \mathbb{R}$. We might refer to $(a\,, \sigma\,, m)$ as the *Lévy–Khintchine triple.*

The constants $-a$ and σ are called the *drift coefficient* and the *diffusion coefficient* of X, respectively. We follow the standard convention of the theory of Lévy processes and rewrite (15.1) slightly differently in the special cases that m is a finite measure. In these cases, we can (and always will) write (15.1) as

$$\Psi(z) = ibz + \frac{\sigma^2 z^2}{2} + \frac{1}{2}\int_{-\infty}^{\infty} \left[1 - \mathrm{e}^{ixz}\right] m(\mathrm{d}z), \tag{15.2}$$

where $b = a + \frac{1}{2}\int_{[-1,1]} x\, m(\mathrm{d}x)$. Characteristic exponents are of interest to us because of the following celebrated result; see Bertoin [3] and Sato [52] for more details.

Theorem 15.1.2 (The Lévy–Khintchine formula)**.** *If X is a Lévy process on $\mathbb{R}$, then there exists a characteristic exponent Ψ such that*

$$\mathrm{E}\exp(izX_t) = \exp(-t\Psi(z)) \tag{15.3}$$

for all $t \geqslant 0$ and $z \in \mathbb{R}$. Conversely, for every characteristic exponent Ψ we can find a probability space on which there exists a Lévy process satisfying (15.3).

Perhaps the three most famous examples of Lévy processes are "uniform motion", "Brownian motion", and "compound Poisson processes" (also known as "continuous-time random walks").

Example 15.1.3 (Uniform motion)**.** If $X_t = \mu t$ for some $\mu \in \mathbb{R}$ and all $t \geqslant 0$, then X is a (non-random) Lévy process whose characteristic exponent satisfies (15.1) with $a := -\mu$ and $\sigma := 0$, and has Lévy measure identically equal to zero.

Example 15.1.4 (Brownian motion)**.** Recall that X is a *Brownian motion* if it is a centered Gaussian process with continuous trajectories and covariance

$$\mathrm{Cov}[X_t\,, X_s] = \min(s\,, t) \text{ for all } s, t \geqslant 0.$$

One can verify directly that a Brownian motion is a Lévy process whose characteristic exponent satisfies (15.1) with $a := 0$ and $\sigma := 1$, and has Lévy measure identically equal to zero.

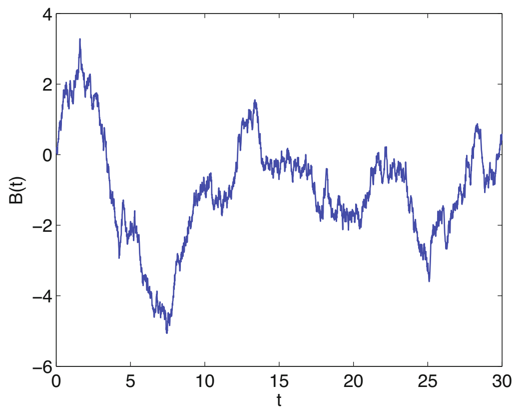


Figure 15.3: A simulation of Brownian motion by time $t = 30$.

Example 15.1.5 (Compound Poisson processes)**.** Suppose that $J_1, J_2, \ldots$ are i.i.d. random variables and let N denote an independent rate-κ Poisson process where $\kappa > 0$ is fixed. Then $X_t := \sum_{j=0}^{N(t)} J_j$, for $t \geqslant 0$, defines a Lévy process, where $J_0 := 0$. Moreover, disintegration and elementary computations with Poisson processes together reveal that

$$\begin{aligned} \mathrm{E}\exp(izX_t) &= \mathrm{E}\left[\{\mathrm{E}\exp(izJ_1)\}^{N(t)}\right] = \exp\left(-t\kappa\left[1-\phi(z)\right]\right) \\ &= \exp\left(-t\kappa\int_{-\infty}^{\infty}\left[1-\mathrm{e}^{ixz}\right]\, m(\mathrm{d}x)\right), \end{aligned}$$

where $m(A) := \mathrm{P}\{J_1 \in A,\ J_1 \neq 0\}$ is manifestly a Lévy measure of finite total mass, and $\phi(z) := \mathrm{E}\exp(izJ_1)$ denotes the characteristic function of J_1. Thus, compound Poisson processes are Lévy processes whose characteristic exponent satisfies (15.2) with $b = \sigma = 0$.

One can combine the previous examples in order to build new Lévy processes as follows: let $\{B_t\}_{t\geqslant 0}$ be a Brownian motion, and $\{C_t\}_{t\geqslant 0}$ an independent compound Poisson process with a Lévy measure m that is finite *per force*. Then for every $a, \sigma \in \mathbb{R}$,

$$X_t := -at + \sigma B_t + C_t, \tag{15.4}$$

for $t \geqslant 0$, is a Lévy process whose characteristic exponent satisfies (15.1). It can be shown that the family of all Lévy processes is the "closure" of those of type (15.4) "in the topology of weak convergence". In other words, Lévy processes of type (15.4) are the building blocks of all Lévy processes on $\mathbb{R}$. Next, thers is an example of a Lévy process that is not of the form (15.4).

Example 15.1.6 (Symmetric stable processes)**.** For every fixed constant $C > 0$, the Borel measure $m(\mathrm{d}x) := C|x|^{-1-\alpha}\,\mathrm{d}x$ is a Lévy measure if and only if $\alpha \in (0\,,2)$. For those values of α,

$$\int_{-\infty}^{\infty} \left[1 - \mathrm{e}^{ixz} - ixz\mathbf{1}_{[-1,1]}(x)\right] m(\mathrm{d}x) = 2C\int_0^\infty \frac{1-\cos(xz)}{x^{1+\alpha}}\,\mathrm{d}x.$$

The integral converges absolutely and is defined in the usual way because $|1 - \cos(\theta)| \leqslant \theta^2$ for all $\theta \in \mathbb{R}$. Moreover, a change of variables shows that the integral is proportional to $|z|^\alpha$. Therefore, we can adjust our choice of C in order to see that

$$\Psi(z) := \kappa|z|^\alpha, \tag{15.5}$$

$z \in \mathbb{R}$, defines the characteristic exponent of a Lévy process whenever $\alpha \in (0\,,2)$ and $\kappa > 0$. The same is true when $\alpha = 2$, though in that case, the Lévy measure needs to be set to zero identically and κ is twice the diffusion coefficient in (15.1). Lévy processes whose characteristic function satisfies (15.5) are said to be *symmetric α-stable.* A symmetric 2-stable Lévy process is a constant multiple of standard Brownian motion. As it turns out, Brownian motion is the only symmetric α-stable Lévy process that has continuous sample trajectories. All other stable processes

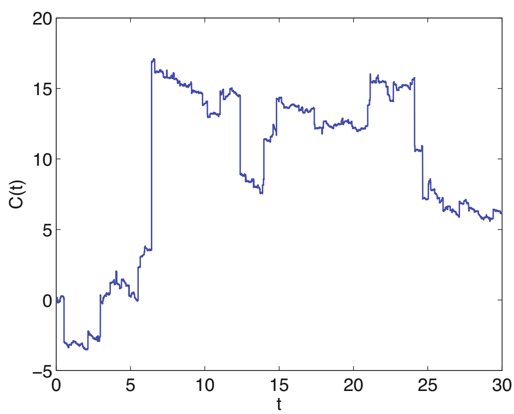


Figure 15.4: A simulation of the Cauchy process by time $t = 30$.

than Brownian motion have pure-jump same functions. Perhaps you can see an example of this pure-jump property in Figure 15.4, which shows a simulation of a symmetric 1-stable Lévy process (a so-called *symmetric Cauchy process*).

Occasionally, we might have use for the following result.

Lemma 15.1.7. $\lim_{|z|\to\infty} \Psi(z)/z^2 = \sigma^2/2$.

When X is a Brownian motion, $\Psi(z) \propto z^2$. The preceding says that, asymptotically, the characteristic exponent of Brownian motion is larger than that of any other Lévy process. The largeness of the characteristic exponent is equivalent to the smallness of the characteristic function; therefore, the preceding is really saying that the tail probabilities of Brownian motion are the smallest among all Lévy processes. It turns out that, for this reason, Brownian motion is the only continuous non-constant Lévy process.

Proof of Lemma 15.1.7. In light of (15.1), we may assume without loss of generality, that $a = \sigma = 0$. Otherwise we replace $\Psi(z)$ by $\Psi(z) - \sigma^2 z^2/2$ henceforth.

Choose and fix $\varepsilon > 0$, and note that $|1 - \mathrm{e}^{-ixz} + ixz| \leqslant x^2 z^2$ for all $x, z \in \mathbb{R}$, thanks to the series expansion of the complex exponential. Therefore,

$$\int_{[-\varepsilon,\varepsilon]} |1 - \mathrm{e}^{-ixz} + ixz|\, m(\mathrm{d}x) \leqslant z^2 \int_{[-\varepsilon,\varepsilon]} x^2\, m(\mathrm{d}x).$$

And

$$\int_{|x|>\varepsilon} |1 - \mathrm{e}^{ixz}|\, m(\mathrm{d}x) \leqslant 2\mu([-\varepsilon\,,\varepsilon]^c),$$

manifestly. Since m is a Lévy measure,

$$m([-\varepsilon\,,\varepsilon]^c) + \int_{[-\varepsilon,\varepsilon]} x^2\, m(\mathrm{d}x) = \int_{-\infty}^{\infty} (1 \wedge x^2)\, m(\mathrm{d}x) < \infty.$$

Therefore, we have proved that

$$\limsup_{|z|\to\infty} \frac{|\Psi(z)|}{z^2} \leqslant \inf_{\varepsilon>0} \int_{[-\varepsilon,\varepsilon]} x^2\, m(\mathrm{d}x),$$

which is zero, thanks to the dominated convergence theorem. □

15.1.2 Lévy processes on $\mathbb{T}$

Usually, one thinks of $\mathbb{T}$ as the quotient $\mathbb{R}/2\pi\mathbb{Z}$. However, it is also helpful to think of the torus as $\mathbb{T} = h(\mathbb{R})$ for the homomorphism $h(x) := x \mod 2\pi$, $x \in \mathbb{R}$. You can think of h as a map that "wraps $\mathbb{R}$ around the torus". It follows from this construction that every Lévy process $\{X_t\}_{t\geqslant 0}$ on $\mathbb{T}$ can be constructed (on some probability space) as follows: $X_t = h(Y_t)$, $t \geqslant 0$, where $Y := \{Y_t\}_{t\geqslant 0}$ is a

Lévy process on $\mathbb{R}$. Thus, for example, if B denotes a Brownian motion on $\mathbb{R}$ then $X_t := h(B_t)$ defines the *Brownian motion on the torus*. In other words, in order to construct X we "wrap B around the torus".

Similarly, all continuous-time random walks on $\mathbb{T}$ are constructed by "wrapping compound Poisson processes around the torus".

15.1.3 Lévy processes on $\mathbb{Z}$

Let X be a Lévy process on $\mathbb{Z}$, and define a sequence of stopping times $T_0 \leqslant T_1 \leqslant T_2 \leqslant \cdots$ as follows: $T_0 := 0$, and

$$T_{n+1} := \inf \{t > T_n \mid X_t \neq X_{t-}\},$$

for all $n \geqslant 0$, where $\inf \varnothing := \infty$. The stopping times $T_1, T_2, \dots$ are the times at which X jumps. Since $\mathbb{Z}$ is discrete, the defining properties of Lévy processes show that either $X_t := 0$ for all $t \geqslant 0$, or $\{X_{T_n}\}_{n=0}^{\infty}$ is a non-degenerate discrete-time random walk on $\mathbb{Z}$. Also as it turns out, the strong Markov property of X (i.e., Proposition 15.0.7) implies that

$$\left\{(T_{n+1} - T_n\, , X_{T_{n+1}} - X_{T_n})\right\}_{n=0}^{\infty}$$

are i.i.d. with T_1 having an exponential distribution. This analysis then implies that *every Lévy process on $\mathbb{Z}$ is a compound Poisson process* (or a continuous-time random walk, if you wish) with values in $\mathbb{Z}$.

15.1.4 Lévy processes on $\mathbb{Z}/n\mathbb{Z}$

Since $\mathbb{Z}/\mathbb{Z}$ is the trivial group, the only Lévy process on $\mathbb{Z}/\mathbb{Z}$ is $X_t := 0$. Therefore, we consider $\mathbb{Z}/n\mathbb{Z}$ where $n \geqslant 2$ for the remainder of this subsection.

We just proved that every Lévy process on $\mathbb{Z}$ is a compound Poisson process. Our justification required only the strong Markov property of X (Proposition 15.0.7) and the fact that the state space of X is $\mathbb{Z}$, which is a discrete LCA group. Thus, it follows from the same line of reasoning that every Lévy process on $\mathbb{Z}/n\mathbb{Z}$ is a compound Poisson process. It might help to hash this out a little more. Therefore, let us consider the simplest non-trivial case in which $n = 2$ and X is a Lévy process on $\mathbb{Z}/2\mathbb{Z}$. Since X is a compound Poisson process we can find i.i.d. random variables $J_1, J_2, \dots$ with values in $\{0\,,1\}$ together with an independent rate-κ Poisson process N (for some $\kappa > 0$) such that

$$X_t = \sum_{j=0}^{N(t)} J_j,$$

for all $t \geqslant 0$, where $J_0 := 0$. The summation symbol refers to the group multiplication in $\mathbb{Z}/2\mathbb{Z}$, of course (that is, addition mod 1). In other words, the Markov process X evolves as follows: at time zero, X is at zero. X remains at zero until the first jump time of the Poisson process N. Then, X jumps to one and stays there until the second jump time of N. At that time X switches to the value 0, and so on. This completely characterizes Lévy processes on $\mathbb{Z}/2\mathbb{Z}$.

Exercise 15.1.8. Characterize all Lévy processes on $\mathbb{Z}/n\mathbb{Z}$ when $n \geqslant 3$. (HINT. You will need $n-1$ jump rates.)

15.2 The semigroup

In order to analyze Lévy processes further we need to cite a number of facts from the theory of Markov processes. See Jacob [27–29] for more details.

We can define a family $\{P_t\}_{t\geqslant 0}$ of linear operators as $(P_tf)(x) := \mathrm{E}\,[f(xX_t)]$, for $t \geqslant 0$, $x \in G$. This is well defined, for example, if $f\colon G \to \mathbb{R}$ is measurable and bounded. But it makes sense more generally still if f is measurable and $\mathrm{E}|f(xX_t)| < \infty$ for all $x \in G$ and $t \geqslant 0$. The Markov property of X ensures that $(P_t[P_sf])(x) = (P_{t+s}f)(x)$ for all $s, t \geqslant 0$ and $x \in G$. That is, we have the semigroup property

$$P_tP_s = P_{t+s}, \tag{15.6}$$

for all $s, t \geqslant 0$, where AB denotes the composition of linear operators A and B, as usual.[1] The family $\{P_t\}_{t\geqslant 0}$ is thus dubbed a *semigroup* of linear operators. Since $X_0 = 0$, we also see that P_0 is the identity operator.

Knowledge of the semigroup $\{P_t\}_{t\geqslant 0}$ amounts to knowing the probability distribution of the process $\{xX_t\}_{t\geqslant 0}$, which we can think of as our Lévy process, started at $x \in G$. Analytically, it turns out that it is slightly more convenient to work with the Lévy process X^{-1} (recall that x^{-1} denotes the group inverse of $x \in G$). Thus, let p_t denote the distribution of X_t^{-1} for all $t \geqslant 0$; that is, $p_t(A) := \mathrm{P}\{X_t^{-1} \in A\}$, $t \geqslant 0$, for every Borel set $A \subseteq G$. Each p_t is a probability measure on G, with $p_0 := \delta_0$. Moreover,

$$(P_tf)(x) = \int_G f(xy^{-1})\, p_t(\mathrm{d}y) = (f * p_t)(x),$$

where "$*$" denotes convolution. In other words, we can (and will) identify every linear operator $f \mapsto P_tf$ with the convolution operator $f \mapsto f * p_t$. In this way, we see that the semigroup property (15.6) is written in equivalent form as $p_{t+s} = p_t * p_s$ for all $s, t \geqslant 0$. Thus, the family $\{p_t\}_{t\geqslant 0}$ (or, equivalently, the operators $\{P_t\}_{t\geqslant 0}$) is known as a *convolution semigroup*.

[1] In fact, (15.6) is just another way to say that X has the Markov property. In this particular context, (15.6) is sometimes known also as the *Chapman–Kolmogorov equation*.

15.3 The Kolmogorov–Fokker–Planck equation

The *Fokker–Planck equation*, or *Kolmogorov's forward equation*, is a description of the evolution of the distribution of X. Somewhat more generally, we wish to know how the function $x \mapsto (P_t f)(x)$ evolves with time. The class of testing function f should be quite rich; in fact, rich enough that we can use our evaluation of $P_t f$ (by an approximation scheme or some such method) to ultimately compute probabilities of the type $\mathrm{P}\{xX_t \in A\}$. For our family of testing functions f, we will choose the space $L^2(G)$, to be concrete. Our first result, Lemma 15.3.1, ensures that the evolution $t \mapsto P_t f$ always takes place in $L^2(G)$ as long as the initial state $P_0 f = f$ is in $L^2(G)$.

First let us observe that, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\|P_t f\|^2_{L^2(G)} &= \int_G \left| \int_G f(xy^{-1})\, p_t(\mathrm{d}y) \right|^2 m_G(\mathrm{d}x) \\ &\leqslant \int_G \int_G |f(xy^{-1})|^2\, p_t(\mathrm{d}y)\, m_G(\mathrm{d}x) = \|f\|^2_{L^2(G)},\end{aligned}$$

for all $f \in L^2(G) \cap L^\infty(G)$. (Recall that m_G is translation invariant, by its very definition.) Therefore, we can continuously extend every P_t to a linear operator acting on $f \in L^2(G)$, which we continue to denote by P_t.

Lemma 15.3.1. *For every $t \geqslant 0$, the map $P_t \colon L^2(G) \to L^2(G)$ is non expanding. Moreover, $t \mapsto \|P_t f\|_{L^2(G)}$ is non increasing for every $f \in L^2(G)$.*

Proof. Since $\|P_t f\|_{L^2(G)} \leqslant \|f\|_{L^2(G)}$ for all $f \in L^2(G) \cap L^\infty(G)$ and $L^\infty(G)$ is dense in $L^2(G)$, it follows that $\|P_t f\|_{L^2(G)} \leqslant \|f\|_{L^2(G)}$ for all $f \in L^2(G)$, by density. Thus, the proof is complete once we prove that $\|P_{t+s} f\|_{L^2(G)} \leqslant \|P_t f\|_{L^2(G)}$ for all $f \in L^2(G)$ and $s, t \geqslant 0$. But this follows from the semigroup property

$$\|P_{t+s} f\|_{L^2(G)} = \|P_s P_t f\|_{L^2(G)} \leqslant \|P_t f\|_{L^2(G)},$$

since $P_t f \in L^2(G)$ for all $f \in L^2(G)$ and P_s is nonexpanding, as we just proved. □

A deeper analysis of Lévy processes on LCA groups hinges on the development of harmonic analysis on LCA groups, which is an interesting topic in its own right. But a complete treatment will distract us too much from our main goals. Instead, we study Lévy processes only on concrete LCA groups that are of interest to us. My hope is that there will be enough examples giving us hints of a large theory. For that larger theory itself see Morris [41] and Rudin [49].

15.3.1 Lévy processes on $\mathbb{R}$

Let X be a Lévy process on $\mathbb{R}$ whose characteristic exponent Ψ is given by (15.1). The semigroup is defined by $(P_t f)(x) = \mathrm{E}[f(x + X_t)]$. Since $P_t f \in L^2(\mathbb{R})$ for all $f \in L^2(\mathbb{R})$ (Lemma 15.3.1), we can try to understand the behavior of $P_t f$ in terms of its action on other functions in $L^2(\mathbb{R})$; namely, we can study the time evolution of $(g\,,P_tf)_{L^2(\mathbb{R})}$ for $g \in L^2(\mathbb{R})$. Since $P_t * f = f * p_t$ when $f \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, and since p_t is the probability distribution of $-X_t$, we apply Parseval's identity in order to see that

$$(g\,,P_tf)_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\hat{f}(z)\mathrm{E}\mathrm{e}^{-izX_t}\,\mathrm{d}z = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\hat{f}(z)\mathrm{e}^{-t\Psi(-z)}\,\mathrm{d}z.$$

Moreover, by density, this identity is valid for all $f \in L^2(\mathbb{R})$.

We wish to differentiate the preceding with respect to the non-negative time variable t. In other words if $s, t \geqslant 0$ are two distinct time points, then we wish to compute the limit, as $s \to t$, of

$$\left(g\,, \frac{P_sf - P_tf}{s-t}\right)_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\hat{f}(z)\left(\frac{\mathrm{e}^{-s\Psi(-z)} - \mathrm{e}^{-t\Psi(-z)}}{s-t}\right)\mathrm{d}z.$$

If we could take the limit inside the integral, then it would follow that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(g\,,P_tf)_{L^2(\mathbb{R})} &= -\frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\mathrm{e}^{-t\Psi(-z)}\hat{f}(z)\Psi(-z)\,\mathrm{d}z \\ &= -\frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\widehat{P_tf}(z)\Psi(-z)\,\mathrm{d}z,\end{aligned} \tag{15.7}$$

for all $g \in L^2(\mathbb{R})$, where the time derivative at $t = 0$ is understood as a right derivative (P_tf is not defined for $t < 0$).

It turns out that the only way to make the preceding rigorous is to restrict attention to a smaller family of functions f than all of $f \in L^2(\mathbb{R})$. With this in mind, define

$$\mathrm{Dom}[\mathcal{G}] := \left\{ h \in L^2(\mathbb{R}) \;\middle|\; \int_{-\infty}^{\infty} |\hat{h}(z)\Psi(z)|^2\,\mathrm{d}z < \infty \right\},$$

where "$\widehat{\ }$" refers to the Fourier transform (in the sense of distribution theory) normalized so that

$$\hat{h}(z) = \int_{-\infty}^{\infty} \exp(ixz)h(x)\,\mathrm{d}x,$$

for all $h \in L^1(\mathbb{R})$. Since $|\exp(-r\Psi(z))| \leqslant 1$ for all $r \geqslant 0$ and $z \in \mathbb{R}$,

$$\left|\frac{\mathrm{e}^{-s\Psi(z)} - \mathrm{e}^{-t\Psi(z)}}{s-t}\right| \leqslant \left|\frac{1 - \mathrm{e}^{-|s-t|\Psi(z)}}{|s-t|}\right| \leqslant |\Psi(z)| = |\Psi(-z)|.$$

Therefore, it follows from the dominated convergence theorem that (15.7) holds whenever $f \in \mathrm{Dom}[\mathcal{G}]$. By the Cauchy–Schwarz inequality, the following defines a linear operator $\mathcal{G}\colon \mathrm{Dom}[\mathcal{G}] \to L^2(\mathbb{R})$:

$$(g\,,\mathcal{G}f)_{L^2(\mathbb{R})} := -\frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{\hat{g}(z)}\hat{f}(z)\Psi(-z)\,\mathrm{d}z.$$

From the preceding statements, it is possible to deduce that if $f \in \mathscr{S}(\mathbb{R})$, then

$$(\mathcal{G}f)(x) := -\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{izx}\overline{\hat{f}(z)}\Psi(-z)\,\mathrm{d}z, \tag{15.8}$$

for almost every $x \in \mathbb{R}$ (or for every $x \in \mathbb{R}$, if we take the preceding as the definition of $\mathcal{G}f$ for $f \in \mathscr{S}(\mathbb{R})$). In any event, $\mathrm{Dom}[\mathcal{G}]$ is the domain of the definition of $\mathcal{G}$, and because of (15.7) and Fubini's theorem we can see that

$$\frac{\mathrm{d}}{\mathrm{d}t}(g\,, P_t f)_{L^2(\mathbb{R})} = (g\,, \mathcal{G}P_t f)_{L^2(\mathbb{R})},$$

for all $g \in L^2(\mathbb{R})$ and $f \in \mathrm{Dom}[\mathcal{G}]$. In other words, we have proved the existence portion of the following fact.

Theorem 15.3.2. *If $f \in \mathrm{Dom}[\mathcal{G}]$, then the 2-parameter function $u_t(x) := (P_t f)(x)$ is the unique weak solution to the evolution equation,*

$$\begin{bmatrix} \dfrac{\mathrm{d}}{\mathrm{d}t}u_t(x) = (\mathcal{G}u_t)(x) & \textit{for } t \geqslant 0 \quad \textit{and} \quad x \in \mathbb{R}, \\ u_0(x) = f(x) & \textit{for } x \in \mathbb{R}. \end{bmatrix} \tag{15.9}$$

Remark 15.3.3. Equation (15.9) is called the *Fokker–Planck equation* as well as *Kolmogorov's forward equation*. We refer to $\mathcal{G}$ and $\mathrm{Dom}[\mathcal{G}]$ as the *generator* of X and the *domain* (of definition) of X. Lemma 15.1.7 implies that $\mathscr{S}(\mathbb{R}) \subset \mathrm{Dom}[\mathcal{G}] \subset L^2(\mathbb{R})$, where $\mathscr{S}(\mathbb{R})$ denotes the collection of all rapidly-decreasing test functions on $\mathbb{R}$. Since $\mathscr{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, it follows that $\mathrm{Dom}[\mathcal{G}]$ is dense in $L^2(\mathbb{R})$ also; i.e., $\mathcal{G}$ is defined densely on $L^2(\mathbb{R})$, as prescribed by (15.8).

Proof of Theorem 15.3.2. We have shown that $u_t(x) = (P_t f)(x)$ is a weak solution. Suppose $v_t(x)$ were another weak solution. That is, suppose $v_t \in \mathrm{Dom}[\mathcal{G}]$ for all $t > 0$, $v_0(x) = f(x)$, and

$$\frac{\mathrm{d}}{\mathrm{d}t}(g\,, v_t)_{L^2(\mathbb{R})} = (g\,, \mathcal{G}v_t)_{L^2(\mathbb{R})}$$

for all $g \in L^2(\mathbb{R})$. Then, $h_t(x) := u_t(x) - v_t(x)$ solves $\dot{h} = \mathcal{G}h$ on $(0\,,\infty) \times \mathbb{R}$, subject to $h_0 \equiv 0$. This is a linear equation; take Fourier transforms in order to see that the unique weak solution is $h_t(x) = 0$. This proves uniqueness. □

Let us work out a family of examples that is relevant to our later needs. First, let us recall that if $f \in \mathrm{Dom}[\mathcal{G}]$ then $\mathcal{G}f \in L^2(\mathbb{R})$ and hence (15.8) and Fubini's theorem together imply that

$$(g\,, \mathcal{G}f)_{L^2(\mathbb{R})} = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{g}(z)\overline{\hat{f}(z)}\Psi(-z)\,\mathrm{d}z,$$

for all $g \in L^2(\mathbb{R})$. Thanks to the Plancherel theorem, the Fourier transform of $\mathcal{G}f$ is

$$\widehat{\mathcal{G}f}(z) = -\hat{f}(z)\Psi(-z), \tag{15.10}$$

$z \in \mathbb{R}$. In other words, we can identify the linear operator $\mathcal{G}$ with a Schwartz distribution, which we continue to write as $\mathcal{G}$, such that $\mathcal{G}f = \mathcal{G} * f$ and $\hat{\mathcal{G}}(z) = -\Psi(-z)$ for all $z \in \mathbb{R}$. We now work out a few examples.

Example 15.3.4 (Uniform motion)**.** Suppose $X_t = at$ for some $a \in \mathbb{R}$ and every $t \geqslant 0$; that is, $\Psi(z) = -iaz$ for all $z \in \mathbb{R}$. In this case,

$$\text{Dom}[\mathcal{G}] = \left\{ f \in L^2(\mathbb{R}) \,\middle|\, \int_{-\infty}^{\infty} |z\hat{f}(z)|^2 \, \mathrm{d}z < \infty \right\},$$

and $(\mathcal{G}f)\hat{}\,(z) = iaz\hat{f}(z)$ thanks to (15.10). Recall that the *weak derivative* of f is a distribution f' such that $(g\,, f')_{L^2(\mathbb{R})} = -(g'\,, f)_{L^2(\mathbb{R})}$ for all rapidly-decreasing test functions g on $\mathbb{R}$. In our setting, we can deduce from the preceding that $\mathcal{G}f = af'$, and

$$\text{Dom}[\mathcal{G}] = \{ f \in L^2(\mathbb{R}) \mid f' \in L^2(\mathbb{R}) \}$$

is the Sobolev space $W^{1,2}(\mathbb{R})$.

Example 15.3.5 (Brownian motion)**.** If X is a Brownian motion on $\mathbb{R}$, then $\Psi(z) = \sigma^2 z^2/2$ for some $\sigma \neq 0$ ($\sigma = 1$ is standard Brownian motion). Thus,

$$\text{Dom}[\mathcal{G}] = \left\{ f \in L^2(\mathbb{R}) \,\middle|\, \int_{-\infty}^{\infty} z^4 |\hat{f}(z)|^2 \, \mathrm{d}z < \infty \right\},$$

and $(\mathcal{G}f)\hat{}\,(z) = -(\sigma^2/2)z^2\hat{f}(z)$. In other words, the generator of our Brownian motion is $\mathcal{G}f = (\sigma^2/2)f''$, where f'' refers to the second (weak) derivative of f, and

$$\text{Dom}[\mathcal{G}] := \left\{ f \in L^2(\mathbb{R}) \mid f'' \in L^2(\mathbb{R}) \right\},$$

which is simply the Sobolev space $W^{2,2}(\mathbb{R})$.

Example 15.3.6 (Zero-drift, zero-diffusion case)**.** Consider a real-valued Lévy process X whose drift and diffusion coefficients $-a$ and σ are both zero. That is,

$$\Psi(z) = \int_{-\infty}^{\infty} \left[1 - \mathrm{e}^{iyz} - iyz\mathbf{1}_{[-1,1]}(y) \right] m(\mathrm{d}y).$$

This and (15.10) together tell us that, for all $f \in \text{Dom}[\mathcal{G}]$,

$$\widehat{\mathcal{G}f}(z) = \int_{-\infty}^{\infty} \hat{f}(z) \left[\mathrm{e}^{-iyz} - 1 + iyz\mathbf{1}_{[-1,1]}(y) \right] m(\mathrm{d}y).$$

Recall that $z \mapsto \hat{f}(z)\mathrm{e}^{-iyz}$ is the Fourier transform of $x \mapsto f(y+x)$, and that $z \mapsto iz\hat{f}(z)$ is the Fourier transform of f'. Therefore,

$$(\mathcal{G}f)(x) = \int_{-\infty}^{\infty} \left[f(x+y) - f(x) - y\mathbf{1}_{[-1,1]}(y) f'(x) \right] m(\mathrm{d}y).$$

Thus, for example, if X is a symmetric α-stable process with $\alpha \neq 2$ (see Example 15.1.6), then $\mathcal{G}$ is seen to be the following "negatively-indexed fractional-derivative operator"

$$(\mathcal{G}f)(x) = C \int_{-\infty}^{\infty} \frac{f(x+y) - f(x)}{|y|^{1+\alpha}} \, \mathrm{d}y. \tag{15.11}$$

We interpret this expression, and others like it in the usual way: the preceding formula yields a well-defined pointwise convergent construction of $\mathscr{G}f$ for all $f \in \mathscr{S}(\mathbb{R})$, and thus, defines $\mathscr{G}$ densely on its domain $\mathrm{Dom}[\mathscr{G}]$.

According to the Lévy–Khintchine formula (Theorem 15.1.2), every Lévy process X on $\mathbb{R}$ can be written as $X_t = D_t + B_t + \Pi_t$, where $D_t = -at$ for some $a \in \mathbb{R}$, B is a Brownian motion with speed σ, and Π is an independent Lévy process with no drift or diffusion. Since $\{P_t\}_{t \geqslant 0}$ is a convolution semigroup, it follows easily that the generator $\mathcal{G}_X$ of X can be decomposed as $\mathcal{G}_X = \mathcal{G}_D + \mathcal{G}_B + \mathcal{G}_\Pi$, on $I := \mathrm{Dom}[\mathcal{G}_D] \cap \mathrm{Dom}[\mathcal{G}_B] \cap \mathrm{Dom}[\mathcal{G}_\Pi]$, notation being obvious. Since rapidly-decreasing test functions are always in I, the preceding examples together show us the following.

Corollary 15.3.7. *Let X be a Lévy process on $\mathbb{R}$ with Lévy–Khintchine triple (a, σ, m). Then, the generator $\mathcal{G}$ of X is defined densely on $L^2(\mathbb{R})$, and satisfies*

$$(\mathcal{G}f)(x) = -af'(x) + \frac{\sigma^2}{2} f''(x) + \int_{-\infty}^{\infty} \left[f(x+y) - f(x) - y \mathbf{1}_{[-1,1]}(y) f'(x) \right] m(\mathrm{d}y),$$

for all $f \in \mathscr{S}(\mathbb{R})$ and $x \in \mathbb{R}$.

Example 15.3.8 (Lévy processes on $\mathbb{Z}$)**.** Let $J_1, J_2, \ldots$ be i.i.d. with values in $\mathbb{Z}$, and N be an independent mean-κ Poisson process for some $\kappa > 0$. Then, we have seen already that $X_t = \sum_{j=0}^{N(t)} J_j$ (with $J_0 := 0$) defines a Lévy process on $\mathbb{Z}$. We have also seen that every Lévy process on $\mathbb{Z}$ is realized in this way. Furthermore, the characteristic exponent of such an X has the form

$$\Psi(z) = \kappa \sum_{x \in \mathbb{Z}} [1 - \mathrm{e}^{-ixz}] \, D(x),$$

where $D \colon \mathbb{Z} \to \mathbb{R}_+$ designates the "displacement probability law" of the random walk $-X$; that is, $D(x) = \mathrm{P}\{J_1 = -x\}$ for all $x \in \mathbb{Z}$. See Example 15.1.5 on page 155 for an equivalent formulation.[2]

According to the theory of Fourier series, for every $f, g \in L^2(\mathbb{Z})$,

$$(g\,, P_t f)_{L^2(\mathbb{Z})} = \frac{1}{2\pi} \int_0^{2\pi} \overline{\hat{g}(z)} \hat{f}(z) \mathrm{E}\mathrm{e}^{-izX_t} \, \mathrm{d}z,$$

[2]It is perhaps worth mentioning that, according to Example 15.1.5, we should actually write $D(x) = \mathrm{P}\{J_1 = -x\,, J_1 \neq 0\}$ and not $D(x) = \mathrm{P}\{J_1 = -x\}$. However, both choices of D are correct because $1 - \mathrm{e}^{-ixz}$ vanishes at $x = 0$.

where "$\widehat{\ }$" denotes the Fourier transform on $\mathbb{Z}$; i.e.,

$$\hat{h}(z) := \sum_{x\in\mathbb{Z}} h(x)\mathrm{e}^{ixz}$$

for all $h \in L^1(\mathbb{Z})$. In particular, since $\mathrm{E}\exp(-izX_t) = \exp(-t\Psi(-z))$, this means that whenever $s, t \geqslant 0$ are distinct and $f, g \in L^2(\mathbb{Z})$,

$$\left(g\,,\frac{P_sf - P_tf}{s-t}\right)_{L^2(\mathbb{R})} = \frac{1}{2\pi}\int_0^{2\pi} \overline{\hat{g}(z)}\hat{f}(z)\left(\frac{\mathrm{e}^{-s\Psi(-z)} - \mathrm{e}^{-t\Psi(-z)}}{s-t}\right)\mathrm{d}z.$$

Define

$$(\mathcal{G}f)(x) := -\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{izx}\overline{\hat{f}(z)}\Psi(-z)\,\mathrm{d}z,$$

for all $f \in \mathrm{Dom}[\mathcal{G}]$ and $x \in \mathbb{Z}$. The integral converges absolutely for every $f \in L^2(\mathbb{Z})$ since Ψ is continuous, whence bounded, on $[0\,,2\pi)$. That is, $\mathrm{Dom}[\mathcal{G}]$ is all of $L^2(\mathbb{Z})$. Moreover, our formula for Ψ yields the following alternative form of $\mathcal{G}$:

$$(\mathcal{G}f)(x) = -\frac{\kappa}{2\pi}\sum_{y\in\mathbb{Z}} D(y)\int_0^{2\pi} \mathrm{e}^{izx}\left[1 - \mathrm{e}^{iyz}\right]\overline{\hat{f}(z)}\,\mathrm{d}z.$$

Since $L^2(\mathbb{Z}) \supset L^1(\mathbb{Z})$, the inversion formula for Fourier series applies as well, and tells us that $(2\pi)^{-1}\int_0^{2\pi}\exp(i\ell z)\overline{\hat{f}(z)}\,\mathrm{d}z = f(\ell)$ for all $\ell \in \mathbb{Z}$. Therefore, it follows that

$$(\mathcal{G}f)(x) = \kappa\sum_{y\in\mathbb{Z}}[f(x+y) - f(x)]\,D(y),$$

for all $x \in \mathbb{Z}$ and $f \in L^2(\mathbb{Z})$.[3] Now we proceed, much as we did for Theorem 15.3.2, in order to deduce the following.

Theorem 15.3.9. *If $f \in L^2(\mathbb{Z})$, then the 2-parameter function $u_t(x) := (P_tf)(x)$ is the unique weak solution to the evolution equation*

$$\left[\begin{array}{ll} \dfrac{\mathrm{d}}{\mathrm{d}t}u_t(x) = (\mathcal{G}u_t)(x) & \text{for } t \geqslant 0 \text{ and } x \in \mathbb{Z}, \\ u_0(x) = f(x) & \text{for } x \in \mathbb{Z}. \end{array}\right.$$

We skip the details, as they are almost exactly the same as their analogous reasons for Theorem 15.3.2. In fact, it is slightly easier to prove Theorem 15.3.9 than to prove Theorem 15.3.2 because we are always considering Fourier integrals on the compact set $\mathbb{T} = [0\,,2\pi)$ here, whereas the Fourier integrals behind

[3] Thus, for example, if X is the continuous-time simple symmetric random walk on $\mathbb{Z}$, then $D = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, whence $(\mathcal{G}f)(x) = \frac{1}{2}(f(x+1) + f(x-1) - 2f(x))$ is the discrete Laplacian of f. We could alternatively derive the generator $\mathcal{G}$ by appealing to the strong Markov property, as we will in the next example.

Theorem 15.3.2 were at times more delicate since they are over the entire real numbers.

If you were not convinced before that under all this lies a general abstract theory of Lévy processes on LCA groups, perhaps you are now. We will not develop that here; instead let us study one final (very simple) example that could very well have been presented first, as it is likely to be well known to many readers already.

Example 15.3.10 (Lévy processes on $\mathbb{Z}/2\mathbb{Z}$)**.** As we saw earlier, the class of all Lévy processes on $\mathbb{Z}/2\mathbb{Z}$ coincides with all random walks on $\{0\,,1\}$. That is, if X is a Lévy process on $\mathbb{Z}/2\mathbb{Z} \cong \{0\,,1\}$ then we can find $\kappa \geqslant 0$ such that X begins at 0 and switches to 1 to 0 to 1, etc., at the jump times of a Poisson process $N := \{N(t)\}_{t\geqslant 0}$ with rate κ. (If $\kappa = 0$, then N is degenerate; i.e., it does not jump.) And conversely, for every $\kappa \geqslant 0$ there is such a random walk. Thus, the family of random walks on $\mathbb{Z}/2\mathbb{Z}$ coincides with the collection of all stationary Markov chains on $\{0\,,1\}$, equivalently, all stationary 2-state automatons. This is a rather simple setting. But we will do the analysis in the style of the preceding examples in order to suggest that there are deeper underlying methods that are tied intimately to the structure of the underlying group (here, $\mathbb{Z}/2\mathbb{Z}$).

One can compute the generator of X in analogy with the preceding examples, by using a suitable form of the "Fourier transform" on $\mathbb{Z}/2\mathbb{Z}$. But perhaps the best definition of the Fourier transform on the LCA group $\mathbb{Z}/2\mathbb{Z}$ is the trivial one: $\hat{f}(z) = f(z)$ for all $z \in \mathbb{Z}/2\mathbb{Z}$. Therefore, we might as well compute the generator using elementary probabilistic ideas.

The family $L^2(\mathbb{Z}/2\mathbb{Z})$ is identified with the collection all real-valued functions on $\{0\,,1\}$. Notice that, as $t \downarrow 0$,

(i) $\mathrm{P}\{N(t) = 0\} = \exp(-\kappa t) = 1 - \kappa t + O(t^2)$;
(ii) $\mathrm{P}\{N(t) = 1\} = \kappa t \exp(-\kappa t) = \kappa t + O(t^2)$; and
(iii) $\mathrm{P}\{N(t) \geqslant 2\} = 1 - \exp(-\kappa t) - \kappa t \exp(-\kappa t) = O(t^2)$.

Since $X_0 = 0$ it follows that, for every $f \in L^2(\mathbb{Z}/2\mathbb{Z})$,

$$\begin{aligned}(P_t f)(0) &= \mathrm{E}\left[f(X_t); N(t) = 0\right] + \mathrm{E}\left[f(X_t); N(t) = 1\right] + \mathrm{E}\left[f(X_t); N(t) \geqslant 2\right] \\ &= f(0)\left(1 - \kappa t + O(t^2)\right) + f(1)\kappa t + O(t^2) \qquad \text{as } t \downarrow 0.\end{aligned}$$

Therefore,

$$(\mathcal{G}f)(0) := \lim_{t\downarrow 0} \frac{(P_t f)(0) - f(0)}{t} = \kappa v(f(1) - f(0)).$$

Similarly,

$$(\mathcal{G}f)(1) := \lim_{t\downarrow 0} \frac{(P_t f)(1) - f(1)}{t} = \kappa(f(0) - f(1)).$$

In other words, $(\mathcal{G}f)(x) = \kappa(f(x+1) - f(x))$, for $x \in \mathbb{Z}/2\mathbb{Z}$, where we recall that "$x + 1$" refers to the group multiplication operation (stated in additive terms).

The linear operator $\mathcal{G}$ is the *generator* of the walk, and its *domain* $\mathrm{Dom}[\mathcal{G}]$ is all of $L^2(\mathbb{Z}/2\mathbb{Z})$.

We have proved most of the following, and leave the remaining few details as exercise for the interested reader.

Theorem 15.3.11. *Let X be a random walk on $\mathbb{Z}/2\mathbb{Z}$ with jump rate $\kappa \geqslant 0$. If $f \in L^2(\mathbb{Z}/2\mathbb{Z})$, then the 2-parameter function $u_t(x) := (P_t f)(x)$ is the unique weak solution to the evolution equation*

$$\left[\begin{array}{ll} \dfrac{\mathrm{d}}{\mathrm{d}t} u_t(x) = (\mathcal{G}u_t)(x) & \text{for } t \geqslant 0 \text{ and } x \in \mathbb{Z}/2\mathbb{Z}, \\ u_0(x) = f(x) & \text{for } x \in \mathbb{Z}/2\mathbb{Z}, \end{array}\right. \tag{15.12}$$

where $(\mathcal{G}h)(x) = \kappa[h(x+1) - h(x)]$ for $x \in \mathbb{Z}/2\mathbb{Z}$ and $h \in L^2(\mathbb{Z}/2\mathbb{Z})$.

One can use the *Kolmogorov–Fokker–Planck equation* (15.12) itself to compute $p_t(x)$ whence $(P_t f)(x)$. For instance, we can set $f := \delta_0$ in order to see that

$$(P_t\delta_0)(x) = \mathrm{E}\delta_0(x + X_t) = \mathrm{P}\{x + X_t = 0\} = p_t(x),$$

and that $P_t\delta_0$ solves (15.12) with $f = \delta_0$. Now,

$$(\mathcal{G}P_t\delta_0)(x) = \kappa[(P_t\delta_0)(x+1) - (P_t\delta_0)(x)] = \kappa[p_t(x+1) - p_t(x)] = \kappa\left[1 - 2p_t(x)\right].$$

It follows from Theorem 15.3.11 that $\mathrm{d}p_t(0)/\mathrm{d}t = (\mathcal{G}p_t)(0) = \kappa[1 - 2p_t(0)]$, subject to the initial condition $p_0(0) = 1$. This is a simple ODE that can be solved explicitly. The solution is

$$p_t(0) = \frac{1 + \mathrm{e}^{-2\kappa t}}{2}, \quad \text{whence also} \quad p_t(1) = \frac{1 - \mathrm{e}^{-2\kappa t}}{2},$$

for all $t \geqslant 0$.

Chapter 16

SPDEs

The "SPDEs" in the title is shorthand for *stochastic partial differential equations* or more appropriately still, *stochastic partial integro-differential equations*. This is the main topic of these lecture notes. Throughout, G denotes an LCA group and X a Lévy process on G with semigroup $\{P_t\}_{t\geqslant 0}$, whose "generator" is denoted by $\mathcal{G}$. By the latter, we mean a linear operator that is defined densely on $L^2(G)$, and satisfies

$$\lim_{t\to\infty} \frac{P_t f - f}{t} = \mathcal{G} f$$

in $L^2(G)$, for every function $f\colon G \to \mathbb{R}$ in the domain $\mathrm{Dom}[\mathcal{G}]$ of the definition of G. Such a generator always exists (see Jacob [27–29]); we have verified this claim for many interesting examples of G.

16.1 A heat equation

For reasons that should by now be familiar, one can then solve the following *Kolmogorov–Fokker–Planck equation* for each $f \in L^2(G)$: $\dot{u}_t(x) = (\mathcal{G}u_t)(x)$ on $(0\,,\infty)\times G$, subject to $u_0 = f$, where $\dot{u}$ denotes $\mathrm{d}u/\mathrm{d}t$ (the weak derivative). This equation is also known as the *linear heat equation for* $\mathcal{G}$, and has a unique weak solution

$$u_t(x) = (P_t f)(x) = (f * p_t)(x).$$

We will be concerned only with Lévy processes whose transition measures $p_t(A) := \mathrm{P}\{-X_t \in A\}$ are nice functions when $t > 0$. Of course, $p_0 = \delta_{\mathrm{e}}$ cannot be a function; specifically,

Convention 16.1.1. We assume from now on that there exists a measurable function $p_t(x)$ of $(t\,,x) \in (0\,,\infty)\times G$ such that

$$p_t(A) = \int_A p_t(x)\, m_G(\mathrm{d}x),$$

for all $t > 0$ and Borel sets $A \subset G$ of finite Haar measure.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_16

We will refer to the functions p_t as the *transition functions* of the Lévy process X.

Example 16.1.2. If G is discrete, then transition functions exist vacuously, and $p_t(x) = \mathrm{P}\{-X_t = x\}$ for all $t \geqslant 0$ and $x \in G$. For more interesting examples, consider the case where $G = \mathbb{R}$ and X is a symmetric α-stable Lévy process for some $\alpha \in (0\,,2]$ ($\alpha = 2$ is Brownian motion). In all cases, we can find $\kappa > 0$ such that

$$\widehat{p}_t(z) = \exp(-t\kappa|z|^\alpha), \tag{16.1}$$

for all $t > 0$ and $z \in \mathbb{R}$ (see Examples 15.1.4 and 15.1.6). Since $\widehat{p}_t \in L^1(\mathbb{R})$ for all $t > 0$, the inversion formula yields

$$p_t(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-ixz-t\kappa|z|^\alpha}\,\mathrm{d}z,$$

for almost every $x \in \mathbb{R}$. The right-hand side is a continuous function of $(t\,,x) \in (0\,,\infty)\times\mathbb{R}$ thanks to the dominated convergence theorem. Therefore, we can *define* $p_t(x)$ as the preceding pointwise, and still retain the property (16.1).

Example 16.1.3. Every Lévy process X on $\mathbb{T} := [0\,,2\pi)$ is obtained by wrapping a Lévy process Y on $\mathbb{R}$ around the torus. Now suppose Y is a Lévy process on $\mathbb{R}$ such that

$$\mathrm{P}\{-Y_t \in A\} = \int_A p_t^Y(y)\,\mathrm{d}y,$$

for all Borel sets $A \subset \mathbb{R}$, where $(t\,,y) \mapsto p_t^Y(y)$ is uniformly continuous and bounded on $[\varepsilon\,,\infty)\times\mathbb{R}$ for every $\varepsilon > 0$. Then, for all Borel sets $B \subset [0\,,2\pi)$,

$$\mathrm{P}\{-X_t \in B\} = \sum_{n=-\infty}^{\infty} \mathrm{P}\left\{-2n\pi - Y_t \in B\right\} = \sum_{n=-\infty}^{\infty}\int_B p_t^Y(y+2n\pi)\,\mathrm{d}y.$$

Thus, using obvious notation, we have the following version of $p_t^X(x)$ for the process X:

$$p_t^X(x) := \sum_{n=-\infty}^{\infty} p_t^Y(x+2n\pi),$$

for every $t > 0$ and $x \in \mathbb{T}$. In this way we can see that a mild decay condition on $y \mapsto p_t^Y(y)$ will ensure that p^X is bounded and continuous uniformly on $[\varepsilon\,,\infty)\times\mathbb{T}$ for every $\varepsilon > 0$. It can be shown that a good enough decay condition in fact holds when Y is a symmetric α-stable Lévy process on $\mathbb{R}$; this fact is elementary when $\alpha = 2$, but requires work when $\alpha \in (0\,,2)$. We omit the details.

Let us choose and fix a finite Borel measure μ on $(0\,,\infty)\times G$ that has "bounded thermal potential"; see Watson [58, 59]. This means that

$$\sup_{x\in G}\sup_{t\in[0,T]}\int_{[0,t]\times G} p_{t-s}(xy^{-1})\,\mu(\mathrm{d}s\,\mathrm{d}y) < \infty$$

for all $T > 0$. Choose and fix a Lipschitz-continuous function $\sigma\colon \mathbb{R} \to \mathbb{R}$ and a function $f \in L^2(G) \cap L^\infty(G)$. Then a more or less standard method from partial differential equations (Duhamel's principle, or more aptly in this case, variation of parameters) shows that the partial integro-differential equation[1]

$$\dot{u}_t(x) = (\mathcal{G}u_t)(x) + \sigma(u_t(x))\mu(t\,, x) \tag{16.2}$$

subject to $u_0(x) = f(x)$, also has a weak solution that can be represented, in integral form, as the unique solution to

$$u_t(x) = (P_t f)(x) + \int_{[0,t]\times G} p_{t-s}(xy^{-1})\sigma(u_s(y))\,\mu(\mathrm{d}s\,\mathrm{d}y). \tag{16.3}$$

Moreover, that solution is the only one which satisfies the additional boundedness condition, $\|u\|_{L^\infty([0,t]\times G,\mu)} < \infty$ for all $T > 0$. Furthermore, standard localization arguments can be applied to show that the preceding continues to hold as long as $f \in L^\infty(G)$. For the sake of simplicity, we will consider only the case that $f \equiv 1$ from now on; this choice simplifies the exposition a little, since in this case $P_t f \equiv 1$, manifestly. And the integral form of the solution to (16.2) also simplifies a little, namely

$$u_t(x) = 1 + \int_{[0,t]\times G} p_{t-s}(xy^{-1})\sigma(u_s(y))\,\mu(\mathrm{d}s\,\mathrm{d}y). \tag{16.4}$$

The SPDEs of this chapter are the same as (16.2), with $f \equiv 1$, except μ is no longer a measure of bounded thermal potential. Rather, μ are replaced by a space-time white noise ξ on G. And we interpret the meaning of the resulting SPDE as the solution to (16.4), if there is one, with μ replaced by ξ. This undertaking at least might make sense, since the integral in (16.4) can in principle be interpreted as a Walsh integral.

We have seen that in many cases Walsh's stochastic integral generalize that of Itô. In similar vein, we will soon see that our SPDEs generalize Itô stochastic differential equations, and our interpretation of an SPDE is a direct generalization, once again *à la* Itô, of the usual interpretation of a one-dimensional Itô stochastic differential equation.

16.2 A parabolic SPDE

Let X be a Lévy process on an LCA group G. We have seen that, in many contexts, X has a generator $\mathcal{G}$, which is a densely-defined linear mapping from its domain $\mathrm{Dom}[\mathcal{G}] \subset L^2(G)$ to $L^2(G)$. Suppose $\sigma\colon \mathbb{R} \to \mathbb{R}$ is a measurable function. We now

[1]When μ is a function, this is fine as is; when it is only a measure, one can make sense of (16.2) in a weak sense. The end result is (16.3).

consider parabolic SPDEs of the following type:

$$\begin{bmatrix} \dot{u}_t(x) = (\mathcal{G}u_t)(x) + \sigma(u_t(x))\xi(t\,,x) & \text{on } (0\,,\infty)\times G, \\ u_0 \equiv 1. \end{bmatrix} \tag{16.5}$$

The preceding has to be interpreted suitably because, as it turns out, the solution will typically not be in the domain of the definition of $\mathcal{G}$ nor $\mathrm{d}/\mathrm{d}t$. This is because ξ is not typically a measure (e.g., for almost all ω). And, of course, ξ is typically not a function, and so $\xi(t\,,x)$ does not make sense.

Still, the integral form (16.4) of the solution continues to make sense, provided that we interpret the $\mathrm{d}\mu$-integral in (16.4) as a Walsh stochastic integral. Therefore, in analogy with the discussion of the preceding section for non-random parabolic equations of the general type (16.5), we would like to say that u is a *solution* to (16.5) in *integral form* if it solves

$$u_t(x) = 1 + \int_{[0,t]\times G} p_{t-s}(xy^{-1})\sigma(u_s(y))\,\xi(\mathrm{d}s\,\mathrm{d}y) \tag{16.6}$$

for all $t > 0$ and $x \in G$, where the integral is understood as a Walsh integral. This is still not quite a fully rigorous definition. Since u and ξ are random, we could interpret the preceding notion of solution in different ways. For instance, a rather weak notion of solution might be to expect u to solve the preceding display for every $t > 0$ and $x \in G$ almost surely; that is, the null set off which (16.6) holds might depend on t and x. A slightly better interpretation is to understand (16.6), in terms of stochastic convolutions, as $u = 1 + (p \circledast \sigma(u))$ a.s., where the identity holds as elements of $\mathscr{W}_\alpha$ for some $\alpha \geqslant 0$; see Section 14.7.

Definition 16.2.1. We say that u is a *mild solution* to (16.5) if there exists $\alpha \geqslant 0$ such that $u \in \mathscr{W}_\alpha$, $\sigma(u) \in \mathscr{W}_\alpha$, and $u = 1 + (p \circledast \sigma(u))$, where the equality holds in $\mathscr{W}_\alpha$.

One can show that a mild solution to (16.5) is a "weak solution" to (16.5) as well. This undertaking hinges on a "stochastic Fubini theorem" which we will not develop here (for related work, see Khoshnevisan [32, §5.3]). Instead, let us present the first fundamental theorem of these lectures.

Theorem 16.2.2. *Suppose*

$$\Upsilon(\alpha) := \int_0^\infty \mathrm{e}^{-\alpha s}\|p_s\|^2_{L^2(G)}\,\mathrm{d}s < \infty \tag{16.7}$$

for some $\alpha > 0$. Then, there exists $\gamma > 0$ and a unique random field $u \in \cap_{\beta>\gamma}\mathscr{W}_\beta$ such that u is a mild solution to (16.5).

We will prove Theorem 16.2.2 at the end of this chapter. First, we would like to understand better the meaning of Condition (16.7), which we call *Dalang's condition*, after the seminal work [12].

Let X' denote an independent copy of X. The process $t \mapsto \bar{X}_t := X_t^{-1} X_t'$ is a Lévy process on G with transition functions

$$\bar{p}_t(x) := \int_G p_t(yx^{-1}) p_t(y)\, m_G(\mathrm{d}y),$$

for all $t > 0, x \in G$. Thus, Dalang's condition (16.7) is often equivalent to the condition that $\bar{X}$ has a finite *α-potential* $\bar{r}_\alpha$ at $x = 0$ for some $\alpha > 0$, where

$$\bar{r}_\alpha(x) := \int_0^\infty \mathrm{e}^{-\alpha t} \bar{p}_t(x)\, \mathrm{d}t,$$

for all $x \in G$; see Foondun–Khoshnevisan [18, 19]. One can use probabilistic potential theory in order to prove that (16.7) is equivalent to the seemingly stronger condition that $\bar{r}_\beta(0) < \infty$ for all $\beta > 0$. In the next few subsections, we will work out examples of this phenomenon. The analysis of the general case requires a good deal of probabilistic potential theory, which is a subject for another lecture series.

Of course, the condition $\bar{r}_0(0) < \infty$ subsumes (16.7). But the former condition is rarely met, whereas (16.7) holds in many natural situations, as we will soon see.

16.2.1 Lévy processes on $\mathbb{R}$

Let X be a Lévy process on $\mathbb{R}$ whose characteristic exponent Ψ satisfies the Lévy–Khintchine representation (15.1). Since $|\mathrm{E}\exp(izX_t)|^2 = \exp(-2t\mathrm{Re}\Psi(z)) \leqslant 1$, it follows that $\mathrm{Re}\Psi(z) \geqslant 0$ for all $z \in \mathbb{R}$.

Suppose that

$$\lim_{|z| \to \infty} \frac{\mathrm{Re}\Psi(z)}{\log |z|} = \infty. \tag{16.8}$$

Then, $\int_{-\infty}^{\infty} |\mathrm{E}\exp(izX_t)|\, \mathrm{d}z < \infty$ and hence,

$$p_t(x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-ixz - t\Psi(z)}\, \mathrm{d}z,$$

$t > 0$, $x \in \mathbb{R}$, is absolutely convergent and defines a version of the transition functions of X, thanks to the inversion formula. Moreover,

$$\bar{p}_t(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-ixz - 2t\mathrm{Re}\Psi(z)}\, \mathrm{d}z \leqslant \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-2t\mathrm{Re}\Psi(z)}\, \mathrm{d}z = \bar{p}_t(0).$$

Therefore, Tonelli's theorem implies that, for all $\alpha > 0$,

$$\bar{r}_\alpha(0) = \sup_{x \in \mathbb{R}} \bar{r}_\alpha(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d}z}{\alpha + 2\mathrm{Re}\Psi(z)}.$$

At this point you should convince yourself that $\bar{r}_\alpha(0) < \infty$ for some $\alpha > 0$ if and only if $\bar{r}_\alpha(0) < \infty$ for all $\alpha > 0$. When X is a symmetric β-stable process (that is, $\Psi(z) \propto |z|^\beta$) then $\beta \in (0\,,2]$, and (16.8) holds. In that case, Dalang's condition (16.7) is equivalent to the condition $\beta \in (1\,,2]$.

The condition $\bar{r}_0(0) < \infty$ is equivalent to

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}z}{\mathrm{Re}\Psi(z)} < \infty. \tag{16.9}$$

Condition (16.9) and Dalang's conditions will never both hold when X is a symmetric stable process ($\Psi(z) \propto |z|^\alpha$ for some $\alpha \in (0\,,2]$). But they do hold if, for example, $\Psi(z) = |z|^\alpha + |z|^\beta$ when $\alpha \in (1\,,2]$ and $\beta \in (0\,,1)$. The resulting Lévy processes are examples of *tempered stable processes*; see Rosiński [48].

We complete the example of linear Lévy processes by saying a few words about Condition (16.8). Recall from (15.1) that

$$\mathrm{Re}\,\Psi(z) = \frac{\sigma^2 z^2}{2} + \frac{1}{2}\int_{-\infty}^{\infty} [1 - \cos(xz)]\, m(\mathrm{d}x)$$

for all $z \in \mathbb{R}$. It follows that if $\sigma \neq 0$ then (16.8) holds automatically; (16.8) can hold also when $\sigma = 0$. To see this, let us suppose that $\sigma = 0$ and $m(\mathrm{d}x) = \dot{m}(|x|)\,\mathrm{d}x$ for a symmetric nonincreasing function $\dot{m}\colon (0\,,\infty) \to \mathbb{R}_+$ satisfying $\int_0^\infty (1 \wedge x^2)\dot{m}(x)\,\mathrm{d}x < \infty$. (Every such function $\dot{m}$ corresponds to a Lévy measure m.) Since $1 - \cos\theta \geqslant \theta^2/2 - \theta^4/24$ for all real θ, it follows that

$$1 - \cos\theta \geqslant \frac{1}{2}\theta^2 - \frac{1}{24}\theta^2 = \frac{11}{24}\theta^2$$

for all $\theta \in [0\,,1]$. Consequently,

$$\begin{aligned}\mathrm{Re}\Psi(z) &\geqslant 2\int_0^\infty (1 - \cos(x|z|))\dot{m}(x)\,\mathrm{d}x \geqslant \frac{11}{12} z^2 \int_{1/2|z|}^{1/|z|} x^2 \dot{m}(x)\,\mathrm{d}x \\ &\geqslant \frac{11}{48}\int_{1/2|z|}^{1/|z|} \dot{m}(x)\,\mathrm{d}x \geqslant \frac{11\dot{m}(1/z)}{96|z|}\end{aligned}$$

for all $z \neq 0$. In particular, (16.8) holds for example when

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon \dot{m}(\varepsilon)}{|\log \varepsilon|} = \infty.$$

The preceding is a "tauberian/abelian" argument in disguise and can also be reversed.

16.2.2 Lévy processes on a denumerable LCA group

If G is countable or finite, for example if $G = \mathbb{Z}$ or $G = \mathbb{Z}/n\mathbb{Z}$ for some integer $n \geqslant 1$, then $\bar{p}_t(x) = \mathrm{P}\{\bar{X}_t = x\}$ is bounded between zero and one. In particular,

$$\bar{r}_\alpha(0) \leqslant \int_0^\infty \mathrm{e}^{-\alpha s}\,\mathrm{d}s = \alpha^{-1} < \infty$$

for all $\alpha > 0$. That is, Dalang's Condition (16.8) holds *a fortiori* and without further restrictions. In this case it is easy to understand exactly when $\bar{r}_0(0)$ is finite as well. Indeed, we note that

$$\|p_t\|_{L^2(G)}^2 = \sum_{x\in G} \mathrm{P}\{-X_t = x\}\mathrm{P}\{-X_t = x\} = \mathrm{P}\{\bar{X}_t = 0\},$$

for all $t > 0$. Therefore,

$$\bar{r}_0(0) = \mathrm{E}\int_0^\infty \mathbf{1}_{\{0\}}(\bar{X}_s)\,\mathrm{d}s,$$

thanks to Tonelli's theorem. The Chung–Fuchs theory [7] of classical random walks now applies, and tells us that $\bar{r}_0(0) < \infty$ if and only if the symmetrized Lévy process $\bar{X}$ is recurrent.

16.2.3 Proof of Theorem 16.2.2

We wish to produce a good solution to $u = 1 + (p \circledast \sigma(u))$.

Recall that constants are always in $\mathscr{W}_0$, whence in $\mathscr{W}_\beta$ for every $\beta > 0$ (see Theorem 14.7.1). In particular, Theorem 14.7.2 implies our result when σ is constant. Therefore, we will consider only the case that σ is not a constant. In that case, the proof uses a fixed-point argument. Namely, define $u_t^{(0)}(x) := 1$, and then iteratively set

$$u_t^{(n+1)}(x) := 1 + \left[p \circledast \sigma(u^{(n)})\right]$$

for all $n \geqslant 0$, provided that the stochastic convolution $p \circledast \sigma(u^{(n)})$ makes sense. We might refer to $u^{(0)}, u^{(1)}, \ldots$ as a *Picard iteration sequence*, if and when it is well defined.

We will use the following lemma in a few spots.

Lemma 16.2.3. *If $v \in \mathscr{W}_\beta$ for some $\beta \geqslant 0$, then $\sigma(v) \in \mathscr{W}_\beta$ as well.*

Proof. Since σ is Lipschitz continuous and non constant,

$$\mathrm{Lip}_\sigma := \sup_{\substack{w,z\in\mathbb{R}\\ w\neq z}} \left|\frac{\sigma(w)-\sigma(z)}{w-z}\right|$$

is strictly positive and finite. In particular, $|\sigma(w)| \leqslant |\sigma(0)| + \mathrm{Lip}_\sigma |w|$, for all $w \in \mathbb{R}$. We may now apply the triangle inequality for the norm $\mathcal{N}_\beta$ in order to see that if $v \in \mathscr{W}_\beta$, then

$$\mathcal{N}_\beta(\sigma(v)) \leqslant |\sigma(0)| + \mathrm{Lip}_\sigma \mathcal{N}_\beta(v) < \infty.$$

There exist elementary random fields $v^{(1)}, v^{(2)}, \ldots$ such that

$$\lim_{n\to\infty} \mathcal{N}_\beta\left(v^{(n)} - v\right) = 0.$$

Since
$$\left|\sigma(v_t^{(n)}(x)) - \sigma(v_t(x))\right| \leqslant \mathrm{Lip}_\sigma \left|v_t^{(n)}(x) - v_t(x)\right|$$
for all $t > 0$ and $x \in G$, it follows that
$$\lim_{n\to\infty} \mathcal{N}_\beta\left(\sigma(v^{(n)}) - \sigma(v)\right) = 0.$$
Since$v^{(n)}$ is an elementary random field, so is $\sigma(v^{(n)})$. Thus, the lemma follows. □

With Lemma 16.2.3 proved and tucked away nicely, we can complete the proof of the theorem.

Proof of Theorem 16.2.2. Since $u^{(0)} \in L^\infty((0\,,\infty)\times G)$, Theorem 14.7.1 ensures that $u^{(0)} \in \mathscr{W}_0$, whence $u^{(0)} \in \mathscr{W}_\beta$ for every $\beta \geqslant 0$. Lemma 16.2.3 shows that $\sigma(u^{(0)}) \in \mathscr{W}_\beta$ for all $\beta \geqslant 0$ as well. Define
$$\gamma := \inf\left\{\alpha > 0 \,\middle|\, \Upsilon(\alpha) > \frac{1}{\mathrm{Lip}_\sigma^2}\right\}.$$
Since σ is non constant, $\gamma \in [0\,,\infty)$.

If $\beta > \gamma$, then we appeal to Theorem 14.7.2 with $\kappa := p$ in order to see that $p \circledast \sigma(u^{(0)}) \in \mathscr{W}_\beta$ as well. Since constants are in $\mathscr{W}_\beta$, it follows also that $u^{(1)} \in \mathscr{W}_\beta$ for every $\beta > \gamma$. The very same reasoning shows us that, whenever there exists an integer $n \geqslant 0$ such that $u^{(n)} \in \mathscr{W}_\beta$ for every $\beta > \gamma$, then also $p \circledast \sigma(v^{(n)}), u^{(n+1)} \in \mathscr{W}_\beta$ for all $\beta > \gamma$.

Next, let us observe that if $n \geqslant 1$ and $\beta > 0$, then
$$\begin{aligned}\mathcal{N}_\beta\left(u^{(n+1)} - u^{(n)}\right) &= \mathcal{N}_\beta\left(p \circledast \left[\sigma(u^{(n)}) - \sigma(u^{(n-1)})\right]\right)\\ &\leqslant \mathcal{N}_\beta\left(\sigma(u^{(n)}) - \sigma(u^{(n-1)})\right)\sqrt{\Upsilon(\beta)}\\ &\leqslant \mathrm{Lip}_\sigma \mathcal{N}_\beta\left(u^{(n)} - u^{(n-1)}\right)\sqrt{\Upsilon(\beta)}.\end{aligned}$$
The first inequality follows from Theorem 14.7.2. The function Υ is continuous and decreasing. Therefore, $\mathrm{Lip}_\sigma\sqrt{\Upsilon(\beta)} < 1$ if $\beta > \gamma$, and hence
$$\sum_n \mathcal{N}_\beta\left(u^{(n+1)} - u^{(n)}\right) < \infty$$
if $\beta > \gamma$. In particular, $n \mapsto u^{(n)}$ is a Cauchy sequence in the Banach space $(W_\beta, \mathcal{N}_\beta)$; therefore, $u := \lim_{n\to\infty} u^{(n)}$ exists in $\mathcal{W}_\beta$ for every $\beta > \gamma$. The $\mathscr{W}_\beta$'s are nested, so u is in $\cap_{\beta>\gamma}\mathscr{W}_\beta$. Finally, note that
$$\begin{aligned}\sum_{n=1}^\infty \mathcal{N}_\beta\left(p \circledast \sigma(u) - p \circledast \sigma(u^{(n)})\right) &= \sum_{n=1}^\infty \mathcal{N}_\beta\left(p \circledast \left[\sigma(u) - \sigma(u^{(n)})\right]\right)\\ &\leqslant \mathrm{Lip}_\sigma \sum_{n=1}^\infty \mathcal{N}_\beta(u - u^{(n)})\sqrt{\Upsilon(\beta)} < \infty.\end{aligned}$$

Therefore, $p \circledast \sigma(u^{(n)})$ converges to $p \circledast \sigma(u)$ in $\mathcal{W}_\beta$ for every $\beta > \gamma$ as well. It follows that u is a fixed point of $u = 1 + p \circledast \sigma(u)$, viewed as an equation on $\cap_{\beta>\gamma}\mathscr{W}_\beta$. Finally, if $v \in \cap_{\beta>\gamma}\mathscr{W}_\beta$ also solves $v = 1 + p \circledast \sigma(v)$, then

$$\mathcal{N}_\beta(u-v) = \mathcal{N}_\beta(p \circledast [\sigma(u) - \sigma(v)]) \leqslant \mathrm{Lip}_\sigma \sqrt{\Upsilon(\beta)}\mathcal{N}_\beta(u-v),$$

for all $\beta \geqslant 0$. If β is sufficiently large, then $\mathcal{N}_\beta(u-v) < \infty$ and $\mathrm{Lip}_\sigma\sqrt{\Upsilon(\beta)} < 1$; therefore, the preceding shows that $\mathcal{N}_\beta(u-v)$ is zero. Uniqueness follows. □

16.3 Examples

We can best understand our SPDE (16.5) via somewhat more concrete examples.

16.3.1 The trivial group

If $G := \{\mathrm{e}\}$ is the trivial group, then the only Lévy process on G is $X_t := \mathrm{e}$ and in fact all functions on G are constants! The transition function of X is $p_t(\mathrm{e}) = 1$, and the generator of X is also computed easily: $P_t f = f$ for all $f \in L^2(G) \cong \mathbb{R}$ and $t \geqslant 0$, whence $(\mathcal{G}f)(\mathrm{e}) := 0$ for all function f on G. Let us write $u_t := u_t(\mathrm{e})$ and $\xi(t) := \xi(t\,,\mathrm{e})$ in order to see that our formal "equation" (16.5) reduces to the following formal equation,

$$\mathrm{d}u_t = \sigma(u_t)\xi(t)\,\mathrm{d}t, \tag{16.10}$$

subject to $u_0 = 1$. The rigorous interpretation of this equation is via the integral formulation $u = 1 + p \circledast \sigma(u) = 1 + \mathbf{1} \circledast \sigma(u)$ of the solution, where $\mathbf{1}(\mathrm{e}) := 1$.

Recall that $B_t := \xi([0\,,t] \times \{\mathrm{e}\})$ defines a Brownian motion and

$$(\mathbf{1} \circledast \sigma(u))_t(\mathrm{e}) = \int_{[0,t]\times\{\mathrm{e}\}} \sigma(u_s(y))\,\xi(\mathrm{d}s\,\mathrm{d}y) = \int_0^t \sigma(u_s)\,\mathrm{d}B_s.$$

This means that (16.5) is described as the solution $\{u_s\}_{s\geqslant 0}$ to

$$u_t = 1 + \int_0^t \sigma(u_s)\,\mathrm{d}B_s \tag{16.11}$$

($t \geqslant 0$). Most people write (16.10) formally as $\mathrm{d}u_t = \sigma(u_t)\,\mathrm{d}B_t$, subject to $u_0 = 1$, and interpret this rigorously as (16.11). The equation (16.11) is called an *Itô stochastic differential equation* (SDE) *with zero drift*. In other words, when G is trivial, our SPDE (16.5) coincides with an arbitrary drift-free Itô stochastic differential equation (with initial value one). One can generalize (16.5) to include all one-dimensional Itô stochastic differential equations. But we will refrain from developing that theory here.

16.3.2 The cyclic group on two elements

It might be tempting to think that since (16.5) codes all drift-free Itô SDEs when G is a group on one element, then (16.5) codes all two-dimensional SDEs when $G = \mathbb{Z}/2\mathbb{Z} \cong \{0\,,1\}$ is an LCA group on two elements. This turns out not to be the case. In fact, something more interesting happens.

Let us write G as $\{0\,,1\}$ in the usual way, so that $B_t^{(i)} := \xi([0\,,t] \times \{i\})$ defines two independent Brownian motions, $i = 1, 2$. Every Lévy process on G corresponds, in a 1-1 fashion, to a non-negative number κ in the following way: X starts at 0 at time zero; at rate κ it jumps to 1; then, at rate κ it jumps to 0; and so on; see Section 15.1.4. The generator of X is described via

$$(\mathcal{G}f)(i) = \kappa\,[f(i+1 \bmod 2) - f(i)]\,,$$

for $i \in \{0\,,1\}$ and $f \in L^2(\{0\,,1\}) \cong \mathbb{R}^2$.[2] The resulting form of (16.5) is usually written as

$$\mathrm{d}u_t(x) = \kappa\,[u_t(x+1 \bmod 2) - u_t(x)] + \sigma(u_t(x))\,\mathrm{d}B_t^{(x)}$$

for $t > 0$ and $x \in \{0\,,1\}$, subject to $u_0 \equiv 1$. Equivalently, we may write the preceding, "in coordinates," as the coupled system of SDEs,

$$\begin{bmatrix} \mathrm{d}u_t(0) = \kappa\,[u_t(1) - u_t(0)]\,\mathrm{d}t + \sigma\,(u_t(0))\,\mathrm{d}B_t^{(0)} \\ \mathrm{d}u_t(1) = \kappa\,[u_t(0) - u_t(1)]\,\mathrm{d}t + \sigma\,(u_t(1))\,\mathrm{d}B_t^{(1)}, \end{bmatrix} \tag{16.12}$$

subject to $u_0 = (1\,,1)$, where we have written $u_t = (u_t(0), u_t(1))$.

Think of $t \mapsto u_t = (u_t(0), u_t(1))$ as a model for the molecular motion of a two-body system, where $u_t(x)$ denotes the position of the particle x at time t. Then, we see that each of the two particles 0 and 1 moves locally according to a standard one-dimensional diffusion $\mathrm{d}x_t = \sigma(x_t)\,\mathrm{d}b_t$, but there is *Ornstein–Uhlenbeck type ("molecular") attraction* between the two particles: for each $x \in \{0\,,1\}$, particle x diffuses independently, but will also move toward particle $x+1 \bmod 2$ with drift proportional to the spatial distance between the two particles.

In order to better see what is going on let us consider the simplest case: $\kappa = 1$ and $\sigma(x) \equiv 1$ for all $x \in \mathbb{R}$. Let us define, for all $t \geqslant 0$,

$$W_t^{(0)} := \frac{B_t^{(0)} + B_t^{(1)}}{\sqrt{2}}, \qquad W_t^{(1)} := \frac{B_t^{(0)} - B_t^{(1)}}{\sqrt{2}}.$$

It is easy to see that $W^{(0)}$ and $W^{(1)}$ are two independent standard Brownian motions. Next, we plan to write u_t in terms of $(W^{(0)}\,, W^{(1)})$. In order to do this, let us define S_t to denote the "separation" between $u_t(0)$ and $u_t(1)$: $S_t := u_t(0) - u_t(1)$

[2] The Hilbert space isometry is, of course, $L^2(\{0\,,1\}) \ni f \mapsto (f(0)\,, f(1)) \in \mathbb{R}^2$.

for all $t \geqslant 0$. It follows from (16.12) that S solves the Itô stochastic differential equation

$$\mathrm{d}S_t = -2S_t\,\mathrm{d}t + \sqrt{2}\mathrm{d}W_t^{(1)},$$

for all $t \geqslant 0$, and subject to $S_0 = u_0(0) - u_0(1) = 0$. The process S is called an *Ornstein–Uhlenbeck process*, and can be explicitly solved in terms of $W^{(1)}$ as follows:

$$S_t = \sqrt{2}\,\mathrm{e}^{-2t}\int_0^t \mathrm{e}^{2s}\,\mathrm{d}W_s^{(1)}.$$

In particular, S and $W^{(0)}$ are independent. At the same time, we can see from (16.12) that

$$u_t(0) + u_t(1) = 2 + \sqrt{2}W_t^{(0)}$$

for all $t \geqslant 0$. Therefore, we can write

$$u_t(0) = \left(\frac{u_t(0)+u_t(1)}{2}\right) + \left(u_t(0) - \frac{u_t(0)+u_t(1)}{2}\right) = 1 + \frac{W_t^{(0)}}{\sqrt{2}} + \frac{S_t}{2}.$$

Similarly,

$$u_t(1) = 1 + \frac{W_t^{(0)}}{\sqrt{2}} - \frac{S_t}{2}.$$

Therefore, we see that $u_t(0)$ and $u_t(1)$ both follow the Brownian motion $1 + 2^{-1/2}W_t^{(0)}$ plus or minus $1/2$ times the Ornstein–Uhlenbeck process S, which is independent of the said Brownian motion. Clearly, $\{S_t\}_{t\geqslant 0}$ is a mean-zero Gaussian process with variance

$$\mathrm{E}(S_t^2) = 2\mathrm{e}^{-4t}\int_0^t \mathrm{e}^{4s}\,\mathrm{d}s \to \frac{1}{2}, \qquad \text{as } t \to \infty.$$

Therefore, at a large time t, $u_t(0)$ behaves as the Brownian motion $2^{-1/2}W_t^{(0)}$ plus an independent mean-zero normal random variable with variance $1/8$, and $u_t(1)$ behaves approximately as the *same* Brownian motion minus the *same* normal random variable. In other words, at large times, the two particles together behave as a two-body molecule: their joint position resembles the position of a single Brownian motion (the bulk motion) plus/minus an independent normal random variable which describes the effect of the "molecular forces" between the two particles. Figure 16.5 depicts a simulation of the positions of the two particles (one in red, the other in blue), except we have started $u_0(0)$ and $u_0(1)$ at 3 and -3, respectively, in order to better see what is going on.

It can be argued similarly that the SPDE (16.5) on $G = \mathbb{Z}/n\mathbb{Z}$ describes the diffusion of an n-body system with Ornstein–Uhlenbeck type attractions. Figure 16.6 shows a simulation of a three-body case. A great deal is known about the ergodic theory of interacting Itô diffusions; see, for example, Cox–Fleischmann–Greven [10], Greven–Hollander [23], and Shiga [53].

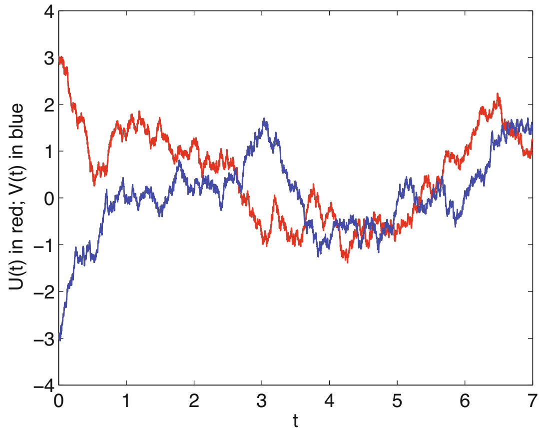


Figure 16.5: A 2-body Brownian motion with molecular attractions. The legend on the y axis is using the notation $U(t) := u_t(0)$, $V(t) := u_t(1)$.

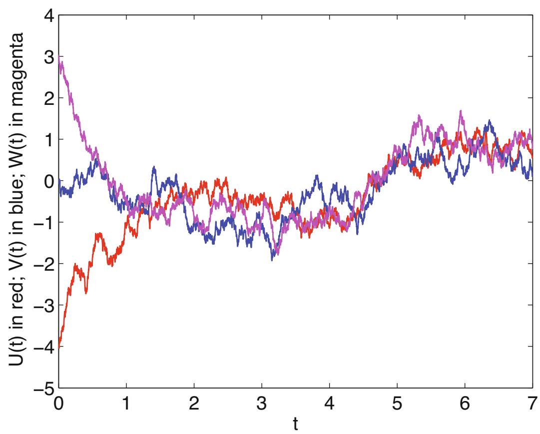


Figure 16.6: A 3-body Brownian motion with homogeneous O-U attractions ($U(t) := u_t(0)$, $V(t) := u_t(1), W(t) := u_t(2)$).

16.3.3 The integer group

If $G = \mathbb{Z}$, then $B_t^{(i)} = \xi([0\,,t]\times\{i\})$ describes the evolution of infinitely many independent Brownian motions, and one can think of (16.5) as another way to write a system of infinitely many interacting particles. The existence and uniqueness of the solution to (16.5), in the case that $G = \mathbb{Z}$, was proved first in Shiga–Shimizu [54].

Recall that the generator of the walk X has the form

$$(\mathcal{G}f)(i) = \kappa \sum_{j=-\infty}^{\infty} [f(i+j) - f(i)]\, D(j),$$

for all $i \in \mathbb{Z}$ and $f \in L^2(\mathbb{Z}) \cong \ell^2$ (the space of square-summable sequences), where $D(i)$ denotes the displacement probabilities[3] of $-X$, and $\kappa \geqslant 0$ is a rate parameter. Then, (16.5) is sometimes also written as the interacting particle system

$$\mathrm{d}u_t(i) = \kappa \sum_{j=-\infty}^{\infty} [u_t(i+j) - u_t(i)]\, D(j)\,\mathrm{d}t + \sigma(u_t(i))\,\mathrm{d}B_t^{(i)},$$

for $i \in \mathbb{Z}$ and $t > 0$, and subject to $u_0(k) = 1$ for all $k \in \mathbb{Z}$.

A noteworthy special case is when X is the simple walk on $\mathbb{Z}$ with jump rate κ. In that case, equal to $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ and hence the generator is

$$(\mathcal{G}f)(i) = \frac{f(i+1) + f(i-1) - 2f(i)}{2} := \frac{1}{2}(\Delta_{\mathbb{Z}} f)(i),$$

for all $i \in \mathbb{Z}$. The linear operator $\Delta_{\mathbb{Z}}$ is called the *discrete Laplacian* on $\mathbb{Z}$. In this particular case, we can think of (16.5) as a coding of the system

$$\mathrm{d}u_t(i) = \frac{\kappa}{2}(\Delta_{\mathbb{Z}} u_t)(i) + \sigma(u_t(i))\,\mathrm{d}B_t^{(i)},$$

subject to $u_0 \equiv 1$. This describes the evolution of infinitely many particles in the following sense: if you think of $u_t(i)$ as the position of the ith particle at time t, then each particle diffuses as $\mathrm{d}x_t = \sigma(x_t)\mathrm{d}b_t$, independently of all other particles, but also for every $i \in \mathbb{Z}$, particle i experiences a drift that pushes it half-way toward the "neighboring particles" $i \pm 1$ (*nearest neighbor Ornstein–Uhlenbeck type attraction*).

16.3.4 The additive reals

One can interpret (16.5) as an SDE with respect to an infinite-dimensional Brownian motion but, instead, I will describe a different intuitive picture that is related to the original one developed by Walsh [57]. For simplicity, consider only the case that the Lévy process X is a Brownian motion on $\mathbb{R}$, with some speed $\kappa \geqslant 0$, so that $\mathcal{G}f = (\kappa/2)f''$ for all f in the Sobolev space $W^{1,2}(\mathbb{R})$; i.e., (16.5) reduces to

$$\dot{u}_t(x) = \frac{\kappa}{2} u_t''(x) + \sigma(u_t(x))\xi(t\,, x),$$

[3] That is, $D(i) = \mathrm{P}\{X_\tau = -i\}$ for the stopping time $\tau := \inf\{s > 0 \mid X_s \neq 0\}$.

subject to $u_0(x) \equiv 1$. Since we may view ξ as an $L^2(\Omega)$-valued measure, it might make sense to think of its "distribution function"

$$B_t(x) := \begin{cases} \xi([0\,,t] \times [0\,,x]) & \text{if } x \geqslant 0, \\ \xi([0\,,t] \times [-x\,,0]) & \text{if } x < 0. \end{cases} \tag{16.13}$$

It is easy to see that, in the sense of distributions,

$$\xi(t\,,x) = \frac{\partial^2}{\partial t \partial x} B_t(x).$$

Thanks to the Wiener isometry (14.2), the two-parameter process B is a mean-zero Gaussian process with covariance

$$\text{Cov}[B_t(x)\,, B_s(y)] = \min(s\,,t)\min(|x|\,,|y|)\mathbf{1}_{(0,\infty)}(xy),$$

and is called the *two-parameter Brownian sheet*[4] on $\mathbb{R}_+ \times \mathbb{R}$. Some people write the SPDE (16.13) in the following way:

$$\dot{u}_t(x) = \frac{\kappa}{2} u_t''(x) + \sigma(u_t(x)) \frac{\partial^2}{\partial t \partial x} B_t(x),$$

subject to $u_0 \equiv 1$. Sometimes, people write the preceding as

$$\dot{u}_t(x) = \frac{\kappa}{2} u_t''(x) + \sigma(u_t(x)) \dot{B}_t(x),$$

where $\dot{B}$ is shorthand for the more cumbersome mixed derivative $\dot{B}'$. But this notation seems to confuse some readers of SPDEs outside probability theory (private communications). I will keep away this terminology in order to avoid the possibility of this sort of confusion.

16.3.5 Higher dimensions

You might find yourself asking why we have examples of $G = \mathbb{R}$ and not $G = \mathbb{R}^n$ for $n > 1$, when it is just as easy to study $\mathbb{R}^n$ as it is to analyse $\mathbb{R}$. This extension can indeed be carried out just as easily as we did it for $G = \mathbb{R}$. The end result is the same as Theorem 16.2.2. Namely, a solution exists provided that $\Upsilon(\alpha) < \infty$ for some $\alpha > 0$, where

$$\Upsilon(\alpha) := \int_0^\infty \mathrm{e}^{-\alpha s} \|p_s\|^2_{L^2(\mathbb{R}^n)} \, \mathrm{d}s.$$

The study of Lévy processes on $\mathbb{R}^n$ is also very much the same as that on $\mathbb{R}$ and, thanks to Plancherel's theorem, it leads to

$$\|p_t\|^2_{L^2(\mathbb{R}^n)} = (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{p}_t(z)|^2 \, \mathrm{d}z = (2\pi)^{-n} \int_{\mathbb{R}^n} \mathrm{e}^{-2t\mathrm{Re}\Psi(z)} \, \mathrm{d}z,$$

[4] To be consistent with some of earlier finite-dimensional examples, some people think of $t \mapsto B_t$ as an "infinite-dimensional Brownian motion."

where Ψ is defined via $\mathrm{E}\exp(iz \cdot X_t) = \exp(-t\Psi(z))$. In other words, Dalang's condition (16.7) on $\mathbb{R}^n$ is the condition that

$$\int_{\mathbb{R}^n} \frac{\mathrm{d}z}{\alpha + 2\mathrm{Re}\Psi(z)} < \infty$$

for some, hence all, $\alpha > 0$. So far, this is almost exactly the same as it was when $n = 1$. However, when $n \geqslant 2$, Lemma 15.1.7 also has its analogue: $|\Psi(z)| = O(\|z\|^2)$ as $\|z\| \to \infty$. Since

$$\int_{\mathbb{R}^n} \frac{\mathrm{d}z}{1 + \|z\|^2} = \infty$$

for $n \geqslant 2$, this means that Dalang's condition never holds in $\mathbb{R}^n$ when $n \geqslant 2$. As it turns out, Dalang's condition is also necessary when σ is a constant; see Dalang [12] and Peszat–Zabczyk [44]. Therefore, we cannot hope to have a general theory of SPDEs of the form (16.5) on $\mathbb{R}^n$ when $n > 1$.

Chapter 17

An Invariance Principle for Parabolic SPDEs

17.1 A central limit theorem

Throughout this chapter we suppose that $J_1, J_2, \ldots$ are independent, identically distributed random variables, with values in $\mathbb{Z}$, and assume that there exist constants $\kappa, \alpha > 0$ such that the characteristic function ϕ of the J_i's satisfies

$$\phi(z) := \mathrm{E}\mathrm{e}^{izJ_1} = 1 - \kappa|z|^\alpha + o\left(|z|^\alpha\right) \quad \text{as} \quad z \to 0. \tag{17.1}$$

Let $N := \{N(t)\}_{t \geqslant 0}$ be an independent Poisson process with rate one, $J_0 := 0$, and define

$$X(t) := \sum_{j=0}^{N(t)} J_j,$$

for $t \geqslant 0$. Then, $\mathrm{E}\mathrm{e}^{izX(t)} = \mathrm{e}^{-t[1-\phi(z)]}$, for all $t \geqslant 0$ and $z \in \mathbb{R}$. It follows readily from this that $X := \{X(t)\}_{t \geqslant 0}$ is a Lévy process (comound Poisson, in fact) with characteristic exponent $\psi(z) = 1 - \phi(z)$. Note, in particular, that for all $z \in \mathbb{R}$, $t \geqslant 0$, and $n \geqslant 1$,

$$\mathrm{E}\exp\left(\frac{izX(nt)}{n^{1/\alpha}}\right) = \left[\mathrm{E}\exp\left(\frac{izX(t)}{n^{1/\alpha}}\right)\right]^n = \exp\left(-nt\left[1 - \phi\left(\frac{z}{n^{1/\alpha}}\right)\right]\right).$$

Therefore, Condition (17.1) ensures that, for all $z \in \mathbb{R}$ and $t \geqslant 0$,

$$\mathrm{E}\exp\left(\frac{izX(nt)}{n^{1/\alpha}}\right) = \exp\left(-nt\left[\frac{\kappa|z|^\alpha}{n} + o(1/n)\right]\right) \to \exp(-t\kappa|z|^\alpha),$$

as $n \to \infty$. Since $z \mapsto \exp(-t\kappa|z|^\alpha)$ is continuous and a pointwise limit of characteristic functions, as we just saw, we can envoke a well-known theorem of Lévy [16, §3.4] to deduce that $z \mapsto \exp(-t\kappa|z|^\alpha)$ is a characteristic function of a random

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_17

variable x_t, and $n^{-1/\alpha}X(nt)$ converges in distribution to x_t as $n \to \infty$. At the same time, another theorem of Lévy [16, Chapter 2, §7] tells us that $z \mapsto \exp(-t\kappa|z|^\alpha)$ is the Fourier transform of a finite (and hence probability) measure if and only if $\alpha \in (0\,,2]$. Therefore, it follows that $n^{-1/\alpha}X(nt)$ converges in distribution to x_t for every $t \geqslant 0$, where $x := \{x_t\}_{t\geqslant 0}$ is a symmetric α-stable Lévy process. The following generalizes these remarks. For reasons that we have mentioned already, we will consider the case that $0 < \alpha \leqslant 2$ only.

Lemma 17.1.1. *If Condition* (17.1) *for some* $\alpha \in (0\,,2]$, *then the finite-dimensional distributions of* $t \mapsto n^{-1/\alpha}X(nt)$ *converge to those of the symmetric* α*-stable Lévy process* X. *That is, for all non-random* $t_k > t_{k-1} > \cdots > t_1 \geqslant 0$, *the random vector*

$$n^{-1/\alpha}\left(X(nt_1),\ldots,X(nt_k)\right)$$

converges in distribution to $(x_{t_1},\ldots,x_{t_k})$.

Proof. We have discussed the case that $k = 1$. Therefore, let us assume without loss of generality that $k \geqslant 2$.

Let $\mathscr{F}_t$ denote the σ-algebra generated by $\{X(r)\}_{r\in[0,t]}$ for all $t \geqslant 0$. Since X is a Lévy process and $X(nt_k) = X(nt_{k-1}) + \{X(nt_k) - X(nt_{k-1})\}$, (17.1) and elementary properties of conditional expectations together yield

$$\begin{aligned}
&\mathrm{E}\left[\exp\left(\frac{iz_kX(nt_k)}{n^{1/\alpha}}\right)\,\middle|\, \mathscr{F}_{t_{k-1}}\right]\\
&\quad= \exp\left(\frac{iz_kX(nt_{k-1})}{n^{1/\alpha}}\right)\mathrm{E}\left[\exp\left(\frac{iz_kX(n[t_k-t_{k-1}])}{n^{1/\alpha}}\right)\right]\\
&\quad= \exp\left(\frac{iz_kX(nt_{k-1})}{n^{1/\alpha}}\right)\exp\left(-n(t_k-t_{k-1})\left[1-\phi\left(\frac{z_k}{n^\alpha}\right)\right]\right)\\
&\quad= \exp\left(\frac{iz_kX(nt_{k-1})}{n^{1/\alpha}}\right)\exp\left(-\kappa(t_k-t_{k-1})|z_k|^\alpha\right),
\end{aligned}$$

almost surely as $n \to \infty$, for all $t \geqslant 0$ and $z_1,\ldots,z_k \in \mathbb{R}$. This and the tower property of conditional expectations together yield the following:

$$\lim_{n\to\infty}\mathrm{E}\exp\left(\sum_{j=1}^k\frac{iz_jX(nt_j)}{n^{1/\alpha}}\right) = \exp\left(-\kappa\sum_{j=1}^{k-1}(t_j-t_{j-1})\sum_{i=j}^k|z_i|^\alpha\right).$$

A similar application of conditional expectations shows that the right-hand side is equal to $\mathrm{E}\exp(\sum_{j=1}^k iz_jx_{t_j})$. Therefore, the result follows from the convergence theorem of Fourier transforms. □

Next, we verify that Condition (17.1) is not vacuous.

Example 17.1.2. Suppose $J_1, J_2, \ldots$ are independent, identically distributed, symmetric random variables in $\mathbb{Z}$ with finite variance σ^2. The Taylor expansion of the cosine reveals that

$$\left|1-\cos\theta-\frac{\theta^2}{2}\right| \leqslant \theta^2 \wedge |\theta|^3,$$

for all θ. By symmetry, $\phi(z) := \mathrm{E}\exp(izJ_1) = \mathrm{E}\cos(zJ_1)$, whence it follows from the preceding that

$$\left|1-\phi(z)-\frac{z^2\sigma^2}{2}\right| \leqslant \mathrm{E}\left(|zJ_1|^2 \wedge |zJ_1|^3\right),$$

which is clearly $o(z^2)$ as $z \to 0$, thanks to the dominated convergence theorem. Therefore, (17.1) holds with $\alpha = 2$ and $\kappa = \sigma^2/2$. Note that κ could, in principle, be any positive real number.

Example 17.1.3. The preceding example shows that Condition (17.1) can hold when $\alpha = 2$. Now we verify that condition for other values of $\alpha \in (0\,,2)$. With this aim in mind, let $J_1, J_2, \ldots$ be i.i.d. random variables with

$$\mathrm{P}\{J_1 = j\} = \frac{1}{2\zeta(1+\alpha)|j|^{1+\alpha}}$$

for $j = \pm 1, \pm 2, \ldots$, where $\zeta(s) := \sum_{k=1}^\infty k^{-s}$ for all $s > 1$. In particular, $J_1, J_2, \ldots$ are symmetric, take values in $\mathbb{Z} \setminus \{0\}$, and satisfy

$$1-\phi(z) = \frac{1}{\zeta(1+\alpha)} \sum_{j=1}^\infty \frac{1-\cos(jz)}{j^{1+\alpha}}$$

for all $z \in \mathbb{R}$, where $\phi(z) := \mathrm{E}\exp(izJ_1) = \mathrm{E}\cos(zJ_1)$. We can write

$$1-\phi(z) = \frac{|z|^\alpha}{\zeta(1+\alpha)} \cdot |z| \sum_{j=1}^\infty \frac{1-\cos(jz)}{(j|z|)^{1+\alpha}} = \frac{|z|^\alpha}{\zeta(1+\alpha)} \cdot \left(\int_0^\infty \frac{1-\cos r}{r^{1+\alpha}}\,\mathrm{d}r + o(1)\right),$$

as $z \to 0$. Therefore, (17.1) holds with $\alpha \in (0\,,2)$ and

$$\kappa := \frac{1}{\zeta(1+\alpha)} \int_0^\infty \frac{1-\cos\theta}{\theta^{1+\alpha}}\,\mathrm{d}\theta = \frac{\pi}{2\zeta(1+\alpha)\Gamma(1+\alpha)\sin(\alpha\pi/2)}.$$

(The latter identity is based on complex function theory.) Thus, it is possible to construct random variables on $\mathbb{Z}$ that satisfy (17.1) for some $\kappa > 0$, regardless of the choice of $\alpha \in (0\,,2)$.

17.2 A local central limit theorem

Let us continue to assume that $J_1, J_2, \ldots$ are symmetric, independent and identically distributed random variables on $\mathbb{Z}$. Let $X := \{X(t)\}_{t\geqslant 0}$ denote the associated continuous-time random walk, as we did in the previous section. Then we can define, for all $\varepsilon > 0$ a continuous-time random walk $X^{(\varepsilon)}$ on $\varepsilon\mathbb{Z}$ by setting $X^{(\varepsilon)}(t) := \varepsilon X(t/\varepsilon^\alpha)$, for $t \geqslant 0$. According to a minor variation of Lemma 17.1.1, if (17.1) holds then the finite-dimensional distributions of $X^{(\varepsilon)}$ converge to those of the symmetric α-stable Lévy process $x := \{x_t\}_{t\geqslant 0}$.[1] Now we ask whether the probability function of $X^{(\varepsilon)}(t)$ can converge to the density function of x_t after suitable rescaling. With this in mind, define for every $\varepsilon > 0$,

$$P_t^{(\varepsilon)}(w) := \mathrm{P}\left\{X^{(\varepsilon)}(t) = w\right\} = \mathrm{P}\left\{X\left(\frac{t}{\varepsilon^\alpha}\right) = \frac{w}{\varepsilon}\right\}$$

for all $t \geqslant 0$ and $w \in \varepsilon\mathbb{Z}$. Clearly, $P_t^{(\varepsilon)}$ defines the transition functions of $X^{(\varepsilon)}$.

Let p_t denote the transition function of the stable process x_t; it might help to recall that $\mathrm{E}f(x_t) = \int_{-\infty}^{\infty} f(y)p_t(-y)\,\mathrm{d}y$ for all Borel functions $f\colon \mathbb{R} \to \mathbb{R}_+$. Since x is symmetric, p_t is also the probability density function of x_t. As it turns out, we will want to know whether one can have

$$\varepsilon^{-1}P_t^{(\varepsilon)}(w) \approx p_t(w).$$

When $t > 0$ is fixed, this sort of question is classical and is an example of a "local CLT" or a "local limit theorem." Many such results are scattered, for example, throughout the classic book by Spitzer [56]. We will need a version however that holds uniformly for all $t > 0$. Of course, one expects $\varepsilon^{-1}P_t^{(\varepsilon)}(w)$ not to approximate $p_t(w)$ very well when $t \approx 0$ because

$$\lim_{t\to 0} P_t^{(\varepsilon)}(w) = \begin{cases} 1 & \text{if} \quad w = 0,\\ 0 & \text{if} \quad w \neq 0,\end{cases}$$

whereas

$$\lim_{t\to 0} p_t(w) = \begin{cases} \infty & \text{if} \quad w = 0,\\ 0 & \text{if} \quad w \neq 0.\end{cases}$$

Thus, our uniform local CLT has to take the size of t into account in some form or another.

Before we state the uniform local CLT that we will need later on, let us state the hypotheses under which that local CLT holds.

Assumption 17.2.1. We assume the following:

(i) $\phi(z) := \mathrm{E}\exp(izJ_1) = 1$ for some $|z| \leqslant \pi$ if and only if $z = 0$;

(ii) there exist $a, \kappa > 0$ and $\alpha \in (1\,, 2]$ such that $\phi(z) = 1 - \kappa|z|^\alpha + o\left(|z|^{a+\alpha}\right)$ as $z \to 0$.

Remark 17.2.2. We would like to say a few things about Assumption 17.2.1. Part (i) of Assumption 17.2.1 is a mild "aperiodicity condition". Part (ii) is a restrictive

[1] Lemma 17.1.1 implies this fact in the case that $\varepsilon^{-\alpha}$ is a positive integer.

form of (17.1). One can show, using Examples 17.1.2 and 17.1.3, that Assumption 17.2.1 is not vacuous; that is, there indeed exist symmetric random variables J_1 on $\mathbb{Z}$ satisfying Assumption 17.2.1.

We have stated Proposition 17.2.3 only when $\alpha > 1$. There are probably versions of this that are valid for $\alpha \in (0\,,1]$. However, we will not need the latter examples for reasons that will manifest themselves in time.

Now we can state and prove our local CLT.

Proposition 17.2.3 (Joseph–Khoshnevisan–Mueller, [30]). *Under Assumption* 17.2.1*, for all $T > 0$ there exist positive and finite constants K, C, ε_0 such that*

$$\sup_{w\in\varepsilon\mathbb{Z}} \left|\varepsilon^{-1}P_t^{(\varepsilon)}(w) - p_t(w)\right| \leqslant C \times \begin{cases} \dfrac{\varepsilon^a|\log\varepsilon|^{(a+\alpha)/\alpha}}{t^{(a+1)/\alpha}} & \text{if } t \geqslant K\varepsilon^\alpha|\log\varepsilon|^{(a+\alpha)/\alpha},\\ t^{-1/\alpha}+\varepsilon^{-1} & \text{otherwise,}\end{cases}$$

uniformly for all $t \in (0\,,T]$ and $\varepsilon \in (0\,,\varepsilon_0)$.

Proof. Recall that x is a Lévy process with characteristic exponent $\Psi(z) = \kappa|z|^\alpha$. By (15.1) and the inversion formula of Fourier transforms,

$$p_t(w) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-izw-\kappa t|z|^\alpha}\,\mathrm{d}z = \frac{1}{\pi}\int_0^\infty \cos(zw)\mathrm{e}^{-\kappa tz^\alpha}\,\mathrm{d}z.$$

In particular,

$$\sup_{w\in\mathbb{R}} p_t(w) = p_t(0) = \frac{c}{t^{1/\alpha}},$$

where $c = \pi^{-1}\int_0^\infty \exp(-\kappa z^\alpha)\,\mathrm{d}z$. Since $P_t^{(\varepsilon)}(w) \leqslant 1$, being a bona fide probability, and

$$|\varepsilon^{-1}P_t^{(\varepsilon)}(w) - p_t(w)| \leqslant \varepsilon^{-1}P_t^{(\varepsilon)}(x) + p_t(w),$$

it follows that

$$\sup_{w\in\varepsilon\mathbb{Z}} \left|\varepsilon^{-1}P_t^{(\varepsilon)}(w) - p_t(w)\right| \leqslant ct^{-1/\alpha} + \varepsilon^{-1},$$

for all $t, \varepsilon > 0$. This yields the second inequality of the theorem.

In order to obtain the first inequality, recall that

$$\mathrm{E}\exp\left(izX(t/\varepsilon^\alpha)\right) = \exp\left(-t\varepsilon^{-\alpha}\left[1-\phi(z)\right]\right),$$

for all $z \in \mathbb{R}$, $t \geqslant 0$, and $\varepsilon > 0$. Therefore, we may apply the inversion theorem of Fourier series in order to deduce the following:

$$\frac{2\pi}{\varepsilon}P_t^{(\varepsilon)}(w) = \int_{-\pi/\varepsilon}^{\pi/\varepsilon} \exp\left(-iwz - \frac{t[1-\phi(\varepsilon z)]}{\varepsilon^\alpha}\right)\mathrm{d}z, \tag{17.2}$$

for all $\varepsilon > 0$, $t \geqslant 0$, and $z \in \mathbb{R}$. Since ϕ is bounded, Assumption 17.2.1 assures us that there exists a finite constant C such that

$$|1 - \phi(v) - \kappa|v|^\alpha| \leqslant C|v|^{\alpha+a},$$

for all $v \in \mathbb{R}$. Consequently, there exists $r_0 \in (0\,,\pi)$ such that

$$1 - \phi(v) \geqslant \frac{\kappa|v|^\alpha}{2} \tag{17.3}$$

whenever $|v| \leqslant r_0$. By symmetry, ϕ is real valued. Define

$$\lambda := \lambda(\varepsilon\,,t) := \left(\frac{(10+2a)|\log\varepsilon|}{\kappa t}\right)^{1/\alpha}.$$

If $t > K\varepsilon^\alpha|\log\varepsilon|^{(a+\alpha)/a}$ for some $K > 20 + 4a$, then

$$\lambda \leqslant \left(\frac{|\log\varepsilon|}{2\varepsilon^\alpha|\log\varepsilon|^{(a+\alpha)/a}}\right)^{1/\alpha} < \frac{r_0}{\varepsilon}$$

uniformly for all $\varepsilon \in (0\,,\varepsilon_0)$, as long as $\varepsilon_0 := \varepsilon_0(r_0)$ is sufficiently small. We fix such a ε_0 for the remainder of this proof. Then we divide the integral in (17.2) into three regions:

$$\frac{2\pi}{\varepsilon} P_t^{(\varepsilon)}(x) = \mathbb{I}_1 + \mathbb{I}_2 + \mathbb{I}_3,$$

where:

$$\begin{aligned}
\mathbb{I}_1 &:= \int_{-\lambda}^{\lambda} \exp\left(-iwz - \frac{t[1-\phi(\varepsilon z)]}{\varepsilon^\alpha}\right) \mathrm{d}z;\\
\mathbb{I}_2 &:= \int_{\substack{z\in\mathbb{R}:\\ \lambda<|z|\leqslant r_0/\varepsilon}} \exp\left(-iwz - y\frac{t[1-\phi(\varepsilon z)]}{\varepsilon^\alpha}\right) \mathrm{d}z; \quad \text{and}\\
\mathbb{I}_3 &:= \int_{\substack{z\in\mathbb{R}:\\ r_0/\varepsilon<|z|\leqslant \pi/\varepsilon}} \exp\left(-iwz - \frac{t[1-\phi(\varepsilon z)]}{\varepsilon^\alpha}\right) \mathrm{d}z.
\end{aligned}$$

By the inversion theorem of Fourier transforms,

$$2\pi p_t(w) = \int_{-\infty}^{\infty} \exp\left(-iwz - \kappa t|z|^\alpha\right) \mathrm{d}z.$$

We plan to show that $\mathbb{I}_1 \approx 2\pi p_t(w)$ and $\mathbb{I}_2, \mathbb{I}_3 \approx 0$, in this order. For our first task, let us note that

$$\begin{aligned}
|\mathbb{I}_1 - 2\pi p_t(w)| &\leqslant \int_{-\lambda}^{\lambda} \left|\exp\left(-\frac{t[1-\phi(\varepsilon z)]}{\varepsilon^\alpha}\right) - \mathrm{e}^{-\kappa t|z|^\alpha}\right| \mathrm{d}z + 2\int_{\lambda}^{\infty} \mathrm{e}^{-\kappa t z^\alpha}\,\mathrm{d}z\\
&= 2\int_0^{\lambda} \mathrm{e}^{-\kappa t z^\alpha}\left|1 - \exp\left(-\frac{t[1-\phi(\varepsilon z)] - \kappa(\varepsilon z)^\alpha}{\varepsilon^\alpha}\right)\right| \mathrm{d}z + 2\int_{\lambda}^{\infty} \mathrm{e}^{-\kappa t z^\alpha}\,\mathrm{d}z.
\end{aligned}$$

Since $0 < \lambda < r_0/\varepsilon < \pi/\varepsilon$, Assumption 17.2.1 implies the existence of finite constants c_1, c_2, c_3 such that, uniformly for all $\varepsilon \in (0\,, \varepsilon_0)$, $z \in (0\,, \lambda)$ and $t \in (0\,, T]$,

$$t\left|\frac{1-\phi(\varepsilon z)-\kappa(\varepsilon z)^\alpha}{\varepsilon^\alpha}\right| \leqslant c_1 t z^{(a+\alpha)}\varepsilon^a \leqslant c_2 t\lambda^{(a+\alpha)}\varepsilon^a$$
$$= c_3\frac{(10+2a)^{(a+\alpha)/\alpha}}{t^{a/\alpha}}\,\varepsilon^a|\log\varepsilon|^{(a+\alpha)/\alpha}.$$

If $t > K\varepsilon^\alpha|\log\varepsilon|^{(a+\alpha)/a}$ for some $K > 20+4a$, then the preceding implies that

$$t\left|\frac{1-\phi(\varepsilon z)-\kappa(\varepsilon z)^\alpha}{\varepsilon^\alpha}\right| \leqslant c_3\frac{(10+2a)^{(a+\alpha)/\alpha}}{K^{a/\alpha}}.$$

Therefore, we may choose (and fix!) $K > 20+4a$ such that

$$t\left|\frac{1-\phi(\varepsilon z)-\kappa(\varepsilon z)^\alpha}{\varepsilon^\alpha}\right| < 1,$$

uniformly for all $\varepsilon \in (0\,, \varepsilon_0)$, $z \in (0\,, \lambda)$ and $t \in (0\,, T]$. Because of the elementary inequality $|1-\exp(-w)| \leqslant 10^5|w|$, valid for all for all $w \in \mathbb{C}$ with $|w| \leqslant 1$, Assumption 17.2.1 implies that

$$|\mathbb{I}_1 - 2\pi p_t(w)|$$
$$\leqslant c_3\frac{(10+2a)^{(a+\alpha)/\alpha}}{t^{a/\alpha}}\,\varepsilon^a|\log\varepsilon|^{(a+\alpha)/\alpha}\int_0^\lambda \mathrm{e}^{-\kappa t z^\alpha}\,\mathrm{d}z + 2\int_\lambda^\infty \mathrm{e}^{-\kappa t z^\alpha}\,\mathrm{d}z.$$

We emphasize that the right-hand side does not depend on $w \in \mathbb{R}$. The first integral is bounded above by

$$\int_0^\infty \exp(-\kappa t z^\alpha)\,\mathrm{d}z = \frac{\Gamma(1/\alpha)}{\alpha(\kappa t)^{1/\alpha}}.$$

By l'Hôpital's rule, there exists a finite constant c_4, depending only on κ and α, such that

$$\int_q^\infty \exp(-\kappa y^\alpha)\,\mathrm{d}y \leqslant c_4\exp(-\kappa q^\alpha/2),$$

for all $q > 0$. Therefore,

$$|\mathbb{I}_1 - 2\pi p_t(w)| \leqslant \frac{c_3(10+2a)^{(a+\alpha)/\alpha}\Gamma(1/\alpha)}{\alpha\kappa^{1/\alpha}t^{(a+1)/\alpha}}\,\varepsilon^a|\log\varepsilon|^{(a+\alpha)/\alpha} + \frac{2c_4\mathrm{e}^{-\kappa t\lambda^\alpha/2}}{t^{1/\alpha}}$$
$$= \text{const}\cdot\left[\frac{\varepsilon^a|\log\varepsilon|^{(a+\alpha)/\alpha}}{t^{(a+1)/\alpha}} + \frac{\varepsilon^{5+a}}{t^{1/\alpha}}\right], \tag{17.4}$$

where "const" denotes a finite constant depending only on κ and α. This shows that $\mathbb{I}_1 \approx 2\pi p_t(w)$.

Next, we prove that $\mathbb{I}_2 \approx 0$ as follows: by (17.3), and the definition of $\mathbb{I}_2$,

$$|\mathbb{I}_2| \leqslant \int_{\substack{z\in\mathbb{R}:\\ \lambda<|z|\leqslant r_0/\varepsilon}} \exp\left(-\frac{t(1-\phi(\varepsilon z))}{\varepsilon^\alpha}\right) \mathrm{d}z \leqslant \int_\lambda^\infty \mathrm{e}^{-\kappa t z^\alpha/2}\,\mathrm{d}z \leqslant \frac{c_4\varepsilon^{5+a}}{t^{1/\alpha}}. \tag{17.5}$$

Finally, we estimate $\mathbb{I}_3$ by noticing that, since ϕ is continuous and real valued (by symmetry), we may appeal to part (i) of Assumption 17.2.1 in order to see that

$$\eta := \sup_{r_0\leqslant|z|\leqslant\pi} \phi(z) \in [0\,,1).$$

Consequently,

$$|\mathbb{I}_3| \leqslant \frac{2\pi}{\varepsilon} \exp\left(-\frac{t(1-\eta)}{\varepsilon^\alpha}\right) \leqslant \frac{c_5}{t^{1/\alpha}} \exp\left(-\frac{t}{c_5\varepsilon^\alpha}\right) \tag{17.6}$$

uniformly for all $t > K\varepsilon^\alpha|\log\varepsilon|^{(a+\alpha)/a}$, for a finite and strictly positive constant c_5 depending only on κ and α.

We can combine (17.4), (17.5), and (17.6) to complete the proof. □

17.3 Particle systems

We now use the notation and results of the previous section in order to establish an "invariance principle" for SPDEs. First, let us establish some notation.

Recall from the previous section that X is a continuous-time symmetric random walk on $\mathbb{Z}$; denote its generator by $\mathscr{G}$. It might help to recall that

$$(\mathscr{G}f)(m) = \sum_{n=-\infty}^{\infty} [f(m+n) - f(m)]D(n),$$

where $D(n) = \mathrm{P}\{J_1 = -n\} = \mathrm{P}\{J_1 = n\}$ for all $n \in \mathbb{Z}$.

Let Ξ denote a space-time white noise on $\mathbb{R}_+ \times \mathbb{Z}$ and consider the SPDE

$$\left[\begin{aligned} &\dot{U}_t(w) = (\mathscr{G}U_t)(w) + \sigma(U_t(w))\,\Xi(t\,,w) \qquad [t>0, w\in\mathbb{Z}],\\ &\text{subject to } U_0(w) \equiv 1 \text{ for all } w\in\mathbb{Z}. \end{aligned}\right. \tag{17.7}$$

Let $B_t(w) := \Xi([0\,,t]\times\{w\})$ and recall that $\{B(w)\}_{w=-\infty}^{\infty}$ is a system of infinitely many independent, identically distributed Brownian motions. Moreover (see §16.3.3), we can write (17.7) as a system of countably-many interacting Itô stochastic differential equations:

$$\left[\begin{aligned} &\mathrm{d}U_t(w) = (\mathscr{G}U_t)(w)\mathrm{d}t + \sigma(U_t(x))\,\mathrm{d}B_t(w) \qquad [t>0, w\in\mathbb{Z}],\\ &\text{subject to } U_0(w) \equiv 1 \text{ for all } w\in\mathbb{Z}. \end{aligned}\right. \tag{17.8}$$

Here and throughtout, we assume that $\sigma\colon \mathbb{R}\to\mathbb{R}$ is non-random and Lipschitz continuous. In this way we may deduce from Theorem 16.2.2 the existence of a unique solution to (17.8).

The equations (17.7) and (17.8) are one and the same; the formulation (17.8) has been studied extensively in the literature on stochastic differential equations, whereas (17.7) is an example of a general stochastic PDE.

We have seen that we may think of (17.8) as a description of a particle system, where $U_t(w)$ describes the position of particle labelled $w\in\mathbb{Z}$ at time $t>0$. Each particle diffuses in space as an Itô diffusion $\mathrm{d}y=\sigma(y)\mathrm{d}b$, where b is a Brownian motion; and particles interact with one another as they feel a kind of "inverse gravitational attraction" to other particles, as described by the linear operator $\mathscr{G}$.

For every $\varepsilon>0$, let π_ε define the canonical map that rescales $\varepsilon\mathbb{Z}$ to $\mathbb{Z}$. That is, for all functions $f\colon \varepsilon\mathbb{Z}\to\mathbb{R}$, and all $w\in\mathbb{Z}$, $(\pi_\varepsilon f)(w):=f(\varepsilon w)$. We may use π_ε to rescale (17.8) to an SPDE on $\mathbb{R}_+\times\varepsilon\mathbb{Z}$, one for every $\varepsilon>0$, as follows:

$$\left[\begin{array}{l}\mathrm{d}U_t^{(\varepsilon)}(w)=\left(\mathscr{G}[\pi_\varepsilon U_t^{(\varepsilon)}]\right)(w)\mathrm{d}t+\sigma\left(U_t^{(\varepsilon)}(w)\right)\mathrm{d}B_t(w/\varepsilon)\quad [t>0, w\in\varepsilon\mathbb{Z}],\\ \text{subject to } U_0(w)\equiv 1 \text{ for all } w\in\varepsilon\mathbb{Z}.\end{array}\right.$$

Our next goal is to study the "fluid limit" of this particle system as $\varepsilon\downarrow 0$. As it is stated, the problem turns out to be ill defined because the system is not scaled appropriately. The correct scaling turns out to be described by the following variation:

$$\left[\begin{array}{l}\mathrm{d}U_t^{(\varepsilon)}(w)=\varepsilon^{-\alpha}\left(\mathscr{G}[\pi_\varepsilon U_t^{(\varepsilon)}]\right)(w)\mathrm{d}t+\varepsilon^{-1/2}\sigma\left(U_t^{(\varepsilon)}(w)\right)\mathrm{d}B_t(w/\varepsilon),\\ \text{for}\quad t>0 \text{ and } w\in\varepsilon\mathbb{Z}, \text{ subject to } U_0(w)\equiv 1 \text{ for all } w\in\varepsilon\mathbb{Z}.\end{array}\right. \tag{17.9}$$

In order to understand what the scaling does, we need to make two remarks: one for the scaling coefficient of $\mathscr{G}$, and one for the coefficient of σ. We do this next, without further ado:

(i) first, we observe that $t\mapsto X_{\lambda t}$ is a continuous-time random walk on the integer lattice, for every $\lambda>0$, and has generator $\lambda\mathscr{G}$. Therefore, the extra scaling factor $\varepsilon^{-\alpha}$ that is put in front of $\mathscr{L}$ amounts to speeding up the underlying random-walk mechanism by a factor of $1/\varepsilon^\alpha$;

(ii) the factor $\varepsilon^{-1/2}$ in front of σ is there to ensure that the scaled noise $\varepsilon^{-1/2}\mathrm{d}B_t(w/\varepsilon)$ corresponds to a system of i.i.d. Brownian motions with variance $\varepsilon^{-1}t$ at time t.

In other words, (i) and (ii) together tell us that should scale the space variable by $1/\varepsilon$ and the time variable by $1/\varepsilon$ for the noise and $1/\varepsilon^\alpha$ for the underlying random walk mechanism (which describes the nature of the gravitational attraction between the particles). Thus, (17.9) is merely a "central-limit-type space-time scaling" of (17.8).

Recall also that x is a symmetric α-stable Lévy process with characteristic exponent $\Psi(z) = \kappa|z|^\alpha$, where κ is the constant of Assumption 17.2.1. Let $\mathcal{G}$ denote the generator of x; that is,

$$(\mathcal{G}f)(w) = C\int_{-\infty}^{\infty} \frac{f(w+y)-f(w)}{|y|^{1+\alpha}}\,\mathrm{d}y,$$

$w \in \mathbb{R}$, for a suitable constant $C \in (0\,,\infty)$ depending only on κ and α; see (15.11). Let ξ denote a space-time white noise on $\mathbb{R}_+ \times \mathbb{R}$ and consider the SPDE,

$$\left[\begin{aligned}&\dot{u}_t(w) = (\mathcal{G}u_t)(w) + \sigma(u_t(w))\xi(t\,,w) \quad [t>0, w\in\mathbb{R}],\\ &\text{subject to } u_0(w)\equiv 1 \text{ for all } w\in\mathbb{R}.\end{aligned}\right. \tag{17.10}$$

Theorem 16.2.2 ensures the existence of a unique solution to (17.9). The next result is the main theorem of these lectures, and shows that (17.10) is the "fluid limit" approximation to (17.9) as $\varepsilon \downarrow 0$. For a precursor to this result, see Funaki [20].

Theorem 17.3.1 (Joseph–Khoshnevisan–Mueller, [30])**.** *Under Assumption* 17.2.1, *there exists a coupling of the solutions to* (17.9) *and* (17.10) (*on the same probability space*) *satisfying the following: for all real numbers $T > 0$, $k \geqslant 2$, and $\rho \in (0, \alpha-1)$,*

$$\sup_{t\in[0,T]}\sup_{w\in\varepsilon\mathbb{Z}} \mathrm{E}\left(\left|U_t^{(\varepsilon)}(w) - u_t(w)\right|^k\right) = O\left(\varepsilon^{\rho k/2}\right) \qquad \text{as } \varepsilon\downarrow 0.$$

The proof is somewhat technical and can be found in [30]. Instead of working out all the details, we content ourselves by merely describing the coupling construction.

Let $(\Omega\,,\mathscr{F},\mathrm{P})$ be a probability space rich enough to support a space-time white noise ξ on $\mathbb{R}_+\times\mathbb{R}$. Of course, we can solve (17.10) uniquely on that probability space. Next, we define

$$B_t^{(\varepsilon)}(w) := \varepsilon^{-1/2}\xi\left([0\,,t]\times[w\varepsilon\,,(w+1)\varepsilon]\right),$$

for $t>0$, $w\in\mathbb{Z}$. Then $\{B^{(\varepsilon)}(w)\}_{w\in\mathbb{Z}}$ is a systerm of i.i.d. Brownian motions, for every fixed $\varepsilon>0$. In particular, Theorem 16.2.2 assures us that the following SPDE has a unique solution for every $\varepsilon>0$:

$$\left[\begin{aligned}&\mathrm{d}V_t^{(\varepsilon)}(w) = \varepsilon^{-\alpha}\left(\mathscr{G}[\pi_\varepsilon V_t^{(\varepsilon)}]\right)(w)\,\mathrm{d}t + \varepsilon^{-1/2}\sigma\left(V_t^{(\varepsilon)}(w)\right)\mathrm{d}B_t^{(\varepsilon)}(w/\varepsilon),\\ &\text{for}\quad t>0 \quad\text{and}\quad w\in\varepsilon\mathbb{Z}, \text{ subject to } V_0^{(\varepsilon)}(w)\equiv 1 \text{ for all } w\in\varepsilon\mathbb{Z}.\end{aligned}\right.$$

By the uniqueness portion of Theorem 16.2.2, the random field $V^{(\varepsilon)}$ has the same finite-dimensional distributions as $U^{(\varepsilon)}$. Moreover, $V^{(\varepsilon)}$ is defined on the same probability space as u. We can write $V^{(\varepsilon)}$ in integrate form as follows:

$$\begin{aligned} V_t^{(\varepsilon)}(w) &= 1 + \varepsilon^{-1/2} \sum_{j=-\infty}^{\infty} \int_0^t P_{t-s}^{(\varepsilon)}(j\varepsilon - w)\sigma\left(V_s^{(\varepsilon)}(j\varepsilon)\right) \mathrm{d}B_s^{(\varepsilon)}(j) \\ &= 1 + \varepsilon^{-1} \int_{[0,t]\times\mathbb{R}} P_{t-s}^{(\varepsilon)}(\varepsilon\lfloor y/\varepsilon\rfloor - w)\sigma\left(V_s^{(\varepsilon)}(\varepsilon\lfloor y/\varepsilon\rfloor)\right) \xi(\mathrm{d}s\,\mathrm{d}y). \end{aligned}$$

If $\varepsilon \approx 0$, then $\varepsilon^{-1}P_{t-s}^{(\varepsilon)}(a) \approx p_t(a)$ in a uniform sense (see Proposition 17.2.3). Thus, using the local CLT, one can then try to prove that $V_t^{(\varepsilon)}(w) \approx \bar{V}_t^{(\varepsilon)}(w)$, where $\bar{V}^{\varepsilon}$ solves

$$\bar{V}_t^{(\varepsilon)}(w) = 1 + \int_{[0,t]\times\mathbb{R}} p_{t-s}(\varepsilon\lfloor y/\varepsilon\rfloor - w)\sigma\left(\bar{V}_s^{(\varepsilon)}(\varepsilon\lfloor y/\varepsilon\rfloor)\right) \xi(\mathrm{d}s\,\mathrm{d}y).$$

The preceding is not an SPDE in itself, but has a unique solution $\bar{V}^{(\varepsilon)}$ for the same sort of reason that Theorem 16.2.2 is true. Now recall that u solves a very similar integral equation, namely,

$$u_t(w) = 1 + \int_{[0,t]\times\mathbb{R}} p_{t-s}(y - w)\sigma(u_s(y))\, \xi(\mathrm{d}s\,\mathrm{d}y).$$

It can be shown that u is Hölder continuous in each variable t and w, and "hence",

$$u_t(w) \approx 1 + \int_{[0,t]\times\mathbb{R}} p_{t-s}(\varepsilon\lfloor y/\varepsilon\rfloor - w)\sigma(u_s(y))\, \xi(\mathrm{d}s\,\mathrm{d}y)$$

when $\varepsilon \approx 0$. By keeping careful track of the errors incurred, one can then show that $u_t(x) \approx \bar{V}_t^{(\varepsilon)}(x)$, whence also $u_t(x) \approx V_t^{(\varepsilon)}(x)$.

Chapter 18

Comparison Theorems

In this chapter we outline very briefly some of the main applications of the invariance principle from the preceding chapter. More specifically, we plan to prove that for a large family of SPDEs: (1) the solution is always positive as long as the initial value is positive; and (2) one can sometimes compute "moment functionals" of solutions to various SPDEs to one another. The material of this chapter is borrowed from Joseph–Khoshnevisan–Mueller [30].

18.1 Positivity

Our first application of Theorem 17.3.1 is a *positivity principle*. Before we describe that, let us digress and talk a little about classical diffusions.

Let $\sigma\colon \mathbb{R} \to \mathbb{R}$ be a non-random Lipschitz-continuous function, and consider a one-dimensional Itô diffusion X that is the solution to the Itô differential equation $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t$, for all $t > 0$, where $X_0 = x_0 \in \mathbb{R}$ is non-random. Let Y denote the solution to the same Itô ODE but with $Y_0 := y_0$ (also non-random). Then, a classical result about one-dimensional diffusions states that if $X_0 \geqslant Y_0$ then $X_t \geqslant Y_t$ a.s. for all $t \geqslant 0$. A particularly important consequence is the following result, due in various levels of generality, to Skorohod, Yamada, and Ikeda and Watanabe; see [25] for the source and previous references.

Proposition 18.1.1. *If $\sigma(0) = 0$ and $X_0 \geqslant 0$, then $X_t \geqslant 0$ a.s. for all $t \geqslant 0$.*

Proof. Let Y be as before, but now with $Y_0 := 0$. I claim that $Y_t = 0$ a.s. for all $t \geqslant 0$. Since $X_t \geqslant Y_t$ a.s. for all $t \geqslant 0$, this claim will prove the proposition. Define $Y_t^{(0)} := Y_0$ for all $t \geqslant 0$, and then iteratively set

$$Y_t^{(n+1)} := Y_0 + \int_0^t \sigma(Y_s^{(n)})\,\mathrm{d}B_s.$$

Since $\sigma(0) = 0$ it follows that $Y_t^{(n)} = 0$ a.s. for all $n \geqslant 0$ and $t \geqslant 0$. As part of our

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_18

construction of solutions to SPDEs, we saw that $\mathrm{E}(|Y_t^{(n)} - Y_t|^2) \to 0$ as $n \to \infty$. Thus it follows that $Y_t = 0$ a.s. for all $t \geqslant 0$, as was asserted. □

Let $\mathscr{G}$ denote the generator of a random walk on $\mathbb{Z}$, as in the previous chapter. Let B denote a field of i.i.d. Brownian motions and define U to be the solution to the interacting system (17.8). That is,

$$\left[\begin{array}{ll} \mathrm{d}U_t(w) = (\mathscr{G}U_t)(w)\mathrm{d}t + \sigma(U_t(x))\,\mathrm{d}B_t(w) & [t>0, w \in \mathbb{Z}], \\ \text{subject to } U_0(w) \equiv 1 \text{ for all } w \in \mathbb{Z}, & \end{array}\right. \tag{18.1}$$

where $\sigma\colon \mathbb{R} \to \mathbb{R}$ is non-random and Lipschitz continuous. The following is a consequence of a more general series of comparison principles for multidimensional Itô diffusions; see Geiß–Manthey [21].

Theorem 18.1.2. *If $\sigma(0) = 0$ then $U_t(w) \geqslant 0$ a.s. for all $t \geqslant 0$ and $w \in \mathbb{Z}$.*

Let us rescale the particle system (18.1) as we did in the previous chapter:

$$\left[\begin{array}{l} \mathrm{d}U_t^{(\varepsilon)}(w) = \varepsilon^{-\alpha}\left(\mathscr{G}[\pi_\varepsilon U_t^{(\varepsilon)}]\right)(w)\mathrm{d}t + \varepsilon^{-1/2}\sigma\left(U_t^{(\varepsilon)}(w)\right)\mathrm{d}B_t(w/\varepsilon), \\ \text{for } t>0 \text{ and } w \in \varepsilon\mathbb{Z}, \text{ subject to } U_0(w) \equiv 1 \text{ for all } w \in \varepsilon\mathbb{Z}. \end{array}\right.$$

Then, Theorem 18.1.2 tells us that if $\sigma(0) = 0$ then $U_t^{(\varepsilon)}(x) \geqslant 0$ a.s. for all $t \geqslant 0$, $\varepsilon > 0$, and $x \in \varepsilon\mathbb{Z}$. This and Theorem 17.3.1 together yield the following.

Theorem 18.1.3. *Let $\mathcal{G}$ denote the generator of a symmetric α-stable Lévy process with $\alpha \in (1\,,2]$. That is, $\mathcal{G}f = \kappa f''$ for some $\kappa > 0$ when $\alpha = 2$ and*

$$(\mathcal{G}f)(w) = C\int_{-\infty}^{\infty} \frac{f(w+y) - f(w)}{|y|^{1+\alpha}}\,\mathrm{d}y,$$

$w \in \mathbb{R}$, for some $C > 0$ when $\alpha \in (1\,,2)$; see (15.11)*. Let ξ denote a space-time white noise on $\mathbb{R}_+ \times \mathbb{R}$ and consider the SPDE,*

$$\left[\begin{array}{ll} \dot{u}_t(w) = (\mathcal{G}u_t)(w) + \sigma(u_t(w))\xi(t\,,w) & [t>0, w \in \mathbb{R}], \\ \textit{subject to } u_0(w) \equiv 1 \textit{ for all } w \in \mathbb{R}. & \end{array}\right.$$

If $\sigma(0) = 0$ then $u_t(x) \geqslant 0$ a.s. for all $t \geqslant 0$ and $x \in \mathbb{R}$.

For related results see, for example, Mueller [42] and Dalang–Khoshnevisan–Mueller–Nualart–Xiao [13, Theorem 5.1].

18.2 The Cox–Fleischmann–Greven inequality

Suppose $\bar{\sigma}\colon \mathbb{R} \to \mathbb{R}$ is non-random and Lipschitz continuous, and let $\bar{U}$ denote the solution to the following system, analogous to (18.1):

$$\left[\begin{array}{ll} \mathrm{d}\bar{U}_t(w) = (\mathscr{G}\bar{U}_t)(w)\mathrm{d}t + \bar{\sigma}(\bar{U}_t(x))\,\mathrm{d}B_t(w) & [t>0, w \in \mathbb{Z}], \\ \text{subject to } \bar{U}_0(w) \equiv 1 \text{ for all } w \in \mathbb{Z}. & \end{array}\right. \tag{18.2}$$

Throughout we assume that $\sigma(0) = \bar\sigma(0) = 0$ so that U and $\bar U$ are a.s.-non-negative random fields; see Theorem 18.1.3.

An important comparison theorem of Cox–Fleischmann–Greven [10] states that, under some mild conditions, whenever $\sigma(u) \leqslant \bar\sigma(u)$ for all $u \geqslant 0$ then various expectation functionals of U are dominated by the same expectation functionals of $\bar U$. In order to state this more precisely, we need some notation.

Let $\mathbb{R}_+^{\mathbb{Z}}$ be the usual collection of all countable non-negative sequences, endowed with the product topology. Recall that $F\colon \mathbb{R}_+^{\mathbb{Z}} \to \mathbb{R}$ is called a *cylinder function* if there exists a finite set $I \subset \mathbb{Z}$ such that $F(x) = F(y)$ for all $x, y \in \mathbb{R}_+^{\mathbb{Z}}$ with $x_k = y_k$ for all $k \notin I$. Let $\mathscr{F}$ denote the collection of all cylinder functions from $\mathbb{R}_+^{\mathbb{Z}}$ to $\mathbb{R}$ such that $\partial_i\partial_j F$ is a measurable and a.e.-positive function for all $i, j \in \mathbb{Z}$. Let $\mathscr{F}_0$ denote the collection of $F \in \mathscr{F}$ such that $x_i \mapsto F(x)$ is nondecreasing (or nonincreasing) for all $i \in \mathbb{Z}$.

Example 18.2.1. Let I denote a finite subset of $\mathbb{Z}$. Then two typical examples of $F \in \mathscr{F}$ are $F(u) := \prod_{j\in I} u_j^{p_j}$ and $F(u) := \exp(-\sum_{j\in I} \lambda_j u_j)$ for $u_i \geqslant 0$, $p_i \in [2\,, \infty)$, and $\lambda_i \in (0\,, \infty)$.

The following compares the moments[1] of $F(U_t)$ to those of $F(\bar U_t)$ for all $F \in \mathscr{F}$.

Theorem 18.2.2 (Cox–Fleischmann–Greven, [10])**.** *Suppose $0 \leqslant \sigma(x) \leqslant \bar\sigma(x)$ for all $x \in \mathbb{R}$ and $\sigma(0) = \bar\sigma(0) = 0$. Then, for all $F \in \mathscr{F}_0$ and $t_1, \ldots, t_n \geqslant 0$,*

$$\mathrm{E}\prod_{j=1}^n F(U_{t_j}) \leqslant \mathrm{E}\prod_{j=1}^n F(\bar U_{t_j}).$$

The preceding holds for every $F \in \mathscr{F}$ when $t_1 = \cdots = t_n$.

This and Theorem 17.3.1 have the following ready consequence for moment comparisons for SPDEs.

Theorem 18.2.3 (Joseph–Khoshnevisan–Mueller, [30])**.** *Suppose $\sigma, \bar\sigma\colon \mathbb{R} \to \mathbb{R}_+$ satisfy $0 \leqslant \sigma(u) \leqslant \bar\sigma(u)$ for all $u \geqslant 0$, and $\sigma(0) = \bar\sigma(0) = 0$. Let u solve*

$$\left[\begin{array}{l} \dot u_t(w) = (\mathcal{G}u_t)(w) + \sigma(u_t(w))\xi(t\,, w) \quad [t > 0, w \in \mathbb{R}], \\ \text{subject to } u_0(w) \equiv 1 \text{ for all } w \in \mathbb{R}, \end{array}\right.$$

and let $\bar u$ solve the same SPDE, but where σ is replaced by $\bar\sigma$. Then, for all $t_1, \ldots, t_n \geqslant 0$,

$$\mathrm{E}\prod_{j=1}^n F(u_{t_j}) \leqslant \mathrm{E}\prod_{j=1}^n F(\bar u_{t_j}),$$

[1] We think of $U_t = \{U_t(x)\}_{x\in\mathbb{Z}}$ as an infinite sequence for every $t \geqslant 0$, so that $F(U_t)$ is a perfectly well-defined process. To be perfectly honest, Cox–Fleischmann–Greven [10] state this theorem under further conditions on σ, $\bar\sigma$, and F. The present result follows easily from theirs and standard approximation arguments.

if either:

(i) $F(u) = \prod_{j\in I} u_j^{p_j}$ *for* $p_j \geqslant 2$ *and* $I \subset \mathbb{Z}$ *finite; or*
(ii) $F(u) = \exp\{-\sum_{j\in I} \lambda_j u_j\}$ *for* $\lambda_j \geqslant 0$.

For an example, consider the case that $\sigma(x) = \lambda x$ for some fixed $\lambda > 0$. The resulting SPDE is called the *parabolic Anderson model.* In the case that $\alpha = 2$, we have

$$\left[\begin{aligned}&\dot{u}_t(w) = \kappa u_t''(w) + \lambda u_t(w)\xi(t\,,w) \quad [t>0, w\in\mathbb{R}],\\ &\text{subject to } u_0(w) \equiv 1 \text{ for all } w \in \mathbb{R},\end{aligned}\right.$$

where $\kappa > 0$ is the viscosity coefficient. It is known that, in this case, there exist finite constants $c_1, \ldots, c_4 > 0$ such that

$$c_1 \exp(c_2\lambda^4 n^3 t) \leqslant \mathrm{E}\left([u_t(x)]^n\right) \leqslant c_3 \exp(c_4\lambda^4 n^3 t),$$

for all $x \in \mathbb{R}$, $n \in [2\,,\infty)$, and $t \geqslant 0$; see, for example, Joseph–Khoshnevisan–Mueller [30].

Now suppose $\tau \colon \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous function and consider the solution v to the SPDE

$$\left[\begin{aligned}&\dot{v}_t(w) = \kappa v_t''(w) + \tau(v_t(w))\xi(t\,,w) \quad [t>0, w\in\mathbb{R}],\\ &\text{subject to } v_0(w) \equiv 1 \text{ for all } w \in \mathbb{R}.\end{aligned}\right.$$

If $\tau(0) = 0$, then $|\tau(x)| \leqslant \mathrm{Lip}_\tau|x|$ for all $x \in \mathbb{R}$, where Lip_τ denotes the Lipschitz constant of τ. Consequently, it follows from Theorem 18.2.3 that there exist finite constants $A, A' > 0$ such that $\mathrm{E}\left([v_t(x)]^n\right) \leqslant A\exp(A'n^3t)$ for all $n \geqslant 2$ and $t > 0$. If, in addition, $\inf_{x\in\mathbb{R}} |\sigma(x)/x| > 0$, then the preceding displayed inequality can be reversed (with different constants). Define

$$\begin{aligned}\gamma(n) &:= \limsup_{t\to\infty} t^{-1} \sup_{x\in\mathbb{R}} \log \mathrm{E}\left([v_t(x)]^n\right),\\ \underline{\gamma}(n) &:= \liminf_{t\to\infty} t^{-1} \inf_{x\in\mathbb{R}} \log \mathrm{E}\left([v_t(x)]^n\right).\end{aligned}$$

The functions γ and $\underline{\gamma}$ are called the *upper* and the *lower Lyapunov exponents* of v, respectively. It follows that, under the present conditions, there exist finite and positive constants B and B' not depending on $\lambda > 0$ and satisfying

$$B\lambda^4 n^3 \leqslant \underline{\gamma}(n) \leqslant \gamma(n) \leqslant B'\lambda^4 n^3,$$

for all $n \geqslant 2$. This shows that the moments of the solution grows very quickly with time. In analogy with the literature on finite-dimensional dynamical systems (see Ruelle [50] for a discussion in the context of turbulence), we may predict then that the solution v to our SPDE ought to be “chaotic”. This is true; see Conus–Joseph–Khoshnevisan [8].

18.3 Slepian's inequality

Let us close these notes by making some remarks on Theorem 18.2.2. We will not prove that theorem here, but it would be a shame not to say anything about that beautiful proof altogether. Let me try and convince you why Theorem 18.2.2 has to be true. I believe the following can in fact be used to construct a proof of Theorem 18.2.2, though the original derivation is slightly different. For related results, see the elegant paper Nourdin–Peccati–Viens [43].

First, one has to believe that the systems (18.1) and (18.2) are locally linear SPDEs. The reason can be gleaned by seeing how Picard iteration works: that method shows basically that a nonlinear SPDE with Lipschitz coefficients is, in fact, a small local perturbation of a linear SPDE with constant coefficients. In the constant coefficient case, σ and $\bar{\sigma}$ are reduced to positive constants and the solutions are Gaussian processes. Thus, Theorem 18.2.2 is reduced to a statement about Gaussian processes. Since the elements of $\mathscr{F}$ and $\mathscr{F}_0$ are cylinder functions, we have in fact a statement about finite-dimensional Gaussian random vectors. That statement is itself very well known, and is due to Slepian [55]. Let us conclude by stating and proving only that statement.

Let $X := (X_1, \ldots, X_n)$ and $Y := (Y_1, \ldots, Y_n)$ be two Gaussian random vectors in $\mathbb{R}^n$ such that $\mathrm{E}X_i = \mathrm{E}Y_i = 0$ for all $i = 1, \ldots, n$. Let Q^X and Q^Y denote the respective covariance matrices of X and Y. That is, $Q^X_{i,j} = \mathrm{E}(X_iX_j)$ and $Q^Y_{i,j} = \mathrm{E}(Y_iY_j)$ for $i, j = 1, \ldots, n$.

Let $C_{\exp}(\mathbb{R}^n)$ denote the collection of all continuous functions $F\colon \mathbb{R}^n \to \mathbb{R}$ growing at most exponentially; that is, $F \in C_{\exp}(\mathbb{R}^n)$ if and only if there exists a finite constant c such that $|F(x)| \leqslant c e^{c\|x\|}$ for all $x \in \mathbb{R}^n$.

Theorem 18.3.1 (Slepian's inequality, [55]). *Suppose $F \in C_{\exp}(\mathbb{R}^n)$ satisfies*

$$\sum\sum_{1\leqslant i,j\leqslant n} \left[Q^X_{i,j} - Q^Y_{i,j}\right] \frac{\partial^2}{\partial x_i \partial x_j} F(x) \geqslant 0 \tag{18.3}$$

for almost all $x \in \mathbb{R}^n$, where the mixed derivatives are understood in the weak sense. Then, $\mathrm{E}[F(X)] \geqslant \mathrm{E}[F(Y)]$.

Proof. Without loss of generality, we may, and will, assume that X and Y are independent.

First consider the case where Q^X and Q^Y are nonsingular, and $F \in \mathscr{S}(\mathbb{R}^n)$ is a rapidly decreasing test function.

Consider the stochastic process $Z := \{Z(t)\}_{t\in[[0,1]}$, defined as

$$Z(t) := \sqrt{t}\, X + \sqrt{1-t}\, Y,$$

for $0 \leqslant t \leqslant 1$. The process of using Z is sometimes called *Gaussian interpolation.* This is in part because, for every fixed $t \in [0\,,1]$, $Z(t)$ is a Gaussian random vector with mean vector zero and covariance matrix $Q(t) := tQ^X + (1-t)Q^Y$. Therefore:

(i) $Z(0) = Y$ and $Z(1) = X$, so we are running a Gaussian process starting at Y and ending at X; and

(ii) the covariance matrix of Z is a convex combination of those of X and Y.

Let μ_t denote the distribution of $Z(t)$ for each $t \in [0\,,1]$. Then every μ_t is a probability measure on $\mathbb{R}^n$ with Fourier transform

$$\widehat{\mu}_t(z) = \mathrm{E}\mathrm{e}^{iz\cdot Z(t)} = \exp\left(-\tfrac{1}{2}z'Q(t)z\right), \text{ for } t \in [0\,,1] \text{ and } z \in \mathbb{R}^n.$$

In particular,

$$\mathrm{E}[F(Z(t))] = \int_{\mathbb{R}^n} F\,\mathrm{d}\mu_t = (2\pi)^{-n}\int_{\mathbb{R}^n} \widehat{F}(z)\mathrm{e}^{-z'Q(t)z/2}\,\mathrm{d}z,$$

thanks to the Parseval identity and the hypotheses on F and Q^X and Q^Y. Thus, we may differentiate to find that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}[F(Z(t))] &= -\frac{1}{2(2\pi)^n}\int_{\mathbb{R}^n}\widehat{F}(z)\mathrm{e}^{-z'Q(t)z/2}z'\left[Q^X - Q^Y\right]z\,\mathrm{d}z\\ &= -\frac{1}{2(2\pi)^n}\sum\sum_{1\leqslant i,j\leqslant n}\left[Q^X_{i,j} - Q^Y_{i,j}\right]\int_{\mathbb{R}^n} z_iz_j\widehat{F}(z)\mathrm{e}^{-z'Q(t)z/2}\,\mathrm{d}z.\end{aligned}$$

The exchange of the derivative and the integral is justified by the assumptions that Q^X and Q^Y are nonsingular and $F \in \mathscr{S}(\mathbb{R}^n)$.[2] Now the Fourier transform of

$$\partial_{i,j}F(x) := \frac{\partial^2 F(x)}{\partial x_i\partial x_j} \qquad \text{is} \qquad \widehat{\partial_{i,j}F}(z) = -z_iz_j\widehat{F}(z),$$

for every $i,j = 1,\ldots,n$ and $z \in \mathbb{R}^n$. Therefore,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}[F(Z(t))] &= \frac{1}{2(2\pi)^n}\sum\sum_{1\leqslant i,j\leqslant n}\left[Q^X_{i,j} - Q^Y_{i,j}\right]\int_{\mathbb{R}^n}\widehat{\partial_{i,j}F}(z)\mathrm{e}^{-z'Q(t)z/2}\,\mathrm{d}z\\ &= \frac{1}{2}\int_{\mathbb{R}^n}\sum\sum_{1\leqslant i,j\leqslant n}\left[Q^X_{i,j} - Q^Y_{i,j}\right]\partial_{i,j}F(x)\,\mu_t(\mathrm{d}x),\end{aligned} \tag{18.4}$$

after a second appeal to the Parseval identity. Since Q^X and Q^Y are non singular, μ_t is mutually absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$. Therefore, thanks to (18.4) and condition (18.3), the map $t \mapsto \mathrm{E}[F(Z(t))]$ is non decreasing, whence $\mathrm{E}[F(Z(0))] \leqslant \mathrm{E}[F(Z(1))]$. This is another way to state the theorem when $F \in \mathscr{S}(\mathbb{R}^n)$ and Q^X and Q^Y are nonsingular.

Next, consider the more general case where F is continuous and satisfies the conditions of the theorem. We continue to assume that Q^X and Q^Y are nonsingular.

[2] In particular, we are using the property that if Q^X and Q^Y are nonsingular, then there exists $\lambda > 0$ such that $z'Q(t)z \geqslant \lambda\|z\|^2$ uniformly for all $t \in [0\,,1]$ and $z \in \mathbb{R}^n$.

Let ψ_ε denote the normal probability density with mean zero and variance $\varepsilon > 0$ for every $\varepsilon > 0$. If $F \in C_{\exp}(\mathbb{R})$ satisfies (18.3), then every function $F_\varepsilon := F * \psi_\varepsilon \in C_{\exp}(\mathbb{R})$ also satisfies (18.3). What we have proved so far shows that $\mathrm{E}[F_\varepsilon(X)] \geqslant \mathrm{E}[F_\varepsilon(Y)]$, for every $\varepsilon > 0$. Let $\varepsilon \downarrow 0$ and appeal to the dominated convergence theorem to see that $\mathrm{E}[F(X)] \geqslant \mathrm{E}[F(Y)]$.

It remains to prove that the nonsingularity hypothesis on Q^X and Q^Y can be dropped altogether. Let W be an independent n-dimensional centered Gaussian vector with identity covariance matrix. What we have proved up to this point shows that

$$\mathrm{E}[F(X + \varepsilon W)] \geqslant \mathrm{E}[F(Y + \varepsilon W)],$$

for every $\varepsilon > 0$. Thanks to continuity and the at-most-exponential growth condition on F,

$$\lim_{\varepsilon \to 0} \mathrm{E}[F(X + \varepsilon W)] = \mathrm{E}[F(X)] \quad \text{and} \quad \lim_{\varepsilon \to 0} \mathrm{E}[F(Y + \varepsilon W)] = \mathrm{E}[F(Y)],$$

by the dominated convergence theorem. This proves the theorem in full generality. $\square$

Chapter 19

A Dash of Color

A great portion of the literature on SPDEs is concerned, in one form or another, with stochastic partial differential equations that are driven by noises showing some form of correlations. As of this time, there seems to be no unified theory of nonlinear SPDEs driven by correlated (henceforth, "colored") noise, except in certain special cases; see, for example, Carmona–Molchanov [5] and Hu–Lu–Nualart [24] and their combined bibliography. By contrast, there is a very general approach to linear SPDEs, thanks to a rich theory of Gaussian processes. I will conclude these notes by describing, very briefly, the theory of linear SPDEs that are driven by quite general Gaussian noises.

19.1 Reproducing kernel Hilbert spaces

In order to study Gaussian noises, we need a notion of "covariance". The correct notion, in our context, is that of covariance operators. This is the topic of this section. Throughout, G denotes an LCA group with Haar measure m_G.

Let $L^1_{\mathrm{loc}}(G)$ denote the class of all *locally integrable* real-valued functions on G. That is, $f \in L^1_{\mathrm{loc}}(G)$ if and only if $f\colon G \to \mathbb{R}$ is measurable and

$$\int_A |f(x)|\, m_G(\mathrm{d}x) < \infty$$

for all compact sets $A \subset G$.

Here and throughout, let K be a self-adjoint linear operator from $C_c(G)$ to $L^1_{\mathrm{loc}}(G)$. Then, by default, the integrals

$$(f_1\,, \mathrm{K}f_2)_{L^2(G)} = \int_G f_1(x)(\mathrm{K}f_2)(x)\, m_G(\mathrm{d}x)$$

are absolutely convergent, as f_1 and f_2 range over all of $C_c(G)$.

Definition 19.1.1. A linear operator $\mathrm{K}\colon C_c(G) \to L^2_{\mathrm{loc}}(G)$ is said to be *positive definite* (on $C_c(G)$) if $(f\,, \mathrm{K}f)_{L^2(G)} \geqslant 0$ for all $f \in C_c(G)$.

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0_19

Suppose K is self-adjoint and positive definite. Since $(f_1,f_2)\mapsto(f_1,\mathrm{K}f_2)_{L^2(G)}$ defines a pre-Hilbertian inner product on $C_c(G)$, general theory tells us that one can construct a centered Gaussian process $\eta := \{\eta(f)\}_{f\in C_c(G)}$ such that

$$\mathrm{Cov}[\eta(f_1)\,,\eta(f_2)] = (f_1\,,\mathrm{K}f_2)_{L^2(G)} \tag{19.1}$$

for all $f_1, f_2 \in C_c(G)$; see Ash–Gardner [2, Chapter 1].

The operator K is the *covariance operator* of the process η. Sometimes, η is called *colored noise*. This is particularly so outside of mathematics.

Example 19.1.2. If K denotes the identity operator – that is, $(\mathrm{K}f)(x) := f(x)$ for all $f \in C_c(G)$ and $x \in G$ – then η is white noise on G, but with its index set restricted to $C_c(G)$.

The following extends Proposition 14.3.2 to the present setting. The next result is a way to say that $\eta\colon C_c(G) \to L^2(\Omega)$ is a *random linear functional.*

Proposition 19.1.3. *$f \mapsto \eta(f)$ is a linear mapping from $C_c(G)$ to $L^2(\Omega)$.*

Proof. The proof uses the same method as did the proof of Proposition 14.3.2. Namely, (19.1) implies that, for all $a \in \mathbb{R}$ and $f \in C_c(G)$,

$$\begin{aligned}\mathrm{E}\left(|\eta(af) - a\eta(f)|^2\right) &= \mathrm{E}\left(|\eta(af)|^2\right) + a^2\mathrm{E}\left(|\eta(f)|^2\right) - 2a\mathrm{Cov}[\eta(af)\,,\eta(f)]\\ &= (af\,,\mathrm{K}[af])_{L^2(G)} + a^2\,(f\,,\mathrm{K}f)_{L^2(G)} - 2a\,(af\,,\mathrm{K}f)_{L^2(G)}\\ &= 0.\end{aligned}$$

Thus, $\eta(af) = a\eta(f)$ a.s. Similarly, one shows that $\eta(f_1+f_2) - (\eta(f_1)+\eta(f_2))$ has zero variance, and hence $\eta(f_1 + f_2) = \eta(f_1) + \eta(f_2)$ a.s. □

Example 19.1.2 and Proposition 19.1.3 together show that, in some sense, colored noise is more general than white noise. But this is not quite true, since we have defined colored noise only as a linear map on $C_c(G)$, whereas white noise was defined on all of $L^2(G)$. One can pursue this matter further in the next section. In the mean time, we associate to K an important Hilbert space that is called the *reproducing kernel Hilbert space* associated with K.

Define $(f_1\,,f_2)_{L^2_{\mathrm{K}}(G)} := (f_1\,,\mathrm{K}f_2)_{L^2(G)}$ for all $f_1, f_2 \in C_c(G)$. Then, it is easy to check that $(\ldots\,,\ldots)_{L^2_{\mathrm{K}}(G)}$ is a pre-Hilbertian inner product on $C_c(G)$. It is natural to define also a corresponding "norm" as

$$\|f\|_{L^2_{\mathrm{K}}(G)} := (f\,,\mathrm{K}f)^{1/2}_{L^2(G)}$$

for all $f \in C_c(G)$. Clearly, $\|af\|_{L^2_{\mathrm{K}}(G)} = |a|\cdot\|f\|_{L^2_{\mathrm{K}}(G)}$ for all $a \in \mathbb{R}$ and $f \in C_c(G)$, and

$$\|f_1 + f_2\|_{L^2_{\mathrm{K}}(G)} \leqslant \|f_1\|_{L^2_{\mathrm{K}}(G)} + \|f_2\|_{L^2_{\mathrm{K}}(G)}$$

for all $f_1, f_2 \in C_c(G)$. In other words, $\|\cdots\|_{L^2_{\mathrm{K}}(G)}$ is a pseudo-norm; however it is not always a bona fide norm, as the following shows.

Exercise 19.1.4. Construct a positive definite, self-adjoint operator $\mathrm{K}\colon C_c(G) \to L^2_{\mathrm{loc}}(G)$ and nonzero function $f \in C_c(G)$ such that $\|f\|_{L^2_{\mathrm{K}}(G)} = 0$. (Hint: Consider $(\mathrm{K}\psi)(x) := \mathbf{1}_S(x)\psi(x)$ for all $\psi \in C_c(G)$ and $x \in G$, for a suitable Borel set $S \subset G$.)

The fact that $\|\cdots\|_{L^2_{\mathrm{K}}(G)}$ is not always a norm should not be a disappointment. In fact, the usual $L^2(\mathbb{R})$-norm is also not a proper norm for the vector space of all measurable functions $f : \mathbb{R} \to \mathbb{R}$ with $\int_{-\infty}^{\infty} |f(x)|^2 \, \mathrm{d}x < \infty$. The way around, here, is as it was in the Lebesgue L^2-theory. In the latter case, we identified two functions $f_1, f_2 \in L^2(\mathbb{R})$ if $\int_{-\infty}^{\infty} |f_1(x) - f_2(x)|^2 \, \mathrm{d}x = 0$; then the $L^2(\mathbb{R})$-norm is rendered a bona fide norm on the space $L^2(\mathbb{R})$ of all resulting equivalence classes.

The same procedure can be applied in the present more general setting: we identify $f_1, f_2 \in C_c(G)$ if $\|f_1 - f_2\|_{L^2_{\mathrm{K}}(G)} = 0$. By default, $\|\cdots\|_{L^2_{\mathrm{K}}(G)}$ is a norm on the resulting equivalence classes, in the usual way. Also, as is usual, we abuse notion slightly and write f for the equivalence class of a function f, etc. The following generalizes classical L^2-spaces of Lebesgue.

Definition 19.1.5. Let the space $L^2_{\mathrm{K}}(G)$ denote the completion of $C_c(G)$ in the norm $\|\cdots\|_{L^2_{\mathrm{K}}(G)}$.

We will endow $L^2_{\mathrm{K}}(G)$ with its canonical inner product $(\cdots, \cdots)_{L^2_{\mathrm{K}}(G)}$ and norm $\|\cdots\|_{L^2_{\mathrm{K}}(G)}$. By its very definition, $L^2_{\mathrm{K}}(G)$ is a Hilbert space. In parts of the literature, $L^2_{\mathrm{K}}(G)$ is called the *reproducing kernel Hilbert space* [RKHS] of the *reproducing kernel* K; see Moore [40] and also Aronszajn [1]. The RKHS of K is important since it is the natural domain of definition of the operator K. By this I mean that we can extend without effort, by density, K to a linear operator from $L^2_{\mathrm{K}}(G)$ to $L^2(G)$.

19.2 Colored noise

Our "colored noise" process η has a natural extension to $L^2_{\mathrm{K}}(G)$. In order to develop that extension, let us first observe that if $\|f_1 - f_2\|_{L^2_{\mathrm{K}}(G)} = 0$ for some $f_1, f_2 \in C_c(G)$, then

$$\mathrm{E}\left(|\eta(f_1) - \eta(f_2)|^2\right) = \|f_1 - f_2\|^2_{L^2_{\mathrm{K}}(G)} = 0,$$

thanks to (19.1). In other words, if $f_1, f_2 \in C_c(G)$ are identified with one another, as elements of $L^2_{\mathrm{K}}(G)$, then $\eta(f_1)$ and $\eta(f_2)$ are identified, as elements of $L^2(\Omega)$, with one another as well. Thus, we obtain a unique extension of $\{\eta(f)\}_{f \in C_c(G)}$ to a process $\{\eta(f)\}_{f \in L^2_{\mathrm{K}}(G)}$ by density. Since $L^2(\Omega)$-limits of Gaussian processes are themselves Gaussian, we arrive at the following extension of Proposition 19.1.3.

Proposition 19.2.1. *The process $\eta := \{\eta(f)\}_{f \in L^2_{\mathrm{K}}(G)}$ is a centered Gaussian process with covariance* $\mathrm{Cov}\,[\eta(f_1), \eta(f_2)] = (f_1, \mathrm{K}f_2)_{L^2_{\mathrm{K}}(G)}$ *for all $f_1, f_2 \in L^2_{\mathrm{K}}(G)$. Furthermore, $f \mapsto \eta(f)$ is a linear isometry from $L^2_{\mathrm{K}}(G)$ to $L^2(\Omega)$.*

Let us conclude this section by mentioning a few examples. The first is a continuation of Example 19.1.2, and shows that this model of colored noise includes white noise.

19.2.1 Example: white noise

Let $\mathrm{K}f := f$ denote the identity operator. Then $L^2_{\mathrm{K}}(G) = L^2(G)$ because $C_c(G)$ is dense in $L^2(G)$ (see Rudin [49, E8, p. 268], for example). Thus, we see that our construction of η as a process indexed by $L^2_{\mathrm{K}}(G)$ is a generalization of white noise on G.

19.2.2 Example: Hilbert–Schmidt covariance

We can associate to every real-valued function $\mathrm{k} \in L^2(G \times G)$ a linear operator K as follows: for all $f \in C_c(G)$ and $x \in G$,

$$(\mathrm{K}f)(x) := \int_G f(y)\mathrm{k}(x\,,y)\, m_G(\mathrm{d}y).$$

The integral converges absolutely for m_G-almost every $x \in G$, thanks to the Cauchy–Schwarz inequality.

Definition 19.2.2. Suppose that k is *symmetric*; i.e., $\mathrm{k}(x\,,y) = \mathrm{k}(y\,,x)$ for almost every $(x\,,y) \in G \times G$. Then, K is called a *Hilbert–Schmidt* operator.

Exercise 19.2.3. Prove that if K is Hilbert–Schmidt, then the RKHS of K includes all of $L^2(G)$. But it is entirely possible that $L^2_{\mathrm{K}}(G)$ contains much more than just $L^2(G)$. For instance, prove that if $\mathrm{k} \in C_c(G \times G)$, then $L^2_{\mathrm{K}}(G)$ contains $L^p(G)$ for all $1 \leqslant p \leqslant \infty$. For a greater challenge, prove that if $\mathrm{k} \in C_c(G \times G)$, then the RKHS of K contains all finite Borel measures on G.

The following proposition ensures that if k is positive definite and K is Hilbert–Schmidt, then K is a covariance operator for some centered Gaussian noise η indexed by the RHKS of K.

Proposition 19.2.4. *The linear operator* K *maps* $C_c(G)$ *to* $C_c(G)$. *Suppose, in addition,* K *is Hilbert–Schmidt and* k *is (real) positive definite; that is,*

$$\int_G m_G(\mathrm{d}x) \int_G m_G(\mathrm{d}y)\ \mathrm{k}(x\,,y)f(x)f(y) \geqslant 0 \tag{19.2}$$

for all $f \in C_c(G)$. *Then,* K *is a self-adjoint, positive definite operator.*

Proof. The dominated convergence theorem ensures that $\mathrm{K}\colon C_c(G) \to C_c(G)$, and if k is symmetric then K is self-adjoint thanks to the Fubini theorem. The positive-definite assertion follows also from Fubini's theorem. □

19.2.3 Example: spatially-homogeneous covariance

Choose and fix a function $\mathrm{k} \in L^1_{\mathrm{loc}}(G)$, and define

$$(\mathrm{K}f)(x) := (\mathrm{k} * f)(x) := \int_G \mathrm{k}(xy^{-1})f(y)\, m_G(\mathrm{d}y)$$

for all $x \in G$ and $f \in C_c(G)$. Then, K is a linear operator from $C_c(G)$ to $L^1_{\mathrm{loc}}(G)$. We suppose, in addition, that k is *symmetric* and positive definite. Positive definite means (19.2), and symmetry of course means that

$$\mathrm{k}(x) = \mathrm{k}(x^{-1}) \tag{19.3}$$

for m_G-a.e. $x \in G$. Then, K is a proper covariance operator. In order to understand what its RKHS might look like, we need to study harmonic analysis on G. This is a topic that we have not discussed in these notes; instead, we have chosen to study harmonic analysis on special examples of G. In keeping with this tradition, we will therefore study a particular example in greater depth.

Example 19.2.5. Suppose $G = \mathbb{R}^n$ for some integer $n \geqslant 1$. Then, K can be written as

$$(\mathrm{K}f)(x) = \int_{\mathbb{R}^n} \mathrm{k}(x-y)f(y)\,\mathrm{d}y$$

for all $x \in \mathbb{R}^n$ and $f \in C_c(\mathbb{R}^n)$. Let $\hat{g}$ denote the Fourier transform of $g \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ in the sense of distributions. The symmetry of k implies that $\hat{\mathrm{k}}$ is real. Indeed, (19.3) implies that

$$\begin{aligned}\left(\hat{\mathrm{k}}\,,\varphi\right)_{L^2(G)} &= (\mathrm{k}\,,\hat{\varphi})_{L^2(G)} = \int_{\mathbb{R}^n} \mathrm{k}(x)\overline{\hat{\varphi}(x)}\,\mathrm{d}x = \int_{\mathbb{R}^n} \mathrm{k}(-x)\overline{\hat{\varphi}(x)}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \mathrm{k}(y)\overline{\hat{\varphi}(-y)}\,\mathrm{d}y = \left(\overline{\hat{\mathrm{k}}}\,,\varphi\right)_{L^2(G)}\end{aligned}$$

for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$. And since k is positive definite, then

$$\begin{aligned}0 &\leqslant \iint_{\mathbb{R}^n\times\mathbb{R}^n} \mathrm{k}(x-y)\varphi(x)\varphi(y)\,\mathrm{d}x\,\mathrm{d}y = (\varphi\,,\mathrm{k}*\varphi)_{L^2(G)} \\ &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} |\hat{\varphi}(z)|^2\,\hat{\mathrm{k}}(z)\,\mathrm{d}z.\end{aligned}$$

for all $\varphi \in \mathscr{S}(\mathbb{R}^n)$. The first equality holds because of the Fubini theorem, and the second holds by the Parseval identity. It follows easily from this that

$$\hat{\mathrm{k}}(x) \geqslant 0 \tag{19.4}$$

for a.e. $x \in \mathbb{R}^n$. In fact, we can trace our way back through the preceding to see that Condition (19.4) on the a.e.-positivity of $\hat{\mathrm{k}}$ is equivalent to the positive-definiteness condition (19.2) on k. A similar application of Parseval's identity shows that $L^2_{\mathrm{K}}(\mathbb{R}^n)$ contains every distribution u on $\mathbb{R}^n$ whose Fourier transform $\hat{u}$ is a measurable function and satisfies

$$\int_{\mathbb{R}^n} |\hat{u}(z)|^2 \, \hat{\mathrm{k}}(z) \, \mathrm{d}z < \infty.$$

Thus, for example, if $\hat{\mathrm{k}} \in L^1(\mathbb{R}^n)$, then $L^2_{\mathrm{K}}(\mathbb{R}^n)$ contains the vector space

$$\mathrm{PM}(\mathbb{R}^n) := \{u \in \mathscr{S}'(\mathbb{R}^n) \mid \hat{u} \in L^\infty(\mathbb{R}^n)\}.$$

Elements of $\mathrm{PM}(\mathbb{R}^n)$ are called *pseudo-measures*; see, for example, Kahane–Salem [31]. It is easy to see that every finite Borel measure, and more generally every signed Borel measure of finite total variation, is in $\mathrm{PM}(\mathbb{R}^n)$.

19.2.4 Example: tensor-product covariance

Suppose G_1 and G_2 are two LCA groups and $f_i \colon G_i \to \mathbb{R}$, $i = 1, 2$. Recall that the *tensor product* of f_1 and f_2 is defined as the function $f_1 \otimes f_2 \colon G_1 \times G_2 \to \mathbb{R}$ given by

$$(f_1 \otimes f_2)(x_1, x_2) := f_1(x_1) \cdot f_2(x_2)$$

for all $(x_1, x_2) \in G_1 \times G_2$. If K_i is a linear operator from $C_c(G_i)$ to $L^1_{\mathrm{loc}}(G_i)$ for $i = 1, 2$, then the tensor product $\mathrm{K}_1 \otimes \mathrm{K}_2$ is the linear operator

$$[(\mathrm{K}_1 \otimes \mathrm{K}_2) f](x_1, x_2) := (\mathrm{K}_1 f_1)(x_1) \cdot (\mathrm{K}_2 f_2)(x_2),$$

for $(x_1, x_2) \in G_1 \times G_2$, defined on all functions of the form $f := f_1 \otimes f_2$ such that $f_i \in C_c(G_i)$ for $i = 1, 2$. The class of such functions f is denoted by $C_c(G_1) \otimes C_c(G_2)$.

Since $C_c(G_1) \otimes C_c(G_2)$ is dense in $C_c(G_1 \times G_2)$, one might expect to be able to extend $\mathrm{K}_1 \otimes \mathrm{K}_2$ to a nice linear operator on all of $C_c(G_1 \times G_2)$ without effort. That is almost always the case, as the following result shows.

Proposition 19.2.6. *If each* K_i *is a bounded linear operator from* $C_c(G_i)$ *to* $C_c(G_i)$, $i = 1, 2$, *then there exists a unique continuous extension*[1] *of* $\mathrm{K}_1 \otimes \mathrm{K}_2$ *to a linear operator from* $C_c(G_1 \times G_2)$ *to* $L^1_{\mathrm{loc}}(G_1 \times G_2)$.

Definition 19.2.7. This extension is the *tensor product* of K_1 and K_2; we continue to denote it by $\mathrm{K}_1 \otimes \mathrm{K}_2$.

For the remainder of this subsection we assume that K_i is a bounded linear operator from $C_c(G_i)$ to itself, for $i = 1, 2$.

We prove Proposition 19.2.6 next. The result is a special case of the classical Tietze extension theorem.

[1] In fact, the extension is a linear map from $C_c(G_1 \times G_2)$ to $C_0(G_1 \times G_2)$, where: (i) $C_0(G_1 \times G_2)$ denotes the space of all continuous real-valued functions on $G_1 \times G_2$ vanishing at the infinity of the Hausdorff–Alexandroff one-point compactification of $G_1 \times G_2$ when $G_1 \times G_2$ is noncompact; and (ii) $C_0(G_1 \times G_2) := C(G_1 \times G_2)$ if $G_1 \times G_2$ is compact.

Proof of Proposition 19.2.6. Since K_i is a bounded linear operator for each $i = 1, 2$, there exist finite constants C_1 and C_2 such that

$$\sup_{z\in G_i} |(\mathrm{K}_i\psi)(z)| \leqslant C_i \sup_{z\in G_i} |\psi(z)|$$

for all $\psi \in C_c(G_i)$, and $i = 1, 2$. Consequently, if $f, g \in C_c(G_1) \otimes C_c(G_2)$, then

$$\begin{aligned}
&\sup_{(x_1,x_2)\in G_1\times G_2} \left|[(\mathrm{K}_1 \otimes \mathrm{K}_2) f](x_1, x_2) - [(\mathrm{K}_1 \otimes \mathrm{K}_2) g](x_1, x_2)\right| \\
&\quad\leqslant \sup_{(x_1,x_2)\in G_1\times G_2} \big|[(\mathrm{K}_1 \otimes \mathrm{K}_2)(f_1 \otimes f_2)](x_1, x_2) \\
&\qquad\qquad - [(\mathrm{K}_1 \otimes \mathrm{K}_2)(g_1 \otimes f_2)](x_1, x_2)\big| \\
&\quad\quad + \sup_{(x_1,x_2)\in G_1\times G_2} \big|[(\mathrm{K}_1 \otimes \mathrm{K}_2)(g_1 \otimes f_2)](x_1, x_2) \\
&\qquad\qquad - [(\mathrm{K}_1 \otimes \mathrm{K}_2)(g_1 \otimes g_2)](x_1, x_2)\big| \\
&\quad\leqslant C_1C_2 \sup_{x_1\in G_1} |f_1(x_1) - g_1(x_1)| \cdot \sup_{x_2\in G_2} |f_2(x_2)| \\
&\quad\quad + C_1C_2 \sup_{x_1\in G_1} |g_1(x_1)| \cdot \sup_{x_2\in G_2} |f_2(x_2) - g_2(x_2)| .
\end{aligned}$$

It follows that, whenever $n \mapsto f^n$ is a Cauchy sequence in $C_c(G_1) \otimes C_c(G_2)$, the sequence $n \mapsto (\mathrm{K}_1 \otimes \mathrm{K}_2)f^n$ is Cauchy in $C_c(G_1 \times G_2)$. The proposition follows from this and the density of $C_c(G_1) \otimes C_c(G_2)$ in $C_c(G_1 \times G_2)$. □

Proposition 19.2.8. *If* K_1 *and* K_2 *are self-adjoint and positive definite, then so is* $\mathrm{K}_1 \otimes \mathrm{K}_2$.

Proof. If $f := f_1 \otimes f_2$ and $g := g_1 \otimes g_2$ are elements of $C_c(G_1) \otimes C_c(G_2)$, then

$$(f\, , [\mathrm{K}_1 \otimes \mathrm{K}_2]\, g)_{L^2(G_1\times G_2)} = (f_1\, , \mathrm{K}_1 f_1)_{L^2(G_1)} \cdot (f_2\, , \mathrm{K}_2 f_2)_{L^2(G_2)} .$$

Therefore, $\mathrm{K}_1 \otimes \mathrm{K}_2$ is positive and self-adjoint, viewed as a linear operator from the vector space $C_c(G_1)\otimes C_c(G_2)$ to itself. The result follows from this fact, thanks to density. □

In other words, Propositions 19.2.6 and 19.2.8 together imply that the tensor product $\mathrm{K}_1 \otimes \mathrm{K}_2$ is a covariance operator. Thus, the remarks of this chapter apply to $\mathrm{K}_1 \otimes \mathrm{K}_2$.

Exercise 19.2.9. Be sure that you can understand how space-time white noise (more precisely, white noise on $\mathbb{R} \times G$) is covered by the preceding tensor-product covariance kernels. (Hint: Consider $\mathrm{K}_1 f := f$ and $\mathrm{K}_2 g := g$.)

19.3 Linear SPDEs with colored-noise forcing

We conclude these notes by studying briefly a linear stochastic PDE that is driven by "space-time colored noise".

To begin, suppose that G is an LCA group. Let K denote a covariance operator on G and suppose Γ is a covariance operator on the additive group $\mathbb{R}$. We will assume in addition that $\mathrm{K}\colon C_c(G) \to C_c(G)$ and $\Gamma\colon C_c(\mathbb{R}) \to C_c(\mathbb{R})$ are bounded linear operators. In this way, Propositions 19.2.6 and 19.2.8 of the previous section ensure the existence of the centered Gaussian noise

$$\eta := \{\eta(f)\}_{f \in L^2_{\Gamma\otimes\mathrm{K}}(\mathbb{R}\times G)},$$

whose covariance operator is $\Gamma \otimes \mathrm{K}$.

As was done earlier, let $\mathcal{G}$ denote the generator of a Lévy process X on G such that the Lévy process X^{-1} has transition densities $\{p_t\}_{t>0}$. Consider the stochastic heat equation,

$$\dot{u} = \mathcal{G}u + \eta, \tag{19.5}$$

on $(0\,,\infty) \times G$ and subject to $u_0 \equiv 0$ for simplicity. As we did earlier, we interpret the preceding SPDE via a formal application of Duhamel's principle. This will (still non-rigorously) lead us to the following candidate expression for the solution to (19.5) (see (16.6), for example):

$$u_t(x) = \int_{(0,t)\times G} p_{t-s}(xy^{-1})\,\eta(\mathrm{d}s\,\mathrm{d}y). \tag{19.6}$$

If the preceding colored-noise Wiener integral makes sense, then we say that (19.5) has a solution u. Else it does not make sense for us to talk about (19.5). Since the preceding integral is a centered Gaussian random variable, the existence of a solution to (19.5) is reduced to the existence of a second moment for $u_t(x)$ for all $t > 0$ and $x \in G$. That problem is now concrete, and has to do with whether or not the transition density function p is in the RKHS of $\Gamma \otimes \mathrm{K}$, as the following shows.

Theorem 19.3.1. *If*

$$(s\,,t) \mapsto p_{t-s}(xy^{-1})\mathbf{1}_{(0,t)}(s) \in L^2_{\Gamma\otimes\mathrm{K}}(\mathbb{R}\times G) \tag{19.7}$$

for all $t > 0$ and $x \in G$, then the stochastic integral in (19.6) *has a finite second moment, and hence the linear SPDE* (19.5) *is well defined.*

Proof. The tensorized vector space $L^2_\Gamma(\mathbb{R}) \otimes L^2_{\mathrm{K}}(G)$ is dense in $L^2_{\Gamma\otimes\mathrm{K}}(\mathbb{R}\times G)$. For all $T\colon \mathbb{R} \to \mathbb{R}$ and $X\colon G \to \mathbb{R}$, and for all $t > 0$ and $x \in G$, define

$$T_t(s) := \mathbf{1}_{[0,t]}(s)T(t-s) \quad \text{and} \quad X_x(y) := X(xy^{-1})$$

for $0 < s < t$, $y \in G$. By density, we can approximate $(s\,,y) \mapsto p_{t-s}(xy^{-1})$ by finite linear combinations of functions of the form $(s\,,y) \mapsto T_t(s)X_x(y) = (T_t \otimes X_x)(s\,,y)$

where $T_t \in L^2_\Gamma(\mathbb{R})$ and $X_x \in L^2_{\mathrm{K}}(G)$ for all $t > 0$ and $x \in G$. Proposition 19.2.1 does the rest, since

$$\|T_t \otimes X_x\|_{L^2_{\Gamma\otimes\mathrm{K}}(\mathbb{R}\times G)} = \|T_t\|_{L^2_\Gamma(\mathbb{R})} \cdot \|X_x\|_{L^2_{\mathrm{K}}(G)} < +\infty,$$

for all $t > 0$ and $x \in G$. □

Though it is not entirely obvious, the present condition (19.7) is the colored-noise extension of the Dalang condition (16.7). Thus, we end with the following exercise, which clarifies the connection.

Exercise 19.3.2. Suppose η is "white in time"; that is, $\Gamma f = f$ for all $f \in C_c(\mathbb{R})$. Suppose also that $p_s \in L^2_{\mathrm{K}}(G)$ for all $s > 0$. Then, prove that (19.7) is equivalent to

$$\int_0^t \|p_s\|^2_{L^2_{\mathrm{K}}(G)} \mathrm{d}s < \infty$$

for all $t > 0$. This condition holds, in particular, if $\int_0^\infty \exp(-\alpha s)\|p_s\|^2_{L^2_{\mathrm{K}}(G)} \mathrm{d}s < \infty$ for some $\alpha > 0$.

Bibliography

[1] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68**(3) (1950), 337–404.

[2] R.B. Ash and M.F. Gardner. Topics in Stochastic Processes. Academic Press Harcourt Brace Jovanovich Publishers, New York, 1975. Probability and Mathematical Statistics, Vol. 27.

[3] J. Bertoin. Lévy Processes. *Cambridge Tracts in Mathematics* **121**. Cambridge University Press, Cambridge, 1996.

[4] E. Carlen and P. Krée. L^p estimates on iterated stochastic integrals. *Ann. Probab.* **19**(1) (1991), 354–368.

[5] R.A. Carmona and S.A. Molchanov. Parabolic Anderson problem and intermittency. *Mem. Amer. Math. Soc.* **108**(518) (1994), viii+125.

[6] P.L. Chow. Stochastic Partial Differential Equations. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[7] K.L. Chung and W.H.J. Fuchs. On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **1951**(6) (1951), 12.

[8] D. Conus, M. Joseph, and D. Khoshnevisan. On the chaotic character of the stochastic heat equation, before the onset of intermittency. *Ann. Probab.* **41**(3B) (2013), 2225–2260.

[9] I. Corwin. The Kardar–Parisi–Zhang equation and universality class. *Random Matrices Theory Appl.* **1**(1) (2012), 76.

[10] J.T. Cox, K. Fleischmann, and A. Greven. Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Related Fields* **105**(4) (1996), 513–528.

[11] G. Da Prato and J. Zabczyk. Stochastic Equations in Infinite Dimensions, *Encyclopedia of Mathematics and its Applications* **44**. Cambridge University Press, Cambridge, 1992.

[12] R.C. Dalang. Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4**(6) (1999), 29 pp. (electronic).

[13] R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, and Y. Xiao. A Minicourse on Stochastic Partial Differential Equations. *Lecture Notes in Mathematics* **1962**. Springer-Verlag, Berlin, 2009.

[14] R.C. Dalang and L. Quer-Sardanyons. Stochastic integrals for spde's: a comparison. *Expo. Math.* **29**(1) (2011), 67–109.

[15] D.A. Dawson and E.A. Perkins. Measure-valued processes and renormalization of branching particle systems. In Stochastic Partial Differential Equations: Six Perspectives. *Math. Surveys Monogr.* **64**, 45–106. Amer. Math. Soc., Providence, RI, 1999.

[16] R. Durrett. Probability: Theory and Examples. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1991.

[17] L.C. Evans. Partial Differential Equations. *Graduate Studies in Mathematics* **19**. American Mathematical Society, Providence, RI, first edition, 1998.

[18] M. Foondun and D. Khoshnevisan. On the stochastic heat equation with spatially-colored random forcing. *Trans. Amer. Math. Soc.* **365**(1) (2013), 409–458.

[19] M. Foondun and D. Khoshnevisan. Corrections and improvements to: "On the stochastic heat equation with spatially-colored random forcing". *Trans. Amer. Math. Soc.* **366**(1) (2014), 561–562.

[20] T. Funaki. Random motion of strings and related stochastic evolution equations. *Nagoya Math. J.* **89** (1983), 129–193.

[21] C. Geiß and R. Manthey. Comparison theorems for stochastic differential equations in finite and infinite dimensions. *Stochastic Process. Appl.* **53**(1) (1994), 23–35.

[22] G. Giacomin, J.L. Lebowitz, and E. Presutti. Deterministic and stochastic hydrodynamical equations arising from simple microscopic model systems. In Stochastic Partial Differential Equations: Six Perspectives. *Math. Surveys Monogr.* **64**, 107–152. Amer. Math. Soc., Providence, RI, 1999.

[23] A. Greven and F. den Hollander. Phase transitions for the long-time behavior of interacting diffusions. *Ann. Probab.* **35**(4) (2007), 1250–1306.

[24] Y. Hu, F. Lu, and D. Nualart. Feynman–Kac formula for the heat equation driven by fractional noise with Hurst parameter $H < 1/2$. *Ann. Probab.* **40**(3) (2012), 1041–1068.

[25] N. Ikeda and S. Watanabe. A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.* **14**(3) (1977), 619–633.

[26] K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo* **20** (1944), 519–524.

[27] N. Jacob. Pseudo Differential Operators and Markov Processes. Vol. I. Imperial College Press, London, 2001.

[28] N. Jacob. Pseudo Differential Operators & Markov Processes. Vol. II. Imperial College Press, London, 2002.

[29] N. Jacob. Pseudo Differential Operators and Markov Processes. Vol. III. Imperial College Press, London, 2005.

[30] M. Joseph, D. Khoshnevisan, and C. Mueller. Strong invariance and noise-comparison principles for some parabolic stochastic PDEs. Preprint, http://arxiv.org/abs/1404.6911, 2014.

[31] J.P. Kahane and R. Salem. Ensembles Parfaits et Séries Trigonométriques. Hermann, Paris, second edition, 1994. With notes by Kahane, T.W. Körner, R. Lyons, and S.W. Drury.

[32] D. Khoshnevisan. Analysis of Stochastic Partial Differential Equations. *CBMS Regional Conference Series in Mathematics* **119**. American Mathematical Society.

[33] D. Khoshnevisan and K. Kim. Non-linear noise excitation and intermittency under high disorder. *Proc. Amer. Math. Soc.*, 2014. Available electronically at http://arxiv.org/abs/1302.1621.

[34] D. Khoshnevisan and K. Kim. Non-linear noise excitation of intermittent stochastic PDEs and the topology of LCA groups. *Ann. Probab.*, 2014. Available electronically at http://arxiv.org/abs/1302.3266.

[35] P. Kotelenez. Stochastic Ordinary and Stochastic Partial Differential Equations. *Stochastic Modelling and Applied Probability* **58**, Springer, New York, 2008.

[36] N.V. Krylov. An analytic approach to SPDEs. In Stochastic Partial Differential Equations: Six Perspectives. *Math. Surveys Monogr.* **64**, 183–242. Amer. Math. Soc., Providence, RI, 1999.

[37] N. Krylov. A brief overview of the L_p-theory of SPDEs. *Theory Stoch. Process.* **14**(2) (2008), 71–78.

[38] N.V. Krylov and B.L. Rozovskii. Stochastic Evolution Equations. In Stochastic Differential Equations: Theory and Applications. *Interdiscip. Math. Sci.* **2**, 1–69. World Sci. Publ., Hackensack, NJ, 2007.

[39] R. Mikulevicius and B.L. Rozovskii. Martingale problems for stochastic PDEs. In Stochastic Partial Differential Equations: Six Perspectives. *Math. Surveys Monogr.* **64**, 243–325. Amer. Math. Soc., Providence, RI, 1999.

[40] E.H. Moore. General Analysis, Part 2: Fundamental Notions of General Analysis. *Mem. Amer. Phil. Soc.* **1** (1939).

[41] S.A. Morris. Pontryagin Duality and the Structure of Locally Compact Abelian Groups. Cambridge University Press, Cambridge-New York-Melbourne, 1977. London Mathematical Society Lecture Note Series, No. 29.

[42] C. Mueller. On the support of solutions to the heat equation with noise. *Stochastics Stochastics Rep.* **37**(4) (1991), 225–245.

[43] I. Nourdin, G. Peccati, and F.G. Viens. Comparison inequalities on Wiener space. *Stochastic Process. Appl.* **124**(4) (2014), 1566–1581.

[44] S. Peszat and J. Zabczyk. Nonlinear stochastic wave and heat equations. *Probab. Theory Related Fields* **116**(3) (2000), 421–443.

[45] C. Prévôt and M. Röckner. A Concise Course on Stochastic Partial Differential Equations. *Lecture Notes in Mathematics* **1905**, Springer, Berlin, 2007.

[46] J. Quastel. Introduction to KPZ. In Current Developments in Mathematics (2011), 125–194. Int. Press, Somerville, MA, 2012.

[47] D. Revuz and M. Yor. Continuous Martingales and Brownian Motion. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer-Verlag, Berlin, 1991.

[48] J. Rosiński. Tempering stable processes. *Stochastic Process. Appl.* **117**(6) (2007), 677–707.

[49] W. Rudin. Fourier Analysis on Groups. Interscience Publishers (a division of John Wiley and Sons), New York-London, 1962.

[50] D. Ruelle. Turbulent dynamical systems. In *Proceedings of the International Congress of Mathematicians* **1–2** (1983), 271–286. PWN, Warsaw, 1984.

[51] M. Sanz-Solé. Malliavin Calculus: With Applications to Stochastic Partial Differential Equations. Fundamental Sciences. EPFL Press, Lausanne, 2005.

[52] K.I. Sato. Lévy Processes and Infinitely Divisible Distributions. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, revised by the author.

[53] T. Shiga. Ergodic theorems and exponential decay of sample paths for certain interacting diffusion systems. *Osaka J. Math.* **29**(4) (1992), 789–807.

[54] T. Shiga and A. Shimizu. Infinite-dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* **20**(3) (1980), 395–416.

[55] D. Slepian. The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.* **41** (1962), 463–501.

[56] F. Spitzer. Principles of Random Walks. The University Series in Higher Mathematics. D. Van Nostrand Co., Inc., Princeton, N.J. Toronto-London, 1964.

[57] J.B. Walsh. An Introduction to Stochastic Partial Differential Equations. In École d'été de Probabilités de Saint-Flour XIV, 1984. *Lecture Notes in Math.* **1180**, 265–439. Springer, Berlin, 1986.

[58] N.A. Watson. Uniqueness and representation theorems for parabolic equations. *J. London Math. Soc.* (2) **8** (1974), 311–321.

[59] N.A. Watson. Green functions, potentials, and the Dirichlet problem for the heat equation. *Proc. London Math. Soc.* (3) **33**(2) (1976), 251–298.

[60] N. Wiener. Differential space. *J. Math. and Physics* **58** (1923), 131–174.

[61] N. Wiener. The Homogeneous Chaos. *Amer. J. Math.* **60**(4) (1938), 897–936.

Index

D. Khoshnevisan, R. Schilling, *From Lévy-Type Processes to Parabolic SPDEs*,
Advanced Courses in Mathematics - CRM Barcelona, DOI 10.1007/978-3-319-34120-0

Advanced Courses in Mathematics – CRM Barcelona (ACM)

Edited by

Enric Ventura, Universitat Politècnica de Catalunya

Since 1995 the Centre de Recerca Matemàtica (CRM) has organised a number of Advanced Courses at the post-doctoral or advanced graduate level on forefront research topics in Barcelona. The books in this series contain revised and expanded versions of the material presented by the authors in their lectures.

■ **Fock, V. / Marshakov, A. / Schaffhauser, F. / Teleman, C. / Wentworth, R.,** Geometry and Quantization of Moduli Spaces (2017)
ISBN 978-3-319-33577-3
This volume is based on four advanced courses held at the Centre de Recerca Matemàtica (CRM), Barcelona. It presents both background information and recent developments on selected topics that are experiencing extraordinary growth within the broad research area of geometry and quantization of moduli spaces. The lectures focus on the geometry of moduli spaces which are mostly associated to compact Riemann surfaces and are presented both from classical and quantum perspectives.

■ **Bolsinov, A. / Morales-Ruiz, J. / Zung, N.,** Geometry and Dynamics of Integrable Systems (2016)
ISBN 978-3-319-33502-5
These lecture notes address three major aspects of integrable systems: obstructions to integrability from differential Galois theory; the description of singularities of integrable systems on the basis of their relation to bi-Hamiltonian systems; and the generalization of integrable systems to the non-Hamiltonian settings. All three sections were written by top experts in their respective fields.
Native to actual problem-solving challenges in mechanics, the topic of integrable systems is currently at the crossroads of several disciplines in pure and applied mathematics, and also has important interactions with physics. The study of integrable systems also actively employs methods from differential geometry. Moreover, it is extremely important in symplectic geometry and Hamiltonian dynamics, and has strong correlations with mathematical physics, Lie theory and algebraic geometry (including mirror symmetry). As such, the book will appeal to experts with a wide range of backgrounds.

■ **Bally, V. / Caramellino, L. / Cont, R.,** Stochastic Integration by Parts and Functional Itô Calculus (2016).
ISBN 978-3-319-27127-9
This volume contains lecture notes from the courses given by Vlad Bally and Rama Cont at the Barcelona Summer School on Stochastic Analysis (July 2012). The notes of the course by Vlad Bally, co-authored with Lucia Caramellino, develop integration by parts formulas in an abstract setting, extending Malliavin's work on abstract Wiener spaces. The results are applied to prove absolute continuity and regularity results of the density for a broad class of random processes. Rama Cont's notes provide an introduction to the Functional Itô Calculus, a nonanticipative functional calculus that extends the classical Itô calculus to path-dependent functionals of stochastic processes. This calculus leads to a new class of path-dependent partial differential equations, termed Functional Kolmogorov Equations, which arise in the study of martingales and forward-backward stochastic differential equations. This book will appeal to both young and senior researchers in probability and stochastic processes, as well as to practitioners in mathematical finance.

■ **Hacking, P. / Laza, R. / Oprea, D.,** Compactifying Moduli Spaces (2015).
ISBN 978-3-0348-0920-7
This book focusses on a large class of objects in moduli theory and provides different perspectives from which compactifications of moduli spaces may be investigated.
Three contributions give an insight on particular aspects of moduli problems. In the first of them, various ways to construct and compactify moduli spaces are presented. In the second, some questions on the boundary of moduli spaces of surfaces are addressed. Finally, the theory of stable quotients is explained, which yields meaningful compactifications of moduli spaces of maps.
Both advanced graduate students and researchers in algebraic geometry will find this book a valuable read.

■ **Llibre, J. / Moeckel, R. / Simó, C.,** Central Configurations, Periodic Orbits, and Hamiltonian Systems (2015).
ISBN 978-3-0348-0932-0

■ **Alexeev, V.,** Moduli of Weighted Hyperplane Arrangements (2015).
ISBN 978-3-0348-0914-6