Lévy Processes and Stochastic Calculus

Second Edition

CAMBRIDGE STUDIES IN ADVANCED MATHEMATICS

All the titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing visit: www.cambridge.org/series/sSeries.asp?code=CSAM

Already published

61 J. D. Dixon *et al. Analytic pro-p groups*
62 R. P. Stanley *Enumerative combinatorics, II*
63 R. M. Dudley *Uniform central limit theorems*
64 J. Jost & X. Li-Jost *Calculus of variations*
65 A. J. Berrick & M. E. Keating *An introduction to rings and modules*
66 S. Morosawa *et al. Holomorphic dynamics*
67 A. J. Berrick & M. E. Keating *Categories and modules with K-theory in view*
68 K. Sato *Lévy processes and infinitely divisible distributions*
69 H. Hida *Modular forms and Galois cohomology*
70 R. Iorio & V. Iorio *Fourier analysis and partial differential equations*
71 R. Blei *Analysis in integer and fractional dimensions*
72 F. Borceux & G. Janelidze *Galois theories*
73 B. Bollobás *Random graphs*
74 R. M. Dudley *Real analysis and probability*
75 T. Sheil-Small *Complex polynomials*
76 C. Voisin *Hodge theory and complex algebraic geometry, I*
77 C. Voisin *Hodge theory and complex algebraic geometry, II*
78 V. Paulsen *Completely bounded maps and operator algebras*
79 F. Gesztesy & H. Holden *Soliton equations and their algebro-geometric solutions, I*
81 S. Mukai *An introduction to invariants and moduli*
82 G. Tourlakis *Lectures in logic and set theory, I*
83 G. Tourlakis *Lectures in logic and set theory, II*
84 R. A. Bailey *Association schemes*
85 J. Carlson, S. Müller-Stach & C. Peters *Period mappings and period domains*
86 J. J. Duistermaat & J. A. C. Kolk *Multidimensional real analysis, I*
87 J. J. Duistermaat & J. A. C. Kolk *Multidimensional real analysis, II*
89 M. C. Golumbic & A. N. Trenk *Tolerance graphs*
90 L. H. Harper *Global methods for combinatorial isoperimetric problems*
91 I. Moerdijk & J. Mrcun *Introduction to foliations and Lie groupoids*
92 J. Kollár, K. E. Smith & A. Corti *Rational and nearly rational varieties*
93 D. Applebaum *Lévy processes and stochastic calculus*
94 B. Conrad *Modular forms and the Ramanujan conjecture*
95 M. Schechter *An introduction to nonlinear analysis*
96 R. Carter *Lie algebras of finite and affine type*
97 H. L. Montgomery & R.C Vaughan *Multiplicative number theory, I*
98 I. Chavel *Riemannian geometry (2nd Edition)*
99 D. Goldfeld *Automorphic forms and L-functions for the group GL(n,R)*
100 M. B. Marcus & J. Rosen *Markov processes, Gaussian processes, and local times*
101 P. Gille & T. Szamuely *Central simple algebras and Galois cohomology*
102 J. Bertoin *Random fragmentation and coagulation Processes*
103 E. Frenkel *Langlands correspondence for loop groups*
104 A. Ambrosetti & A. Malchiodi *Nonlinear analysis and semilinear elliptic problems*
105 T. Tao & V. H. Vu *Additive combinatorics*
106 E. B. Davies *Linear operators and their spectra*
107 K. Kodaira *Complex analysis*
108 T. Ceccherini-Silberstein, F. Scarabotti & F. Tolli *Harmonic analysis on finite groups*
109 H. Geiges *An introduction to contact topology*
110 J. Faraut *Analysis on Lie groups*
111 E. Park *Complex topological K-theory*
112 D. W. Stroock *Partial differential equations for probabilists*
113 A. Kirillov *An introduction to Lie groups and Lie algebras*
114 F. Gesztesy *et al. Soliton equations and their algebro-geometric solutions, II*
115 A. Iserles *A first course in the numerical analysis of differential equations (2nd Edition)*

Lévy Processes and Stochastic Calculus

Second Edition

DAVID APPLEBAUM

University of Sheffield

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521738651

First edition published 2004

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-73865-1 hardback

To Jill

And lest I should be exalted above measure through the abundance of revelations, there was given to me a thorn in the flesh, a messenger of Satan to buffet me, lest I should be exalted above measure.

Second Epistle of St Paul to the Corinthians, Chapter 12

The more we jump – the more we get – if not more quality, then at least more variety. James Gleick *Faster*

Contents

Preface to Second Edition

It is four years since the first version of this book appeared and there has continued to be intense activity focused on Lévy processes and related areas. One way of gauging this is to look at the number of books and monographs which have appeared in this time. Regarding fluctuation theory of Lévy processes, there is a new volume by A. Kyprianou [221] and the St Flour lectures of R. Doney [96]. From the point of view of interactions with analysis, N. Jacob has published the third and final volume of his impressive trilogy [182]. Applications to finance has continued to be a highly active and fast moving area and there are two new books here – a highly comprehensive and thorough guide by R. Cont and P. Tankov [81] and a helpful introduction aimed at practioners from W.Schoutens [329]. There have also been new editions of classic texts by Jacod and Shiryaev [183] and Protter [298].

Changes to the present volume are of two types. On the one hand there was the need to correct errors and typos and also to make improvements where this was appropriate. In this respect, I am extremely grateful to all those readers who contacted me with remarks and suggestions. In particular I would like to thank Fangjun Xu, who is currently a first-year graduate student at Nanzai University, who worked through the whole book with great zeal and provided me with an extremely helpful list of typos and mistakes. Where there were more serious errors, he took the trouble to come up with his own proofs, all of which were correct.

I have also included some new material, particularly where I think that the topics are important for future work. These include the following. Chapter 1 now has a short introductory section on regular variation which is the main tool in the burgeoning field of 'heavy tailed modelling'. In Chapter 2, there is additional material on bounded variation Lévy processes and on the existence of moments for Lévy processes. Chapter 4 includes new estimates on moments of Lévy-type stochastic integrals which have recently been obtained by H. Kunita [218]. In

Chapter 5, I have replaced the proof of the Itô and martingale representation theorem which was previously given only in the Brownian motion case, with one that works for general Lévy processes. I then develop the theory of multiple Wiener–Itô integrals (again in the general context) and apply the martingale representation theorem to prove Itô's result on chaos decomposition. I have also included a short introduction to Malliavin calculus, albeit only in the Brownian case, as this is now an area of intense activity which extends from quite abstract path space analysis and geometry to option pricing. As it is quite extensively dealt with in Cont and Tankov [81] and Schoutens [329], I resisted the temptation to include more material on mathematical finance with one exception – a natural extension of the Black–Scholes pde to include jump terms now makes a brief entrance on to the stage. In Chapter 6, the rather complicated proof of continuity of solutions of SDEs with respect to their initial conditions has been replaced by a new streamlined version due to Kunita [218] and employing his estimates on stochastic integrals mentioned above. There is also a new section on Lyapunov exponents for SDEs which opens the gates to the study of their asymptotic stability. Once again it is a pleasure to thank Fangjun Xu who carefully read and commented on all of this material. The statutory free copy of the book will be small recompense for his labours. I would also like to thank N.H. Bingham and H. Kunita for helpful remarks and my student M. Siakalli for some beneficial discussions. Cambridge University Press have continued to offer superb support and I would once again like to thank my editor David Tranah and all of his staff, particularly Peter Thompson who took great pains in helping me navigate through the elaborate system of CUP-style LaTeX.

Of course the ultimate responsibility for any typos and more serious errors is mine. Readers are strongly encouraged to continue to send them to me at d.applebaum@sheffield.ac.uk. They will be posted on my website at http://www.applebaum.staff.shef.ac.uk/.

Preface

The aim of this book is to provide a straightforward and accessible introduction to stochastic integrals and stochastic differential equations driven by Lévy processes.

Lévy processes are essentially stochastic processes with stationary and independent increments. Their importance in probability theory stems from the following facts:

- they are analogues of random walks in continuous time;
- they form special subclasses of both semimartingales and Markov processes for which the analysis is on the one hand much simpler and on the other hand provides valuable guidance for the general case;
- they are the simplest examples of random motion whose sample paths are right-continuous and have a number (at most countable) of random jump discontinuities occurring at random times, on each finite time interval.
- they include a number of very important processes as special cases, including Brownian motion, the Poisson process, stable and self-decomposable processes and subordinators.

Although much of the basic theory was established in the 1930s, recent years have seen a great deal of new theoretical development as well as novel applications in such diverse areas as mathematical finance and quantum field theory. Recent texts that have given systematic expositions of the theory have been Bertoin [39] and Sato [323]. Samorodnitsky and Taqqu [319] is a bible for stable processes and related ideas of self-similarity, while a more applications-oriented view of the stable world can be found in Uchaikin and Zolotarev [350]. Analytic features of Lévy processes are emphasised in Jacob [179, 180]. A number of new developments in both theory and applications are surveyed in the volume [26].

Stochastic calculus is motivated by the attempt to understand the behaviour of systems whose evolution in time $X = (X(t), t \geq 0)$ contains both deterministic and random noise components. If X were purely deterministic then three centuries of calculus have taught us that we should seek an infinitesimal description of the way X changes in time by means of a differential equation

$$\frac{dX(t)}{dt} = F(t, X(t))dt.$$

If randomness is also present then the natural generalisation of this is a stochastic differential equation:

$$dX(t) = F(t, X(t))dt + G(t, X(t))dN(t),$$

where $(N(t), t \geq 0)$ is a 'driving noise'.

There are many texts that deal with the situation where $N(t)$ is a Brownian motion or, more generally, a continuous semimartingale (see e.g. Karatzas and Shreve [200], Revuz and Yor [306], Kunita [215]). The only volumes that deal systematically with the case of general (not necessarily continuous) semimartingales are Protter [298], Jacod and Shiryaev [183], Métivier [262] and, more recently, Bichteler [47]; however, all these make heavy demands on the reader in terms of mathematical sophistication. The approach of the current volume is to take $N(t)$ to be a Lévy process (or a process that can be built from a Lévy process in a natural way). This has two distinct advantages:

- The mathematical sophistication required is much less than for general semimartingales; nonetheless, anyone wanting to learn the general case will find this a useful first step in which all the key features appear within a simpler framework.
- Greater access is given to the theory for those who are only interested in applications involving Lévy processes.

The organisation of the book is as follows. Chapter 1 begins with a brief review of measure and probability. We then meet the key notions of infinite divisibility and Lévy processes. The main aim here is to get acquainted with the concepts, so proofs are kept to a minimum. The chapter also serves to provide orientation towards a number of interesting theoretical developments in the subject that are not essential for stochastic calculus.

In Chapter 2, we begin by presenting some of the basic ideas behind stochastic calculus, such as filtrations, adapted processes and martingales. The main aim is to give a martingale-based proof of the Lévy–Itô decomposition of an arbitrary Lévy process into Brownian and Poisson parts. We then meet the important idea of interlacing, whereby the path of a Lévy process is obtained as the almost-sure limit of a sequence of Brownian motions with drift interspersed with jumps of random size appearing at random times.

Chapter 3 aims to move beyond Lévy processes to study more general Markov processes and their associated semigroups of linear mappings. We emphasise, however, that the structure of Lévy processes is the paradigm case and this is exhibited both through the Courrège formula for the infinitesimal generator of Feller processes and the Beurling–Deny formula for symmetric Dirichlet forms. This chapter is more analytical in flavour than the rest of the book and makes extensive use of the theory of linear operators, particularly those of pseudo-differential type. Readers who lack background in this area can find most of what they need in the chapter appendix.

Stochastic integration is developed in Chapter 4. A novel aspect of our approach is that Brownian and Poisson integration are unified using the idea of a martingale-valued measure. At first sight this may strike the reader as technically complicated but, in fact, the assumptions that are imposed ensure that the development remains accessible and straightforward. A highlight of this chapter is the proof of Itô's formula for Lévy-type stochastic integrals.

The first part of Chapter 5 deals with a number of useful spin-offs from stochastic integration. Specifically, we study the Doléans-Dade stochastic exponential, Girsanov's theorem and its application to change of measure, the Cameron–Martin formula and the beginnings of analysis in Wiener space and martingale representation theorems. Most of these are important tools in mathematical finance and the latter part of the chapter is devoted to surveying the application of Lévy processes to option pricing, with an emphasis on the specific goal of finding an improvement to the celebrated but flawed Black–Scholes formula generated by Brownian motion. At the time of writing, this area is evolving at a rapid pace and we have been content to concentrate on one approach using hyperbolic Lévy processes that has been rather well developed. We have included, however, a large number of references to alternative models.

Finally, in Chapter 6, we study stochastic differential equations driven by Lévy processes. Under general conditions, the solutions of these are Feller processes and so we gain a concrete class of examples of the theory developed in Chapter 3. Solutions also give rise to stochastic flows and hence generate random dynamical systems.

The book naturally falls into two parts. The first three chapters develop the fundamentals of Lévy processes with an emphasis on those that are useful in stochastic calculus. The final three chapters develop the stochastic calculus of Lévy processes.

Each chapter closes with some brief historical remarks and suggestions for further reading. I emphasise that these notes are only indicative; no attempt has been made at a thorough historical account, and in this respect I apologise to any readers who feel that their contribution is unjustly omitted. More thorough historical notes in relation to Lévy processes can be found in the chapter notes to Sato [323], and for stochastic calculus with jumps see those in Protter [298].

This book requires background knowledge of probability and measure theory (such as might be obtained in a final-year undergraduate mathematics honours programme), some facility with real analysis and a smattering of functional analysis (particularly Hilbert spaces). Knowledge of basic complex variable theory and some general topology would also be an advantage, but readers who lack this should be able to read on without too much loss. The book is designed to be suitable for underpinning a taught masters level course or for independent study by first-year graduate students in mathematics and related programmes. Indeed, the two parts would make a nice pair of linked half-year modules. Alternatively, a course could also be built from the core of the book, Chapters 1, 2, 4 and 6. Readers with a specific interest in finance can safely omit Chapter 3 and Section 6.4 onwards, while analysts who wish to deepen their understanding of stochastic representations of semigroups might leave out Chapter 5.

A number of exercises of varying difficulty are scattered throughout the text. I have resisted the temptation to include worked solutions, since I believe that the absence of these provides better research training for graduate students. However, anyone having persistent difficulty in solving a problem may contact me by e-mail or otherwise.

I began my research career as a mathematical physicist and learned modern probability as part of my education in quantum theory. I would like to express my deepest thanks to my teachers Robin Hudson, K.R. Parthasarathy and Luigi Accardi for helping me to develop the foundations on which later studies have been built. My fascination with Lévy processes began with my attempt to understand their wonderful role in implementing cocycles by means of annihilation, creation and conservation processes associated with the free quantum field, and this can be regarded as the starting point for quantum stochastic calculus. Unfortunately, this topic lies outside the scope of this volume but interested readers can consult Parthasarathy [291], pp. 152–61 or Meyer [267], pp. 120–1.

My understanding of the probabilistic properties of Lévy processes has deepened as a result of work in stochastic differential equations with jumps over the past 10 years, and it is a great pleasure to thank my collaborators Hiroshi Kunita, Serge Cohen, Anne Estrade, Jiang-Lun Wu and my student Fuchang Tang for many joyful and enlightening discussions. I would also like to thank

René Schilling for many valuable conversations concerning topics related to this book. It was he who taught me about the beautiful relationship with pseudo-differential operators, which is described in Chapter 3. Thanks are also due to Jean Jacod for clarifying my understanding of the concept of predictability and to my colleague Tony Sackfield for advice about Bessel functions.

Earlier versions of this book were full of errors and misunderstandings and I am enormously indebted to Nick Bingham, Tsukasa Fujiwara, Fehmi Özkan and René Schilling, all of whom devoted the time and energy to read extensively and criticize early drafts. Some very helpful comments were also made by Krishna Athreya, Ole Barndorff-Nielsen, Uwe Franz, Vassili Kolokoltsov, Hiroshi Kunita, Martin Lindsay, Nikolai Leonenko, Carlo Marinelli (particularly with regard to LaTeX) and Ray Streater. Nick Bingham also deserves a special thanks for providing me with a valuable tutorial on English grammar. Many thanks are also due to two anonymous referees employed by Cambridge University Press. The book is greatly enriched thanks to their perceptive observations and insights.

In March 2003, I had the pleasure of giving a course, partially based on this book, at the University of Greifswald, as part of a graduate school on quantum independent increment processes. My thanks go to the organisers, Michael Schürmann and Uwe Franz, and all the participants for a number of observations that have improved the manuscript.

Many thanks are also due to David Tranah and the staff at Cambridge University Press for their highly professional yet sensitive management of this project.

Despite all this invaluable assistance, some errors surely still remain and the author would be grateful to be e-mailed about these at dba@maths.ntu.ac.uk. Corrections received after publication will be posted on his website http://www.scm.ntu.ac.uk/dba/.[1]

[1] Note added in second edition. This website is no longer active. The relevant address is now http://www.applebaum.staff.shef.ac.uk/

Overview

It can be very useful to gain an intuitive feel for the behaviour of Lévy processes and the purpose of this short introduction is to try to develop this. Of necessity, our mathematical approach here is somewhat naive and informal – the structured, rigorous development begins in Chapter 1.

Suppose that we are given a probability space $(\Omega, \mathcal{F}, P)$. A Lévy process $X = (X(t), t \geq 0)$ taking values in $\mathbb{R}^d$ is essentially a stochastic process having stationary and independent increments; we always assume that $X(0) = 0$ with probability 1. So:

- each $X(t) : \Omega \to \mathbb{R}^d$;
- given any selection of distinct time-points $0 \leq t_1 < t_2 < \cdots < t_n$, the random vectors $X(t_1), X(t_2) - X(t_1), X(t_3) - X(t_2), \ldots, X(t_n) - X(t_{n-1})$ are all independent;
- given any two distinct times $0 \leq s < t < \infty$, the probability distribution of $X(t) - X(s)$ coincides with that of $X(t - s)$.

The key formula in this book from which so much else flows, is the magnificent *Lévy–Khintchine formula*, which says that any Lévy process has a specific form for its characteristic function. More precisely, for all $t \geq 0$, $u \in \mathbb{R}^d$,

$$\mathbb{E}(e^{i(u,X(t))}) = e^{t\eta(u)} \tag{0.1}$$

where

$$\eta(u) = i(b,u) - \frac{1}{2}(u,au) + \int_{\mathbb{R}^d - \{0\}} \left[e^{i(u,y)} - 1 - i(u,y)\chi_{0<|y|<1}(y)\right]\nu(dy). \tag{0.2}$$

In this formula $b \in \mathbb{R}^d$, a is a positive definite symmetric $d \times d$ matrix and ν is a Lévy measure on $\mathbb{R}^d - \{0\}$, so that $\int_{\mathbb{R}^d - \{0\}} \min\{1, |y|^2\}\nu(dy) < \infty$. If you have not seen it before, (0.2) will look quite mysterious to you, so we need to try to extract its meaning.

First suppose that $a = \nu = 0$; then (0.1), just becomes $\mathbb{E}(e^{i(u,X(t))}) = e^{it(u,b)}$, so that $X(t) = bt$ is simply deterministic motion in a straight line. The vector b determines the velocity of this motion and is usually called the *drift*.

Now suppose that we also have $a \neq 0$, so that (0.1) takes the form $\mathbb{E}(e^{i(u,X(t))}) = \exp\{t[i(b,u) - \frac{1}{2}(u,au)]\}$. We can recognise this as the characteristic function of a Gaussian random variable $X(t)$ having mean vector tb and covariant matrix ta. In fact we can say more about this case: the process $(X(t), t \geq 0)$ is a *Brownian motion with drift*, and such processes have been extensively studied for over 100 years. In particular, the sample paths $t \to X(t)(\omega)$ are continuous (albeit nowhere differentiable) for almost all $\omega \in \Omega$. The case $b = 0$, $a = I$ is usually called *standard Brownian motion*.

Now consider the case where we also have $\nu \neq 0$. If ν is a finite measure we can rewrite (0.2) as

$$\eta(u) = i(b', u) - \tfrac{1}{2}(u, au) + \int_{\mathbb{R}^d - \{0\}} (e^{i(u,y)} - 1)\nu(dy),$$

where $b' = b - \int_{0<|y|<1} y\nu(dy)$. We will take the simplest possible form for ν, i.e. $\nu = \lambda\delta_h$ where $\lambda > 0$ and δ_h is a Dirac mass concentrated at $h \in \mathbb{R}^d - \{0\}$.

In this case we can set $X(t) = b't + \sqrt{a}B(t) + N(t)$, where $B = (B(t), t \geq 0)$ is a standard Brownian motion and $N = (N(t), t \geq 0)$ is an independent process for which

$$\mathbb{E}(e^{i(u,N(t))}) = \exp\left[\lambda t(e^{i(u,h)} - 1)\right].$$

We can now recognise N as a Poisson process of intensity λ taking values in the set $\{nh, n \in \mathbb{N}\}$, so that $P(N(t) = nh) = e^{-\lambda t}[(\lambda t)^n/n!]$ and $N(t)$ counts discrete events that occur at the random times $(T_n, n \in \mathbb{N})$. Our interpretation of the paths of X in this case is now as follows. X follows the path of a Brownian motion with drift from time zero until the random time T_1. At time T_1 the path has a jump discontinuity of size $|h|$. Between T_1 and T_2 we again see Brownian motion with drift, and there is another jump discontinuity of size $|h|$ at time T_2. We can continue to build the path in this manner indefinitely.

The next stage is to take $\nu = \sum_{i=1}^{m} \lambda_i \delta_{h_i}$, where $m \in \mathbb{N}$, $\lambda_i > 0$ and $h_i \in \mathbb{R}^d - \{0\}$, for $1 \leq i \leq m$. We can then write

$$X(t) = b't + \sqrt{a}B(t) + N_1(t) + \cdots + N_m(t),$$

where $N_1, \ldots, N_m$ are independent Poisson processes (which are also independent of B); each N_i has intensity λ_i and takes values in the set $\{nh_i, n \in \mathbb{N}\}$ where $1 \leq i \leq m$. In this case, the path of X is again a Brownian motion with drift, interspersed with jumps taking place at random times. This time, though, each jump size may be any of the m numbers $|h_1|, \ldots, |h_m|$.

In the general case where ν is finite, we can see that we have passed to the limit in which jump sizes take values in the full continuum of possibilities, corresponding to a continuum of Poisson processes. So a Lévy process of this type is a Brownian motion with drift interspersed with jumps of arbitrary size. Even when ν fails to be finite, if we have $\int_{0<|x|<1} |x|\nu(dx) < \infty$ a simple exercise in using the mean value theorem shows that we can still make this interpretation.

The most subtle case of the Lévy–Khintchine formula (0.2) is when $\int_{0<|x|<1} |x|\nu(dx) = \infty$ but $\int_{0<|x|<1} |x|^2\nu(dx) < \infty$. Thinking analytically, $e^{i(u,y)} - 1$ may no longer be ν-integrable but

$$e^{i(u,y)} - 1 - i(u, y)\chi_{0<|y|<1}(y)$$

always is. Intuitively, we may argue that the measure ν has become so fine that it is no longer capable of distinguishing small jumps from drift. Consequently it is necessary to amalgamate them together under the integral term. Despite this subtlety, it is still possible to interpret the general Lévy process as a Brownian motion with drift b interspersed with 'jumps' of arbitrary size, provided we recognise that at the microscopic level tiny jumps and short bursts of drift are treated as one. A more subtle discussion of this, and an account of the phenomenon of 'creep', can be found at the end of Section 2.4. We will see in Chapter 2 that the path can always be constructed as the limit of a sequence of terms, each of which is a Brownian motion with drift interspersed with bona fide jumps.

When $\nu < \infty$, we can write the sample-path decomposition directly as

$$X(t) = bt + \sqrt{a}B(t) + \sum_{0 \leq s \leq t} \Delta X(s), \tag{0.3}$$

where $\Delta X(s)$ is the jump at time s (e.g. if $\nu = \lambda\delta_h$ then $\Delta X(s) = 0$ or h). Instead of dealing directly with the jumps it is more convenient to count the times at which the jumps occur, so for each Borel set A in $\mathbb{R}^d - \{0\}$ and for each $t \geq 0$ we define

$$N(t, A) = \#\{0 \leq s \leq t; \Delta X(s) \in A\}.$$

This is an interesting object: if we fix t and A then $N(t, A)$ is a random variable; however, if we fix $\omega \in \Omega$ and $t \geq 0$ then $N(t, \cdot)(\omega)$ is a measure. Finally, if we fix A with $\nu(A) < \infty$ then $(N(t, A), t \geq 0)$ is a Poisson process with intensity $\nu(A)$.

When $\nu < \infty$, we can write

$$\sum_{0 \leq s \leq t} \Delta X(s) = \int_{\mathbb{R}-\{0\}} x N(t, dx).$$

(Readers might find it helpful to consider first the simple case where $\nu = \sum_{i=1}^{m} \lambda_i \delta_{h_i}$.)

In the case of general ν, the delicate analysis whereby small jumps and drift become amalgamated leads to the celebrated *Lévy–Itô decomposition*,

$$X(t) = bt + \sqrt{a}B(t) + \int_{0<|x|<1} x\big[N(t, dx) - t\nu(dx)\big] + \int_{|x|\geq 1} x N(t, dx).$$

Full proofs of the Lévy–Khintchine formula and the Lévy–Itô decomposition are given in Chapters 1 and 2.

Let us return to the consideration of standard Brownian motion $B = (B(t), t \geq 0)$. Each $B(t)$ has a Gaussian density

$$p_t(x) = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{2t}\right)$$

and, as was first pointed out by Einstein [106], this satisfies the diffusion equation

$$\frac{\partial p_t(x)}{\partial t} = \tfrac{1}{2}\Delta p_t(x),$$

where Δ is the usual Laplacian in $\mathbb{R}^d$. More generally, suppose that we want to build a solution $u = (u(t, x), t \geq 0, x \in \mathbb{R}^d)$ to the diffusion equation that has a fixed initial condition $u(0, x) = f(x)$ for all $x \in \mathbb{R}^d$, where f is a bounded continuous function on $\mathbb{R}^d$. We then have

$$u(t, x) = \int_{\mathbb{R}^d} f(x + y)p_t(y)dy = \mathbb{E}(f(x + B(t))). \tag{0.4}$$

The modern way of thinking about this utilises the powerful machinery of operator theory. We define $(T_t f)(x) = u(t, x)$; then $(T_t, t \geq 0)$ is a one-parameter semigroup of linear operators on the Banach space of bounded continuous functions. The semigroup is completely determined by its infinitesimal generator

Δ, so that we may formally write $T_t = e^{t\Delta}$ and note that, from the diffusion equation,

$$\Delta f = \frac{d}{dt}(T_t f)\Big|_{t=0}$$

for all f where this makes sense.

This circle of ideas has a nice physical interpretation. The semigroup or, equivalently, its infinitesimal version – the diffusion equation – gives a deterministic macroscopic description of the effects of Brownian motion. We see from (0.4) that to obtain this we must average over all possible paths of the particle that is executing Brownian motion. We can, of course, get a microscopic description by forgetting about the semigroup and just concentrating on the process $(B(t), t \geq 0)$. The price we have to pay for this is that we can no longer describe the dynamics deterministically. Each $B(t)$ is a random variable, and any statement we make about it can only be expressed as a probability. More generally, as we will see in Chapter 6, we have a dichotomy between solutions of stochastic differential equations, which are microscopic and random, and their averages, which solve partial differential equations and are macroscopic and deterministic.

The first stage in generalising this interplay of concepts is to replace Brownian motion by a general Lévy process $X = (X(t), t \geq 0)$. Although X may not in general have a density, we may still obtain the semigroup by $(T(t)f)(x) = \mathbb{E}(f(X(t) + x))$, and the infinitesimal generator then takes the more general form

$$\begin{aligned}(Af)(x) &= b^i(\partial_i f)(x) + \tfrac{1}{2}a^{ij}(\partial_i \partial_j f)(x) \\ &\quad + \int_{\mathbb{R}^d - \{0\}} [f(x+y) - f(x) - y^i(\partial_i f)(x)\chi_{0<|y|<1}(y)]\nu(dy).\end{aligned} \tag{0.5}$$

In fact this structure is completely determined by the Lévy–Khinchine formula, and we have the following important correspondences:

- drift $\longleftrightarrow$ first-order differential operator
- diffusion $\longleftrightarrow$ second-order differential operator
- jumps $\longleftrightarrow$ superposition of difference operators

This enables us to read off our intuitive description of the path from the form of the generator, and this is very useful in more general situations where we no longer have a Lévy–Khinchine formula. The formula (0.5) is established in Chapter 3, and we will also derive an alternative representation using pseudo-differential operators.

More generally, the relationship between stochastic processes and semigroups extends to a wider class of Markov processes $Y = (Y(t), t \geq 0)$, and here the semigroup is given by conditioning:

$$(T_t f)(x) = \mathbb{E}\big(f(Y(t))|Y(0) = x\big).$$

Under certain general conditions that we will describe in Chapter 3, the generator is of the Courrège form

$$\begin{aligned}(Af)(x) &= c(x)f(x) + b^i(x)(\partial_i f)(x) + a^{ij}(x)(\partial_i\partial_j f)(x)\\ &\quad + \int_{\mathbb{R}^d - \{x\}} [f(y) - f(x) - \phi(x,y)(y^i - x^i)(\partial_i f)(x)]\mu(x, dy).\end{aligned} \tag{0.6}$$

Note the similarities between equations (0.5) and (0.6). Once again there are drift, diffusion and jump terms, however, these are no longer fixed in space but change from point to point. There is an additional term, controlled by the function c, that corresponds to killing (we could also have included this in the Lévy case), and the function ϕ is simply a smoothed version of the indicator function that effects the cut-off between large and small jumps.

Under certain conditions, we can generalise the Lévy–Itô decomposition and describe the process Y as the solution of a stochastic differential equation

$$\begin{aligned}dY(t) &= b(Y(t-))dt + \sqrt{a(Y(t-))}dB(t)\\ &\quad + \int_{|x|<1} F(Y(t-), x)\big[N(dt, dx) - dt\nu(dx)\big]\\ &\quad + \int_{|x|\geq 1} G(Y(t-), x)N(dt, dx).\end{aligned} \tag{0.7}$$

The kernel $\mu(x, \cdot)$ appearing in (0.6) can be expressed in terms of the Lévy measure ν and the coefficients F and G. This is described in detail in Chapter 6.

To make sense of the stochastic differential equation (0.7), we must rewrite it as an integral equation, which means that we must give meaning to *stochastic* *integrals* such as

$$\int_0^t U(s)dB(s) \quad \text{and} \quad \int_0^t \int_{0<|x|<1} V(s,x)(N(ds,dx) - ds\nu(dx))$$

for suitable U and V. The usual Riemann–Stieltjes or Lebesgue–Stieltjes approach no longer works for these objects, and we need to introduce some extra structure. To model the flow of information with time, we introduce a filtration $(\mathcal{F}_t, t \geq 0)$ that is an increasing family of sub-σ-algebras of $\mathcal{F}$, and we say that a process U is adapted if each $U(t)$ is $\mathcal{F}_t$-measurable for each $t \geq 0$. We then define

$$\int_0^t U(s)dB(s) = \lim_{n\to\infty} \sum_{j=1}^{m_n} U\left(t_j^{(n)}\right)\left[B\left(t_{j+1}^{(n)}\right) - B\left(t_j^{(n)}\right)\right]$$

where $0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{m_n}^{(n)} = t$ is a sequence of partitions of $[0,t]$ whose mesh tends to zero as $n \to \infty$. The key point in the definition is that for each term in the summand, $U(t_j^{(n)})$ is fixed in the past while the increment $B(t_{j+1}^{(n)}) - B(t_j^{(n)})$ extends into the future. If a Riemann–Stieltjes theory were possible, we could evaluate $U(x_j^{(n)})$ at an arbitrary point for which $t_j^{(n)} < x_j^{(n)} < t_{j+1}^{(n)}$. The other integral,

$$\int_0^t \int_{0<|x|<1} V(s,x)\big[N(ds,dx) - ds\nu(dx)\big],$$

is defined similarly.

This definition of a stochastic integral has profound implications. In Chapter 4, we will explore the properties of a class of Lévy-type stochastic integrals that take the form

$$Y(t) = \int_0^t G(s)ds + \int_0^t F(s)dB(s) + \int_0^t \int_{0<|x|<1} H(s,x)\big[N(ds,dx) - ds\nu(dx)\big] + \int_0^t K(s,x)N(ds,dx)$$

and, for convenience, we will take $d = 1$ for now. In the case where F, H and K are identically zero and f is a differentiable function, the chain rule from differential calculus gives $f(Y(t)) = \int_0^t f'(Y(s))G(s)ds$, which we can write more succinctly as $df(Y(t)) = f'(Y(t))G(t)dt$. This formula breaks down for

Lévy-type stochastic integrals, and in its place we get the famous *Itô formula*,

$$
\begin{aligned}
& df(Y(t)) \\
& \quad = f'(Y(t))G(t)dt + f'(Y(t))F(t)dB(t) + \tfrac{1}{2}f''(Y(t))F(t)^2 dt \\
& \qquad + \int_{|x|\geq 1} [f(Y(t-) + K(t,x)) - f(Y(t-))]N(dt,dx) \\
& \qquad + \int_{0<|x|<1} [f(Y(t-) + H(t,x)) - f(Y(t-))](N(dt,dx) - \nu(dx)dt) \\
& \qquad + \int_{0<|x|<1} \Big[f(Y(t-) + H(t,x)) - f(Y(t-)) \\
& \qquad - H(t,x)f'(Y(t-))\Big]\nu(dx)dt.
\end{aligned}
$$

If you have not seen this before, think of a Taylor series expansion in which $dB(t)^2$ behaves like dt and $N(dt,dx)^2$ behaves like $N(dt,dx)$. Alternatively, you can wait for the full development in Chapter 4. Itô's formula is the key to the wonderful world of stochastic calculus. It lies behind the extraction of the Courrège generator (0.6) from equation (0.7). It also has many important applications including option pricing, the Black–Scholes formula and attempts to replace the latter using more realistic models based on Lévy processes. This is all revealed in Chapter 5, but now the preview is at an end and it is time to begin the journey . . .

Notation

Throughout this book, we will deal extensively with random variables taking values in the Euclidean space $\mathbb{R}^d$, where $d \in \mathbb{N}$. We recall that elements of $\mathbb{R}^d$ are vectors $x = (x_1, x_2, \ldots, x_d)$ with each $x_i \in \mathbb{R}$ for $1 \leq i \leq d$. The inner product in $\mathbb{R}^d$ is denoted by (x, y) where $x, y \in \mathbb{R}^d$, so that

$$(x, y) = \sum_{i=1}^{d} x_i y_i.$$

This induces the Euclidean norm $|x| = (x, x)^{1/2} = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}$. We will use the Einstein summation convention throughout this book, wherein summation is understood with respect to repeated upper and lower indices, so for example if $x, y \in \mathbb{R}^d$ and $A = (A^i_j)$ is a $d \times d$ matrix then

$$A^i_j x_i y^j = \sum_{i,j=1}^{d} A^i_j x_i y^j = (x, Ay).$$

We say that such a matrix is *positive definite* if $(x, Ax) \geq 0$ for all $x \in \mathbb{R}^d$ and *strictly positive definite* if the inequality can be strengthened to $(x, Ax) > 0$ for all $x \in \mathbb{R}^d$, with $x \neq 0$ (note that some authors call these ‘non-negative definite’ and ‘positive definite’, respectively). The transpose of a matrix A will always be denoted A^{T}. The determinant of a square matrix is written as $\det(A)$ and its trace as $\mathrm{tr}(A)$. The identity matrix will always be denoted I.
The set of all $d \times d$ real-valued matrices is denoted $M_d(\mathbb{R})$.

If $S \subseteq \mathbb{R}^d$ then its orthogonal complement is $S^{\perp} = \{x \in \mathbb{R}^d; (x, y) = 0$ for all $y \in S\}$.

The open ball of radius r centred at x in $\mathbb{R}^d$ is denoted $B_r(x) = \{y \in \mathbb{R}^d; |y - x| < r\}$ and we will always write $\hat{B} = B_1(0)$. The sphere in $\mathbb{R}^d$ is the

Cambridge University Press
978-0-521-73865-1 - Lévy Processes and Stochastic Calculus: Second Edition
David Applebaum
Frontmatter
More information

$(d-1)$-dimensional submanifold, denoted S^{d-1}, defined by $S^{d-1} = \{x \in \mathbb{R}^d; |x| = 1\}$.

We sometimes write $\mathbb{R}^+ = [0, \infty)$.

The sign of $u \in \mathbb{R}$ is denoted $\text{sgn}(u)$ so that $\text{sgn}(u) = (u/|u|)$ if $u \neq 0$, with $\text{sgn}(0) = 0$.

For $z \in \mathbb{C}$, $\Re(z)$ and $\Im(z)$ denote the real and imaginary parts of z, respectively.

The complement of a set A will always be denoted A^c and $\bar{A}$ will mean closure in some topology. If f is a mapping between two sets A and B, we denote its range as $\text{Ran}(f) = \{y \in B;\ y = f(x) \text{ for some } x \in A\}$.

For $1 \leq n \leq \infty$, we write $C^n(\mathbb{R}^d)$ to denote the set of all n-times differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$, all of whose derivatives are continuous. The jth first-order partial derivative of $f \in C^1(\mathbb{R}^d)$ at $x \in \mathbb{R}^d$ will sometimes be written $(\partial_j f)(x)$. Similarly, if $f \in C^2(\mathbb{R}^d)$, we write

$$(\partial_i \partial_j f)(x) \quad \text{for} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

When $d = 1$ and $f \in C^n(\mathbb{R})$, we sometimes write

$$f^{(r)}(x) \quad \text{for} \quad \frac{d^r f}{dx^r}(x),$$

where $1 \leq r \leq n$.

Let $\mathcal{H}$ be a real inner product space, equipped with the inner product $\langle \cdot, \cdot \rangle$ and associated norm $||x|| = \langle x, x \rangle^{1/2}$, for each $x \in \mathcal{H}$. We will frequently have occasion to use the *polarisation identity*

$$\langle x, y \rangle = \tfrac{1}{4}(||x+y||^2 - ||x-y||^2),$$

for each $x, y \in \mathcal{H}$.

For $a, b \in \mathbb{R}$, we will use $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$.

We will occasionally use Landau notation, according to which $(o(n), n \in \mathbb{N})$ is any real-valued sequence for which $\lim_{n\to\infty}(o(n)/n) = 0$ and $(O(n), n \in \mathbb{N})$ is any non-negative sequence for which $\limsup_{n\to\infty}(O(n)/n) < \infty$. Functions $o(t)$ and $O(t)$ are defined similarly. If $f, g : \mathbb{R} \to \mathbb{R}$ and $a \in \mathbb{R} \cup \{\infty\}$, then by $f \sim g$ as $x \to a$ we mean $\lim_{x\to a}\left[f(x)/g(x)\right] = 1$.

If $f : \mathbb{R}^d \to \mathbb{R}$ then by $\lim_{s\uparrow t} f(s) = l$ we mean $\lim_{s\to t,\, s<t} f(s) = l$. Similarly, $\lim_{s\downarrow t} f(s) = l$ means $\lim_{s\to t,\, s>t} f(s) = l$.