Stochastic Mechanics
Random Media
Signal Processing and Image Synthesis
Mathematical Economics and Finance
Stochastic Optimization
Stochastic Control
Stochastic Models in Life Sciences

Stochastic Modelling and Applied Probability

(Formerly: Applications of Mathematics)

21

Stochastic Modelling and Applied Probability

formerly entitled Applications of Mathematics

---

1 Fleming/Rishel, **Deterministic and Stochastic Optimal Control** (1975)
2 Marchuk, **Methods of Numerical Mathematics,** Second Edition (1982)
3 Balakrishnan, **Applied Functional Analysis,** Second Edition (1981)
4 Borovkov, **Stochastic Processes in Queueing Theory** (1976)
5 Liptser/Shiryaev, **Statistics of Random Processes I: General Theory** (2000) first ed. 1977
6 Liptser/Shiryaev, **Statistics of Random Processes II: Applications** (2000) first ed. 1978
7 Vorob'ev, **Game Theory: Lectures for Economists and Systems Scientists** (1977)
8 Shiryaev, **Optimal Stopping Rules** (1978)
9 Ibragimov/Rozanov, **Gaussian Random Processes** (1978)
10 Wonham, **Linear Multivariable Control: A Geometric Approach,** Third Edition (1985)
11 Hida, **Brownian Motion** (1980)
12 Hestenes, **Conjugate Direction Methods in Optimization** (1980)
13 Kallianpur, **Stochastic Filtering Theory** (1980)
14 Krylov, **Controlled Diffusion Processes** (1980)
15 Prabhu, **Stochastic Storage Processes: Queues, Insurance Risk, and Dams** (1980)
16 Ibragimov/Has'minskii, **Statistical Estimation: Asymptotic Theory** (1981)
17 Cesari, **Optimization: Theory and Applications** (1982)
18 Elliott, **Stochastic Calculus and Applications** (1982)
19 Marchuk/Shaidourov, **Difference Methods and Their Extrapolations** (1983)
20 Hijab, **Stabilization of Control Systems** (1986)
21 Protter, **Stochastic Integration and Differential Equations** (1990)
22 Benveniste/Métivier/Priouret, **Adaptive Algorithms and Stochastic Approximations** (1990)
23 Kloeden/Platen, **Numerical Solution of Stochastic Differential Equations** (1992) corrected 3rd printing 1999
24 Kushner/Dupuis, **Numerical Methods for Stochastic Control Problems in Continuous Time** (1992)
25 Fleming/Soner, **Controlled Markov Processes and Viscosity Solutions** (1993)
26 Baccelli/Brémaud, **Elements of Queueing Theory** (1994)
27 Winkler, **Image Analysis, Random Fields and Dynamic Monte Carlo Methods** (1995)
28 Kalpazidou, **Cycle Representations of Markov Processes** (1995)
29 Elliott/Aggoun/Moore, **Hidden Markov Models: Estimation and Control** (1995)
30 Hernández-Lerma/Lasserre, **Discrete-Time Markov Control Processes** (1995)
31 Devroye/Györfi/Lugosi, **A Probabilistic Theory of Pattern Recognition** (1996)
32 Maitra/Sudderth, **Discrete Gambling and Stochastic Games** (1996)
33 Embrechts/Klüppelberg/Mikosch, **Modelling Extremal Events for Insurance and Finance** (1997) corrected 2nd printing 1999
34 Duflo, **Random Iterative Models** (1997)
35 Kushner/Yin, **Stochastic Approximation Algorithms and Applications** (1997)
36 Musiela/Rutkowski, **Martingale Methods in Financial Modelling** (1997, 2nd. ed. 2005)
37 Yin, **Continuous-Time Markov Chains and Applications** (1998)
38 Dembo/Zeitouni, **Large Deviations Techniques and Applications** (1998)
39 Karatzas, **Methods of Mathematical Finance** (1998)
40 Fayolle/Iasnogorodski/Malyshev, **Random Walks in the Quarter-Plane** (1999)
41 Aves/Jensen, **Stochastic Models in Reliability** (1999)
42 Hernandez-Lerma/Lasserre, **Further Topics on Discrete-Time Markov Control Processes** (1999)

(continued after index)

Philip E. Protter

Stochastic Integration and Differential Equations

Second Edition, Version 2.1

Springer

Author

Philip E. Protter

Cornell University
School of Operations Research and Industrial Engineering
Rhodes Hall
14853 Ithaca, NY, USA
e-mail: pep4@cornell.edu

Managing Editors

B. Rozovskii
Denney Research Building 308
Center for Applied Mathematical Sciences
University of Southern California
1042 West Thirty-Sixth Place
Los Angeles, CA 90089, USA
rozovski@math.usc.edu

M. Yor
Laboratoire de Probabilités et Modèles Aléatoires
Université de Paris VI
175, rue du Chevaleret
75013 Paris, France

Cover illustration: Cover pattern courtesy of Rick Durrett, Cornell University, Ithaca, New York.

Mathematics Subject Classification (2000):
Primary: 60H05, 60H10, 60H20
Secondary: 60G07, 60G17, 60G44, 60G51

2nd Edition. Corrected 3rd Printing 2005

Library of Congress Cataloging-in-Publication Data

Protter, Philip E.
Stochastic integration and differential equations / Philip E. Protter.-- 2nd ed.
p. cm. -- (Applications of mathematics, ISSN 0172-4568 ; 21)
Includes bibliographical references and indexes.
ISBN 978-3-642-05560-7 ISBN 978-3-662-10061-5 (eBook)
DOI 10.1007/978-3-662-10061-5
1. Stochastic integrals. 2. Martingales (Mathematics) 3. Stochastic differential equations. I. Title. II. Series.

QA274.22.P76 2004
519.2--dc22

2003059169

ISSN 0172-4568
ISBN 978-3-642-05560-7

springeronline.com

Originally published by Springer Berlin Heidelberg New York in 2005
Softcover reprint of the hardcover 2nd edition 2005

Cover Design: Erich Kirchner, Heidelberg
Typesetting by the Author using a Springer TEX macro package
Printed on acid-free paper 41/3142-543210

To Diane and Rachel

Preface to the Second Edition

It has been thirteen years since the first edition was published, with its subtitle "a new approach." While the book has had some success, there are still almost no other books that use the same approach. (See however the recent book by K. Bichteler [17].) There are nevertheless of course other extant books, many of them quite good, although the majority still are devoted primarily to the case of continuous sample paths, and others treat stochastic integration as one of many topics. Examples of alternative texts which have appeared since the first edition of this book are: [34], [46], [90], [115], [202], [196], [224], [235], and [248]. While the subject has not changed much, there have been new developments, and subjects we thought unimportant in 1990 and did not include, we now think important enough either to include or to expand in this book.

The most obvious changes in this edition are that we have added exercises at the end of each chapter, and we have also added Chap. VI which introduces the expansion of filtrations. However we have also completely rewritten Chap. III. In the first edition we followed an elementary approach which was P. A. Meyer's original approach before the methods of Doléans-Dade. In order to remain friends with Freddy Delbaen, and also because we now agree with him, we have instead used the modern approach of predictability rather than naturality. However we benefited from the new proof of the Doob-Meyer Theorem due to R. Bass, which ultimately uses only Doob's quadratic martingale inequality, and in passing reveals the role played by totally inaccessible stopping times. The treatment of Girsanov's theorem now includes the case where the two probability measures are not necessarily equivalent, and we include the Kazamaki-Novikov theorems. We have also added a section on compensators, with examples. In Chap. IV we have expanded our treatment of martingale representation to include the Jacod-Yor Theorem, and this has allowed us to use the Emery-Azéma martingales as a class of examples of martingales with the martingale representation property. Also, largely because of the Delbaen-Schachermayer theory of the fundamental theorems of mathematical finance, we have included the topic of sigma martingales. In Chap. V we added a section which includes some useful results about the solutions of stochastic differential equations, inspired by the review of the first edition by E. Pardoux [207]. We have also made small changes throughout the book; for instance we have included specific examples of Lévy processes and their corresponding Lévy measures, in Sect. 4 of Chap. I.

The exercises are gathered at the end of the chapters, in no particular order. Some of the (presumed) harder problems we have designated with a star (*), and occasionally we have used two stars (**). While of course many of the problems are of our own creation, a significant number are theorems or lemmas taken from research papers, or taken from other books. We do not attempt to ascribe credit, other than listing the sources in the bibliography, primarily because they have been gathered over the past decade and often we don't remember from where they came. We have tried systematically to refrain from relegating a needed lemma as an exercise; thus in that sense the exercises are independent from the text, and (we hope) serve primarily to illustrate the concepts and possible applications of the theorems.[1]

Last, we have the pleasant task of thanking the numerous people who helped with this book, either by suggesting improvements, finding typos and mistakes, alerting me to references, or by reading chapters and making comments. We wish to thank patient students both at Purdue University and Cornell University who have been subjected to preliminary versions over the years, and the following individuals: C. Beneš, R. Cont, F. Diener, M. Diener, R. Durrett, T. Fujiwara, K. Giesecke, L. Goldberg, R. Haboush, J. Jacod, H. Kraft, K. Lee, J. Ma, J. Mitro, J. Rodriguez, K. Schürger, D. Sezer, J. A. Trujillo Ferreras, R. Williams, M. Yor, and Yong Zeng. Th. Jeulin, K. Shimbo, and Yan Zeng gave extraordinary help, and my editor C. Byrne gives advice and has patience that is impressive. Over the last decade I have learned much from many discussions with Darrell Duffie, Jean Jacod, Tom Kurtz, and Denis Talay, and this no doubt is reflected in this new edition. Finally, I wish to give a special thanks to M. Kozdron who hastened the appearance of this book through his superb help with LaTeX, as well as his own advice on all aspects of the book.

This postscript concerns the Corrected Second Edition. Since the appearance of the second edition, Marc Yor has read the book with care and made many suggestions which have been incorporated in this corrected edition. Many are subtle, but without doubt the reader will benefit greatly from them, and we wish to thank him for this gift. I am also grateful for help received from others, including K. Asrat, K. Shimbo, and Y. Zeng.

Ithaca, NY | *Philip Protter*
February 2005

[1] Solutions of some of the exercises are posted on the author's web page, URL http://www.orie.cornell.edu/~protter/books.html (July, 2004).

Preface to the First Edition

The idea of this book began with an invitation to give a course at the Third Chilean Winter School in Probability and Statistics, at Santiago de Chile, in July, 1984. Faced with the problem of teaching stochastic integration in only a few weeks, I realized that the work of C. Dellacherie [44] provided an outline for just such a pedagogic approach. I developed this into a series of lectures (Protter [217]), using the work of K. Bichteler [16], E. Lenglart [158] and P. Protter [218], as well as that of Dellacherie. I then taught from these lecture notes, expanding and improving them, in courses at Purdue University, the University of Wisconsin at Madison, and the University of Rouen in France. I take this opportunity to thank these institutions and Professor Rolando Rebolledo for my initial invitation to Chile.

This book assumes the reader has some knowledge of the theory of stochastic processes, including elementary martingale theory. While we have recalled the few necessary martingale theorems in Chap. I, we have not provided proofs, as there are already many excellent treatments of martingale theory readily available (e.g., Breiman [25], Dellacherie-Meyer [47, 48], or Ethier-Kurtz [74]). There are several other texts on stochastic integration, all of which adopt to some extent the usual approach and thus require the general theory. The books of Elliott [66], Kopp [138], Métivier [174], Rogers-Williams [226] and to a much lesser extent Letta [162] are examples. The books of McKean [169], Chung-Williams [34], and Karatzas-Shreve [129] avoid the general theory by limiting their scope to Brownian motion (McKean) and to continuous semimartingales.

Our hope is that this book will allow a rapid introduction to some of the deepest theorems of the subject, without first having to be burdened with the beautiful but highly technical "general theory of processes."

Many people have aided in the writing of this book, either through discussions or by reading one of the versions of the manuscript. I would like to thank J. Azema, M. Barlow, A. Bose, M. Brown, C. Constantini, C. Dellacherie, D. Duffie, M. Emery, N. Falkner, E. Goggin, D. Gottlieb, A. Gut, S. He, J. Jacod, T. Kurtz, J. de Sam Lazaro, R. Léandre, E. Lenglart, G. Letta,

S. Levantal, P. A. Meyer, E. Pardoux, H. Rubin, T. Sellke, R. Stockbridge, C. Stricker, P. Sundar, and M. Yor. I would especially like to thank J. San Martin for his careful reading of the manuscript in several of its versions.

Svante Janson read the entire manuscript in several versions, giving me support, encouragement, and wonderful suggestions, all of which improved the book. He also found, and helped to correct, several errors. I am extremely grateful to him, especially for his enthusiasm and generosity.

The National Science Foundation provided partial support throughout the writing of this book.

I wish to thank Judy Snider for her cheerful and excellent typing of several versions of this book.

Philip Protter

Contents

Introduction

In this book we present a new approach to the theory of modern stochastic integration. The novelty is that we define a semimartingale as a stochastic process which is a "good integrator" on an elementary class of processes, rather than as a process that can be written as the sum of a local martingale and an adapted process with paths of finite variation on compacts: This approach has the advantage over the customary approach of not requiring a close analysis of the structure of martingales as a prerequisite. This is a significant advantage because such an analysis of martingales itself requires a highly technical body of knowledge known as "the general theory of processes." Our approach has a further advantage of giving traditionally difficult and non-intuitive theorems (such as Stricker's Theorem) transparently simple proofs. We have tried to capitalize on the natural advantage of our approach by systematically choosing the simplest, least technical proofs and presentations. As an example we have used K. M. Rao's proofs of the Doob-Meyer decomposition theorems in Chap. III, rather than the more abstract but less intuitive Doléans-Dade measure approach.

In Chap. I we present preliminaries, including the Poisson process, Brownian motion, and Lévy processes. Naturally our treatment presents those properties of these processes that are germane to stochastic integration.

In Chap. II we define a semimartingale as a good integrator and establish many of its properties and give examples. By restricting the class of integrands to adapted processes having left continuous paths with right limits, we are able to give an intuitive Riemann-type definition of the stochastic integral as the limit of sums. This is sufficient to prove many theorems (and treat many applications) including a change of variables formula ("Itô's formula").

Chapter III is devoted to developing a minimal amount of "general theory" in order to prove the Bichteler-Dellacherie Theorem, which shows that our "good integrator" definition of a semimartingale is equivalent to the usual one as a process X having a decomposition $X = M + A$, into the sum of a local martingale M and an adapted process A having paths of finite variation on compacts. Nevertheless most of the theorems covered en route (Doob-

Meyer, Meyer-Girsanov) are themselves key results in the theory. The core of the whole treatment is the Doob-Meyer decomposition theorem. We have followed the relatively recent proof due to R. Bass, which is especially simple for the case where the martingale jumps only at totally inaccessible stopping times, and in all cases uses no mathematical tool deeper than Doob's quadratic martingale inequality. This allows us to avoid the detailed treatment of natural processes which was ubiquitous in the first edition, although we still use natural processes from time to time, as they do simplify some proofs.

Using the results of Chap. III we extend the stochastic integral by continuity to predictable integrands in Chap. IV, thus making the stochastic integral a Lebesgue-type integral. We use predictable integrands to develop a theory of martingale representation. The theory we develop is an L^2 theory, but we also prove that the dual of the martingale space $\mathcal{H}^1$ is BMO and then prove the Jacod-Yor Theorem on martingale representation, which in turn allows us to present a class of examples having both jumps and martingale representation. We also use predictable integrands to give a presentation of semimartingale local times.

Chapter V serves as an introduction to the enormous subject of stochastic differential equations. We present theorems on the existence and uniqueness of solutions as well as stability results. Fisk-Stratonovich equations are presented, as well as the Markov nature of the solutions when the differentials have Markov-type properties. The last part of the chapter is an introduction to the theory of flows, followed by moment estimates on the solutions, and other minor but useful results. Throughout Chap. V we have tried to achieve a balance between maximum generality and the simplicity of the proofs.

Chapter VI provides an introduction to the theory of the expansion of filtrations (known as "grossissements de filtrations" in the French literature). We present first a theory of initial expansions, which includes Jacod's Theorem. Jacod's Theorem gives a sufficient condition for semimartingales to remain semimartingales in the expanded filtration. We next present the more difficult theory of progressive expansion, which involves expanding filtrations to turn a random time into a stopping time, then analyzing what happens to the semimartingales of the first filtration when considered in the expanded filtration. Last, we give an application of these ideas to time reversal.

I

Preliminaries

1 Basic Definitions and Notation

We assume as given a complete probability space $(\Omega, \mathcal{F}, P)$. In addition we are given a *filtration* $(\mathcal{F}_t)_{0 \leq t \leq \infty}$. By a filtration we mean a family of σ-algebras $(\mathcal{F}_t)_{0 \leq t \leq \infty}$ that is increasing, i.e., $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$. For convenience, we will usually write $\mathbb{F}$ for the filtration $(\mathcal{F}_t)_{0 \leq t \leq \infty}$.

Definition. A filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is said to satisfy the **usual hypotheses** if

(i) $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$;
(ii) $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$, all t, $0 \leq t < \infty$; that is, the filtration $\mathbb{F}$ is *right continuous*.

We always assume that the usual hypotheses hold.

Definition. A random variable $T : \Omega \to [0, \infty]$ is a **stopping time** if the event $\{T \leq t\} \in \mathcal{F}_t$, every t, $0 \leq t \leq \infty$.

One important consequence of the right continuity of the filtration is the following theorem.

Theorem 1. *The event $\{T < t\} \in \mathcal{F}_t$, $0 \leq t \leq \infty$, if and only if T is a stopping time.*

Proof. Since $\{T \leq t\} = \bigcap_{t+\varepsilon>u>t}\{T < u\}$, any $\varepsilon > 0$, we have $\{T \leq t\} \in \bigcap_{u>t} \mathcal{F}_u = \mathcal{F}_t$, so T is a stopping time. For the converse, $\{T < t\} = \bigcup_{t>\varepsilon>0}\{T \leq t - \varepsilon\}$, and $\{T \leq t - \varepsilon\} \in \mathcal{F}_{t-\varepsilon}$, hence also in $\mathcal{F}_t$. $\square$

A **stochastic process** X on $(\Omega, \mathcal{F}, P)$ is a collection of $\mathbb{R}$-valued or $\mathbb{R}^d$-valued random variables $(X_t)_{0 \leq t < \infty}$. The process X is said to be **adapted** if $X_t \in \mathcal{F}_t$ (that is, is $\mathcal{F}_t$ measurable) for each t. We must take care to be precise about the concept of equality of two stochastic processes.

Definition. Two stochastic processes X and Y are **modifications** if $X_t = Y_t$ a.s., each t. Two processes X and Y are **indistinguishable** if a.s., for all t, $X_t = Y_t$.

If X and Y are *modifications* there exists a null set, N_t, such that if $\omega \notin N_t$, then $X_t(\omega) = Y_t(\omega)$. The null set N_t depends on t. Since the interval $[0,\infty)$ is uncountable the set $N = \bigcup_{0 \le t < \infty} N_t$ could have any probability between 0 and 1, and it could even be non-measurable. If X and Y are *indistinguishable*, however, then there exists one null set N such that if $\omega \notin N$, then $X_t(\omega) = Y_t(\omega)$, for all t. In other words, the functions $t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ are the same for all $\omega \notin N$, where $P(N) = 0$. The set N is in $\mathcal{F}_t$, all t, since $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$. The functions $t \mapsto X_t(\omega)$ mapping $[0,\infty)$ into $\mathbb{R}$ are called the **sample paths** of the stochastic process X.

Definition. A stochastic process X is said to be **càdlàg** if it a.s. has sample paths which are right continuous, with left limits. Similarly, a stochastic process X is said to be **càglàd** if it a.s. has sample paths which are left continuous, with right limits. (The nonsensical words *càdlàg* and *càglàd* are acronyms from the French for *continu à droite, limité à gauche* and *continu à gauche, limité à droite*, respectively.)

Theorem 2. *Let X and Y be two stochastic processes, with X a modification of Y. If X and Y have right continuous paths a.s., then X and Y are indistinguishable.*

Proof. Let A be the null set where the paths of X are not right continuous, and let B be the analogous set for Y. Let $N_t = \{\omega : X_t(\omega) \neq Y_t(\omega)\}$, and let $N = \bigcup_{t \in \mathbb{Q}} N_t$, where $\mathbb{Q}$ denotes the rationals in $[0,\infty)$. Then $P(N) = 0$. Let $M = A \cup B \cup N$, and $P(M) = 0$. We have $X_t(\omega) = Y_t(\omega)$ for all $t \in \mathbb{Q}$, $\omega \notin M$. If t is not rational, let t_n decrease to t through $\mathbb{Q}$. For $\omega \notin M$, $X_{t_n}(\omega) = Y_{t_n}(\omega)$, each n, and $X_t(\omega) = \lim_{n\to\infty} X_{t_n}(\omega) = \lim_{n\to\infty} Y_{t_n}(\omega) = Y_t(\omega)$. Since $P(M) = 0$, X and Y are indistinguishable. $\square$

Corollary. Let X and Y be two stochastic processes which are càdlàg. If X is a modification of Y, then X and Y are indistinguishable.

Càdlàg processes provide natural examples of stopping times.

Definition. Let X be a stochastic process and let Λ be a Borel set in $\mathbb{R}$. Define

$$T(\omega) = \inf\{t > 0 : X_t \in \Lambda\}.$$

Then T is called the **hitting time** of Λ for X.

Theorem 3. *Let X be an adapted càdlàg stochastic process, and let Λ be an open set. Then the hitting time of Λ is a stopping time.*

Proof. By Theorem 1 it suffices to show that $\{T < t\} \in \mathcal{F}_t$, $0 \le t < \infty$. But

$$\{T < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t)} \{X_s \in \Lambda\},$$

since Λ is open and X has right continuous paths. Since $\{X_s \in \Lambda\} = X_s^{-1}(\Lambda) \in \mathcal{F}_s$, the result follows. $\square$

Theorem 4. *Let X be an adapted càdlàg stochastic process, and let Λ be a closed set. Then the random variable*

$$T(\omega) = \inf\{t > 0 : X_t(\omega) \in \Lambda \ \text{ or } \ X_{t-}(\omega) \in \Lambda\}$$

is a stopping time.

Proof. By $X_{t-}(\omega)$ we mean $\lim_{s\to t,\, s<t} X_s(\omega)$. Let $A_n = \{x : d(x, \Lambda) < 1/n\}$, where $d(x, \Lambda)$ denotes the distance from a point x to Λ. Then A_n is an open set and

$$\{T \le t\} = \{X_t \in \Lambda \text{ or } \ X_{t-} \in \Lambda\} \cup \{\bigcap_n \bigcup_{s \in \mathbb{Q}\cap[0,t)} \{X_s \in A_n\}\}. \qquad \square$$

It is a very deep result that the hitting time of a *Borel set* is a stopping time. We do not have need of this result.

The next theorem collects elementary facts about stopping times; we leave the proof to the reader.

Theorem 5. *Let S, T be stopping times. Then the following are stopping times:*

(i) $S \wedge T = \min(S, T)$;
(ii) $S \vee T = \max(S, T)$;
(iii) $S + T$;
(iv) αS, *where* $\alpha > 1$.

The σ-algebra $\mathcal{F}_t$ can be thought of as representing all (theoretically) observable events up to and including time t. We would like to have an analogous notion of events that are observable before a random time.

Definition. Let T be a stopping time. The **stopping time σ-algebra $\mathcal{F}_T$** is defined to be

$$\{\Lambda \in \mathcal{F} : \Lambda \cap \{T \le t\} \in \mathcal{F}_t, \text{ all } t \ge 0\}.$$

The previous definition is not especially intuitive. However it does well represent "knowledge" up to time T, as the next theorem illustrates.

Theorem 6. *Let T be a finite stopping time. Then $\mathcal{F}_T$ is the smallest σ-algebra containing all càdlàg adapted processes sampled at T. That is,*

$$\mathcal{F}_T = \sigma\{X_T; X \text{ all adapted càdlàg processes}\}.$$

Proof. Let $\mathcal{G} = \sigma\{X_T; X$ all adapted càdlàg processes$\}$. Let $\Lambda \in \mathcal{F}_T$. Then $X_t = 1_\Lambda 1_{\{t \ge T\}}$[1] is a càdlàg process, and $X_T = 1_\Lambda$. Hence $\Lambda \in \mathcal{G}$, and $\mathcal{F}_T \subset \mathcal{G}$.

---

[1] 1_A is the indicator function of $A : 1_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$

Next let X be an adapted càdlàg process. We need to show X_T is $\mathcal{F}_T$ measurable. Consider $X(s,\omega)$ as a function from $[0,\infty)\times\Omega$ into $\mathbb{R}$. Construct $\varphi : \{T\le t\}\to[0,\infty)\times\Omega$ by $\varphi(\omega)=(T(\omega),\omega)$. Then since X is adapted and càdlàg, we have $X_T = X\circ\varphi$ is a measurable mapping from $(\{T\le t\},\mathcal{F}_t\cap\{T\le t\})$ into $(\mathbb{R},\mathcal{B})$, where $\mathcal{B}$ are the Borel sets of $\mathbb{R}$. Therefore

$$\{\omega : X(T(\omega),\omega)\in B\}\cap\{T\le t\}$$

is in $\mathcal{F}_t$, and this implies $X_T\in\mathcal{F}_T$. Therefore $\mathcal{G}\subset\mathcal{F}_T$. □

We leave it to the reader to check that if $S\le T$ a.s., then $\mathcal{F}_S\subset\mathcal{F}_T$, and the less obvious (and less important) fact that $\mathcal{F}_S\cap\mathcal{F}_T=\mathcal{F}_{S\wedge T}$.

If X and Y are càdlàg, then $X_t=Y_t$ a.s. each t implies that X and Y are indistinguishable, as we have already noted. Since fixed times are stopping times, obviously if $X_T=Y_T$ a.s. for each finite stopping time T, then X and Y are indistinguishable. If X is càdlàg, let ΔX denote the process $\Delta X_t = X_t - X_{t-}$. Then ΔX is not càdlàg, though it is adapted and for a.a. ω, $t\mapsto\Delta X_t=0$ except for at most countably many t. We record here a useful result.

Theorem 7. *Let X be adapted and càdlàg. If $\Delta X_T 1_{\{T<\infty\}}=0$ a.s. for each stopping time T, then ΔX is indistinguishable from the zero process.*

Proof. It suffices to prove the result on $[0,t_0]$ for $0<t_0<\infty$. The set $\{t : |\Delta X_t|>0\}$ is countable a.s. since X is càdlàg. Moreover

$$\{t : |\Delta X_t|>0\}=\bigcup_{n=1}^{\infty}\{t : |\Delta X_t|>\frac{1}{n}\}$$

and the set $\{t : |\Delta X_t|>1/n\}$ must be finite for each n, since $t_0<\infty$. Using Theorem 4 we define stopping times for each n inductively as follows:

$$T^{n,1}=\inf\{t>0 : |\Delta X_t|>\frac{1}{n}\}$$
$$T^{n,k}=\inf\{t>T^{n,k-1} : |\Delta X_t|>\frac{1}{n}\}.$$

Then $T^{n,k}>T^{n,k-1}$ a.s. on $\{T^{n,k-1}<\infty\}$. Moreover,

$$\{|\Delta X_t|>0\}=\bigcup_{n,k}\{|\Delta X_{T^{n,k}}1_{\{T^{n,k}<\infty\}}|>0\},$$

where the right side of the equality is a countable union. The result follows. □

Corollary. Let X and Y be adapted and càdlàg. If for each stopping time T, $\Delta X_T 1_{\{T<\infty\}}=\Delta Y_T 1_{\{T<\infty\}}$ a.s., then ΔX and ΔY are indistinguishable.

A much more general version of Theorem 7 is true, but it is a very deep result which uses Meyer's "section theorems," and we will not have need of it. See, for example, Dellacherie [43] or Dellacherie-Meyer [47].

A fundamental theorem of measure theory that we will need from time to time is known as the Monotone Class Theorem. Actually there are several such theorems, but the one given here is sufficient for our needs.

Definition. A **monotone vector space** $\mathcal{H}$ on a space Ω is defined to be a collection of bounded, real-valued functions f on Ω satisfying the three conditions:

(i) $\mathcal{H}$ is a vector space over $\mathbb{R}$;
(ii) $1_\Omega \in \mathcal{H}$ (i.e., constant functions are in $\mathcal{H}$); and
(iii) if $(f_n)_{n\geq 1} \subset \mathcal{H}$, and $0 \leq f_1 \leq f_2 \leq \cdots \leq f_n \leq \cdots$, and $\lim_{n\to\infty} f_n = f$, and f is bounded, then $f \in \mathcal{H}$.

Definition. A collection $\mathcal{M}$ of real functions defined on a space Ω is said to be **multiplicative** if $f,\ g \in \mathcal{M}$ implies that $fg \in \mathcal{M}$.

For a collection of real-valued functions $\mathcal{M}$ defined on Ω, we let $\sigma\{\mathcal{M}\}$ denote the space of functions defined on Ω which are measurable with respect to the σ-algebra on Ω generated by $\{f^{-1}(\Lambda); \Lambda \in \mathcal{B}(\mathbb{R}), f \in \mathcal{M}\}$.

Theorem 8 (Monotone Class Theorem). *Let $\mathcal{M}$ be a multiplicative class of bounded real-valued functions defined on a space Ω, and let $\mathcal{A} = \sigma\{\mathcal{M}\}$. If $\mathcal{H}$ is a monotone vector space containing $\mathcal{M}$, then $\mathcal{H}$ contains all bounded, $\mathcal{A}$ measurable functions.*

Theorem 8 is proved in Dellacherie-Meyer [47, page 14] with the additional hypothesis that $\mathcal{H}$ is closed under uniform convergence. This extra hypothesis is unnecessary, however, since *every monotone vector space is closed under uniform convergence.* (See Sharpe [233, page 365].)

2 Martingales

In this section we give, mostly without proofs, only the essential results from the theory of continuous time martingales. The reader can consult any of a large number of texts to find excellent proofs; for example Dellacherie-Meyer [48], or Ethier-Kurtz [74]. Also, recall that we will always assume as given a filtered, complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq\infty}$ is assumed to be right continuous.

Definition. A real-valued, adapted process $X = (X_t)_{0\leq t<\infty}$ is called a **martingale**(resp. **supermartingale, submartingale**) with respect to the filtration $\mathbb{F}$ if

(i) $X_t \in L^1(dP)$; that is, $E\{|X_t|\} < \infty$;
(ii) if $s \leq t$, then $E\{X_t|\mathcal{F}_s\} = X_s$, a.s. (resp. $E\{X_t|\mathcal{F}_s\} \leq X_s$, resp. $\geq X_s$).

Note that martingales are only defined on $[0,\infty)$; that is, for finite t and not $t=\infty$. It is often possible to extend the definition to $t=\infty$.

Definition. A martingale X is said to be **closed** by a random variable Y if $E\{|Y|\}<\infty$ and $X_t = E\{Y|\mathcal{F}_t\}$, $0\le t<\infty$.

A random variable Y closing a martingale is not necessarily unique. We give a necessary and sufficient condition for a martingale to be closed (as well as a construction for closing it) in Theorem 12.

Theorem 9. *Let X be a supermartingale. The function $t\mapsto E\{X_t\}$ is right continuous if and only if there exists a modification Y of X which is càdlàg. Such a modification is unique.*

By uniqueness we mean up to indistinguishability. Our standing assumption that the "usual hypotheses" are satisfied is used implicitly in the statement of Theorem 9. Also, note that the process Y is, of course, also a supermartingale. Theorem 9 is proved using Doob's upcrossing inequalities. If X is a martingale then $t\mapsto E\{X_t\}$ is constant, and hence it has a right continuous modification.

Corollary. If $X=(X_t)_{0\le t<\infty}$ is a martingale then there exists a unique modification Y of X which is càdlàg.

Since all martingales have right continuous modifications, *we will always assume that we are taking the right continuous version*, without any special mention. Note that it follows from this corollary and Theorem 2 that a right continuous martingale is càdlàg.

Theorem 10 (Martingale Convergence Theorem). *Let X be a right continuous supermartingale such that $\sup_{0\le t<\infty} E\{|X_t|\}<\infty$. Then the random variable $Y=\lim_{t\to\infty} X_t$ a.s. exists, and $E\{|Y|\}<\infty$. Moreover if X is a martingale closed by a random variable Z, then Y also closes X and $Y=E\{Z|\bigvee_{0\le t<\infty}\mathcal{F}_t\}$.*[2]

A condition known as uniform integrability is sufficient for a martingale to be closed.

Definition. A family of random variables $(U_\alpha)_{\alpha\in A}$ is **uniformly integrable** if

$$\lim_{n\to\infty}\sup_\alpha \int_{\{|U_\alpha|\ge n\}} |U_\alpha|dP = 0.$$

Theorem 11. *Let $(U_\alpha)_{\alpha\in A}$ be a subset of L^1. The following are equivalent:*

(i) $(U_\alpha)_{\alpha\in A}$ *is uniformly integrable.*
(ii) $\sup_{\alpha\in A} E\{|U_\alpha|\}<\infty$, *and for every $\varepsilon>0$ there exists $\delta>0$ such that $\Lambda\in\mathcal{F}$, $P(\Lambda)\le\delta$, imply $E\{|U_\alpha 1_\Lambda|\}<\varepsilon$.*

[2] $\bigvee_{0\le t<\infty}\mathcal{F}_t$ denotes the smallest σ-algebra generated by $(\mathcal{F}_t)$, all t, $0\le t<\infty$.

(iii) There exists a positive, increasing, convex function $G(x)$ defined on $[0,\infty)$ such that $\lim_{x\to\infty}\frac{G(x)}{x}=+\infty$ and $\sup_\alpha E\{G\circ|U_\alpha|\}<\infty$.

The assumption that G is convex is not needed for the implications (iii) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i).

Theorem 12. *Let X be a right continuous martingale which is uniformly integrable. Then $Y=\lim_{t\to\infty}X_t$ a.s. exists, $E\{|Y|\}<\infty$, and Y closes X as a martingale.*

Theorem 13. *Let X be a (right continuous) martingale. Then $(X_t)_{t\geq 0}$ is closed if and only if $(X_t)_{t\geq 0}$ is uniformly integrable, and if and only if $Y=\lim_{t\to\infty}X_t$ exists a.s., $E\{|Y|\}<\infty$, and $(X_t)_{0\leq t\leq\infty}$ is a martingale, where $X_\infty=Y$.*

If X is a uniformly integrable martingale, then X_t converges to $X_\infty=Y$ in L^1 as well as almost surely. The next theorem we use only once (in the proof of Theorem 28), but we give it here for completeness. The notation $(X_n)_{n\leq 0}$ refers to a process indexed by the non-positive integers: $\ldots, X_{-2}, X_{-1}, X_0$.

Theorem 14 (Backwards Convergence Theorem). *Let $(X_n)_{n\leq 0}$ be a martingale. Then $\lim_{n\to-\infty}X_n=E\{X_0|\bigcap_{n=-\infty}^{0}\mathcal{F}_n\}$ a.s. and in L^1.*

A less probabilistic interpretation of martingales uses Hilbert space theory. Let $Y\in L^2(\Omega,\mathcal{F},P)$. Since $\mathcal{F}_t\subseteq\mathcal{F}$, the spaces $L^2(\Omega,\mathcal{F}_t,P)$ form a family of Hilbert subspaces of $L^2(\Omega,\mathcal{F},P)$. Let ${}^{\pi_t}Y$ denote the Hilbert space projection of Y onto $L^2(\Omega,\mathcal{F}_t,P)$.

Theorem 15. *Let $Y\in L^2(\Omega,\mathcal{F},P)$. The process $X_t={}^{\pi_t}Y$ is a uniformly integrable martingale.*

Proof. It suffices to show $E\{Y|\mathcal{F}_t\}={}^{\pi_t}Y$. The random variable $E\{Y|\mathcal{F}_t\}$ is the unique $\mathcal{F}_t$ measurable r.v. such that $\int_A YdP=\int_A E\{Y|\mathcal{F}_t\}dP$, for any event $A\in\mathcal{F}_t$. We have $\int_A YdP=\int_A {}^{\pi_t}YdP+\int_A(Y-{}^{\pi_t}Y)dP$. But $\int_A(Y-{}^{\pi_t}Y)dP=\int 1_A(Y-{}^{\pi_t}Y)dP$. Since $1_A\in L^2(\Omega,\mathcal{F}_t,P)$, and $(Y-{}^{\pi_t}Y)$ is in the orthocomplement of $L^2(\Omega,\mathcal{F}_t,P)$, we have $\int 1_A(Y-{}^{\pi_t}Y)dP=0$, and thus by uniqueness $E\{Y|\mathcal{F}_t\}={}^{\pi_t}Y$. Since $\|{}^{\pi_t}Y\|_{L^2}\leq\|Y\|_{L^2}$, by part (iii) of Theorem 11 we have that X is uniformly integrable (take $G(x)=x^2$). □

The next theorem is one of the most useful martingale theorems for our purposes. A supermartingale X is closed by a random variable X_∞ if $X_t\geq E\{X_\infty|\mathcal{F}_t\}$ for each $t\geq 0$, with $X_\infty\in L^1$. Note that a nonnegative supermartingale can always be closed by $X_\infty=0$.

Theorem 16 (Doob's Optional Sampling Theorem). *Let X be a right continuous martingale (respectively a supermartingale), which is closed by a random variable X_∞. Let S and T be two stopping times such that $S\leq T$ a.s. Then X_S and X_T are integrable and*

$$X_S=(\geq)E\{X_T|\mathcal{F}_S\}\quad a.s.$$

Theorem 16 has a general version for supermartingales, if we take the stopping times *bounded.*

Theorem 17. *Let X be a right continuous supermartingale (resp. martingale), and let S and T be two bounded stopping times such that $S \leq T$ a.s. Then X_S and X_T are integrable and*

$$X_S \geq E\{X_T|\mathcal{F}_S\} \quad a.s. \ (resp. \ =).$$

If T is a stopping time, then so is $t \wedge T = \min(t, T)$, for each $t \geq 0$.

Definition. Let X be a stochastic process and let T be a random time. X^T is said to be the **process stopped** at T if $X_t^T = X_{t \wedge T}$.

Note that if X is adapted and càdlàg and if T is a stopping time, then

$$X_t^T = X_{t \wedge T} = X_t 1_{\{t<T\}} + X_T 1_{\{t \geq T\}}$$

is also adapted. A martingale stopped at a stopping time is still a martingale, as the next theorem shows; see its corollary.

Theorem 18. *Let X be a uniformly integrable right continuous martingale, and let T be a stopping time. Then $X^T = (X_{t \wedge T})_{0 \leq t \leq \infty}$ is also a uniformly integrable right continuous martingale.*

Proof. X^T is clearly right continuous. By Theorem 16

$$\begin{aligned} X_{t \wedge T} &= E\{X_T|\mathcal{F}_{t \wedge T}\} \\ &= E\{X_T 1_{\{T<t\}} + X_T 1_{\{T \geq t\}}|\mathcal{F}_{t \wedge T}\} \\ &= X_T 1_{\{T<t\}} + E\{X_T 1_{\{T \geq t\}}|\mathcal{F}_{t \wedge T}\}. \end{aligned}$$

However for $H \in \mathcal{F}_t$ we have $H 1_{\{T \geq t\}} \in \mathcal{F}_T$. Thus,

$$= X_T 1_{\{T<t\}} + E\{X_T|\mathcal{F}_t\} 1_{\{T \geq t\}}.$$

Therefore

$$\begin{aligned} X_{t \wedge T} &= X_T 1_{\{T<t\}} + E\{X_T|\mathcal{F}_t\} 1_{\{T \geq t\}} \\ &= E\{X_T|\mathcal{F}_t\}, \end{aligned}$$

since $X_T 1_{\{T<t\}}$ is $\mathcal{F}_t$ measurable. Thus X^T is a uniformly integrable $\mathcal{F}_t$ martingale by Theorem 13. □

Observe that the difficulty in Theorem 18 is to show that X^T is a martingale for the filtration $(\mathcal{F}_t)_{0 \leq t \leq \infty}$. It is a trivial consequence of Theorem 16 that $X^T = X_{t \wedge T}$ is a martingale for the filtration $(\mathcal{G}_t)_{0 \leq t \leq \infty}$ given by $\mathcal{G}_t = \mathcal{F}_{t \wedge T}$. The next corollary to Theorem 18 follows easily.

Corollary. Let M be a martingale, and T a finite valued stopping time. Then M^T, the martingale stopped at T, is still a martingale.

Corollary. Let Y be an integrable random variable and let S, T be stopping times. Then

$$\begin{aligned} E\{E\{Y|\mathcal{F}_S\}|\mathcal{F}_T\} &= E\{E\{Y|\mathcal{F}_T\}|\mathcal{F}_S\} \\ &= E\{Y|\mathcal{F}_{S\wedge T}\}. \end{aligned}$$

Proof. Let $Y_t = E\{Y|\mathcal{F}_t\}$. Then Y^T is a uniformly integrable martingale and

$$\begin{aligned} Y_{S\wedge T} = Y_S^T &= E\{Y_T|\mathcal{F}_S\} \\ &= E\{E\{Y|\mathcal{F}_T\}|\mathcal{F}_S\}. \end{aligned}$$

Interchanging the roles of T and S yields

$$\begin{aligned} Y_{S\wedge T} = Y_T^S &= E\{Y_S|\mathcal{F}_T\} \\ &= E\{E\{Y|\mathcal{F}_S\}|\mathcal{F}_T\}. \end{aligned}$$

Finally, $E\{Y|\mathcal{F}_{S\wedge T}\} = Y_{S\wedge T}$. □

The next inequality is elementary, but indispensable.

Theorem 19 (Jensen's Inequality). *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex, and let X and $\varphi(X)$ be integrable random variables. For any σ-algebra $\mathcal{G}$,*

$$\varphi \circ E\{X|\mathcal{G}\} \leq E\{\varphi(X)|\mathcal{G}\}.$$

Corollary 1. Let X be a martingale, and let φ be convex such that $\varphi(X_t)$ is integrable, $0 \leq t < \infty$. Then $\varphi(X)$ is a submartingale. In particular, if M is a martingale, then $|M|$ is a submartingale.

Corollary 2. Let X be a submartingale and let φ be convex, non-decreasing, and such that $\varphi(X_t)_{0\leq t<\infty}$ is integrable. Then $\varphi(X)$ is also a submartingale.

We end our review of martingale theory with Doob's inequalities; the most important is when $p = 2$.

Theorem 20. *Let X be a positive submartingale. For all $p > 1$, with q conjugate to p (i.e., $\frac{1}{p} + \frac{1}{q} = 1$), we have*

$$\|\sup_t X_t\|_{L^p} \leq q \sup_t \|X_t\|_{L^p}.$$

For a real valued process, we let X^* denote $\sup_s |X_s|$. Note that if M is a martingale with $M_\infty \in L^2$, then $|M|$ is a positive submartingale, and taking $p = 2$ we have

$$E\{(M^*)^2\} \leq 4E\{M_\infty^2\}.$$

This last inequality is called **Doob's maximal quadratic inequality**.

An elementary but useful result concerning martingales is the following.

Theorem 21. *Let $X = (X_t)_{0\le t\le\infty}$ be an adapted process with càdlàg paths. Suppose $E\{|X_T|\} < \infty$ and $E\{X_T\} = 0$ for any stopping time T, finite or not. Then X is a uniformly integrable martingale.*

Proof. Let $0 \le u < \infty$, and let $\Lambda \in \mathcal{F}_u$. Let

$$u_\Lambda = \begin{cases} u, & \text{if } \omega \in \Lambda, \\ \infty, & \text{if } \omega \notin \Lambda. \end{cases}$$

Then u_Λ is a stopping time and therefore $E\{X_{u_\Lambda}\} = 0$ by hypothesis. Moreover,

$$\begin{aligned} E\{X_u 1_\Lambda\} + E\{X_\infty 1_{\Lambda^c}\} &= 0 \\ \text{But, } E\{X_\infty 1_\Lambda\} + E\{X_\infty 1_{\Lambda^c}\} &= 0, \\ E\{X_u 1_\Lambda\} &= E\{X_\infty 1_\Lambda\} \end{aligned}$$

which yields $X_u = E\{X_\infty|\mathcal{F}_u\}$, which gives that $X = (X_u)_{u\ge 0}$ is a uniformly integrable martingale. □

Definition. A martingale X with $X_0 = 0$ and $E\{X_t^2\} < \infty$ for each $t > 0$ is called a **square integrable martingale**. If X is a uniformly integrable martingale (so that $X_t = E\{X_\infty|\mathcal{F}_t\}$), and $E\{X_\infty^2\} < \infty$ as well, then X is called an L^2 **martingale**.

Clearly, any L^2 martingale is also a square integrable martingale. Note further that if X is a uniformly integrable martingale then $\lim_{t\to\infty} E\{X_t^2\} = E\{X_\infty^2\} \le \infty$ and it is finite if X is an L^2 martingale. See Exercise 32 of Chap. II, and also Sect. 3 of Chap. IV.

3 The Poisson Process and Brownian Motion

The Poisson process and Brownian motion are two fundamental examples in the theory of continuous time stochastic processes. The Poisson process is the simpler of the two, and we begin with it. We recall that we assume given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses.

Let $(T_n)_{n\ge 0}$ be a strictly increasing sequence of positive random variables. We always take $T_0 = 0$ a.s. Recall that the indicator function $1_{\{t\ge T_n\}}$ is defined as

$$1_{\{t\ge T_n\}} = \begin{cases} 1, & \text{if } t \ge T_n(\omega), \\ 0, & \text{if } t < T_n(\omega). \end{cases}$$

Definition. The process $N = (N_t)_{0\le t\le\infty}$ defined by

$$N_t = \sum_{n\ge 1} 1_{\{t\ge T_n\}}$$

with values in $\mathbb{N}\cup\{\infty\}$ where $\mathbb{N} = \{0, 1, 2, \dots\}$ is called the **counting process** associated to the sequence $(T_n)_{n\ge 1}$.

If we set $T = \sup_n T_n$, then

$$[T_n, \infty) = \{N \geq n\} = \{(t, \omega) : N_t(\omega) \geq n\}$$

as well as

$$[T_n, T_{n+1}) = \{N = n\}, \quad \text{and} \quad [T, \infty) = \{N = \infty\}.$$

The random variable T is the **explosion time** of N. If $T = \infty$ a.s., then N is a counting process *without explosions*. For $T = \infty$, note that for $0 \leq s < t < \infty$ we have

$$N_t - N_s = \sum_{n \geq 1} 1_{\{s < T_n \leq t\}}.$$

The increment $N_t - N_s$ counts the number of random times T_n that occur between the fixed times s and t.

As we have defined a counting process it is not necessarily adapted to the filtration $\mathbb{F}$. Indeed, we have the following.

Theorem 22. *A counting process N is adapted if and only if the associated random variables $(T_n)_{n \geq 1}$ are stopping times.*

Proof. If the $(T_n)_{n \geq 0}$ are stopping times (with $T_0 = 0$ a.s.), then the event

$$\{N_t = n\} = \{\omega : T_n(\omega) \leq t < T_{n+1}(\omega)\} \in \mathcal{F}_t,$$

for each n. Thus $N_t \in \mathcal{F}_t$ and N is adapted. If N is adapted, then $\{T_n \leq t\} = \{N_t \geq n\} \in \mathcal{F}_t$, each t, and therefore T_n is a stopping time. □

Note that a counting process without explosions has right continuous paths with left limits; hence a counting process without explosions is càdlàg.

Definition. An adapted counting process N is a **Poisson process** if

(i) for any s, t, $0 \leq s < t < \infty$, $N_t - N_s$ is independent of $\mathcal{F}_s$;
(ii) for any s, t, u, v, $0 \leq s < t < \infty$, $0 \leq u < v < \infty$, $t - s = v - u$, then the distribution of $N_t - N_s$ is the same as that of $N_v - N_u$.

Properties (i) and (ii) are known respectively as *increments independent of the past*, and *stationary increments*.

Theorem 23. *Let N be a Poisson process. Then*

$$P(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!},$$

$n = 0, 1, 2, \ldots$, for some $\lambda \geq 0$. That is, N_t has the Poisson distribution with parameter λt. Moreover, N is continuous in probability[3] *and does not have explosions.*

[3] N is continuous in probability means that for $t > 0$, $\lim_{u \to t} N_u = N_t$ where the limit is taken in probability.

Proof. The proof of Theorem 23 is standard and is often given in more elementary courses (cf., e.g., Çinlar [35, page 71]). We sketch it here.

Step 1. For all $t \geq 0$, $P(N_t = 0) = e^{-\lambda t}$, for some constant $\lambda \geq 0$.

Since $\{N_t = 0\} = \{N_s = 0\} \cap \{N_t - N_s = 0\}$ for $0 \leq s < t < \infty$ by the independence of the increments,

$$\begin{aligned} P(N_t = 0) &= P(N_s = 0)P(N_t - N_s = 0) \\ &= P(N_s = 0)P(N_{t-s} = 0), \end{aligned}$$

by the stationarity of the increments. Let $\alpha(t) = P(N_t = 0)$. We have $\alpha(t) = \alpha(s)\alpha(t-s)$, for all $0 \leq s < t < \infty$. Since $\alpha(t)$ can be easily seen to be right continuous in t, we deduce that either $\alpha(t) = 0$ for all $t \geq 0$ or

$$\alpha(t) = e^{-\lambda t} \quad \text{for some } \lambda \geq 0.$$

If $\alpha(t) = 0$ it would follow that $N_t(\omega) = \infty$ a.s. for *all* t which would contradict that N is a counting process. Note that $\lim_{u \to t} P(|N_u - N_t| > \varepsilon) = \lim_{u \to t} P(|N_{u-t}| > \varepsilon) = \lim_{v \to 0} P(N_v > \varepsilon) = \lim_{v \to 0} 1 - e^{-\lambda v} = 0$; hence N is continuous in probability.

Step 2. $P(N_t \geq 2)$ is $o(t)$. (That is, $\lim_{t \to 0} \frac{1}{t} P(N_t \geq 2) = 0$.)

Let $\beta(t) = P(N_t \geq 2)$. Since the paths of N are non-decreasing, β is also non-decreasing. One readily checks that showing $\lim_{t \to 0} \frac{1}{t}\beta(t) = 0$ is equivalent to showing that $\lim_{n \to \infty} n\beta(\frac{1}{n}) = 0$. Divide $[0, 1]$ into n subintervals of equal length, and let S_n denote the number of subintervals containing at least two arrivals. By the independence and stationarity of the increments S_n is the sum of n i.i.d. zero-one valued random variables, and hence has a Binomial distribution (n,p), where $p = \beta(\frac{1}{n})$. Therefore $E\{S_n\} = np = n\beta(\frac{1}{n})$.

Since N is a counting process, we know the arrival times are *strictly* increasing; that is, $T_n < T_{n+1}$ a.s. Since $S_n \leq N_1$, if $E\{N_1\} < \infty$ we can use the Dominated Convergence Theorem to conclude $\lim_{n \to \infty} n\beta(\frac{1}{n}) = \lim_{n \to \infty} E\{S_n\} = 0$. (That $E\{N_1\} < \infty$ is a consequence of Theorem 34, established in Sect. 4).

Also note that $E\{N_1\} < \infty$ implies $N_1 < \infty$ a.s. and hence there are no explosions before time 1. This implies for fixed ω, for n sufficiently large no subinterval has more than one arrival (otherwise there would be an explosion). Hence, $\lim_{n \to \infty} S_n(\omega) = 0$ a.s.

Step 3. $\lim_{t \to 0} \frac{1}{t} P\{N_t = 1\} = \lambda$.

Since $P\{N_t = 1\} = 1 - P\{N_t = 0\} - P\{N_t \geq 2\}$, it follows that

$$\lim_{t \to 0} \frac{1}{t} P\{N_t = 1\} = \lim_{t \to 0} \frac{1 - e^{-\lambda t} + o(t)}{t} = \lambda.$$

Step 4. Conclusion.

We write $\varphi(t) = E\{\alpha^{N_t}\}$, for $0 \le \alpha \le 1$. Then for $0 \le s < t < \infty$, the independence and stationarity of the increments implies that $\varphi(t+s) = \varphi(t)\varphi(s)$ which in turn implies that $\varphi(t) = e^{t\psi(\alpha)}$. But

$$\begin{aligned}\varphi(t) &= \sum_{n=0}^{\infty} \alpha^n P(N_t = n) \\ &= P(N_t = 0) + \alpha P(N_t = 1) + \sum_{n=2}^{\infty} \alpha^n P(N_t = n),\end{aligned}$$

and $\psi(\alpha) = \varphi'(0)$, the derivative of φ at 0. Therefore

$$\begin{aligned}\psi(\alpha) &= \lim_{t\to 0} \frac{\varphi(t)-1}{t} = \lim_{t\to 0}\left\{\frac{P(N_t=0)-1}{t} + \frac{\alpha P(N_t=1)}{t} + \frac{1}{t}o(t)\right\} \\ &= -\lambda + \lambda\alpha.\end{aligned}$$

Therefore $\varphi(t) = e^{-\lambda t + \lambda\alpha t}$, hence

$$\varphi(t) = \sum_{n=0}^{\infty} \alpha^n P(N_t = n) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n \alpha^n}{n!}.$$

Equating coefficients of the two infinite series yields

$$P(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!},$$

for $n = 0, 1, 2, \ldots$. □

Definition. The parameter λ associated to a Poisson process by Theorem 23 is called the **intensity**, or **arrival rate**, of the process.

Corollary. A Poisson process N with intensity λ satisfies

$$\begin{aligned}&E\{N_t\} = \lambda t, \\ &\text{Variance}(N_t) = \text{Var}(N_t) = \lambda t.\end{aligned}$$

The proof is trivial and we omit it.

There are other, equivalent definitions of the Poisson process. For example, a counting process N without explosion can be seen to be a Poisson process if for all s, t, $0 \le s < t < \infty$, $E\{N_t\} < \infty$ and

$$E\{N_t - N_s | \mathcal{F}_s\} = \lambda(t-s).$$

Theorem 24. *Let N be a Poisson process with intensity λ. Then $N_t - \lambda t$ and $(N_t - \lambda t)^2 - \lambda t$ are martingales.*

Proof. Since λt is non-random, the process $N_t - \lambda t$ has mean zero and independent increments. Therefore

$$E\{N_t - \lambda t - (N_s - \lambda s)|\mathcal{F}_s\} = E\{N_t - \lambda t - (N_s - \lambda s)\} = 0,$$

for $0 \le s < t < \infty$. The analogous statement holds for $(N_t - \lambda t)^2 - \lambda t$. □

Definition. Let H be a stochastic process. The **natural filtration** of H, denoted $\mathbb{F}^0 = (\mathcal{F}_t^0)_{0 \le t < \infty}$, is defined by $\mathcal{F}_t^0 = \sigma\{H_s; s \le t\}$. That is, $\mathcal{F}_t^0$ is the smallest filtration that makes H adapted.

Note that natural filtrations are *not* assumed to contain all the P-null sets of $\mathcal{F}$.

Theorem 25. *Let N be a counting process. The natural filtration of N is right continuous.*

Proof. Let $E = [0, \infty]$ and $\mathcal{B}$ be the Borel sets of E, and let Γ be the path space given by

$$\Gamma = (\prod_{s \in [0,\infty)} E_s, \bigotimes_{s \in [0,\infty)} \mathcal{B}_s).$$

Define the maps $\pi_t : \Omega \to \Gamma$ by

$$\pi_t(\omega) = s \mapsto N_{s \wedge t}(\omega).$$

Thus the range of π_t is contained in the set of functions constant after t. The σ-algebra $\mathcal{F}_t^0$ is also generated by the single function space-valued random variable π_t.

Let Λ be an event in $\bigcap_{n \ge 1} \mathcal{F}^0_{t+\frac{1}{n}}$. Then there exists a set $A_n \in \bigotimes_{s \in [0,\infty)} \mathcal{B}_s$ such that $\Lambda = \{\pi_{t+\frac{1}{n}} \in A_n\}$. Next set $W_n = \{\pi_t = \pi_{t+\frac{1}{n}}\}$. For each ω, there exists an n such that $s \mapsto N_s(\omega)$ is constant on $[t, t+\frac{1}{n}]$; therefore $\Omega = \bigcup_{n \ge 1} W_n$, where W_n is an increasing sequence of events. Therefore

$$\begin{aligned}
\Lambda &= \lim_n (W_n \cap \Lambda) \\
&= \lim_n (W_n \cap \{\pi_{t+\frac{1}{n}} \in A_n\}) \\
&= \lim_n (W_n \cap \{\pi_t \in A_n\}) \\
&= \lim_n \{\pi_t \in A_n\},
\end{aligned}$$

which implies $\Lambda \in \mathcal{F}_t^0$. We conclude $\bigcap_{n \ge 1} \mathcal{F}^0_{t+\frac{1}{n}} \subset \mathcal{F}_t^0$, which implies they are equal. □

We next turn our attention to the Brownian motion process. Recall that we are assuming as given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ that satisfies the usual hypotheses.

Definition. An adapted process $B = (B_t)_{0 \le t < \infty}$ taking values in $\mathbb{R}^n$ is called an **n-dimensional $\mathbb{F}$ Brownian motion** if

(i) for $0 \le s < t < \infty$, $B_t - B_s$ is independent of $\mathcal{F}_s$ (*increments are independent of the past*);
(ii) for $0 < s < t$, $B_t - B_s$ is a Gaussian random variable with mean zero and variance matrix $(t-s)C$, for a given, non-random matrix C.

The Brownian motion *starts at* x if $P(B_0 = x) = 1$.

Often one takes the filtration $\mathbb{F}$ to be the filtration generated by the Brownian motion itself: that is, for each $t \ge 0$, $\mathcal{F}_t = \sigma\{B_s; s \le t\} \vee \mathcal{N}$, where $\mathcal{N}$ are the null sets of $\mathcal{F}$. In this case we call it simply a **Brownian motion** without the modifier of the filtration $\mathbb{F}$. (It could also be called an "intrinsic Brownian motion" in analogy with Lévy processes, but this is not used in practice.) The existence of Brownian motion is proved using a path-space construction, together with Kolmogorov's Extension Theorem. It is simple to check that a Brownian motion is a martingale as long as $E\{|B_0|\} < \infty$. Therefore by Theorem 9 there exists a version which has right continuous paths, a.s. Actually, more is true.

Theorem 26. *Let B be a Brownian motion. Then there exists a modification of B which has continuous paths a.s.*

Theorem 26 is often proved in textbooks on probability theory (e.g., Breiman [25]). It can also be proved as an elementary consequence of Kolmogorov's Lemma (Theorem 73 of Chap. IV). *We will always assume that we are using the version of Brownian motion with continuous paths.* We will also assume, unless stated otherwise, that C is the identity matrix. We then say that a Brownian motion B with continuous paths, with $C = I$ the identity matrix, and with $B_0 = x$ for some $x \in \mathbb{R}^n$, is a **standard Brownian motion**. Note that for an $\mathbb{R}^n$ standard Brownian motion B, writing $B_t = (B_t^1, \dots, B_t^n)$, $0 \le t < \infty$, then each B^i is an $\mathbb{R}^1$ Brownian motion with continuous paths, and the B^i's are independent.

We have already observed that a Brownian motion B with $E\{|B_0|\} < \infty$ is a martingale. Another important elementary observation is the following.

Theorem 27. *Let $B = (B_t)_{0 \le t < \infty}$ be a one dimensional standard Brownian motion with $B_0 = 0$. Then $M_t = B_t^2 - t$ is a martingale.*

Proof. $E\{M_t\} = E\{B_t^2 - t\} = 0$. Also

$$E\{M_t - M_s | \mathcal{F}_s\} = E\{B_t^2 - B_s^2 - (t-s) | \mathcal{F}_s\},$$

and

$$E\{B_t B_s | \mathcal{F}_s\} = B_s E\{B_t | \mathcal{F}_s\} = B_s^2,$$

since B is a martingale with B_s, $B_t \in L^2$. Therefore

$$\begin{aligned} E\{M_t - M_s | \mathcal{F}_s\} &= E\{B_t^2 - 2B_tB_s + B_s^2 - (t-s) | \mathcal{F}_s\} \\ &= E\{(B_t - B_s)^2 - (t-s) | \mathcal{F}_s\} \\ &= E\{(B_t - B_s)^2\} - (t-s) \\ &= 0, \end{aligned}$$

due to the independence of the increments from the past. □

Theorem 28. *Let π_n be a sequence of partitions of $[a, a+t]$. Suppose $\pi_m \subset \pi_n$ if $m > n$ (that is, the sequence is a refining sequence). Suppose moreover that $\lim_{n\to\infty} \operatorname{mesh}(\pi_n) = 0$. Let $\pi_n B = \sum_{t_i \in \pi_n} (B_{t_{i+1}} - B_{t_i})^2$. Then $\lim_{n\to\infty} \pi_n B = t$ a.s., for a standard Brownian motion B.*

Proof. We first show convergence in mean square. We have

$$\begin{aligned} \pi_n B - t &= \sum_{t_i \in \pi_n} \{(B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\} \\ &= \sum_i Y_i, \end{aligned}$$

where Y_i are independent random variables with zero means. Therefore

$$E\{(\pi_n B - t)^2\} = E\{(\sum_i Y_i)^2\} = \sum_i E\{Y_i^2\}.$$

Next observe that $(B_{t_{i+1}} - B_{t_i})^2/(t_{i+1} - t_i)$ has the distribution of Z^2, where Z is Gaussian with mean 0 and variance 1. Therefore

$$\begin{aligned} E\{(\pi_n B - t)^2\} &= E\{(Z^2 - 1)^2\} \sum_{t_i \in \pi_n} (t_{i+1} - t_i)^2 \\ &\leq E\{(Z^2 - 1)^2\} \operatorname{mesh}(\pi_n) t, \end{aligned}$$

which tends to 0 as n tends to ∞. This establishes L^2 convergence (and hence convergence in probability as well).

To obtain the a.s. convergence we use the Backwards Martingale Convergence Theorem (Theorem 14). Define

$$N_n(\omega) = \pi_{-n} B = \sum_{t_i \in \pi_{-n}} (B_{t_{i+1}}(\omega) - B_{t_i}(\omega))^2,$$

for $n = -1, -2, -3, \ldots$. Then it is straightforward (though notationally messy) to show that

$$E\{N_n | N_{n-1}, N_{n-2}, \ldots\} = N_{n-1}.$$

Therefore N_n is a martingale relative to $\mathcal{G}_n = \sigma\{N_k, k \leq n\}$, $n = -1, -2, \ldots$. By Theorem 14 we deduce $\lim_{n\to-\infty} N_n = \lim_{n\to\infty} \pi_n B$ exists a.s., and since $\pi_n B$ converges to t in L^2, we must have $\lim_{n\to\infty} \pi_n B = t$ a.s. as well. □

Comments. As noted in the proofs, the proof is simple (and half as long) if we conclude only L^2 convergence (and hence convergence in probability), instead of a.s. convergence. Also, we can avoid the use of the Backwards Martingale Convergence Theorem (Theorem 14) in the second half of the proof if we add the hypothesis that $\sum_n \text{mesh}(\pi_n) < \infty$. The result then follows, after having proved the L^2 convergence, by using the Borel-Cantelli Lemma and Chebsyshev's inequality. Furthermore to conclude only L^2 convergence we do not need the hypothesis that the sequence of partitions be refining.

Theorem 28 can be used to prove that the paths of Brownian motion are of unbounded variation on compacts. It is this fact that is central to the difficulties in defining an integral with respect to Brownian motion (and martingales in general).

Theorem 29. *For almost all ω, the sample paths $t \mapsto B_t(\omega)$ of a standard Brownian motion B are of unbounded variation on any interval.*

Proof. Let $A = [a, b]$ be an interval. The variation of paths of B is defined to be

$$V_A(\omega) = \sup_{\pi \in \mathcal{P}} \sum_{t_i \in \pi} |B_{t_{i+1}} - B_{t_i}|$$

where $\mathcal{P}$ are all finite partitions of $[a, b]$. Suppose $P(V_A < \infty) > 0$. Let π_n be a sequence of refining partitions of $[a, b]$ with $\lim_n \text{mesh}(\pi_n) = 0$. Then by Theorem 28 on $\{V_A < \infty\}$,

$$\begin{aligned} b - a &= \lim_{n \to \infty} \sum_{t_i \in \pi_n} (B_{t_{i+1}} - B_{t_i})^2 \\ &\leq \lim_{n \to \infty} \sup_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}| \sum_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}| \\ &\leq \lim_{n \to \infty} \sup_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}| V_A \\ &= 0, \end{aligned}$$

since $\sup_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}|$ tends to 0 a.s. as $\text{mesh}(\pi_n)$ tends to 0 by the a.s. uniform continuity of the paths on A. Since $b - a \leq 0$ is absurd, by Theorem 27 we conclude $V_A = \infty$ a.s. Since the null set can depend on the interval $[a, b]$, we only consider intervals with rational endpoints a, b with $a < b$. Such a collection is countable, and since any interval $(a, b) = \bigcup_{n=1}^{\infty} [a_n, b_n]$ with a_n, b_n rational, we can omit the dependence of the null set on the interval. □

We conclude this section by observing that not only are the increments of standard Brownian motion independent, they are also stationary. Thus Brownian motion is a Lévy process (as is the Poisson process), and the theorems of Sect. 4 apply to it. In particular, by Theorem 31 of Sect. 4, we can conclude that the *completed natural filtration of standard Brownian motion is right continuous*.

4 Lévy Processes

The Lévy processes, which include the Poisson process and Brownian motion as special cases, were the first class of stochastic processes to be studied in the modern spirit (by the French mathematician Paul Lévy). They still provide prototypic examples for Markov processes as well as for semimartingales. Most of the results of this section hold for $\mathbb{R}^n$-valued processes; for notational simplicity, however, we will consider only $\mathbb{R}$-valued processes.[4] Once again we recall that we are assuming given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses.

Definition. An adapted process $X = (X_t)_{t \geq 0}$ with $X_0 = 0$ a.s. is a **Lévy process** if

(i) X has *increments independent of the past*; that is, $X_t - X_s$ is independent of $\mathcal{F}_s$, $0 \leq s < t < \infty$; and
(ii) X has *stationary increments*; that is, $X_t - X_s$ has the same distribution as X_{t-s}, $0 \leq s < t < \infty$; and
(iii) X_t is *continuous in probability*; that is, $\lim_{t \to s} X_t = X_s$, where the limit is taken in probability.

Note that it is not necessary to involve the filtration $\mathbb{F}$ in the definition of a Lévy process. Here is a (less general) alternative definition; to distinguish the two, we will call it an *intrinsic Lévy process.*

Definition. A process $X = (X_t)_{t \geq 0}$ with $X_0 = 0$ a.s. is an **intrinsic Lévy process** if

(i) X has *independent increments*; that is, $X_t - X_s$ is independent of $X_v - X_u$ if $(u, v) \cap (s, t) = \emptyset$; and
(ii) X has *stationary increments*; that is, $X_t - X_s$ has the same distribution as $X_v - X_u$ if $t - s = v - u > 0$; and
(iii) X_t is *continuous in probability.*

Of course, an intrinsic Lévy process is a Lévy process for its minimal (completed) filtration.

If we take the Fourier transform of each X_t we get a function $f(t, u) = f_t(u)$ given by

$$f_t(u) = E\{e^{iuX_t}\},$$

where $f_0(u) = 1$, and $f_{t+s}(u) = f_t(u) f_s(u)$, and $f_t(u) \neq 0$ for every (t, u). Using the (right) continuity in probability we conclude $f_t(u) = \exp\{-t\psi(u)\}$, for some continuous function $\psi(u)$ with $\psi(0) = 0$. (Bochner's Theorem can be used to show the converse. If ψ is continuous, $\psi(0) = 0$, and if for all $t \geq 0$, $f_t(u) = e^{-t\psi(u)}$ satisfies $\sum_{i,j} \alpha_i \overline{\alpha}_j f_t(u_i - u_j) \geq 0$, for all finite

[4] $\mathbb{R}^n$ denotes n-dimensional Euclidean space. $\mathbb{R}_+ = [0, \infty)$ denotes the non-negative real numbers.

$(u_1, \ldots, u_n; \alpha_1, \ldots, \alpha_n)$, then there exists a Lévy process corresponding to f.)

In particular it follows that if X is a Lévy process then for each $t > 0$, X_t has an infinitely divisible distribution. Inversely it can be shown that for each infinitely divisible distribution μ there exists a Lévy process X such that μ is the distribution of X_1.

Theorem 30. *Let X be a Lévy process. There exists a unique modification Y of X which is càdlàg and which is also a Lévy process.*

Proof. Let $M_t^u = \frac{e^{iuX_t}}{f_t(u)}$. For each fixed u in $\mathbb{Q}$, the rationals in $\mathbb{R}$, the process $(M_t^u)_{0 \leq t < \infty}$ is a complex-valued martingale (relative to $\mathbb{F}$).

We first show that the paths of X cannot explode a.s. For any real u, $(M_t^u)_{t \geq 0}$ is a (complex-valued) martingale and thus for a.a. ω the functions $t \mapsto M_t^u(\omega)$ and $t \mapsto e^{iuX_t(\omega)}$, with $t \in \mathbb{Q}_+$, are the restrictions to $\mathbb{Q}_+$ of càdlàg functions. Let

$$\Lambda = \{(\omega, u) \in \Omega \times \mathbb{R} : e^{iuX_t(\omega)}, t \in \mathbb{Q}_+, \text{ is not the restriction of a càdlàg function}\}.$$

One can check that Λ is a measurable set. Furthermore, we have seen that $\int 1_\Lambda(\omega, u) P(d\omega) = 0$, each $u \in \mathbb{R}$. By Fubini's Theorem

$$\int \int_{-\infty}^{\infty} 1_\Lambda(\omega, u) du P(d\omega) = \int_{-\infty}^{\infty} \int 1_\Lambda(\omega, u) P(d\omega) du = 0,$$

hence we conclude that for a.a. ω the function $t \mapsto e^{iuX_t(\omega)}$, $t \in \mathbb{Q}_+$ is the restriction of a càdlàg function for almost all $u \in \mathbb{R}$. We can now conclude that the function $t \mapsto X_t(\omega)$, $t \in \mathbb{Q}_+$, is the restriction of a càdlàg function for every such ω, with the help of the lemma that follows the proof of this theorem.

Next set $Y_t(\omega) = \lim_{s \in \mathbb{Q}_+, s \downarrow t} X_s(\omega)$ for all ω in the projection onto Ω of $\{\Omega \times \mathbb{R}\} \setminus \Lambda$ and $Y_t = 0$ on Λ, all t. Since $\mathcal{F}_t$ contains all the P-null sets of $\mathcal{F}$ and $(\mathcal{F}_t)_{0 \leq t < \infty}$ is right continuous, $Y_t \in \mathcal{F}_t$. Since X is continuous in probability, $P\{Y_t \neq X_t\} = 0$, hence Y is a modification of X. It is clear that Y is a Lévy process as well. □

The next lemma was used in the proof of Theorem 30. Although it is a pure analysis lemma, we give a proof using probability theory.

Lemma. Let x_n be a sequence of real numbers such that e^{iux_n} converges as n tends to ∞ for almost all $u \in \mathbb{R}$. Then x_n converges to a finite limit.

Proof. We will verify the following Cauchy criterion: x_n converges if for any increasing sequences n_k and m_k, then $\lim_{k \to \infty} x_{n_k} - x_{m_k} = 0$. Let U be a random variable which has the uniform distribution on $[0, 1]$. For any real t, by hypothesis a.s. $e^{itUx_{n_k}}$ and $e^{itUx_{m_k}}$ converge to the same limit. Therefore,

$$\lim_{k\to\infty} e^{itU(x_{n_k}-x_{m_k})} = 1 \quad \text{a.s.}$$

so that the characteristic functions converge,

$$\lim_{k\to\infty} E\{e^{it(x_{n_k}-x_{m_k})U}\} = 1,$$

for all $t \in \mathbb{R}$. Consequently $(x_{n_k} - x_{m_k})U$ converges to zero in probability, whence $\lim_{k\to\infty} x_{n_k} - x_{m_k} = 0$, as claimed. $\square$

We will henceforth *always assume* that we are using the (unique) càdlàg version of any given Lévy process. Lévy processes provide us with examples of filtrations that satisfy the "usual hypotheses," as the next theorem shows.

Theorem 31. *Let X be a Lévy process and let $\mathcal{G}_t = \mathcal{F}_t^0 \vee \mathcal{N}$, where $(\mathcal{F}_t^0)_{0\le t<\infty}$ is the natural filtration of X, and $\mathcal{N}$ are the P-null sets of $\mathcal{F}$. Then $(\mathcal{G}_t)_{0\le t<\infty}$ is right continuous.*

Proof. We must show $\mathcal{G}_{t+} = \mathcal{G}_t$, where $\mathcal{G}_{t+} = \bigcap_{u>t} \mathcal{G}_u$. Note that since the filtration $\mathcal{G}$ is increasing, it suffices to show that $\mathcal{G}_t = \bigcap_{n\ge 1} \mathcal{G}_{t+\frac{1}{n}}$. Thus, we can take countable limits and it follows that if $s_1, \ldots, s_n \le t$, then for $(u_1, \ldots, u_n)$

$$\begin{aligned} E\{e^{i(u_1 X_{s_1}+\cdots+u_n X_{s_n})}|\mathcal{G}_t\} &= E\{e^{i(u_1 X_{s_1}+\cdots+u_n X_{s_n})}|\mathcal{G}_{t+}\} \\ &= e^{i(u_1 X_{s_1}+\cdots+u_n X_{s_n})}. \end{aligned}$$

For $v_1, \ldots, v_n > t$ and $(u_1, \ldots, u_n)$, we give the proof for $n = 2$ for notational convenience. Therefore let $z > v > t$, and suppose given u_1 and u_2. We have

$$\begin{aligned} E\{e^{i(u_1 X_v + u_2 X_z)}|\mathcal{G}_{t+}\} &= \lim_{w\downarrow t} E\{e^{i(u_1 X_v + u_2 X_z)}|\mathcal{G}_w\} \\ &= \lim_{w\downarrow t} E\{e^{iu_1 X_v} \frac{e^{iu_2 X_z}}{f_z(u_2)} f_z(u_2)|\mathcal{G}_w\} \\ &= \lim_{w\downarrow t} E\{e^{iu_1 X_v} \frac{e^{iu_2 X_v}}{f_v(u_2)} f_z(u_2)|\mathcal{G}_w\}, \end{aligned}$$

using that $M_v^{u_2} = \frac{e^{iu_2 X_v}}{f_v(u_2)}$ is a martingale. Combining terms the above becomes

$$= \lim_{w\downarrow t} E\{e^{i(u_1+u_2)X_v} f_{z-v}(u_2)|\mathcal{G}_w\}$$

and the same martingale argument yields

$$\begin{aligned} &= \lim_{w\downarrow t} e^{i(u_1+u_2)X_w} f_{v-w}(u_1+u_2) f_{z-v}(u_2) \\ &= e^{i(u_1+u_2)X_t} f_{v-t}(u_1+u_2) f_{z-v}(u_2) \\ &\ \ \vdots \\ &= E\{e^{i(u_1 X_v + u_2 X_z)}|\mathcal{G}_t\}. \end{aligned}$$

It follows that $E\{e^{i\Sigma u_j X_{s_j}}|\mathcal{G}_{t+}\} = E\{e^{i\Sigma u_j X_{s_j}}|\mathcal{G}_t\}$ for all $(s_1, \ldots, s_n)$ and all $(u_1, \ldots, u_n)$, whence $E\{Z|\mathcal{G}_{t+}\} = E\{Z|\mathcal{G}_t\}$ for every bounded $Z \in \bigvee_{0 \le s < \infty} \mathcal{F}_s^0$. This implies $\mathcal{G}_{t+} = \mathcal{G}_t$ except possibly for events of probability zero. However since both σ-algebras contain $\mathcal{N}$, we conclude $\mathcal{G}_{t+} = \mathcal{G}_t$ for each $t \ge 0$. □

The next theorem shows that a Lévy process "renews itself" at stopping times.

Theorem 32. . *Let X be a Lévy process and let T be a stopping time. On the set $\{T < \infty\}$ the process $Y = (Y_t)_{0 \le t < \infty}$ defined by $Y_t = X_{T+t} - X_T$ is a Lévy process adapted to $\mathcal{H}_t = \mathcal{F}_{T+t}$, Y is independent of $\mathcal{F}_T$ and Y has the same distribution as X.*

Proof. First assume T is bounded. Let $A \in \mathcal{F}_T$ and let $(u_1, \ldots, u_n; t_0, \ldots, t_n)$ be given with u_j in a countable dense set (for example the rationals $\mathbb{Q}$) and $t_j \in \mathbb{R}_+$, t_j increasing with j.

Recall that $M_t^{u_j} = \frac{e^{iu_j X_t}}{f_t(u_j)}$ is a martingale, where $f_t(u_j) = E\{e^{iu_j X_t}\}$. Then

$$
\begin{aligned}
&E\left\{1_A \exp\{i \sum_j u_j (X_{T+t_j} - X_{T+t_{j-1}})\}\right\} \\
&= E\left\{1_A \prod_j \frac{M_{T+t_j}^{u_j}}{M_{T+t_{j-1}}^{u_j}} \frac{f_{T+t_j}(u_j)}{f_{T+t_{j-1}}(u_j)}\right\} \\
&= P(A) \prod_j f_{t_j - t_{j-1}}(u_j)
\end{aligned}
$$

by applying the Optional Sampling Theorem (Theorem 16) n times. Note that this shows the independence of $Y_t = X_{T+t} - X_T$ from $\mathcal{F}_T$ as well as showing that Y has independent and stationary increments and *that the distribution of Y is the same as that of X.*

If T is not bounded, we let $T^n = \min(T, n) = T \wedge n$. The formula is valid for $\Lambda_n = A \cap \{T \le n\}$ when $A \in \mathcal{F}_T$, since then $\Lambda_n \in \mathcal{F}_{T \wedge n}$. Taking limits and using the Dominated Convergence Theorem we see that our formula holds for unbounded T as well, for events $\Lambda = A \cap \{T < \infty\}$, $A \in \mathcal{F}_T$. This gives the result. □

Since a standard Brownian motion is a Lévy process, Theorem 32 gives us *a fortiori* the strong Markov property for Brownian motion. This allows us to establish a pretty result for Brownian motion, known as the *reflection principle.* Let $B = (B_t)_{t \ge 0}$ denote a standard Brownian motion, $B_0 = 0$ a.s., and let $S_t = \sup_{0 \le s \le t} B_s$, the *maximum process* of Brownian motion. Since B is continuous, $S_t = \sup_{0 \le u \le t, u \in \mathbb{Q}} B_u$, where $\mathbb{Q}$ denotes the rationals; hence S_t is an adapted process with non-decreasing paths.

Theorem 33 (Reflection Principle for Brownian Motion). *Let $B = (B_t)_{t\geq 0}$ be standard Brownian motion ($B_0 = 0$ a.s.) and $S_t = \sup_{0\leq s\leq t} B_s$, the Brownian maximum process. For $y \geq 0$, $z > 0$,*

$$P(B_t < z - y; S_t \geq z) = P(B_t > y + z).$$

Proof. Let $T = \inf\{t > 0 : B_t = z\}$. Then T is a stopping time by Theorem 4, and $P(T < \infty) = 1$. We next define a new process X by

$$X_t = B_t 1_{\{t<T\}} + (2B_T - B_t)1_{\{t\geq T\}}.$$

The process X is the Brownian motion B up to time T, and after time T it is the Brownian motion B "reflected" about the constant level z. Since B and $-B$ have the same distribution, it follows from Theorem 32 that X is also a standard Brownian motion.

Next we define

$$R = \inf\{t > 0 : X_t = z\}.$$

Then clearly

$$P(R \leq t; X_t < z - y) = P(T \leq t; B_t < z - y),$$

since (R, X) and (T, B) have the same distribution. However we also have that $R = T$ identically, whence

$$\{R \leq t; X_t < z - y\} = \{T \leq t; B_t > z + y\}$$

by the construction of X. Therefore

$$P(T \leq t; B_t > z + y) = P(T \leq t; B_t < z - y). \tag{$*$}$$

The left side of $(*)$ equals

$$P(S_t \geq z; B_t > z + y) = P(B_t > z + y),$$

where the last equality is a consequence of the containment $\{S_t \geq z\} \supset \{B_t > z + y\}$. Also the right side of $(*)$ equals $P(S_t \geq z; B_t < z - y)$. Combining these yields

$$P(S_t \geq z; B_t < z - y) = P(B_t > z + y),$$

which is what was to be proved. □

We also have a reflection principle for Lévy processes. See Exercises 30 and 31.

Corollary. Let $B = (B_t)_{t\geq 0}$ be standard Brownian motion ($B_0 = 0$ a.s.) and $S_t = \sup_{0<s\leq t} B_s$. For $z > 0$,

$$P(S_t > z) = 2P(B_t > z).$$

Proof. Take $y = 0$ in Theorem 33. Then

$$P(B_t < z; S_t \geq z) = P(B_t > z).$$

Adding $P(B_t > z)$ to both sides and noting that $\{B_t > z\} = \{B_t > z\} \cap \{S_t \geq z\}$ yields the result since $P(B_t = z) = 0$. □

A Lévy process is càdlàg, and hence the only type of discontinuities it can have is jump discontinuities. Letting $X_{t-} = \lim_{s \uparrow t} X_s$, the left limit at t, we define

$$\Delta X_t = X_t - X_{t-},$$

the jump at t. If $\sup_t |\Delta X_t| \leq C < \infty$ a.s., where C is a non-random constant, then we say that X has **bounded jumps**.

Our next result states that a Lévy process with bounded jumps has finite moments of all orders. This fact was used in Sect. 3 (Step 2 of the proof of Theorem 23) to show that $E\{N_1\} < \infty$ for a Poisson process N.

Theorem 34. *Let X be a Lévy process with bounded jumps. Then $E\{|X_t|^n\} < \infty$ for all $n = 1, 2, 3, \ldots$.*

Proof. Let C be a (non-random) bound for the jumps of X. Define the stopping times

$$T_1 = \inf\{t : |X_t| \geq C\}$$
$$\vdots$$
$$T_{n+1} = \inf\{t > T_n : |X_t - X_{T_n}| \geq C\}.$$

Since the paths are right continuous, the stopping times $(T_n)_{n \geq 1}$ form a strictly increasing sequence. Moreover $|\Delta X_T| \leq C$ by hypothesis for *any* stopping time T. Therefore $\sup_s |X_s^{T_n}| \leq 2nC$ by recursion. Theorem 32 implies that $T_n - T_{n-1}$ is independent of $\mathcal{F}_{T_{n-1}}$ and also that the distribution of $T_n - T_{n-1}$ is the same as that of T_1.

The above implies that

$$E\{e^{-T_n}\} = (E\{e^{-T_1}\})^n = \alpha^n,$$

for some α, $0 \leq \alpha < 1$. But also

$$P\{|X_t| > 2nC\} \leq P\{T_n < t\} \leq \frac{E\{e^{-T_n}\}}{e^{-t}} \leq e^t \alpha^n,$$

which implies that X_t has an exponential moment and hence moments of all orders. □

We next turn our attention to an analysis of the *jumps of a Lévy process.* Let Λ be a Borel set in $\mathbb{R}$ bounded away from 0 (that is, $0 \notin \overline{\Lambda}$, where $\overline{\Lambda}$ is the closure of Λ). For a Lévy process X we define the random variables

$$T^1_\Lambda = \inf\{t > 0 : \Delta X_t \in \Lambda\}$$

$$\vdots$$

$$T^{n+1}_\Lambda = \inf\{t > T^n_\Lambda : \Delta X_t \in \Lambda\}.$$

Since X has càdlàg paths and $0 \notin \overline{\Lambda}$, the reader can readily check that $\{T^n_\Lambda \geq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$ and therefore each T^n_Λ is a stopping time. Moreover $0 \notin \overline{\Lambda}$ and càdlàg paths further imply $T^1_\Lambda > 0$ a.s. and that $\lim_{n\to\infty} T^n_\Lambda = \infty$ a.s. We define

$$N^\Lambda_t = \sum_{0<s\leq t} 1_\Lambda(\Delta X_s) = \sum_{n=1}^{\infty} 1_{\{T^n_\Lambda \leq t\}}$$

and observe that N^Λ is a counting process without an explosion. It is straightforward to check that for $0 \leq s < t < \infty$,

$$N^\Lambda_t - N^\Lambda_s \in \sigma\{X_u - X_v; s \leq v < u \leq t\},$$

and therefore $N^\Lambda_t - N^\Lambda_s$ is independent of $\mathcal{F}_s$; that is, N^Λ has independent increments. Note further that $N^\Lambda_t - N^\Lambda_s$ is the number of jumps that $Z_u = X_{s+u} - X_s$ has in Λ, $0 \leq u \leq t - s$. By the stationarity of the distributions of X, we conclude $N^\Lambda_t - N^\Lambda_s$ has the same distribution as N^Λ_{t-s}. Therefore N^Λ is a counting process with stationary and independent increments. *We conclude that N^Λ is a Poisson process.* Let $\nu(\Lambda) = E\{N^\Lambda_1\}$ be the parameter of the Poisson process N^Λ ($\nu(\Lambda) < \infty$ by the proof of Theorem 34).

Theorem 35. *The set function $\Lambda \mapsto N^\Lambda_t(\omega)$ defines a σ-finite measure on $\mathbb{R} \setminus \{0\}$ for each fixed (t, ω). The set function $\nu(\Lambda) = E\{N^\Lambda_1\}$ also defines a σ-finite measure on $\mathbb{R} \setminus \{0\}$.*

Proof. The set function $\Lambda \mapsto N^\Lambda_t(\omega)$ is simply a counting measure: $\mu(\Lambda) = \{\text{number of } s \leq t : \Delta X_s(\omega) \in \Lambda\}$. It is then clear that ν is also a measure. □

Definition. The measure ν defined by

$$\nu(\Lambda) = E\{N^\Lambda_1\} = E\{\sum_{0<s\leq 1} 1_\Lambda(\Delta X_s)\}$$

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

We wish to investigate further the role the Lévy measure plays in governing the jumps of X. To this end we establish a preliminary result. We let $N_t(\omega, dx)$ denote the random measure of Theorem 35. Since $N_t(\omega, dx)$ is a counting measure, the next result is obvious.

Theorem 36. *Let Λ be a Borel set of $\mathbb{R}$, $0 \notin \overline{\Lambda}$, f Borel and finite on Λ. Then*

$$\int_\Lambda f(x) N_t(\omega, dx) = \sum_{0<s\leq t} f(\Delta X_s) 1_\Lambda(\Delta X_s).$$

Just as we showed that N_t^Λ has independent and stationary increments, we have the following consequence.

Corollary. Let Λ be a Borel set of $\mathbb{R}$ with $0 \notin \overline{\Lambda}$, and let f be Borel and finite on Λ. Then

$$\int_\Lambda f(x) N_t(\cdot, dx)$$

is a Lévy process.

For a given set Λ (as always, $0 \notin \overline{\Lambda}$), we defined the **associated jump process** to be

$$J_t^\Lambda = \sum_{0<s\leq t} \Delta X_s 1_\Lambda(\Delta X_s).$$

By Theorem 36 and its corollary we conclude that

$$J_t^\Lambda = \int_\Lambda x N_t(\cdot, dx).$$

Hence, J_t^Λ is a Lévy process itself, it *is* defined, and $J_t^\Lambda < \infty$ a.s., each $t \geq 0$.

Theorem 37. *Given Λ, $0 \notin \overline{\Lambda}$, the process $X_t - J_t^\Lambda$ is a Lévy process.*

Proof. It is clear that we need only check the independence and stationarity of the increments. But

$$X_t - J_t^\Lambda - (X_s - J_s^\Lambda) = X_t - X_s - \sum_{s<u\leq t} \Delta X_u 1_\Lambda(\Delta X_u)$$

which is clearly $\sigma\{X_v - X_u; s \leq u < v \leq t\}$ measurable, and due to the stationarity of the increments of X it has the same law as $X_{t-s} - J_{t-s}^\Lambda$. □

We are now in a position to consider

$$Y_t^a = X_t - \sum_{0<s\leq t} \Delta X_s 1_{\{|\Delta X_s| \geq a\}},$$

for some constant $a > 0$. The advantage of doing this is that Y^a then has jumps bounded by a, and hence has finite moments of all orders (Theorem 34). We can choose any $a > 0$; we arbitrarily choose $a = 1$. Note that

$$\begin{aligned} Y_t^1 &= X_t - J_t^{(-\infty,-1]\cup[1,\infty)} \\ &= X_t - \int_{|x|\geq 1} x N_t(\cdot, dx). \end{aligned}$$

The next theorem gives an interpretation of the Lévy measure as the expected rate at which the jumps of the Lévy process fall in a given set.

Theorem 38. *Let Λ be Borel with $0 \notin \overline{\Lambda}$. Let ν be the Lévy measure of X, and let $f1_\Lambda \in L^2(d\nu)$. Then*

$$E\{\int_\Lambda f(x)N_t(\cdot, dx)\} = t\int_\Lambda f(x)\nu(dx)$$

and also

$$E\{(\int_\Lambda f(x)N_t(\cdot, dx) - t\int_\Lambda f(x)\nu(dx))^2\} = t\int_\Lambda f(x)^2\nu(dx).$$

Proof. First let $f = \sum_j a_j 1_{\Lambda_j}$, a simple function. Then

$$\begin{aligned} E\{\sum_j a_j N_t^{\Lambda_j}\} &= \sum a_j E\{N_t^{\Lambda_j}\} \\ &= t\sum_j a_j \nu(\Lambda_j), \end{aligned}$$

since $N_t^{\Lambda_j}$ is a Poisson process with parameter $\nu(\Lambda_j)$. The first equality follows easily.

For the second equality, let $M_t^i = N_t^{\Lambda_i} - t\nu(\Lambda_i)$. The M_t^i are L^p martingales, all $p \geq 1$, by the proof of Theorem 34. Moreover, $E\{M_t^i\} = 0$. Suppose Λ_i, Λ_j are disjoint. We have

$$E\{M_t^i M_t^j\} = E\{\sum_k (M_{t_{k+1}}^i - M_{t_k}^i)\sum_\ell (M_{t_{\ell+1}}^j - M_{t_\ell}^j)\}$$

for any partition $0 = t_0 < t_1 < \cdots < t_n = t$. Using the martingale property we have

$$E\{M_t^i M_t^j\} = E\{\sum_k (M_{t_{k+1}}^i - M_{t_k}^i)(M_{t_{k+1}}^j - M_{t_k}^j)\}.$$

Using the inequality $|ab| \leq a^2 + b^2$, we have

$$\sum_k (M_{t_{k+1}}^i - M_{t_k}^i)(M_{t_{k+1}}^j - M_{t_k}^j) \leq \sum_k (M_{t_{k+1}}^i - M_{t_k}^i)^2 + \sum_k (M_{t_{k+1}}^j - M_{t_k}^j)^2.$$

However $\sum_k (M_{t_{k+1}}^i - M_{t_k}^i)^2 \leq (N_t^{\Lambda_i})^2 + \nu(\Lambda_i)^2 t^2$; therefore the sums are dominated by an integrable random variable. Since M_t^i and M_t^j have paths of finite variation on $[0, t]$ it is easy to deduce that if we take a sequence $(\pi_n)_{n \geq 1}$ of partitions where the mesh tends to 0 we have

$$\lim_{n\to\infty} \sum_{t_k, t_{k+1} \in \pi_n} (M_{t_{k+1}}^i - M_{t_k}^i)(M_{t_{k+1}}^j - M_{t_k}^j) = \sum_{0<s\leq t} \Delta M_s^i \Delta M_s^j.$$

Using Lebesgue's Dominated Convergence Theorem we conclude

$$E\{M_t^i M_t^j\} = E\{\sum_{0<s\leq t} \Delta M_s^i \Delta M_s^j\} = 0;$$

the expectation above is 0 because Λ_i and Λ_j are disjoint, implying that M^i and M^j jump at different times. The second equality is now easy to verify for simple functions. For general f, let f_n be a sequence of simple functions such that $f_n 1_\Lambda$ converges to f in $L^2(d\nu)$, and the result follows. □

Remark. The first statement in Theorem 38 remains true if $f1_\Lambda \in L^1(d\nu)$. See Exercise 28.

Corollary. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and vanish in a neighborhood of 0. Then

$$E\{\sum_{0<s\leq t} f(\Delta X_s)\} = t\int_{-\infty}^{\infty} f(x)\nu(dx).$$

Proof. We need only combine Theorem 38 with Theorem 36. □

Theorem 39. *Let Λ_1, Λ_2 be two disjoint Borel sets with $0 \notin \overline{\Lambda}_1$, $0 \notin \overline{\Lambda}_2$. Then the two processes*

$$J_t^1 = \sum_{0<s\leq t} \Delta X_s 1_{\Lambda_1}(\Delta X_s)$$

$$J_t^2 = \sum_{0<s\leq t} \Delta X_s 1_{\Lambda_2}(\Delta X_s)$$

are independent Lévy processes.

Proof. By Theorem 36 and its corollary we have that J^1 and J^2 are Lévy processes. To show they are independent, we begin by forming for u, v in $\mathbb{R}$,

$$C_t^u = \frac{e^{iuJ_t^1}}{E\{e^{iuJ_t^1}\}} - 1$$

$$D_t^v = \frac{e^{ivJ_t^2}}{E\{e^{ivJ_t^2}\}} - 1.$$

Then C^u and D^v are both martingales, with $E\{C_t^u\} = E\{D_t^v\} = 0$. As in the proof of Theorem 38, let π_n: $0 = t_0 < t_1 < \cdots < t_n = t$ be a sequence of partitions of $[0,t]$ with $\lim_n \mathrm{mesh}(\pi_n) = 0$. Then

$$\begin{aligned} E\{C_t^u D_t^v\} &= E\{\sum_k (C_{t_{k+1}}^u - C_{t_k}^u)\sum_\ell (D_{t_{\ell+1}}^v - D_{t_\ell}^v)\} \\ &= E\{\sum_k (C_{t_{k+1}}^u - C_{t_k}^u)(D_{t_{k+1}}^v - D_{t_k}^v)\}. \end{aligned}$$

Since C^u and D^v have paths of finite variation on compacts, it follows by letting $\mathrm{mesh}(\pi_n)$ tend to 0, that

$$E\{C_t^u D_t^v\} = E\{\sum_{0<s\leq t} \Delta C_s^u \Delta D_s^v\}.$$

The expectation above equals zero because C^u and D^v jump at different times, due to the void intersection of Λ_1 and Λ_2.

We conclude that $E\{C_t^u D_t^v\} = 0$, and thus

$$E\{e^{iuJ_t^1} e^{ivJ_t^2}\} = E\{e^{iuJ_t^1}\} E\{e^{ivJ_t^2}\},$$

which in turn implies, because of the independence and stationarity of the increments that

$$E\{e^{i(u_1 J_{t_1}^1 + u_2(J_{t_2}^1 - J_{t_1}^1) + \cdots + u_n(J_{t_n}^1 - J_{t_{n-1}}^1))} e^{i(v_1 J_{t_1}^2 + \cdots + v_n(J_{t_n}^2 - J_{t_{n-1}}^2))}\}$$
$$= E\{e^{iu_1 J_{t_1}^1 + i\sum_{j=2}^n u_j(J_{t_j}^1 - J_{t_{j-1}}^1)}\} E\{e^{iv_1 J_{t_1}^2 + i\sum_{j=2}^n v_j(J_{t_j}^2 - J_{t_{j-1}}^2)}\}.$$

This is enough to give independence. □

The preceding results combine to yield the following useful theorem, which is one of the fundamental results about Lévy processes.

Theorem 40. *Let X be a Lévy process. Then $X_t = Y_t + Z_t$, where Y, Z are Lévy processes, Y is a martingale with bounded jumps, $Y_t \in L^p$ for all $p \geq 1$ and Z has paths of finite variation on compacts.*

Proof. Let $J_t = \sum_{0<s\leq t} \Delta X_s 1_{\{|\Delta X_s| \geq 1\}}$. Since X has càdlàg paths, for each fixed ω the function $s \mapsto X_s(\omega)$ has only finitely many jumps bigger than one on $[0,t]$. Therefore J has paths of finite variation on compacts. J is also a Lévy process by Theorem 36 and its corollary. The process $W = X - J$ is also a Lévy process (Theorem 37), and W has jumps bounded by one. We know therefore that, $n \geq 1$, $E\{|W_t|^n\}$ exists (Theorem 34), and the stationary increments of W implies $E\{W_t\} = \alpha t$, for $\alpha = E\{W_1\}$. (Recall $E\{W_0\} = 0$.) We set $Y_t = W_t - E\{W_t\}$. Then Y has independent increments and mean 0; it is a martingale. Setting $Z_t = J_t + \alpha t$ completes the proof. □

While Theorem 40 is the most important result about Lévy processes from the standpoint of stochastic integration, the next two theorems provide a better understanding of Lévy processes themselves.

Theorem 41. *Let X be a Lévy process with jumps bounded by a. That is, $\sup_s |\Delta X_s| \leq a$ a.s. Let $Z_t = X_t - E\{X_t\}$. Then Z is a martingale and $Z_t = Z_t^c + Z_t^d$ where Z^c is a martingale with continuous paths, Z^d is a martingale,*

$$Z_t^d = \int_{\{|x|\leq a\}} x(N_t(\cdot, dx) - t\nu(dx)),$$

and Z^c and Z^d are independent Lévy processes.

Proof. Z has mean zero and independent increments so it is a martingale, as well as a Lévy process. For a given set Λ we define

$$M_t^\Lambda = \int_\Lambda x N_t(\cdot, dx) - t \int_\Lambda x\nu(dx)$$
$$= \sum_{0<s\le t} \Delta X_s 1_\Lambda(\Delta X_s) - t \int_\Lambda x\nu(dx).$$

For this proof we take $a = 1$. Let $\Lambda_k = \{\frac{1}{k+1} < |x| \le \frac{1}{k}\}$. Then M^{Λ_k} are pairwise independent Lévy processes and martingales (Theorem 39). Set $M^n = \sum_{k=1}^n M^{\Lambda_k}$. Then the martingales $Z - M^n$ and M^n are independent by an argument similar to the one in the proof of Theorem 39. Moreover $\mathrm{Var}(Z_t) = \mathrm{Var}(Z_t - M_t^n) + \mathrm{Var}(M_t^n)$ where $\mathrm{Var}(X)$ denotes the variance of a random variable X. Therefore $\mathrm{Var}(M_t^n) \le \mathrm{Var}(Z_t) < \infty$ for all n. We deduce that M_t^n is Cauchy in L^2 and hence converges in L^2 as n tends to ∞ to a martingale Z_t^d, and $Z - M^n$ also converges to a martingale Z^c. Using Doob's maximal quadratic inequality (Theorem 20), we can find a subsequence converging a.s., uniformly in t on compacts, which permits the conclusion that Z^c has continuous paths. The independence of Z^d and Z^c follows from the independence of M^n and $Z - M^n$, for every n. □

Note that a consequence of the convergence of M_t^n to Z_t^d in L^2 in the proof of Theorem 41 is that the integral $\int_{[-1,0)\cup(0,1]} x^2\nu(dx)$ is finite. Note that this improves a bit on the conclusion in Theorem 38.

We recall that for a set Λ, $0 \notin \overline{\Lambda}$, the process $N_t^\Lambda = \int_\Lambda N_t(\cdot, dx)$ is a Poisson process with parameter $\nu(\Lambda)$, and thus $N_t^\Lambda - t\nu(\Lambda)$ is a martingale.

Definition. Let N be a Poisson process with parameter λ. Then $N_t - \lambda t$ is called a **compensated Poisson process**.

Theorem 41 can be interpreted as saying that a Lévy process with bounded jumps decomposes into the sum of a continuous martingale Lévy process and a martingale which is a mixture of compensated Poisson processes. It is not hard to show that $E\{e^{iuZ_t^c}\} = e^{-t\sigma^2 u^2/2}$, which implies that Z^c must be a Brownian motion. The full decomposition theorem then follows easily. We state it here without proof (consult Bertoin [14], Bretagnolle [28], Feller [75], or Jacod-Shiryaev [115] for a proof).

Theorem 42 (Lévy Decomposition Theorem). *Let X be a Lévy process. Then X has a decomposition*

$$X_t = B_t + \int_{\{|x|<1\}} x(N_t(\cdot, dx) - t\nu(dx))$$
$$+ tE\{X_1 - \int_{\{|x|\ge 1\}} xN_1(\cdot, dx)\} + \int_{\{|x|\ge 1\}} xN_t(\cdot, dx)$$
$$= B_t + \int_{\{|x|<1\}} x(N_t(\cdot, dx) - t\nu(dx)) + \alpha t + \sum_{0<s\le t} \Delta X_s 1_{\{|\Delta X_s|\ge 1\}}$$

where B is a Brownian motion; for any set Λ, $0 \notin \overline{\Lambda}$, $N_t^\Lambda = \int_\Lambda N_t(\cdot, dx)$ is a Poisson process independent of B; N_t^Λ is independent of N_t^Γ if Λ and Γ are