Philip E. Protter

Stochastic Integration and Differential Equations

Second Edition

Springer

Author

Philip E. Protter

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

Managing Editors

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

Cover pattern by courtesy of Rick Durrett (Cornell University, Ithaca)

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

ISSN 0172-4568

ISBN 3-540-00313-4 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science + Business Media GmbH

http://www.springer.de

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

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 [15].) 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: [32], [44], [87], [110], [186], [180], [208], [216], and [226]. 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 [191]. 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.

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.

Ithaca, NY
August 2003

Philip Protter

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 [42] provided an outline for just such a pedagogic approach. I developed this into a series of lectures (Protter [201]), using the work of K. Bichteler [14], E. Lenglart [145] and P. Protter [202], 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 [23], Dellacherie-Meyer [45, 46], or Ethier-Kurtz [71]). 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 [63], Kopp [130], Métivier [158], Rogers-Williams [210] and to a much lesser extent Letta [148] are examples. The books of McKean [153], Chung-Williams [32], and Karatzas-Shreve [121] 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. Leandre, 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, and 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 $\mathcal{F}_t$. □

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, limites à gauche* and *continu à gauche, limites 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 a **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*

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

is a stopping time.

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

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

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 processes sampled at T. That is,*

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

Proof. Let $\mathcal{G} = \sigma\{X_T; X \text{ all adapted càdlàg processes}\}$. 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:

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

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 [41] or Dellacherie-Meyer [45].

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 the 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 [45, 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 [215, 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 [46], or Ethier-Kurtz [71]. 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 $X_t = E\{Y|\mathcal{F}_t\}$, $0 \leq t < \infty$.

A random variable Y closing a martingale is not necessarily unique. We give a 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 $X = (X_t)_{0 \leq t < \infty}$ 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, $\sup_{0 \leq t < \infty} E\{|X_t|\} < \infty$. 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 $Y = E\{Z| \bigvee_{0 \leq t < \infty} \mathcal{F}_t\}$.*[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

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

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* $\Lambda \in \mathcal{F}$, $P(\Lambda) \leq \delta$, *imply* $E\{|U_\alpha 1_\Lambda|\} < \varepsilon$.

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

(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 uniformly integrable 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: $\dots, X_{-2}, X_{-1}, X_0$.

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 $\int_A Y dP = \int_A E\{Y|\mathcal{F}_t\} dP$, for any event $A \in \mathcal{F}_t$. We have $\int_A Y dP = \int_A {}^{\pi_t}Y dP + \int_A (Y - {}^{\pi_t}Y) dP$. 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 $G(x) = x^2$). $\square$

The next theorem is one of the most useful martingale theorems for our purposes.

Theorem 16 (Doob's Optional Sampling Theorem). *Let X be a right continuous martingale, 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*

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

Theorem 16 has a similar version for supermartingales.

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*

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

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.

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 $\mathcal{G}_t = \mathcal{F}_{t \wedge T}$.

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*

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

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 $X = (X_t)_{0\leq t\leq\infty}$ 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 $0 \leq s < t < \infty$, and let $\Lambda \in \mathcal{F}_s$. Let

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

Then u_Λ are stopping times for all $u \geq s$. Moreover

$$\begin{aligned}\int_\Lambda X_u dP &= \int X_{u_\Lambda} dP - \int_{\Omega\setminus\Lambda} X_\infty dP \\ &= -\int_{\Omega\setminus\Lambda} X_\infty dP\end{aligned}$$

since $E\{X_{u_\Lambda}\} = 0$ by hypothesis, for $u \geq s$. Thus for $\Lambda \in \mathcal{F}_s$ and $s < t$, $E\{X_t 1_\Lambda\} = E\{X_s 1_\Lambda\} = -E\{X_\infty 1_{\Omega\setminus\Lambda}\}$, which implies $E\{X_t|\mathcal{F}_s\} = X_s$, and X is a martingale, $0 \leq t \leq \infty$. □

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 $E\{X_\infty^2\} < \infty$ as well, then X is called an L^2 **martingale**.

Clearly, any L^2 martingale is also a square integrable martingale. See also Sect. 3 of Chap. IV.

3 The Poisson Process and Brownian Motion

The Poisson process and Brownian motion are the 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 $(T_n)_{n\geq 0}$ be a strictly increasing sequence of positive random variables. We always take $T_0 = 0$ a.s. Recall that the indicator function $1_{\{t\geq T_n\}}$ is defined as

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

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

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

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

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.*

Proof. The proof of Theorem 23 is standard and is often given in more elementary courses (cf., e.g., Çinlar [33, page 71]). We sketch it here.

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

[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.

Since $\{N_t = 0\} = \{N_s = 0\} \cap \{N_t - N_s = 0\}$ for $0 \le s < t < \infty$ 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 $0 \le s < t < \infty$. Since $\alpha(t)$ can be easily seen to be right continuous in t, we deduce that either $\alpha(t) = 0$ for all $t \ge 0$ or

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

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

Let $\beta(t) = P(N_t \ge 2)$. 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 $S_n \le N_1$, 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 $P\{N_t = 1\} = 1 - P\{N_t = 0\} - P\{N_t \ge 2\}$, 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 $\Lambda_n \in \bigotimes_{s\in[0,\infty)} \mathcal{B}_s$ such that $\Lambda = \{\pi_{t+\frac{1}{n}} \in \Lambda_n\}$. 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

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

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 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 $P(B_0 = x) = 1$.

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 [23]). It can also be proved as an elementary consequence of Kolmogorov's Lemma (Theorem 72 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

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

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* $\lim_{n\to\infty} \mathrm{mesh}(\pi_n) = 0$. *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 $n = -1, -2, -3, \dots$. Then it is straightforward (though notationally messy) to show that

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

Therefore N_n is a martingale relative to $\mathcal{G}_n = \sigma\{N_k, k \leq n\}$, $n = -1, -2, \dots$. 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 $\sum_n \operatorname{mesh}(\pi_n) < \infty$. 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 $\lim_n \operatorname{mesh}(\pi_n) = 0$. 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 $\operatorname{mesh}(\pi_n)$ 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.

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

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. An 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 $(u_1, \ldots, u_n; \alpha_1, \ldots, \alpha_n)$, 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 $(\mathcal{F}_t)_{0 \le t < \infty}$ 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. □

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 $s_1, \dots, s_n \le t$, then for $(u_1, \dots, u_n)$

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

For $v_1, \dots, v_n > t$ and $(u_1, \dots, u_n)$, 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 $(s_1, \dots, s_n)$ and all $(u_1, \dots, u_n)$, 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 $(u_1, \dots, u_n; t_0, \dots, t_n)$ 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 $M_t^{u_j} = \frac{e^{iu_jX_t}}{f_t(u_j)}$ is a martingale, where $f_t(u_j) = E\{e^{iu_jX_t}\}$. 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 $B = (B_t)_{t \ge 0}$ be standard Brownian motion ($B_0 = 0$ a.s.) and $S_t = \sup_{0 \le s \le t} B_s$, the Brownian maximum process. For $y \ge 0$, $z > 0$,*

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

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

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

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

$$T_\Lambda^1 = \inf\{t > 0 : \Delta X_t \in \Lambda\}$$
$$\vdots$$
$$T_\Lambda^{n+1} = \inf\{t > T_\Lambda^n : \Delta X_t \in \Lambda\}.$$

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

$$N_t^\Lambda = \sum_{0<s\leq t} 1_\Lambda(\Delta X_s) = \sum_{n=1}^{\infty} 1_{\{T_\Lambda^n \leq t\}}$$

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

$$N_t^\Lambda - N_s^\Lambda \in \sigma\{X_u - X_v; s \leq v < u \leq t\},$$

and therefore $N_t^\Lambda - N_s^\Lambda$ is independent of $\mathcal{F}_s$; that is, N^Λ has independent increments. Note further that $N_t^\Lambda - N_s^\Lambda$ 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 $N_t^\Lambda - N_s^\Lambda$ has the same distribution as N_{t-s}^Λ. Therefore N^Λ is a counting process with stationary and independent increments. *We conclude that N^Λ is a Poisson process.* Let $\nu(\Lambda) = E\{N_1^\Lambda\}$ be the parameter of the Poisson process N^Λ ($\nu(\Lambda) < \infty$ by the proof of Theorem 34).

Theorem 35. *The set function $\Lambda \mapsto N_t^\Lambda(\omega)$ defines a σ-finite measure on $\mathbb{R} \setminus \{0\}$ for each fixed (t, ω). The set function $\nu(\Lambda) = E\{N_1^\Lambda\}$ also defines a σ-finite measure on $\mathbb{R} \setminus \{0\}$.*

Proof. The set function $\Lambda \mapsto N_t^\Lambda(\omega)$ 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

$$\nu(\Lambda) = E\{N_1^\Lambda\} = E\{\sum_{0<s\leq 1} 1_\Lambda(\Delta X_s)\}$$

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 $f 1_\Lambda \in L^1(d\nu)$. 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. $\square$

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

$$E\{e^{i(u_1J_{t_1}^1+u_2(J_{t_2}^1-J_{t_1}^1)+\cdots+u_n(J_{t_n}^1-J_{t_{n-1}}^1))}e^{i(v_1J_{t_1}^2+\cdots+v_n(J_{t_n}^2-J_{t_{n-1}}^2))}\}$$
$$= E\{e^{iu_1J_{t_1}^1+i\sum_{j=2}^n u_j(J_{t_j}^1-J_{t_{j-1}}^1)}\}E\{e^{iv_1J_{t_1}^2+i\sum_{j=2}^n v_j(J_{t_j}^2-J_{t_{j-1}}^2)}\}.$$

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

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

For this proof we take $a = 1$. Let $\Lambda_k = \{\frac{1}{k+1} < |x| \leq \frac{1}{k}\}$. 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 $\mathrm{Var}(M_t^n) \leq \mathrm{Var}(Z_t) < \infty$ 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 [12], Bretagnolle [25], Feller [72], or Jacod-Shiryaev [110] for a proof).

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

$$\begin{aligned} X_t &= B_t + \int_{\{|x|<1\}} x(N_t(\cdot, dx) - t\nu(dx)) \\ &\quad + tE\{X_1 - \int_{\{|x|\geq 1\}} xN_1(\cdot, dx)\} + \int_{\{|x|\geq 1\}} xN_t(\cdot, dx) \\ &= B_t + \int_{\{|x|<1\}} x(N_t(\cdot, dx) - t\nu(dx)) + \alpha t + \sum_{0<s\leq t} \Delta X_s 1_{\{|\Delta X_s|\geq 1\}} \end{aligned}$$

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 disjoint; N_t^Λ has parameter $\nu(\Lambda)$; and $\nu(dx)$ is a measure on $\mathbb{R}\setminus\{0\}$ such that $\int \min(1, x^2)\nu(dx) < \infty$.

Finally we observe that Theorem 42 gives us a formula (known as the **Lévy-Khintchine formula**) for the Fourier transform of a Lévy process.

Theorem 43. *Let X be a Lévy process with Lévy measure ν. Then*

$$E\{e^{iuX_t}\} = e^{-t\psi(u)},$$

where

$$\psi(u) = \frac{\sigma^2}{2}u^2 - i\alpha u + \int_{\{|x|\geq 1\}} (1 - e^{iux})\nu(dx) + \int_{\{|x|<1\}} (1 - e^{iux} + iux)\nu(dx).$$

Moreover given ν, σ^2, α, the corresponding Lévy process is unique in distribution.

The Lévy-Khintchine formula has another, more common expression. (For a complete discussion of the various formulae of Lévy and Khintchine for the case of infinitely divisible distributions, see Feller [72, pages 558–565].) Indeed, in the n-dimensional case we can express the Lévy-Khintchine formulas as follows (letting $(\cdot,\cdot)$ denote the standard inner product for $\mathbb{R}^n$).

Theorem 44. *Let X be a Lévy process in $\mathbb{R}^n$, with $X_0 = 0$. Then there exists a convolution semigroup of probability measures on $\mathbb{R}^n$ such that $\mathcal{L}(X_t) = \mu_t$. The characteristic function of μ_t is $e^{-t\psi(u)}$, where ψ is given in terms of a positive definite $n \times n$ matrix Σ and a* **Lévy measure** *ν on $\mathbb{R}^n$ such that*

(i) $\nu(\Lambda) < \infty$ if Λ is Borel on $\mathbb{R}^n$ and bounded away from 0;
(ii) $\int_{|x|\leq 1} |x|^2\nu(dx) < \infty$ and $\nu\{0\} = 0$.

Also there exists a bounded, continuous **centering function** *$\eta : \mathbb{R}^n \to \mathbb{R}^n$ such that $\lim_{x\to 0} \frac{\eta(x)-x}{|x|^2} = 0$, and a constant $\alpha \in \mathbb{R}^n$ such that*

$$\psi(u) = -i(u,\alpha) + \frac{1}{2}(\Sigma u, u) - \int e^{i(u,x)} - 1 - i(u,\eta(x))\nu(dx).$$

A nice proof of the above theorem can be found in Bertoin [12, pages 13–15]. Several observations concerning this theorem are in order. First, one still has (in the n-dimensional case) the key property that for $f : \mathbb{R}^n \to \mathbb{R}^n$, bounded, vanishing in a neighborhood of 0, that

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

Second, in the case $n = 1$, one often takes $\eta(x) = \frac{x}{1+x^2}$, possibly at the cost of changing the constant α. In the next section (Sect. 6) we discuss semigroups of Markov processes. The semigroup P_t for the Lévy process X has the form $P_tf(x) = \int f(x+y)\mu_t(dy)$, and one can show in fact that the "infinitesimal generator" of the semigroup P_t is, for $f \in b\mathcal{C}^2(\mathbb{R}^n)$,

$$
\begin{aligned}
Af(x) &= \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} \\
&= -(\nabla f(x), \alpha) - \sum_{j,k=1}^{n} \sigma_{jk} \frac{\partial^2 f}{\partial x_j \partial x_k}(x) \\
&\quad + \int f(x+y) - f(x) - (\nabla f(x), \eta(y))\nu(dy).
\end{aligned}
$$

Note further that if f is constant in a neighborhood of x, then

$$
\lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} = \int f(x+y) - f(x)\nu(dy).
$$

We next consider some examples of Lévy processes. For the examples we take $n = 1$.

Example. If a Lévy process X has continuous paths, then ν is identically 0, and X must be Brownian motion with drift. That is, $X_t = \sigma B_t + \alpha t$, for constants α, σ, are the only Lévy processes with continuous paths.

Example. Assume that ν is a finite measure. That is, $\nu(\mathbb{R}^n) < \infty$, and that $\sigma^2 = 0$. Then X is a **compound Poisson process** with jump arrival intensity $\lambda = \nu(\mathbb{R}^n)$. That is, let $N_t = \sum_{i=1}^{\infty} 1_{\{t \geq T_i\}}$ be a Poisson process with arrival intensity λ. Let $Z_t = \sum_{i=1}^{\infty} U_i 1_{\{t\geq T_i\}}$ where the sequence $(U_j)_{j\geq 1}$ are i.i.d., independent of the sequence $(T_j)_{j\geq 1}$, and $\mathcal{L}(U_1) = \frac{1}{\lambda}\nu$. Then one easily verifies $E\{e^{iuZ_t}\} = e^{-t\psi(u)}$, where $\psi(u) = \int(1 - e^{iux})\nu(dx)$.

Example. Let $\nu(dx) = \sum_{k=1}^{\infty} \alpha_k \varepsilon_{\beta_k}(dx)$, where $\varepsilon_{\beta_k}(dx)$ denotes the positive point mass measure at $\beta_k \in \mathbb{R}$ of size one. Assume $\sum_{k=1}^{\infty} \beta_k^2 \alpha_k < \infty$ and $\sigma^2 = 0$. Let N^k be a sequence of independent Poisson processes with parameters α_k. Then $X_t = \sum_{k=1}^{\infty} \beta_k (N_t^k - \alpha_k t)$ has ν as its Lévy measure.

Example. Let X be a Lévy process which is a **Gamma process**. That is, $\mathcal{L}(X_t) = \Gamma(\alpha, \beta)$, with $\beta \geq 0$ and $\alpha = at$. More precisely, the density for X_t is $f_t(x) = \frac{\beta^{at}}{\Gamma(at)} x^{ta-1} e^{-\beta x} 1_{\{x>0\}}$. In this case $\sigma^2 = 0$, and taking $a = \beta = 1$ gives $\phi_t(u) = E\{e^{iuZ_t}\} = (1 - iu)^{-t}$. (See, for example, page 43 of Jacod-Protter [109] or p. 73 of Bertoin [12].) Then $\phi_t(u) = e^{-t\ln(1-iu)} = e^{t\psi(u)}$. But

$$
\psi'(u) = i\frac{1}{(1-iu)} = i\phi_1(u) = iE\{e^{iuZ_1}\} = i\int_0^{\infty} e^{iux} e^{-x} dx.
$$

Integrating yields

$$
\psi(u) = \int_0^{\infty} \frac{e^{iux} - 1}{x} e^{-x} dx
$$

which implies, by the Lévy-Khintchine formula, that $\nu(dx) = \frac{1}{x} e^{-x} 1_{\{x>0\}} dx$.

Example. Our last example is that of **stable processes**. We begin by recalling what a stable distribution is. Extending the notion of the Central Limit Theorem, consider the normalized sums

$$S_n = \frac{\sum_{i=1}^n X_i}{\gamma_n} - \beta_n, \quad \gamma_n > 0.$$

We say that a random variable X has a **stable law** if it is the limit in distribution as $n \to \infty$ of sums of the above form. One can then prove (see, for example, Breiman [23, page 200]) that either X has a normal distribution, or there exists a number $\alpha, 0 < \alpha < 2$, called the **index** of the stable law, and also constants $m_1 \geq 0$, $m_2 \geq 0$, and δ such that $E\{e^{iuX}\} = e^{\psi(u)}$ and

$$\psi(u) = iu\delta + m_1 \int_0^\infty (e^{iux} - 1 - \frac{iux}{1+x^2}) \frac{1}{x^{1+\alpha}} dx$$
$$+ m_2 \int_{-\infty}^0 (e^{iux} - 1 - \frac{iux}{1+x^2}) \frac{1}{|x|^{1+\alpha}} dx.$$

One says that a stable distribution is *symmetric* if $m_1 = m_2$. By analogy we define a **symmetric stable process** (in the case $n = 1$) to be a Lévy process where the Lévy measure has the form $\nu(dx) = \frac{1}{|x|^{(1+\alpha)}} dx$, with $0 < \alpha < 2$, and the constant σ^2 in the Lévy-Khintchine formula is 0, which, of course, implies that no Brownian component is present in the process. Note that such stable processes have nice *scaling properties*. If α is the index, then the law of $(X_t)_{t \geq 0}$ is the same as the law of $(\beta^{(-\frac{1}{\alpha})} X_{\beta t})_{t \geq 0}$. Consult Bertoin [12] for many other properties of stable processes.

5 Why the Usual Hypotheses?

We defined the usual hypotheses at the very beginning of this book, and we saw with Lévy processes that the natural filtration of a Lévy process, once completed, satisfies *a fortiori* right continuity, and thus the usual hypotheses hold (see Theorem 31). This phenomenon is true much more generally. Indeed it is true for all "reasonable" strong Markov processes, as well as for all counting processes (cf., Theorem 25). In this section we will show that it is true for a very large class of Markov processes. We first begin with some definitions.

Definition 1. Let X be an $\mathbb{R}^d$-valued process, with X adapted to a filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq \infty}$. X is **Markov with respect to** $\mathbb{F}$ if $\mathcal{F}_t$ and $\sigma\{X_s; s \geq t\}$ are conditionally independent, given X_t.

Definition 2. The process X in Definition 1 is **Markov with respect to** $\mathbb{F}$ if for any random variable $Y \in b\sigma\{X_s; s \geq t\}$, $E\{Y|\mathcal{F}_t\} = E\{Y|X_t\}$, a.s., all $t \geq 0$.

The following theorem has a straightforward proof.

Theorem 45. *Definitions 1 and 2 are equivalent. Moreover X is Markov with respect to $\mathbb{F}$ if and only if $E\{f(X_u)|\mathcal{F}_t\} = E\{f(X_u)|X_t\}$ a.s. for all $u \geq t \geq 0$.*

Let X be a càdlàg process and let X be Markov with respect to $\mathbb{F}$. Since X is Markov, we can associate with it a *Markov transition function* on $(\mathbb{R}^d, \mathcal{B})$, where $\mathcal{B}$ denotes the Borel sets of $\mathbb{R}^d$. This is a family $(P_{s,t})_{s,t\in\mathbb{R}_+}$ of Markov kernels on $(\mathbb{R}^d, \mathcal{B})$ such that

$$P_{s,t}P_{t,u} = P_{s,u}$$

for each triple (s,t,u) with $s < t < u$. We then have the fundamental relationship

$$E\{f(X_t)|\mathcal{F}_s\} = P_{s,t}(X_s, f) \quad \text{a.s.}$$

Note that, by using indicator functions such as $1_A = f$, we have

$$P(X_t \in A|\mathcal{F}_s) = P_{s,t}(X_s, A) \quad \text{a.s.}$$

Also note that the above implies the equality

$$P(X_t \in A|X_s) = P_{s,t}(X_s, A) \quad \text{a.s.}$$

which, of course, is weaker. That is, the converse implication is not true in general. The most important special case is Markov processes which are *time homogeneous*. In this case we have a transition semigroup P_t given by

$$P_{t-s} = P_{s,t},$$

for $s \leq t$.

Definition. Let $\mathcal{C}_0$ denote continuous functions vanishing at infinity. A Markov process is a **Feller process** if its transition semigroup $(P_t)_{t\geq 0}$ has the following two properties.

(a) For each $f \in \mathcal{C}_0$ and each $t > 0$, $P_t f \in \mathcal{C}_0$.
(b) For each $f \in \mathcal{C}_0$ and each $x \in \mathbb{R}^d$, $P_t f(x) \to f(x)$ as $t \to 0$.

Examples of Feller processes include Brownian motion and the Poisson process. More generally, one can show that diffusions[5] and all Lévy processes are Feller processes.

Theorem 46. *Let X be a càdlàg Feller process for a filtration $\mathbb{F}$. Then it is also Markov with respect to the filtration $\mathbb{F}_+ = (\mathcal{F}_{t+})_{0\leq t\leq\infty}$.*

Proof. Let $t > 0$ be fixed, and assume $f \in \mathcal{C}_0$. Let $Y_t = f(X_t)$. For $0 \leq s \leq t$, let

$$Y_s = E\{f(X_t)|\mathcal{F}_s\} = E\{f(X_t)|X_s\} = P_{t-s}f(X_s).$$

Then Y is a martingale for $0 \leq s \leq t$, and it is càdlàg, because $P_t f$ is continuous, and X is càdlàg. Therefore, Y is a martingale with respect to $\mathbb{F}_+$ too, which means that $E\{f(X_t)|\mathcal{F}_{s+}\} = Y_s$, from which we deduce that $E\{f(X_t)|\mathcal{F}_{s+}\} = E\{f(X_t)|X_s\}$, and we are done. □

[5] This depends, of course, on how one defines a diffusion; the definition is not yet standardized in the literature.

We now define P^μ as a probability measure such that X is Markov with respect to P^μ and also $\mathcal{L}(X_0) = \mu$ under P^μ, where $\mathcal{L}(X_0)$ denotes the distribution, or "law," of X_0. Let $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{0\le t\le\infty}$ be the filtration where $\mathcal{F}_t^\mu$ denotes the P^μ completion of $\mathcal{F}_t^0$, $\mathcal{F}_t^0 = \sigma\{X_s : s \le t\}$, and $\mathcal{F}^0 = \sigma\{X_s : s \ge 0\}$. Also, let $\mathcal{F}_t = \bigcap_{\{\|\mu\|=1\}} \mathcal{F}_t^\mu$.

Theorem 47. *If X is Markov with respect to P^μ, or alternatively Markov with respect to P^μ for all μ with $\|\mu\| = 1$, then $\mathbb{F}^\mu$ and $\mathbb{F}$ are, respectively, right continuous.*

Proof. We prove the result for $\mathbb{F}^\mu$, from which the result for $\mathbb{F}$ follows. Let $\mathcal{H}$ denote all $h \in \mathbf{b}\mathcal{F}^\mu$ such that

$$E\{h|\mathcal{F}_t^\mu\} = E\{h|\mathcal{F}_{t+}^\mu\}, \quad P^\mu\text{-a.s.}$$

Then $\mathcal{H} \supset \{H = f(X_s),\ f \in \mathbf{b}\mathcal{B}\}$, which is obvious for $s \le t$, and for $s > t$ it is just Theorem 46. Then $\mathcal{H}$ contains all H of the form $H = \prod_{i=1}^n f_i(X_{s_i})$, for any $f \in \mathbf{b}\mathcal{B}$. Using the Monotone Class Theorem, $\mathcal{H} \supset \mathbf{b}\mathcal{F}^0$. By completion, $\mathcal{H} \supset \mathbf{b}\mathcal{F}^\mu$. Therefore if $h \in \mathcal{F}_{t+}^\mu$, we have $E\{h|\mathcal{F}_t^\mu\} = E\{h|\mathcal{F}_{t+}^\mu\} = h$, P^μ-a.s. □

The property that the natural filtration of a Feller process is right continuous when completed is not restricted to Feller processes. Indeed, it can also be shown to be true for more general classes of processes, such as *Hunt processes*[6]. However it is simply two elementary properties which yield the right continuity of the filtration in a general setting.

Definition. Let X be a càdlàg Markov process for a filtration $\mathbb{F}$, which we take to be completed and therefore can assume to be right continuous, in view of Theorem 47. Let T be a stopping time. We say that X **verifies the strong Markov property** at T if

$$E\{f(X_{T+s})1_{\{T<\infty\}}|\mathcal{F}_T\} = P_s(X_T, f)1_{\{T<\infty\}} \quad \text{a.s.}$$

We say that X is **strong Markov** if the above holds for *every* stopping time T.

Suppose now that the semigroup (P_t) is *Borel.* That is, if $f \in \mathcal{B}$, then $P_t f \in \mathcal{B}$ as well. Suppose that for every probability measure μ on $(\mathbb{R}^d, \mathcal{B})$, there exists a probability space $(\Omega, \mathcal{G}, Q)$ and a càdlàg Markov process $(Y_t)_{t\ge0}$ defined on this space with $\mathcal{L}(X_0) = \mu$. Suppose further that every such càdlàg Markov process Y is strong Markov. Then the filtrations $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{0\le t\le\infty}$ and $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le\infty}$ can be shown to be right continuous in a manner analogous to the previous proofs.

---

[6] See, for example, Sharpe [215, page 219] for a definition of a Hunt process.

6 Local Martingales

Recall that we assume as given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ *satisfying the usual hypotheses.* For a process X and a stopping time T we further recall that X^T denotes the *stopped process*

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

Definition. An adapted, càdlàg process X is a **local martingale** if there exists a sequence of increasing stopping times, T_n, with $\lim_{n\to\infty} T_n = \infty$ a.s. such that $X_{t\wedge T_n} 1_{\{T_n>0\}}$ is a uniformly integrable martingale for each n. Such a sequence (T_n) of stopping times is called a **fundamental sequence**.

Example. Clearly any càdlàg martingale is a local martingale (take $T_n \equiv n$).

Example. We give an example of a local martingale which is not a martingale. Let $(B_t)_{0\leq t<\infty}$ be a Brownian motion in $\mathbb{R}^3$ with $B_0 = x$, where $x \neq 0$. Let $u(y) = \|y\|^{-1}$, a superharmonic function on $\mathbb{R}^3$. As a consequence of Itô's formula (cf., Theorems 32 and 41 of Chap. II) one can show that $X_t = u(B_t)$ is a positive supermartingale (indeed, it is even a uniformly integrable supermartingale). Next let $T_n = \inf\{t > 0 : \|B_t\| \leq 1/n\}$. Outside of the ball of radius $1/n$ centered at the origin the function u is harmonic. Again by Itô's formula one can show that this implies $u(B_{t\wedge T_n})$ is a martingale. Since $u(B_{t\wedge T_n})$ is bounded by n it is uniformly integrable. On the other hand, if $B_0 = x \neq 0$ then $\lim_{t\to\infty} E\{u(B_t)\} = 0$ as is easily seen by a calculation, while $E\{u(B_0)\} = \|x\|^{-1}$. Since the expectations of a martingale are constant, $u(B_t)$ is not a martingale. (For more on this example, see Sect. 6 of Chap. II, following Corollary 4 of Theorem 27, as well as Exercise 20 in Chap. II.)

The reason we multiply $X_{t\wedge T_n}$ by $1_{\{T_n>0\}}$ is to relax the integrability condition on X_0. This is useful, for example, in the consideration of stochastic integral equations with a non-integrable initial condition.

Definition. A stopping time T **reduces** a process M if M^T is a uniformly integrable martingale.

Theorem 48. *Let M, N be local martingales and let S and T be stopping times.*

(a) If T reduces M and $S \leq T$ a.s., then S reduces M.
(b) The sum $M + N$ is also a local martingale.
(c) If S, T both reduce M, then $S \vee T$ also reduces M.
(d) The processes M^T, $M^T 1_{\{T>0\}}$ are local martingales.
(e) Let X be a càdlàg process and let T_n be a sequence of stopping times increasing to ∞ a.s. such that $X^{T_n} 1_{\{T_n>0\}}$ is a local martingale for each n. Then X is a local martingale.

Proof. (a) follows from the Optional Sampling Theorem (Theorem 16) and Theorem 13. For (b), if $(S_n)_{n\geq 1}$, $(T_n)_{n\geq 1}$ are fundamental sequences for M

and N, respectively, then $S_n \wedge T_n$ is a fundamental sequence for $M+N$. For (c), let $X_t = M_t - M_0$. Then $X^{S\vee T} = X^S + X^T - X^{S\wedge T}$ is a uniformly integrable martingale. But

$$|M_0|1_{\{S\vee T>0\}} \le |M_0|1_{\{S>0\}} + |M_0|1_{\{T>0\}},$$

which is in L^1, therefore $X^{S\vee T}+M_0 1_{\{S\vee T>0\}} = M^{S\vee T}1_{\{S\vee T>0\}}$ is a uniformly integrable martingale. The proof of (d) consists of the observation that if $(T_n)_{n\geq 1}$ is a fundamental sequence for M, then it is also one for M^T and $M^T 1_{\{T>0\}}$ by the Optional Sampling Theorem (Theorem 16). For (e), we have $X^{T_n}1_{\{T_n>0\}} = M^n$ is a local martingale for each n. For fixed n we know there exists a fundamental sequence $U^{n,k}$ increasing to ∞ a.s. as k tends to ∞. For each n, choose $k = k(n)$ such that $P(U^{n,k(n)} < T_n \wedge n) < 2^{-n}$. Then $\lim_n U^{n,k(n)} = \infty$ a.s., and $U^{n,k(n)} \wedge T_n$ reduces X for each n. We take $R^m = \max(U^{1,k(1)} \wedge T_1, \ldots, U^{m,k(m)} \wedge T_m)$, and each R^m reduces X by (c), the R^m are increasing, and $\lim R^m = \infty$ a.s. Therefore X is a local martingale. □

Corollary. Local martingales form a vector space.

We will often need to know that a reduced local martingale, M^T, is in L^p and not simply uniformly integrable.

Definition. Let X be a stochastic process. A **property π is said to hold locally** if there exists a sequence of stopping times $(T_n)_{n\geq 1}$ increasing to ∞ a.s. such that $X^{T_n}1_{\{T_n>0\}}$ has property π, each $n \geq 1$.

We see from Theorem 48(e) that a process which is locally a local martingale is also a local martingale. Other examples of local properties that arise frequently are *locally bounded* and *locally square integrable.* Theorems 49 and 50 are consequences of Theorem 48(e).

Theorem 49. *Let X be a process which is locally a square integrable martingale. Then X is a local martingale.*

The next theorem shows that the traditional "uniform integrability" assumption in the definition of local martingale is not really necessary.

Theorem 50. *Let M be adapted, càdlàg and let $(T_n)_{n\geq 1}$ be a sequence of stopping times increasing to ∞ a.s. If $M^{T_n}1_{\{T_n>0\}}$ is a martingale for each n, then M is a local martingale.*

It is often of interest to determine when a local martingale is actually a martingale. A simple condition involves the maximal function. Recall that $X_t^* = \sup_{s\leq t}|X_s|$ and $X^* = \sup_s |X_s|$.

Theorem 51. *Let X be a local martingale such that $E\{X_t^*\} < \infty$ for every $t \geq 0$. Then X is a martingale. If $E\{X^*\} < \infty$, then X is a uniformly integrable martingale.*

Proof. Let $(T_n)_{n\geq 1}$ be a fundamental sequence of stopping times for X. If $s \leq t$, then $E\{X_{t\wedge T_n}|\mathcal{F}_s\} = X_{s\wedge T_n}$. The Dominated Convergence Theorem yields $E\{X_t|\mathcal{F}_s\} = X_s$. If $E\{X^*\} < \infty$, since each $|X_t| \leq X^*$, it follows that $(X_t)_{t\geq 0}$ is uniformly integrable. □

Note that in particular a bounded local martingale is a uniformly integrable martingale. Other sufficient conditions for a local martingale to be a martingale are given in Corollaries 3 and 4 to Theorem 27 in Chap. II, and Kazamaki's and Novikov's criteria (Theorems 40 and 41 of Chap. III) establish the result for the important special cases of continuous local martingales.

7 Stieltjes Integration and Change of Variables

Stochastic integration with respect to semimartingales can be thought of as an extension of path-by-path Stieltjes integration. We present here the essential elementary ideas of Stieltjes integration appropriate to our interests. We assume the reader is familiar with the Lebesgue theory of measure and integration on $\mathbb{R}_+$.

Definition. Let $A = (A_t)_{t\geq 0}$ be a càdlàg process. A is an **increasing process** if the paths of $A : t \to A_t(\omega)$ are non-decreasing for almost all ω. A is called a **finite variation process** (FV) if almost all of the paths of A are of finite variation on each compact interval of $\mathbb{R}_+$.

Let A be an increasing process. Fix an ω such that $t \mapsto A_t(\omega)$ is right continuous and non-decreasing. This function induces a measure $\mu_A(\omega, ds)$ on $\mathbb{R}_+$. If f is a bounded, Borel function on $\mathbb{R}_+$, then $\int_0^t f(s)\mu_A(\omega, ds)$ is well-defined for each $t > 0$. We denote this integral by $\int_0^t f(s)dA_s(\omega)$. If $F_s = F(s,\omega)$ is bounded and jointly measurable, we can define, ω-by-ω, the integral $I(t,\omega) = \int_0^t F(s,\omega)dA_s(\omega)$. I is right continuous in t and jointly measurable.

Proceeding analogously for A an FV process (except that the induced measure $\mu_A(\omega, ds)$ can have negative measure; that is, it is a *signed measure*), we can define a jointly measurable integral

$$I(t,\omega) = \int_0^t F(s,\omega)dA_s(\omega)$$

for F bounded and jointly measurable.

Let A be an FV process. We define

$$|A|_t = \sup_{n\geq 1} \sum_{k=1}^{2^n} |A_{\frac{tk}{2^n}} - A_{\frac{t(k-1)}{2^n}}|. \qquad (*)$$

Then $|A|_t < \infty$ a.s., and it is an increasing process.

Definition. For A an FV process, the **total variation process**, $|A| = (|A|_t)_{t\geq 0}$, is the increasing process defined in $(*)$ above.

Notation. Let A be an FV process and let F be a jointly measurable process such that $\int_0^t F(s,\omega)dA_s(\omega)$ exists and is finite for all $t > 0$, a.s. We let

$$(F \cdot A)_t(\omega) = \int_0^t F(s,\omega)dA_s(\omega)$$

and we write $F \cdot A$ to denote the process $F \cdot A = (F \cdot A_t)_{t\geq 0}$. We will also write $\int_0^t F_s|dA_s|$ for $(F \cdot |A|)_t$. The next result is an absolute continuity result for Stieltjes integrals.

Theorem 52. *Let A, C be adapted, strictly increasing processes such that $C - A$ is also an increasing process. Then there exists a jointly measurable, adapted process H (defined on $(0,\infty)$) such that $0 \leq H \leq 1$ and*

$$A = H \cdot C$$

or equivalently

$$A_t = \int_0^t H_s dC_s.$$

Proof. If μ and ν are two Borel measures on $\mathbb{R}_+$ with $\mu \ll \nu$, then we can set

$$\alpha(s,t) = \begin{cases} \frac{\mu((s,t])}{\nu((s,t])}, & \text{if } \nu((s,t]) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Defining h and k by $h(t) = \liminf_{s\uparrow t} \alpha(s,t)$, $k(t) = \limsup_{t\downarrow s} \alpha(s,t)$, then h and k are both Borel measurable, and moreover they are each versions of the Radon-Nikodym derivative. That is,

$$d\mu = h\,d\nu, \qquad d\mu = k\,d\nu.$$

To complete the proof it suffices to show that we can follow the above procedure in a (t,ω) measurable fashion. With the convention $\frac{0}{0} = 0$, it suffices to define

$$H(t,\omega) = \liminf_{r\uparrow 1, r\in\mathbb{Q}_+} \frac{(A(t,\omega) - A(rt,\omega))}{(C(t,\omega) - C(rt,\omega))};$$

such an H is clearly adapted since both A and C are. □

Corollary. Let A be an FV process. There exists a jointly measurable, adapted process H, $-1 \leq H \leq 1$, such that

$$|A| = H \cdot A \qquad \text{and} \qquad A = H \cdot |A|$$

or equivalently

$$|A|_t = \int_0^t H_s dA_s \qquad \text{and} \qquad A_t = \int_0^t H_s|dA_s|.$$

Proof. We define $A_t^+ = \frac{1}{2}(|A|_t + A_t)$ and $A_t^- = \frac{1}{2}(|A|_t - A_t)$. Then A^+ and A^- are both increasing processes, and $|A| - A^+$ and $|A| - A^-$ are also increasing processes. By Theorem 52 there exist processes H^+ and H^- such that $A_t^+ = \int_0^t H_s^+ |dA_s|$, $A_t^- = \int_0^t H_s^- |dA_s|$. It then follows that $A_t = A_t^+ - A_t^- = \int_0^t (H_s^+ - H_s^-)|dA_s|$. Let $H_t \equiv H_t^+ - H_t^-$ and suppose H^+ and H^- are defined as in the proof of Theorem 52. Except for a P-null set, for a given ω it is clear that $|H_s(\omega)| = 1$ $dA_s(\omega)$ almost all s. Considering $H \cdot A$, we have

$$\begin{aligned}\int_0^t H_s dA_s &= \int_0^t H_s d(\int_0^s H_u |dA_u|) \\ &= \int_0^t H_s H_s |dA_s| = \int_0^t 1 |dA_s| = |A|_t.\end{aligned}$$

This completes the proof. □

When the integrand process H has *continuous paths*, the Stieltjes integral $\int_0^t H_s dA_s$ is also known as the **Riemann-Stieltjes integral** (for fixed ω). In this case we can define the integral as the limit of approximating sums. Such a result is proved in elementary textbooks on real analysis (e.g., Protter-Morrey [195, pages 316, 317]).

Theorem 53. *Let A be an FV process and let H be a jointly measurable process such that a.s. $s \mapsto H(s, \omega)$ is continuous. Let π_n be a sequence of finite random partitions of $[0, t]$ with $\lim_{n\to\infty} \mathrm{mesh}(\pi_n) = 0$. Then for $T_k \le S_k \le T_{k+1}$,*

$$\lim_{n\to\infty} \sum_{T_k, T_{k+1} \in \pi_n} H_{S_k}(A_{T_{k+1}} - A_{T_k}) = \int_0^t H_s dA_s \ a.s.$$

We next prove a change of variables formula when the *FV* process is *continuous*. Itô's formula (Theorem 32 of Chap. II) is a generalization of this result.

Theorem 54 (Change of Variables). *Let A be an FV process with continuous paths, and let f be such that its derivative f' exists and is continuous. Then $(f(A_t))_{t \ge 0}$ is an FV process and*

$$f(A_t) - f(A_0) = \int_0^t f'(A_s) dA_s.$$

Proof. For fixed ω, the function $s \mapsto f'(A_s(\omega))$ is continuous on $[0, t]$ and hence bounded. Therefore the integral $\int_0^t f'(A_s) dA_s$ exists. Fix t and let π_n be a sequence of partitions of $[0, t]$ with $\lim_{n\to\infty} \mathrm{mesh}(\pi_n) = 0$. Then

$$\begin{aligned}f(A_t) - f(A_0) &= \sum_{t_k, t_{k+1} \in \pi_n} \{f(A_{t_{k+1}}) - f(A_{t_k})\} \\ &= \sum_k f'(A_{S_k})(A_{t_{k+1}} - A_{t_k}),\end{aligned}$$

by the Mean Value Theorem, for some S_k, $t_k \le S_k \le t_{k+1}$. The result now follows by taking limits and Theorem 53. □

Comment. We will see in Chap. II that the sums

$$\sum_{t_k \in \pi_n[0,t]} f(A_{t_k})(A_{t_{k+1}} - A_{t_k})$$

converge in probability to

$$\int_0^t f(A_{s-})dA_s$$

for a continuous function f and an FV process A. This leads to the more general change of variables formula, valid for any FV process A, and $f \in C^1$, namely

$$f(A_t) - f(A_0) = \int_0^t f'(A_{s-})dA_s + \sum_{0<s\le t} \{f(A_s) - f(A_{s-}) - f'(A_{s-})\Delta A_s\}.$$

The next corollary explains why Theorem 54 is known as a change of variables formula. The proof is an immediate application of Theorem 54.

Corollary. Let g be continuous, and let $h(t) = \int_0^t g(u)du$. Let A be an FV process with continuous paths. Then

$$\begin{aligned} h(A_t) - h(A_0) &= \int_{A_0}^{A_t} g(u)du \\ &= \int_0^t g(A_s)dA_s. \end{aligned}$$

We conclude this section with an example.

Example. Let N be a Poisson process with parameter λ. Then $M_t = N_t - \lambda t$, the **compensated Poisson process**, is a martingale, as well as an FV process. For a bounded (say), jointly measurable process H, we have

$$\begin{aligned} I_t = \int_0^t H_s dM_s = \int_0^t H_s d(N_s - \lambda s) &= \int_0^t H_s dN_s - \lambda \int_0^t H_s ds \\ &= \sum_{n=1}^{\infty} H_{T_n} 1_{\{t \ge T_n\}} - \lambda \int_0^t H_s ds \end{aligned}$$

where $(T_n)_{n\ge1}$ are the arrival times of the Poisson process N. Now suppose the process H is bounded, adapted, and has continuous sample paths. For $0 \le s < t < \infty$, we then have

$$
\begin{aligned}
E\{I_t - I_s|\mathcal{F}_s\} &= E\{\int_s^t H_u dM_u|\mathcal{F}_s\} \\
&= E\{\lim_{n\to\infty} \sum_{t_k,t_{k+1}\in\pi_n} H_{t_k}(M_{t_{k+1}} - M_{t_k})|\mathcal{F}_s\} \\
&= \lim_{n\to\infty} \sum_{t_k,t_{k+1}\in\pi_n} E\{E\{H_{t_k}(M_{t_{k+1}} - M_{t_k})|\mathcal{F}_{t_k}\}|\mathcal{F}_s\} \\
&= 0.
\end{aligned}
$$

The interchange of limits can be justified by the Dominated Convergence Theorem. We conclude that the integral process I is a martingale. This fact, that the stochastic Stieltjes integral of an adapted, bounded, continuous process with respect to a martingale is again a martingale, is true in much greater generality. We shall treat this systematically in Chap. II.

8 Naïve Stochastic Integration Is Impossible

In Sect. 7 we saw that for an FV process A, and a continuous integrand process H, we could express the integral $\int_0^t H_s dA_s$ as the limit of sums (Theorem 53). The Brownian motion process, B, however, has paths of infinite variation on compacts. In this section we demonstrate, with the aid of the Banach-Steinhaus Theorem[7], some of the difficulties that are inherent in trying to extend the notion of Stieltjes integration to processes that have paths of unbounded variation, such as Brownian motion. For the reader's convenience we recall the Banach-Steinhaus Theorem.

Theorem 55. *Let X be a Banach space and let Y be a normed linear space. Let $\{T_\alpha\}$ be a family of bounded linear operators from X into Y. If for each $x \in X$ the set $\{T_\alpha x\}$ is bounded, then the set $\{\|T_\alpha\|\}$ is bounded.*

Let us put aside stochastic processes for the moment. Let $x(t)$ be a right continuous function on $[0,1]$, and let π_n be a refining sequence of dyadic rational partitions of $[0,1]$ with $\lim_{n\to\infty} \text{mesh}(\pi_n) = 0$. We ask the question "What conditions on x are needed so that the sums

$$
S_n = \sum_{t_k,t_{k+1}\in\pi_n} h(t_k)(x(t_{k+1}) - x(t_k)) \qquad (*)
$$

converge to a finite limit as $n \to \infty$ for all continuous functions h?" From Theorem 53 of Sect. 7 we know that x of finite variation is sufficient. However, it is also necessary.

Theorem 56. *If the sums S_n of $(*)$ converge to a limit for every continuous function h then x is of finite variation.*

[7] The Banach-Steinhaus Theorem is also known as the Principle of Uniform Boundedness.

Proof. Let X be the Banach space of continuous functions equipped with the supremum norm. Let Y be $\mathbb{R}$, equipped with absolute value as the norm. For $h \in X$, let

$$T_n(h) = \sum_{t_k, t_{k+1} \in \pi_n} h(t_k)(x(t_{k+1}) - x(t_k)).$$

For each fixed n it is simple to construct an h in X such that $h(t_k) = \operatorname{sign}\{x(t_{k+1}) - x(t_k)\}$, and $\|h\| = 1$. For such an h we have

$$T_n(h) = \sum_{t_k, t_{k+1} \in \pi_n} |x(t_{k+1}) - x(t_k)|.$$

Therefore

$$\|T_n\| \geq \sum_{t_k, t_{k+1} \in \pi_n} |x(t_{k+1}) - x(t_k)|,$$

each n, and $\sup_n \|T_n\| \geq$ total variation of x. On the other hand for each $h \in X$ we have $\lim_{n\to\infty} T_n(h)$ exists and therefore $\sup_n \|T_n(h)\| < \infty$. The Banach-Steinhaus Theorem then implies that $\sup_n \|T_n\| < \infty$, hence the total variation of x is finite. □

Returning to stochastic processes, we might hope to circumvent the limitations imposed by Theorem 56 by appealing to convergence in probability. That is, if $X(s, \omega)$ is right continuous (or even continuous), can we have the sums

$$\sum_{t_k, t_{k+1} \in \pi_n[0,1]} H_{t_k}(X_{t_{k+1}} - X_{t_k}) \qquad (**)$$

converging to a limit *in probability* for every continuous process H? Unfortunately the answer is that X must still have paths of finite variation, a.s. The reason is that one can make the procedure used in the proof of Theorem 56 measurable in ω, and hence a subsequence of the sums in $(**)$ can be made to converge a.s. to $+\infty$ on the set where X is not of finite variation. If this set has positive probability, the sums cannot converge in probability.

The preceding discussion makes it appear impossible to develop a coherent notion of a stochastic integral $\int_0^t H_s dX_s$ when X is a process with paths of infinite variation on compacts; for example a Brownian motion. Nevertheless this is precisely what we will do in Chap. II.

Bibliographic Notes

The basic definitions and notation presented here have become fundamental to the modern study of stochastic processes, and they can be found many places, such as Dellacherie-Meyer [45], Doob [55], and Jacod-Shiryaev [110]. Theorem 3 is true in much greater generality. For example the hitting time of a Borel set is a stopping time. This result is very difficult, and proofs can be found in Dellacherie [41] or [42].

The résumé of martingale theory consists of standard theorems. The reader does not need results from martingale theory beyond what is presented here. Those proofs not given can be found in many places, for example Breiman [23], Dellacherie-Meyer [46], or Ethier-Kurtz [71].

The Poisson process and Brownian motion are the two most important stochastic processes for the theory of stochastic integration. Our treatment of the Poisson process follows Çinlar [33]. Theorem 25 is in Bremaud [24], and is due to J. de Sam Lazaro. The facts about Brownian motion needed for the theory of stochastic integration are the only ones presented here. A good source for more detailed information on Brownian motion is Revuz-Yor [208], or Hida [88].

Lévy processes (processes with stationary and independent increments) are a crucial source of examples for the theory of semimartingales and stochastic integrals. Indeed in large part the theory is abstracted from the properties of these processes. There do not seem to be many presentations of Lévy processes concerned with their properties which are relevant to stochastic integration beyond that of Jacod-Shiryaev [110]. However, one can consult the books of Bertoin [12], Rogers-Williams [210], and Revuz-Yor [208]. Our approach is inspired by Bretagnolle [25]. Local martingales were first proposed by K. Itô and S. Watanabe [102] in order to generalize the Doob-Meyer decomposition. Standard Stieltjes integration applied to finite variation stochastic processes was not well known before the fundamental work of Meyer [171]. Finally, the idea of using the Banach-Steinhaus Theorem to show that naïve stochastic integration is impossible is due to Meyer [178].

Exercises for Chapter I

For all of these exercises, we assume as given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses.

Exercise 1. Let S, T be stopping times, $S \leq T$ a.s. Show $\mathcal{F}_S \subset \mathcal{F}_T$.

Exercise 2. Give an example where S, T are stopping times, $S \leq T$, but $T - S$ is not a stopping time.

Exercise 3. Let $(T_n)_{n\geq 1}$ be a sequence of stopping times. Show that $\sup_n T_n$, $\inf_n T_n$, $\limsup_{n\to\infty} T_n$, and $\liminf_{n\to\infty} T_n$ are all stopping times.

Exercise 4. Suppose $(T_n)_{n\geq 1}$ is a sequence of stopping times decreasing to a random time T. Show that T is a stopping time, and moreover that $\mathcal{F}_T = \bigcap_n \mathcal{F}_{T_n}$.

Exercise 5. Let $p > 1$ and let M^n be a sequence of continuous martingales (that is, martingales whose paths are continuous, a.s.) with $M^n_\infty \in L^p$, each n. Suppose $M^n_\infty \to X$ in L^p and let $M_t = E\{X|\mathcal{F}_t\}$.

(a) Show $M_t \in L^p$, all $t \geq 0$.

(b) Show M is a continuous martingale.

Exercise 6. Let $N = (N_t)_{t\geq 0}$ be Poisson with intensity λ. Suppose λt is an integer. Show that $E\{|N_t - \lambda t|\} = \frac{2(\lambda t)^{\lambda t} e^{-\lambda t}}{(\lambda t-1)!}$.

Exercise 7. Let $N = (N_t)_{t\geq 0}$ be Poisson with intensity λ. Show that N is continuous in L^2 (and hence in probability), but of course N does not have continuous sample paths.

Exercise 8. Let $B_t = (B_t^1, B_t^2, B_t^3)$ be three dimensional Brownian motion (that is, the B^i are i.i.d. one dimensional Brownian motions, $1 \leq i \leq 3$). Let $\tau_\alpha = \sup_t\{\|B_t\| \leq \alpha\}$. The time τ_α is known as a **last exit time**. Show that if $\alpha > 0$, then τ_α is not a stopping time for the natural filtration of B.

Exercise 9. Let $(N_t^i)_{t\geq 0}$ be a sequence of i.i.d. Poisson processes with parameter $\lambda = 1$. Let $M_t^i = \frac{1}{i}(N_t^i - t)$, and let $M = (\sum_{i=1}^{\infty} M_t^i)_{t\geq 0}$.

(a) Show that M is well-defined (i.e., that the series converges) in the L^2 sense.
*(b) Show that for any $t > 0$, $\sum_{s\leq t} \Delta M_s = \infty$ a.s., where $\Delta M_s = M_s - M_{s-}$ is the jump of M at time s.

Exercise 10. Let $(N_t^i)_{t\geq 0}$ and $(L_t^j)_{t\geq 0}$be two independent sequences of i.i.d. Poisson processes of parameters $\lambda = 1$. Let $M_t^i = \frac{1}{i}(N_t^i - L_t^i)$.

(a) Show that M is a martingale and that M changes only by jumps.
*(b) Show that $\sum_{s\leq t} |\Delta M_s| = \infty$ a.s., and $t > 0$.

Exercise 11. Show that a compound Poisson process is a Lévy process.

Exercise 12. Let Z be a compound Poisson process, with $E\{|U_1|\} < \infty$. Show that $Z_t - E\{U_1\}\lambda t$ is a martingale.

Exercise 13. Let $Z = (Z_t)_{t\geq 0}$ be a Lévy process which is a martingale, but Z has no Brownian component (that is, the constant $\sigma^2 = 0$ in the Lévy-Khintchine formula), and a finite Lévy measure ν. Let $\lambda = \nu(\mathbb{R})$. Show that Z is a compensated compound Poisson process with arrival intensity λ and i.i.d. jumps with distribution $\mu = \frac{1}{\lambda}\nu$.

***Exercise 14.** Let $0 < T_1 < T_2 < \cdots < T_n < \cdots$ be any increasing sequence of finite-valued stopping times, increasing to ∞. Let $N_t = \sum_{i=1}^{\infty} 1_{\{t\geq T_i\}}$, and let $(U_i)_{i\geq 1}$ be independent, and also the $(U_i)_{i\geq 1}$ are independent of the process N. Show that if $\sup_i E\{|U_i|\} < \infty$, all i, and $E\{U_i\} = 0$, all i, then $Z_t = \sum_{i=1}^{\infty} U_i 1_{\{t\geq T_i\}}$ is a martingale. (*Note:* We are not assuming that N is a Poisson process here. Furthermore, the random variables $(U_i)_{i\geq 1}$, while independent, are not necessarily identically distributed.)

Exercise 15. Let Z be a Lévy process which is a martingale, but with no Brownian component and with a Lévy measure of the form

$$\nu(dx) = \sum_{k=1}^{\infty} \alpha_k \varepsilon_{\beta_k}(dx),$$

where $\varepsilon_{\beta_k}(dx)$ denotes point mass at $\beta_k \in \mathbb{R}$ of size one, with $\sum_{k=1}^{\infty} \beta_k^2 \alpha_k < \infty$. Show that Z is of the form

$$Z_t = \sum_{k=1}^{\infty} \beta_k (N_t^k - \alpha_k t)$$

where $(N_t^k)_{t \geq 0}$ is an independent sequence of Poisson processes with parameters $(\alpha_k)_{k \geq 1}$. Verify that Z is an L^2 martingale.

Exercise 16. Let $B = (B_t)_{0 \leq t \leq 1}$ be one dimensional Brownian motion on $[0,1]$. Let $W_t = B_{1-t} - B_1$, $0 \leq t \leq 1$. Show that W has the same distribution as does B, and that W is a Brownian motion for $0 \leq t \leq 1$ for its natural filtration.

Exercise 17. Let X be any process with càdlàg paths. Set $\varepsilon > 0$.

(a) Show that on $[0,t]$ there are only a finite number of jumps larger than ε. That is, if $N_t = \sum_{s \leq t} 1_{\{|\Delta X_s| \geq \varepsilon\}}$, where $\Delta X_s = X_s - X_{s-}$, then $P(N_t < \infty) = 1$, each $t > 0$.
(b) Conclude that X can have only a countable number of jumps on $[0,t]$.

Exercise 18. Let Z be a Lévy process, and let $\varepsilon > 0$. Let $\Delta Z_s = Z_s - Z_{s-}$, and set $J_t^{\varepsilon} = \sum_{s \leq t} \Delta Z_s 1_{\{|\Delta Z_s| \geq \varepsilon\}}$. Show that $Z - J^{\varepsilon}$ and J^{ε} are independent Lévy processes.

Exercise 19. Let X be an adapted process with continuous paths. Show that X is locally bounded.

***Exercise 20.** Give an example of an adapted process X with $X_0 = 0$ that has càdlàg paths which are *not* locally bounded.

Exercise 21. (This exercise shows that 2 is the best constant in Doob's quadratic martingale inequality.) Let $(\Omega, \mathcal{F}, P)$ be Lebesgue measure on $(0,1)$ with the Borel sets. Define the filtration

$$\mathcal{F}_t = \{\Lambda \in \mathcal{F} : (t,1) \subset \Lambda, \text{ or } \Lambda \subset (0,t]\}, \quad 0 < t < 1,$$

and $\mathcal{F}_t = \mathcal{F}$ for $t \geq 1$. Intuitively, with this filtration we know all about what is going on during $(0,t]$, but nothing about what is happening during $(t,1)$.

(a) Show that for $Y \in L^1$, the martingale $M_t = E\{Y|\mathcal{F}_t\}$ is given by

$$M_t(\omega) = \begin{cases} Y(\omega), & 0 < \omega \leq t, \\ \dfrac{1}{1-t}\displaystyle\int_t^1 Y(s)ds, & t < \omega < 1. \end{cases}$$

(b) Let $Y(\omega) = (1-\omega)^{-\alpha}$ for some α, $0 < \alpha < 1/2$. Show that $Y \in L^2$ and also if $M_t = E\{Y|\mathcal{F}_t\}$, then

$$M_t(\omega) = \frac{1}{(1-\alpha)}\frac{1}{(1-t)^\alpha} = \frac{1}{1-\alpha}Y(t), \quad t < \omega < 1.$$

(c) Show $\sup_t M_t = \frac{1}{1-\alpha}Y$, ω-by-ω, and deduce

$$\|\sup_t M\|_{L^2} = \frac{1}{1-\alpha}\|Y\|_{L^2} = \frac{1}{1-\alpha}\|M_\infty\|_{L^2}.$$

Note that $\lim_{\alpha\to 1/2}(1-\alpha)^{-1} = 2$.

Exercise 22. (This exercise gives an example of a martingale with each path bounded but which nevertheless is not locally bounded.) Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be as in Exercise 21. Let $Y(\omega) = \frac{1}{\sqrt{\omega}}$, and $M_t = E\{Y|\mathcal{F}_t\}$.

(a) Show that

$$M_t(\omega) = \frac{1}{\sqrt{\omega}}1_{(0,t]}(\omega) + \frac{2}{1+\sqrt{t}}1_{(t,1)}(\omega).$$

(b) Let T be a stopping time, T not identically 0. Show that $T(\omega) \geq \omega$ on some ω-interval $(0, \varepsilon)$ with $\varepsilon > 0$.
(c) Use (b) to show $E\{M_T^2\} \geq \int_0^\varepsilon \frac{1}{s}ds = \infty$, and deduce that M cannot be locally an L^2 martingale.
(d) Conclude that M has bounded paths but that $M_T 1_{\{T>0\}}$ is unbounded for every stopping time T not identically 0.

***Exercise 23.** Find an example of a local martingale M and finite stopping times S, T with $S \leq T$ a.s. such that $E\{|M_T||\mathcal{F}_S\} < \infty$, but $M_S \neq E\{M_T|\mathcal{F}_S\}$ a.s. (*Hint:* Begin by showing that the equality would imply that every positive local martingale is a martingale.)

Exercise 24. (This exercise helps to clarify the difference between stopping time σ-algebras of the two forms $\mathcal{G}_{T-}$ and $\mathcal{G}_T$.) Let Z be a Lévy process, and let $\mathcal{G}_t = \sigma\{Z_s; s \leq t\} \vee \mathcal{N}$, where $\mathcal{N}$ are all the null sets in the given complete probability space $(\Omega, \mathcal{F}, P)$. For a stopping time T define $\mathcal{G}_{T-} = \sigma\{A \cap (t < T); A \in \mathcal{G}_t\}$; at time 0, set $\mathcal{G}_{0-} = \mathcal{G}_0$. For a process X, let $X^{T-} = X_t 1_{\{t<T\}} + X_{T-}1_{\{t\geq T\}}$, and take $X_{0-} = 0$.

(a) Show $\mathcal{G}_{T-} = \sigma\{T, Z^{T-}\} \vee \mathcal{N}$.
(b) Show $\mathcal{G}_T = \sigma\{T, Z^T\} \vee \mathcal{N}$.
(c) Show $\mathcal{G}_T = \sigma\{\mathcal{G}_{T-}, Z_T\}$ if T is a.s. finite.
(d) Conclude that for a bounded time T if $Z_T = Z_{T-}$ a.s., then $\mathcal{G}_T = \mathcal{G}_{T-}$.

(Note that if $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ is the completed natural filtration of a Brownian motion, then $\mathcal{G}_T = \mathcal{G}_{T-}$ for all bounded stopping times.)

Exercise 25. Let Z be a Lévy process (or any process which is continuous in probability). Show that the probability Z jumps at a given fixed time t is zero.

Exercise 26. Use Exercises 24 and 25 to show that with the assumptions of Exercise 24, for any fixed time t, $\mathcal{G}_{t-} = \mathcal{G}_t$, giving a continuity property of the filtration for fixed times (since by Theorem 31 we always have $\mathcal{G}_{t+} = \mathcal{G}_t$), but of course not for stopping times.

Exercise 27. Let S, T be two stopping times with $S \leq T$. Show $\mathcal{F}_{S-} \subset \mathcal{F}_{T-}$. Moreover if $(T_n)_{n\geq 1}$ is any increasing sequence of stopping times with $\lim_{n\to\infty} T_n = T$, show that

$$\mathcal{F}_{T-} = \bigvee_n \mathcal{F}_{T_n-}.$$

Exercise 28. Prove the first equality of Theorem 38 when $f1_A \in L^1(d\nu)$.

****Exercise 29.** Show that if Z is a Lévy process and a local martingale, then Z is a martingale. That is, *all Lévy process local martingales are actually martingales.*

Exercise 30 (reflection principle for Lévy processes). Let Z be a **symmetric Lévy process** with $Z_0 = 0$. That is, Z and $-Z$ have the same distribution. Let $S_t = \sup_{s\leq t} Z_s$. Show that

$$P(S_t \geq z; Z_t < z - y) \leq P(Z_t > z + y).$$

Exercise 31. Let Z be a symmetric Lévy process with $Z_0 = 0$ as in Exercise 30. Show that

$$P(S_t \geq z) \leq 2P(Z_t \geq z).$$

II

Semimartingales and Stochastic Integrals

1 Introduction to Semimartingales

The purpose of the theory of *stochastic integration* is to give a reasonable meaning to the idea of a differential to as wide a class of stochastic processes as possible. We saw in Sect. 8 of Chap. I that using Stieltjes integration on a path-by-path basis excludes such fundamental processes as Brownian motion, and martingales in general. Markov processes also in general have paths of unbounded variation and are similarly excluded. Therefore we must find an approach more general than Stieltjes integration.

We will define stochastic integrals first as a limit of sums. *A priori* this seems hopeless, since even by restricting our integrands to *continuous* processes we saw as a consequence of Theorem 56 of Chap. I that the differential must be of finite variation on compacts. However an analysis of the proof of Theorem 56 offers some hope. In order to construct a function h such that $h(t_k) = \operatorname{sign}(x(t_{k+1}) - x(t_k))$, we need to be able to "see" the trajectory of x on $(t_k, t_{k+1}]$. The idea of K. Itô was to restrict the integrands to those that could not see into the future increments, namely adapted processes.

The foregoing considerations lead us to define the stochastic processes that will serve as differentials as those that are "good integrators" on an appropriate class of adapted processes. We will, as discussed in Chap. I, assume that we are given a filtered, complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the *usual hypotheses.*

Definition. A process H is said to be **simple predictable** if H has a representation

$$H_t = H_0 1_{\{0\}}(t) + \sum_{i=1}^{n} H_i 1_{(T_i, T_{i+1}]}(t)$$

where $0 = T_1 \leq \cdots \leq T_{n+1} < \infty$ is a finite sequence of stopping times, $H_i \in \mathcal{F}_{T_i}$ with $|H_i| < \infty$ a.s., $0 \leq i \leq n$. The collection of simple predictable processes is denoted $\mathbf{S}$.

Note that we can take $T_1 = T_0 = 0$ in the above definition, so there is no "gap" between T_0 and T_1. We can topologize $\mathbf{S}$ by uniform convergence in (t, ω), and we denote $\mathbf{S}$ endowed with this topology by $\mathbf{S}_u$. We also write $\mathbf{L}^0$ for the space of finite-valued random variables *topologized by convergence in probability.*

Let X be a stochastic process. An operator, I_X, induced by X should have two fundamental properties to earn the name "integral." The operator I_X should be linear, and it should satisfy some version of the Bounded Convergence Theorem. A particularly weak form of the Bounded Convergence Theorem is that the uniform convergence of processes H^n to H implies only the convergence in probability of $I_X(H^n)$ to $I_X(H)$.

Inspired by the above considerations, for a given process X we define a linear mapping $I_X : \mathbf{S} \to \mathbf{L}^0$ by letting

$$I_X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i (X_{T_{i+1}} - X_{T_i}),$$

where $H \in \mathbf{S}$ has the representation

$$H_t = H_0 1_{\{0\}} + \sum_{i=1}^{n} H_i 1_{(T_i, T_{i+1}]}.$$

Since this definition is a path-by-path definition for the step functions $H(\omega)$, it does not depend on the choice of the representation of H in $\mathbf{S}$.

Definition. A process X is a **total semimartingale** if X is càdlàg, adapted, and $I_X : \mathbf{S}_u \to \mathbf{L}^0$ is continuous.

Recall that for a process X and a stopping time T, the notation X^T denotes the process $(X_{t \wedge T})_{t \geq 0}$.

Definition. A process X is called a **semimartingale** if, for each $t \in [0, \infty)$, X^t is a total semimartingale.

With our definition of a semimartingale, the second fundamental property we want (bounded convergence) holds. We postpone consideration of examples of semimartingales to Sect. 3.

2 Stability Properties of Semimartingales

We state a sequence of theorems giving some of the stability results which are particularly simple.

Theorem 1. *The set of (total) semimartingales is a vector space.*

Proof. This is immediate from the definition. □

Theorem 2. *If Q is a probability and absolutely continuous with respect to P, then every (total) P semimartingale X is a (total) Q semimartingale.*

Proof. Convergence in P-probability implies convergence in Q-probability. Thus the theorem follows from the definition of X. □

Theorem 3. *Let $(P_k)_{k\geq 1}$ be a sequence of probabilities such that X is a P_k semimartingale for each k. Let $R = \sum_{k=1}^{\infty} \lambda_k P_k$, where $\lambda_k \geq 0$, each k, and $\sum_{k=1}^{\infty} \lambda_k = 1$. Then X is a semimartingale under R as well.*

Proof. Suppose $H^n \in \mathbf{S}$ converges uniformly to $H \in \mathbf{S}$. Since X is a P_k semimartingale for all P_k, $I_X(H^n)$ converges to $I_X(H)$ in probability for every P_k. This then implies $I_X(H^n)$ converges to $I_X(H)$ under R. □

Theorem 4 (Stricker's Theorem). *Let X be a semimartingale for the filtration $\mathbb{F}$. Let $\mathbb{G}$ be a subfiltration of $\mathbb{F}$, such that X is adapted to the $\mathbb{G}$ filtration. Then X is a $\mathbb{G}$ semimartingale.*

Proof. For a filtration $\mathbb{H}$, let $\mathbf{S}(\mathbb{H})$ denote the simple predictable processes for the filtration $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$. In this case we have $\mathbf{S}(\mathbb{G})$ is contained in $\mathbf{S}(\mathbb{F})$. The theorem is then an immediate consequence of the definition. □

Theorem 4 shows that we can always shrink a filtration and preserve the property of being a semimartingale (as long as the process X is still adapted), since we are shrinking as well the possible integrands; this, in effect, makes it "easier" for the process X to be a semimartingale. Expanding the filtration, therefore, should be—and is—a much more delicate issue. Expansion of filtrations is considered in much greater detail in Chap. VI. We present here an elementary but useful result. Recall that we are given a filtered space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses.

Theorem 5 (Jacod's Countable Expansion). *Let $\mathcal{A}$ be a collection of events in $\mathcal{F}$ such that if A_α, $A_\beta \in \mathcal{A}$ then $A_\alpha \cap A_\beta = \emptyset$, $(\alpha \neq \beta)$. Let $\mathcal{H}_t$ be the filtration generated by $\mathcal{F}_t$ and $\mathcal{A}$. Then every $(\mathbb{F}, P)$ semimartingale is an $(\mathbb{H}, P)$ semimartingale also.*

Proof. Let $A_n \in \mathcal{A}$. If $P(A_n) = 0$, then A_n and A_n^c are in $\mathcal{F}_0$ by hypothesis. We assume, therefore, that $P(A_n) > 0$. Note that there can be at most a countable number of $A_n \in \mathcal{A}$ such that $P(A_n) > 0$. If $\Lambda = \bigcup_{n\geq 1} A_n$ is the union of all $A_n \in \mathcal{A}$ with $P(A_n) > 0$, we can also add Λ^c to $\mathcal{A}$ without loss of generality. Thus we can assume that $\mathcal{A}$ is a countable partition of Ω with $P(A_n) > 0$ for every $A_n \in \mathcal{A}$. Define a new probability Q_n by $Q_n(\cdot) = P(\cdot|A_n)$, for A_n fixed. Then $Q_n \ll P$, and X is a $(\mathbb{F}, Q_n)$ semimartingale by Theorem 2. If we enlarge the filtration $\mathbb{F}$ by all the $\mathcal{F}$ measurable events that have Q_n-probability 0 or 1, we get a larger filtration $\mathbb{J}^n = (\mathcal{J}_t^n)_{t\geq 0}$, and X is a $(\mathbb{J}^n, Q_n)$ semimartingale. Since $Q_n(A_m) = 0$ or 1 for $m \neq n$, we have $\mathcal{F}_t \subset \mathcal{H}_t \subset \mathcal{J}_t^n$, for $t \geq 0$, and for all n. By Stricker's Theorem (Theorem 4) we conclude that X is an $(\mathbb{H}, Q_n)$ semimartingale. Finally, we have $dP = \sum_{n\geq 1} P(A_n) dQ_n$, where X is an $(\mathbb{H}, Q_n)$ semimartingale for each n. Therefore by Theorem 3 we conclude X is an $(\mathbb{H}, P)$ semimartingale. □

Corollary. Let $\mathcal{A}$ be a finite collection of events in $\mathcal{F}$, and let $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ be the filtration generated by $\mathcal{F}_t$ and $\mathcal{A}$. Then every $(\mathbb{F}, P)$ semimartingale is an $(\mathbb{H}, P)$ semimartingale also.

Proof. Since $\mathcal{A}$ is finite, one can always find a (finite) partition Π of Ω such that $\mathcal{H}_t = \mathcal{F}_t \vee \Pi$. The corollary then follows by Theorem 5. □

Note that if $B = (B_t)_{t\geq 0}$ is a Brownian motion for a filtration $(\mathcal{F}_t)_{t\geq 0}$, by Theorem 5 we are able to add, in a certain manner, an infinite number of "future" events to the filtration and B will no longer be a martingale, but it will stay a semimartingale. This has interesting implications in finance theory (the theory of continuous trading). See for example Duffie-Huang [60].

The corollary of the next theorem states that being a semimartingale is a "local" property; that is, a local semimartingale is a semimartingale. We get a stronger result by stopping at T_n- rather than at T_n in the next theorem. A process X is stopped at $T-$ if $X_t^{T-} = X_t 1_{\{0\leq t<T\}} + X_{T-}1_{\{t\geq T\}}$, where $X_{0-} = 0$.

Theorem 6. *Let X be a càdlàg, adapted process. Let (T_n) be a sequence of positive r.v. increasing to ∞ a.s., and let (X^n) be a sequence of semimartingales such that for each n, $X^{T_n-} = (X^n)^{T_n-}$. Then X is a semimartingale.*

Proof. We wish to show X^t is a total semimartingale, each $t > 0$. Define $R_n = T_n 1_{\{T_n\leq t\}} + \infty 1_{\{T_n>t\}}$. Then

$$P\{|I_{X^t}(H)| \geq c\} \leq P\{|I_{(X^n)^t}(H)| \geq c\} + P(R_n < \infty).$$

But $P(R_n < \infty) = P(T_n \leq t)$, and since T_n increases to ∞ a.s., $P(T_n \leq t) \to 0$ as $n \to \infty$. Thus if H^k tends to 0 in $\mathbf{S}_u$, given $\varepsilon > 0$, we choose n so that $P(R_n < \infty) < \varepsilon/2$, and then choose k so large that $P\{|I_{(X^n)^t}(H^k)| \geq c\} < \varepsilon/2$. Thus, for k large enough, $P\{|I_{X^t}(H^k)| \geq c\} < \varepsilon$. □

Corollary. Let X be a process. If there exists a sequence (T_n) of stopping times increasing to ∞ a.s., such that X^{T_n} (or $X^{T_n}1_{\{T_n>0\}}$) is a semimartingale, each n, then X is also a semimartingale.

3 Elementary Examples of Semimartingales

The elementary properties of semimartingales established in Sect. 2 will allow us to see that many common processes are semimartingales. For example, the Poisson process, Brownian motion, and more generally all Lévy processes are semimartingales.

Theorem 7. *Each adapted process with càdlàg paths of finite variation on compacts (of finite total variation) is a semimartingale (a total semimartingale).*

Proof. It suffices to observe that $|I_X(H)| \leq \|H\|_u \int_0^\infty |dX_s|$, where $\int_0^\infty |dX_s|$ denotes the Lebesgue-Stieltjes total variation and $\|H\|_u = \sup_{(t,w)} |H(t,w)|$. □

Theorem 8. *Each L^2 martingale with càdlàg paths is a semimartingale.*

Proof. Let X be an L^2 martingale with $X_0 = 0$, and let $H \in \mathbf{S}$. Using Doob's Optional Sampling Theorem and the L^2 orthogonality of the increments of L^2 martingales, it suffices to observe that

$$\begin{aligned} E\{(I_X(H))^2\} &= E\{(\sum_{i=0}^n H_i(X_{T_{i+1}} - X_{T_i}))^2\} = E\{\sum_{i=0}^n H_i^2(X_{T_{i+1}} - X_{T_i})^2\} \\ &\leq \|H\|_u^2 E\{\sum_{i=0}^n (X_{T_{i+1}} - X_{T_i})^2\} = \|H\|_u^2 E\{\sum_{i=0}^n (X_{T_{i+1}}^2 - X_{T_i}^2)\} \\ &= \|H\|_u^2 E\{X_{T_{n+1}}^2\} \leq \|H\|_u^2 E\{X_\infty^2\}. \end{aligned}$$ □

Corollary 1. Each càdlàg, locally square integrable local martingale is a semimartingale.

Proof. Apply Theorem 8 together with the corollary to Theorem 6. □

Corollary 2. A local martingale with continuous paths is a semimartingale.

Proof. Apply Corollary 1 together with Theorem 51 in Chap. I. □

Corollary 3. The Wiener process (that is, Brownian motion) is a semimartingale.

Proof. The Wiener process B_t is a martingale with continuous paths if B_0 is integrable. It is always a continuous local martingale. □

Definition. We will say an adapted process X with càdlàg paths is **decomposable** if it can be decomposed $X_t = X_0 + M_t + A_t$, where $M_0 = A_0 = 0$, M is a locally square integrable martingale, and A is càdlàg, adapted, with paths of finite variation on compacts.

Theorem 9. *A decomposable process is a semimartingale.*

Proof. Let $X_t = X_0 + M_t + A_t$ be a decomposition of X. Then M is a semimartingale by Corollary 1 of Theorem 8, and A is a semimartingale by Theorem 7. Since semimartingales form a vector space (Theorem 1) we have the result. □

Corollary. A Lévy process is a semimartingale.

Proof. By Theorem 40 of Chap. I we know that a Lévy process is decomposable. Theorem 9 then gives the result. □

Since Lévy processes are prototypic strong Markov processes, one may well wonder if all $\mathbb{R}^n$-valued strong Markov processes are semimartingales. Simple examples, such as $X_t = B_t^{1/3}$, where B is standard Brownian motion, show this is not the case (while this example is simple, the proof that X is not a semimartingale is not elementary[1]). However if one is willing to "regularize" the Markov process by a transformation of the space (in the case of this example using the "scale function" $S(x) = x^3$), "most reasonable" strong Markov processes are semimartingales. Indeed, **Dynkin's formula**, which states that if f is in the domain of the infinitesimal generator G of the strong Markov process Z, then the process

$$M_t = f(Z_t) - f(Z_0) - \int_0^t Gf(Z_s)ds$$

is well-defined and is a local martingale, hints strongly that if the domain of G is rich enough, the process Z is a semimartingale. In this regard see Sect. 7 of Çinlar, Jacod, Protter, and Sharpe [34].

4 Stochastic Integrals

In Sect. 1 we defined semimartingales as adapted, càdlàg processes that acted as "good integrators" on the simple predictable processes. We now wish to enlarge the space of processes we can consider as integrands. In Chap. IV we will consider a large class of processes, namely those that are "predictably measurable" and have appropriate finiteness properties. Here, however, by keeping our space of integrands small—yet large enough to be interesting—we can keep the theory free of technical problems, as well as intuitive.

A particularly nice class of processes for our purposes is the class of adapted processes with left continuous paths that have right limits (the French acronym would be càglàd).

Definition. We let $\mathbb{D}$ denote the space of adapted processes with càdlàg paths, $\mathbb{L}$ denote the space of adapted processes with càglàd paths (left continuous with right limits) and $\mathbf{b}\mathbb{L}$ denote processes in $\mathbb{L}$ with bounded paths.

We have previously considered $\mathbf{S}_u$, the space of simple, predictable processes endowed with the topology of uniform convergence; and $\mathbf{L}^0$, the space of finite-valued random variables topologized by convergence in probability. We need to consider a third type of convergence.

---

[1] See Theorem 71 of Chap. IV which proves a similar assertion for $X_t = |B_t|^\alpha$, $0 < \alpha < 1/2$.

Definition. A sequence of processes $(H^n)_{n\geq 1}$ converges to a process H **uniformly on compacts in probability** (abbreviated *ucp*) if, for each $t > 0$, $\sup_{0\leq s\leq t} |H_s^n - H_s|$ converges to 0 in probability.

We write $H_t^* = \sup_{0\leq s\leq t} |H_s|$. Then if $Y^n \in \mathbb{D}$ we have $Y^n \to Y$ in *ucp* if $(Y^n - Y)_t^*$ converges to 0 in probability for each $t > 0$. We write $\mathbb{D}_{ucp}$, $\mathbb{L}_{ucp}$, $\mathbf{S}_{ucp}$ to denote the respective spaces endowed with the *ucp* topology. We observe that $\mathbb{D}_{ucp}$ is a metrizable space; indeed, a compatible metric is given, for X, $Y \in \mathbb{D}$, by

$$d(X,Y) = \sum_{n=1}^{\infty} \frac{1}{2^n} E\{\min(1, (X-Y)_n^*)\}.$$

The above metric space $\mathbb{D}_{ucp}$ is complete. For a semimartingale X and a process $H \in \mathbf{S}$, we have defined the appropriate notion $I_X(H)$ of a stochastic integral. The next result is key to extending this definition.

Theorem 10. *The space* $\mathbf{S}$ *is dense in* $\mathbb{L}$ *under the ucp topology.*

Proof. Let $Y \in \mathbb{L}$. Let $R_n = \inf\{t : |Y_t| > n\}$. Then R_n is a stopping time and $Y^n = Y^{R_n} 1_{\{R_n > 0\}}$ are in $\mathbf{b}\mathbb{L}$ and converge to Y in *ucp*. Thus $\mathbf{b}\mathbb{L}$ is dense in $\mathbb{L}$. Without loss we now assume $Y \in \mathbf{b}\mathbb{L}$. Define Z by $Z_t = \lim_{\substack{u\to t\\ u>t}} Y_u$. Then $Z \in \mathbb{D}$. For $\varepsilon > 0$, define

$$\begin{aligned} T_0^\varepsilon &= 0 \\ T_{n+1}^\varepsilon &= \inf\{t : t > T_n^\varepsilon \text{ and } |Z_t - Z_{T_n^\varepsilon}| > \varepsilon\}. \end{aligned}$$

Since Z is càdlàg, the T_n^ε are stopping times increasing to ∞ a.s. as n increases. Let $Z^\varepsilon = \sum_n Z_{T_n^\varepsilon} 1_{[T_n^\varepsilon, T_{n+1}^\varepsilon)}$, for each $\varepsilon > 0$. Then Z^ε are bounded and converge uniformly to Z as ε tends to 0. Let $U^\varepsilon = Y_0 1_{\{0\}} + \sum_n Z_{T_n^\varepsilon} 1_{(T_n^\varepsilon, T_{n+1}^\varepsilon]}$. Then $U^\varepsilon \in \mathbf{b}\mathbb{L}$ and the preceding implies U^ε converges uniformly on compacts to $Y_0 1_{\{0\}} + Z_- = Y$.

Finally, define

$$Y^{n,\varepsilon} = Y_0 1_{\{0\}} + \sum_{i=1}^{n} Z_{T_i^\varepsilon} 1_{(T_i^\varepsilon \wedge n, T_{i+1}^\varepsilon \wedge n]}$$

and this can be made arbitrarily close to $Y \in \mathbf{b}\mathbb{L}$ by taking ε small enough and n large enough. □

We defined a semimartingale X as a process that induced a continuous operator I_X from $\mathbf{S}_u$ into $\mathbf{L}^0$. This I_X maps *processes* into *random variables* (note that X_{T_i} is a random variable). We next define an operator (the stochastic integral operator), induced by X, that will map *processes* into *processes* (note that X^{T_i} is a process).

Definition. For $H \in \mathbf{S}$ and X a càdlàg process, define the (linear) mapping $J_X : \mathbf{S} \to \mathbb{D}$ by

$$J_X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i (X^{T_{i+1}} - X^{T_i})$$

for H in $\mathbf{S}$ with the representation

$$H = H_0 1_{\{0\}} + \sum_{i=1}^{n} H_i 1_{(T_i, T_{i+1}]},$$

$H_i \in \mathcal{F}_{T_i}$ and $0 = T_0 \leq T_1 \leq \cdots \leq T_{n+1} < \infty$ stopping times.

Definition. For $H \in \mathbf{S}$ and X an adapted càdlàg process, we call $J_X(H)$ the **stochastic integral of H with respect to X.**

We use interchangeably three notations for the stochastic integral:

$$J_X(H) = \int H_s dX_s = H \cdot X.$$

Observe that $J_X(H)_t = I_{X^t}(H)$. Indeed, I_X plays the role of a definite integral. For $H \in \mathbf{S}$, $I_X(H) = \int_0^\infty H_s dX_s$.

Theorem 11. *Let X be a semimartingale. Then the mapping $J_X : \mathbf{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous.*

Proof. Since we are only dealing with convergence on compact sets, without loss of generality we take X to be a total semimartingale. First suppose H^k in $\mathbf{S}$ tends to 0 uniformly and is uniformly bounded. We will show $J_X(H^k)$ tends to 0 *ucp*. Let $\delta > 0$ be given and define stopping times T^k by

$$T^k = \inf\{t : |(H^k \cdot X)_t| \geq \delta\}.$$

Then $H^k 1_{[0,T^k]} \in \mathbf{S}$ and tends to 0 uniformly as k tends to ∞. Thus for every t,

$$\begin{aligned} P\{(H^k \cdot X)_t^* > \delta\} &\leq P\{|H^k \cdot X_{T^k \wedge t}| \geq \delta\} \\ &= P\{|(H^k 1_{[0,T^k]} \cdot X)_t| \geq \delta\} \\ &= P\{|I_X(H^k 1_{[0,T^k \wedge t]})| \geq \delta\} \end{aligned}$$

which tends to 0 by the definition of total semimartingale.

We have just shown that $J_X : \mathbf{S}_u \to \mathbb{D}_{ucp}$ is continuous. We now use this to show $J_X : \mathbf{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous. Suppose H^k goes to 0 *ucp*. Let $\delta > 0$, $\varepsilon > 0$, $t > 0$. We have seen that there exists η such that $\|H\|_u \leq \eta$ implies $P(J_X(H)_t^* > \delta) < \frac{\varepsilon}{2}$. Let $R_k = \inf\{s : |H_s^k| > \eta\}$, and set $\widetilde{H}^k = H^k 1_{[0,R_k]} 1_{\{R_k > 0\}}$. Then $\widetilde{H}^k \in \mathbf{S}$ and $\|\widetilde{H}^k\|_u \leq \eta$ by left continuity. Since $R^k \geq t$ implies $(\widetilde{H}^k \cdot X)_t^* = (H^k \cdot X)_t^*$, we have

$$\begin{aligned} P((H^k \cdot X)_t^* > \delta) &\leq P((\widetilde{H}^k \cdot X)_t^* > \delta) + P(R^k < t) \\ &\leq \frac{\varepsilon}{2} + P((H^k)_t^* > \eta) \\ &< \varepsilon \end{aligned}$$

if k is large enough, since $\lim_{k\to\infty} P((H^k)_t^* > \eta) = 0$. □

We have seen that when X is a semimartingale, the integration operator J_X is continuous on $\mathbf{S}_{ucp}$, and also that $\mathbf{S}_{ucp}$ is dense in $\mathbb{L}_{ucp}$. Hence we are able to extend the linear integration operator J_X from $\mathbf{S}$ to $\mathbb{L}$ by continuity, since $\mathbb{D}_{ucp}$ is a complete metric space.

Definition. Let X be a semimartingale. The continuous linear mapping $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ obtained as the extension of $J_X : \mathbf{S} \to \mathbb{D}$ is called the **stochastic integral**.

The preceding definition is rich enough for us to give an *immediate example* of a surprising stochastic integral. First recall that if a process $(A_t)_{t\geq 0}$ has continuous paths of finite variation with $A_0 = 0$, then the Riemann-Stieltjes integral of $\int_0^t A_s dA_s$ yields the formula (see Theorem 54 of Chap. I)

$$\int_0^t A_s dA_s = \frac{1}{2}A_t^2. \tag{$*$}$$

Let us now consider a standard Brownian motion $B = (B_t)_{t\geq 0}$ with $B_0 = 0$. The process B does not have paths of finite variation on compacts, but it *is* a semimartingale. Let (π_n) be a refining sequence of partitions of $[0, \infty)$ with $\lim_{n\to\infty} \mathrm{mesh}(\pi_n) = 0$. Let $B_t^n = \sum_{t_k\in\pi_n} B_{t_k} 1_{(t_k,t_{k+1}]}$. Then $B^n \in \mathbb{L}$ for each n. Moreover, B^n converges to B in *ucp*. Fix $t \geq 0$ and assume that t is a partition point of each π_n. Then

$$J_B(B^n)_t = \sum_{\substack{t_k\in\pi_n \\ t_k<t}} B_{t_k}(B^{t_{k+1}} - B^{t_k})$$

and

$$\begin{aligned} J_B(B)_t = \lim_{n\to\infty} J_B(B^n)_t &= \lim_n \sum_{t_k\in\pi_n, t_k<t} B_{t_k}(B_{t_{k+1}} - B_{t_k}) \\ &= \lim_n \sum_{t_k\in\pi_n, t_k<t} \Big\{\frac{1}{2}(B_{t_{k+1}} + B_{t_k})(B_{t_{k+1}} - B_{t_k}) \\ &\qquad - \frac{1}{2}(B_{t_{k+1}} - B_{t_k})(B_{t_{k+1}} - B_{t_k})\Big\} \\ &= \frac{1}{2}B_t^2 - \frac{1}{2}\lim_n \sum_{t_k\in\pi_n} (B_{t_{k+1}} - B_{t_k})^2 \end{aligned}$$

where we have used telescoping sums. The last term on the right, however, converges a.s. to t, by Theorem 28 of Chap. I. We therefore conclude that

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t, \tag{**}$$

a formula distinctly different from the Riemann-Stieltjes formula (*). The change of variables formula (Theorem 32) presented in Sect. 7, will give a deeper understanding of the difference in formulas (*) and (**).

5 Properties of Stochastic Integrals

Throughout this paragraph X will denote a semimartingale and H will denote an element of $\mathbb{L}$. Recall that the stochastic integral defined in Sect. 4 will be denoted by the three notations $J_X(H) = H \cdot X = \int H_s dX_s$. Evaluating these processes at t, we have

$$H \cdot X_t = \int_0^t H_s dX_s = \int_{[0,t]} H_s dX_s.$$

To exclude 0 in the integral we write

$$\int_{0+}^t H_s dX_s = \int_{(0,t]} H_s dX_s.$$

The integral $\int_0^\infty H_s dX_s$ is defined to be $\lim_{t\to\infty} \int_0^t H_s dX_s$ when the limit exists. Note that $\int_0^t H_s dX_s = H_0 X_0 + \int_{0+}^t H_s dX_s$. For a process $Y \in \mathbb{D}$, we recall that $\Delta Y_t = Y_t - Y_{t-}$, the jump at t. Also since $Y_{0-} = 0$ we have $\Delta Y_0 = Y_0$. Recall further that for a process Z and stopping time T, we let Z^T denote the stopped process $(Z_t^T) = Z_{t\wedge T}$. We will establish in this section several elementary properties of the stochastic integral. These properties will help us to understand the integral and to examine examples at the end of the section. Recall that two processes Y and Z are **indistinguishable** if $P\{\omega : t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ are the same functions$\} = 1$.

Theorem 12. *Let T be a stopping time. Then $(H \cdot X)^T = H1_{[0,T]} \cdot X = H \cdot (X^T)$.*

Theorem 13. *The jump process $\Delta(H\cdot X)_s$ is indistinguishable from $H_s(\Delta X_s)$.*

Proofs. Both properties are clear when $H \in \mathbf{S}$, and they follow when $H \in \mathbb{L}$ by passing to the limit. Indeed, we take convergence in *ucp* for each t, then a.s. convergence by taking a subsequence, and then we choose the rationals (which of course is countable and dense, so the union of null sets is still a null set), and then use the path regularity to get indistinguishability. □

Let Q denote another probability law, and let $H_Q \cdot X$ denote the stochastic integral of H with respect to X computed under the law Q.

Theorem 14. *Let $Q \ll P$. Then $H_Q \cdot X$ is Q indistinguishable from $H_P \cdot X$.*

Proof. Note that by Theorem 2, X is a Q semimartingale. The theorem is clear if $H \in \mathbf{S}$, and it follows for $H \in \mathbb{L}$ by passage to the limit in the *ucp* topology, since convergence in P-probability implies convergence in Q-probability. □

Theorem 15. *Let P_k be a sequence of probabilities such that X is a P_k semimartingale for each k. Let $R = \sum_{k=1}^{\infty} \lambda_k P_k$ where $\lambda_k \geq 0$, each k, and $\sum_{k=1}^{\infty} \lambda_k = 1$. Then $H_R \cdot X = H_{P_k} \cdot X$, P_k-a.s., for all k such that $\lambda_k > 0$.*

Proof. If $\lambda_k > 0$ then $P_k \ll R$, and the result follows by Theorem 14. Note that by Theorem 3 we know that X is an R semimartingale. □

Corollary. Let P and Q be any probabilities and suppose X is a semimartingale relative to both P and Q. Then there exists a process $H \cdot X$ which is a version of both $H_P \cdot X$ and $H_Q \cdot X$.

Proof. Let $R = (P+Q)/2$. Then $H_R \cdot X$ is such a process by Theorem 15. □

Theorem 16. *Let $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ be another filtration such that H is in both $\mathbb{L}(\mathbb{G})$ and $\mathbb{L}(\mathbb{F})$, and such that X is also a $\mathbb{G}$ semimartingale. Then $H_{\mathbb{G}} \cdot X = H_{\mathbb{F}} \cdot X$.*

Proof. $\mathbb{L}(\mathbb{G})$ denotes left continuous processes adapted to the filtration $\mathbb{G}$. As in the proof of Theorem 10, we can construct a sequence of processes H^n converging to H where the construction of the H^n depends only on H. Thus $H^n \in \mathbf{S}(\mathbb{G}) \cap \mathbf{S}(\mathbb{F})$ and converges to H in *ucp*. Since the result is clear for $\mathbf{S}$, the full result follows by passing to the limit. □

Remark. While Theorem 16 is a simple result in this context, it is far from simple if the integrands are predictably measurable processes, rather than processes in $\mathbb{L}$. See the comment following Theorem 33 of Chap. IV.

The next two theorems are especially interesting because they show—at least for integrands in $\mathbb{L}$—that the stochastic integral agrees with the path-by-path Lebesgue-Stieltjes integral, whenever it is possible to do so.

Theorem 17. *If the semimartingale X has paths of finite variation on compacts, then $H \cdot X$ is indistinguishable from the Lebesgue-Stieltjes integral, computed path-by-path.*

Proof. The result is evident for $H \in \mathbf{S}$. Let $H^n \in \mathbf{S}$ converge to H in *ucp*. Then there exists a subsequence n_k such that $\lim_{n_k \to \infty} (H^{n_k} - H)^*_t = 0$ a.s., and the result follows by interchanging limits, justified by the uniform a.s. convergence. □

Theorem 18. *Let X, $\overline{X}$ be two semimartingales, and let H, $\overline{H} \in \mathbb{L}$. Let $A = \{\omega : H_\cdot(\omega) = \overline{H}_\cdot(\omega)$ and $X_\cdot(\omega) = \overline{X}_\cdot(\omega)\}$, and let $B = \{\omega : t \mapsto X_t(\omega)$ is of finite variation on compacts$\}$. Then $H \cdot X = \overline{H} \cdot \overline{X}$ on A, and $H \cdot X$ is equal to a path-by-path Lebesgue-Stieltjes integral on B.*

Proof. Without loss of generality we assume $P(\Lambda) > 0$. Define a new probability law Q by $Q(\Lambda) = P(\Lambda|A)$. Then under Q we have that H and $\overline{H}$ as well as X and $\overline{X}$ are indistinguishable. Thus $H_Q \cdot X = \overline{H}_Q \cdot \overline{X}$, and hence $H \cdot X = \overline{H} \cdot \overline{X}$ P-a.s. on A by Theorem 14, since $Q \ll P$.

As for the second assertion, if $B = \Omega$ the result is merely Theorem 17. Define R by $R(\Lambda) = P(\Lambda|B)$, assuming without loss that $P(B) > 0$. Then $R \ll P$ and $B = \Omega$, R-a.s. Hence $H_R \cdot X$ equals the Lebesgue-Stieljes integral R-a.s. by Theorem 17, and the result follows by Theorem 14. □

The preceding theorem and following corollary are known as the **local behavior of the integral.**

Corollary. With the notation of Theorem 18, let S, T be two stopping times with $S < T$. Define

$$\begin{aligned} C &= \{\omega : H_t(\omega) = \overline{H}_t(\omega); X_t(\omega) = \overline{X}_t(\omega); S(\omega) < t \le T(\omega)\} \\ D &= \{\omega : t \mapsto X_t(\omega) \text{ is of finite variation on } S(\omega) < t < T(\omega)\}. \end{aligned}$$

Then $H \cdot X^T - H \cdot X^S = \overline{H} \cdot \overline{X}^T - \overline{H} \cdot \overline{X}^S$ on C and $H \cdot X^T - H \cdot X^S$ equals a path-by-path Lebesgue-Stieltjes integral on D.

Proof. Let $Y_t = X_t - X_{t\wedge S}$. Then $H \cdot Y = H \cdot X - H \cdot X^S$, and Y does not change the set $[0, S]$, which is evident, or which—alternatively—can be viewed as an easy consequence of Theorem 18. One now applies Theorem 18 to Y^T to obtain the result. □

Theorem 19 (Associativity). *The stochastic integral process $Y = H \cdot X$ is itself a semimartingale, and for $G \in \mathbb{L}$ we have*

$$G \cdot Y = G \cdot (H \cdot X) = (GH) \cdot X.$$

Proof. Suppose we know $Y = H \cdot X$ is a semimartingale. Then $G \cdot Y = J_Y(G)$. If G, H are in $\mathbf{S}$, then it is clear that $J_Y(G) = J_X(GH)$. The associativity then extends to $\mathbb{L}$ by continuity.

It remains to show that $Y = H \cdot X$ is a semimartingale. Let (H^n) be in $\mathbf{S}$ converging in *ucp* to H. Then $H^n \cdot X$ converges to $H \cdot X$ in *ucp*. Thus there exists a subsequence (n_k) such that $H^{n_k} \cdot X$ converges a.s. to $H \cdot X$.

Let $G \in \mathbf{S}$ and let $Y^{n_k} = H^{n_k} \cdot X$, $Y = H \cdot X$. The Y^{n_k} are semimartingales converging pointwise to the process Y. For $G \in \mathbf{S}$, $J_Y(G)$ is defined for any process Y; so we have

$$\begin{aligned} J_Y(G) = G \cdot Y = \lim_{n_k \to \infty} G \cdot Y^{n_k} &= \lim_{n_k \to \infty} G \cdot (H^{n_k} \cdot X) \\ &= \lim_{n_k \to \infty} (GH^{n_k}) \cdot X \end{aligned}$$

which equals $\lim_{n_k \to \infty} J_X(GH^{n_k}) = J_X(GH)$, since X is a semimartingale. Therefore $J_Y(G) = J_X(GH)$ for $G \in \mathbf{S}$.

Let G^n converge to G in $\mathbf{S}_u$. Then $G^n H$ converges to GH in $\mathbb{L}_{ucp}$, and since X is a semimartingale, $\lim_{n\to\infty} J_Y(G^n) = \lim_{n\to\infty} J_X(G^n H) = J_X(GH) = J_Y(G)$. This implies Y^t is a total semimartingale, and so $Y = H \cdot X$ is a semimartingale. □

Theorem 19 shows that *the property of being a semimartingale is preserved by stochastic integration* . Also by Theorem 17 if the semimartingale X is an FV process, then the stochastic integral agrees with the Lebesgue-Stieltjes integral, and by the theory of Lebesgue-Stieltjes integration we are able to conclude the stochastic integral is an FV process also. That is, *the property of being an FV process is preserved by stochastic integration for integrands in* $\mathbb{L}$.[2]

One may well ask if other properties are preserved by stochastic integration; in particular, are the stochastic integrals of martingales and local martingales still martingales and local martingales? Local martingales are indeed preserved by stochastic integration, but we are not yet able easily to prove it. Instead we show that *locally square integrable local martingales are preserved by stochastic integration for integrands in* $\mathbb{L}$.

Theorem 20. *Let X be a locally square integrable local martingale, and let $H \in \mathbb{L}$. Then the stochastic integral $H \cdot X$ is also a locally square integrable local martingale.*

Proof. We have seen that a locally square integrable local martingale is a semimartingale (Corollary 1 of Theorem 8), so we can formulate $H \cdot X$. Without loss of generality, assume $X_0 = 0$. Also, if T^k increases to ∞ a.s. and $(H \cdot X)^{T^k}$ is a locally square integrable local martingale for each k, it is simple to check that $H \cdot X$ itself is one. Thus without loss we assume X is a square integrable martingale. By stopping H, we may further assume H is bounded, by ℓ. Let $H^n \in \mathbf{S}$ be such that H^n converges to H in ucp. We can then modify H^n, call it $\widetilde{H}^n$, such that $\widetilde{H}^n$ is bounded by ℓ, $\widetilde{H}^n \in \mathbf{S}$, and $\widetilde{H}^n$ converges uniformly to H in probability on $[0, t]$. Since $\widetilde{H}^n \in \mathbf{bS}$, one can check that $\widetilde{H}^n \cdot X$ is a martingale. Moreover

$$
\begin{aligned}
E\{(\widetilde{H}^n \cdot X)_t^2\} &= E\{\sum_{i=1}^{k_n} (\widetilde{H}_i^n (X_t^{T_{i+1}} - X_t^{T_i}))^2\} \\
&\le \ell^2 E\{\sum_{i=1}^{k_n} (X_{T_{i+1}}^2 - X_{T_i}^2)\} \\
&\le \ell^2 E\{X_\infty^2\},
\end{aligned}
$$

and hence $(\widetilde{H}^n \cdot X)_t$ are uniformly bounded in L^2 and thus uniformly integrable. Passing to the limit then shows both that $H \cdot X$ is a martingale and that it is square integrable. □

[2] See Exercise 43 in Chap. IV which shows this is not true in general.

In Theorem 29 of Chap. III we show the more general result that if M is a local martingale and $H \in \mathbb{L}$, then $H \cdot M$ is again a local martingale.

A classical result from the theory of Lebesgue measure and integration (on $\mathbb{R}$) is that a bounded, measurable function f mapping an interval $[a, b]$ to $\mathbb{R}$ is Riemann integrable if and only if the set of discontinuities of f has Lebesgue measure zero (e.g., Kingman and Taylor [127, page 129]). Therefore we cannot hope to express the stochastic integral as a limit of sums unless the integrands have reasonably smooth sample paths. The spaces $\mathbb{D}$ and $\mathbb{L}$ consist of processes which jump at most countably often. As we will see in Theorem 21, this is smooth enough.

Definition. Let σ denote a finite sequence of finite stopping times:

$$0 = T_0 \leq T_1 \leq \cdots \leq T_k < \infty.$$

The sequence σ is called a **random partition**. A sequence of random partitions σ_n,

$$\sigma_n : T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n$$

is said to **tend to the identity** if

(i) $\lim_n \sup_k T_k^n = \infty$ a.s.; and
(ii) $\|\sigma_n\| = \sup_k |T_{k+1}^n - T_k^n|$ converges to 0 a.s.

Let Y be a process and let σ be a random partition. We define the process Y *sampled* at σ to be

$$Y^\sigma \equiv Y_0 1_{\{0\}} + \sum_k Y_{T_k} 1_{(T_k, T_{k+1}]}.$$

It is easy to check that

$$\int Y_s^\sigma \, dX_s = Y_0 X_0 + \sum_i Y_{T_i} (X^{T_{i+1}} - X^{T_i}),$$

for any semimartingale X, any process Y in $\mathbf{S}$, $\mathbb{D}$, or $\mathbb{L}$.

Theorem 21. *Let X be a semimartingale, and let Y be a process in $\mathbb{D}$ or in $\mathbb{L}$. Let (σ_n) be a sequence of random partitions tending to the identity. Then the processes $\int_{0+}^t Y_s^{\sigma_n} dX_s = \sum_i Y_{T_i^n}(X_t^{T_{i+1}^n} - X_t^{T_i^n})$ tend to the stochastic integral $(Y_-) \cdot X$ in ucp.*

Proof. (The notation Y_- means the process whose value at s is given by $(Y_-)_s = \lim_{u \to s,\, u<s} Y_u$; also, $(Y_-)_0 = 0$, by convention.) We prove the theorem for the case where Y is càdlàg, the other case being analogous. Also without loss of generality we can take $X_0 = 0$. Y càdlàg implies $Y_- \in \mathbb{L}$. Let $Y^k \in \mathbf{S}$ such that Y^k converges to Y_- (*ucp*). We have

$$\int (Y_- - Y^{\sigma_n})_s dX_s = \int (Y_- - Y^k)_s dX_s + \int (Y^k - (Y_+^k)^{\sigma_n})_s dX_s$$
$$+ \int ((Y_+^k)^{\sigma_n} - Y^{\sigma_n})_s dX_s$$

where Y_+^k denotes the càdlàg version of Y^k. The first term on the right side equals $J_X(Y_- - Y^k)$, and since J_X is continuous in $\mathbb{L}_{ucp}$ and since $Y_- - Y^k \to 0$, we have $\int (Y_- - Y^k)_s dX_s \to 0$ (ucp). The same reasoning applies to the third term, for fixed n, as $k \to \infty$. Indeed, the convergence to 0 of $((Y_+^k)^{\sigma_n} - Y^{\sigma_n})$ as $k \to \infty$ is uniform in n.

It remains to consider the middle term on the right side above. Since the Y^k are simple predictable we can write the stochastic integrals in closed form, and since X is right continuous the integrals (for fixed (k,ω)) $\int (Y^k - (Y_+^k)^{\sigma_n})_s dX_s$ tend to 0 as $n \to \infty$. Thus one merely chooses k so large that the first and third terms are small, and then for fixed k, the middle term can be made small for large enough n. □

We consider here another example. We have already seen at the end of Sect. 4 that if B is a standard Wiener process, then

$$\int_0^t B_s dB_s = \frac{1}{2} B_t^2 - \frac{1}{2} t,$$

showing that the semimartingale calculus does not formally generalize the Riemann-Stieltjes calculus.

We have seen as well that the stochastic integral agrees with the Lebesgue-Stieltjes integral when possible (Theorems 17 and 18) and that the stochastic integral also preserves the martingale property (Theorem 20; at least for locally square integrable local martingales). The following example shows that the restriction of integrands to $\mathbb{L}$ is not as innocent as it may seem if we want to have both of these properties.

Example. Let $M_t = N_t - \lambda t$, a compensated Poisson process (and hence a martingale with $M_t \in L^p$ for all $t \geq 0$ and all $p \geq 1$). Let $(T_i)_{i \geq 1}$ be the jump times of M. Let $H_t = 1_{[0,T_1)}(t)$. Then $H \in \mathbb{D}$. The Lebesgue-Stieltjes integral is

$$\begin{aligned} \int_0^t H_s dM_s &= \int_0^t H_s dN_s - \lambda \int_0^t H_s ds \\ &= \sum_{i=1}^{\infty} H_{T_i} 1_{\{t \geq T_i\}} - \lambda \int_0^t H_s ds \\ &= -\lambda (t \wedge T_1). \end{aligned}$$

This process is not a martingale. We conclude that the space of integrands cannot be expanded even to $\mathbb{D}$, in general, and preserve the structure of the theory already established.[3]

6 The Quadratic Variation of a Semimartingale

The quadratic variation process of a semimartingale, also known as the bracket process, is a simple object that nevertheless plays a fundamental role.

Definition. Let X, Y be semimartingales. The **quadratic variation process** of X, denoted $[X,X] = ([X,X]_t)_{t\geq 0}$, is defined by

$$[X,X] = X^2 - 2\int X_- dX$$

(recall that $X_{0-} = 0$). The **quadratic covariation** of X, Y, also called the **bracket process** of X, Y, is defined by

$$[X,Y] = XY - \int X_- dY - \int Y_- dX.$$

It is clear that the operation $(X,Y) \to [X,Y]$ is bilinear and symmetric. We therefore have a **polarization identity**

$$[X,Y] = \frac{1}{2}([X+Y,X+Y] - [X,X] - [Y,Y]).$$

The next theorem gives some elementary properties of $[X,X]$. (*X is assumed to be a given semimartingale throughout this section*).

Theorem 22. *The quadratic variation process of X is a càdlàg, increasing, adapted process. Moreover it satisfies the following.*

(i) $[X,X]_0 = X_0^2$ and $\Delta[X,X] = (\Delta X)^2$.
(ii) If σ_n is a sequence of random partitions tending to the identity, then

$$X_0^2 + \sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \to [X,X]$$

with convergence in ucp, where σ_n is the sequence $0 = T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots \leq T_{k_n}^n$ and where T_i^n are stopping times.
(iii) If T is any stopping time, then $[X^T, X] = [X, X^T] = [X^T, X^T] = [X,X]^T$.

[3] Ruth Williams has commented to us that this example would be more convincing if M were itself a semimartingale, H bounded and in $\mathbb{D}$, and $\int_0^t H_s dM_s$ were not a semimartingale. Such a construction is carried out in [1].

Proof. X is càdlàg, adapted, and so also is $\int X_- dX$ by its definition; thus $[X, X]$ is càdlàg, adapted as well. Recall the property of the stochastic integral: $\Delta(X_- \cdot X) = X_- \Delta X$. Then

$$\begin{aligned}(\Delta X)_s^2 &= (X_s - X_{s-})^2 = X_s^2 - 2X_s X_{s-} + X_{s-}^2 \\ &= X_s^2 - X_{s-}^2 + 2X_{s-}(X_{s-} - X_s) \\ &= \Delta(X^2)_s - 2X_{s-}(\Delta X_s),\end{aligned}$$

from which part (i) follows.

For part (ii), by replacing X with $\widetilde{X} = X - X_0$, we may assume $X_0 = 0$. Let $R_n = \sup_i T_i^n$. Then $R_n < \infty$ a.s. and $\lim_n R_n = \infty$ a.s., and thus by telescoping series

$$(X^2)^{R_n} = \sum_i \{(X^2)^{T_{i+1}^n} - (X^2)^{T_i^n}\}$$

converges *ucp* to X^2. Moreover, the series $\sum_i X_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})$ converges in *ucp* to $\int X_- dX$ by Theorem 21, since X is càdlàg. Since $b^2 - a^2 - 2a(b-a) = (b-a)^2$, and since $X_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n}) = X^{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})$, we can combine the two series convergences above to obtain the result. Finally, note that if $s < t$, then the approximating sums in part (ii) include more terms (all non-negative), so it is clear that $[X, X]$ is non-decreasing. (Note that, *a priori*, one only has $[X, X]_s \leq [X, X]_t$ a.s., with the null set depending on s and t; it is the property that $[X, X]$ has càdlàg paths that allows one to eliminate the dependence of the null set on s and t.) Part (iii) is a simple consequence of part (ii). □

An immediate consequence of Theorem 22 is the observation that if B is a Brownian motion, then $[B, B]_t = t$, since in Theorem 28 of Chap. I we showed the a.s. convergence of sums of the form in part (ii) of Theorem 22 when the partitions are refining.

Another consequence of Theorem 22 is that if X is a semimartingale with continuous paths of finite variation, then $[X, X]$ is the constant process equal to X_0^2. To see this one need only observe that

$$\begin{aligned}\sum (X^{T_{i+1}^n} - X^{T_i^n})^2 &\leq \sup_i |X^{T_{i+1}^n} - X^{T_i^n}| \sum_i |X^{T_{i+1}^n} - X^{T_i^n}| \\ &\leq \sup_i |X^{T_{i+1}^n} - X^{T_i^n}| V,\end{aligned}$$

where V is the total variation. Therefore the sums tend to 0 as $\|\sigma_n\| \to 0$. Theorem 22 has several more consequences which we state as corollaries.

Corollary 1. The bracket process $[X, Y]$ of two semimartingales has paths of finite variation on compacts, and it is also a semimartingale.

Proof. By the polarization identity $[X,Y]$ is the difference of two increasing processes, hence its paths are of finite variation. Moreover, the paths are clearly càdlàg, and the process is adapted. Hence by Theorem 7 it is a semimartingale. □

Corollary 2 (Integration by Parts). Let X, Y be two semimartingales. Then XY is a semimartingale and

$$XY = \int X_- dY + \int Y_- dX + [X,Y].$$

Proof. The formula follows trivially from the definition of $[X,Y]$. That XY is a semimartingale follows from the formula, Theorem 19, and Corollary 1 above. □

In the integration by parts formula above, we have $(X_-)_0 = (Y_-)_0 = 0$. Hence evaluating at 0 yields

$$X_0Y_0 = (X_-)_0Y_0 + (Y_-)_0X_0 + [X,Y]_0.$$

Since $[X,Y]_0 = \Delta X_0 \Delta Y_0 = X_0Y_0$, the formula is valid. Without the convention that $(X_-)_0 = 0$, we could have written the formula

$$X_tY_t = \int_{0+}^t X_{s-}dY_s + \int_{0+}^t Y_{s-}dX_s + [X,Y]_t.$$

Corollary 3. All semimartingales on a given filtered probability space form an algebra.

Proof. Since semimartingales form a vector space, Corollary 2 shows they form an algebra. □

A theorem analogous to Theorem 22 holds for $[X,Y]$ as well as $[X,X]$. It can be proved analogously to Theorem 22, or more simply by polarization. We omit the proof.

Theorem 23. *Let X and Y be two semimartingales. Then the bracket process $[X,Y]$ satisfies the following.*

(i) $[X,Y]_0 = X_0Y_0$ and $\Delta[X,Y] = \Delta X \Delta Y$.
(ii) If σ_n is a sequence of random partitions tending to the identity, then

$$[X,Y] = X_0Y_0 + \lim_{n\to\infty} \sum_i (X^{T^n_{i+1}} - X^{T^n_i})(Y^{T^n_{i+1}} - Y^{T^n_i}),$$

where convergence is in ucp, and where σ_n is the sequence $0 = T_0^n \le T_1^n \le \cdots \le T_i^n \le \cdots \le T^n_{k_n}$, with T_i^n stopping times.
(iii) If T is any stopping time, then $[X^T,Y] = [X,Y^T] = [X^T,Y^T] = [X,Y]^T$.

We next record a real analysis theorem from the Lebesgue-Stieltjes theory of integration. It can be proved via the Monotone Class Theorem.

Theorem 24. *Let α, β, γ be functions mapping $[0,\infty)$ to $\mathbb{R}$ with $\alpha(0) = \beta(0) = \gamma(0) = 0$. Suppose α, β, γ are all right continuous, α is of finite variation, and β and γ are each increasing. Suppose further that for all s, t with $s \leq t$, we have*

$$|\int_s^t d\alpha_u| \leq (\int_s^t d\beta_u)^{\frac{1}{2}}(\int_s^t d\gamma_u)^{\frac{1}{2}}.$$

Then for any measurable functions f, g we have

$$\int_s^t |fg||d\alpha| \leq (\int_s^t f^2 d\beta)^{\frac{1}{2}}(\int_s^t g^2 d\gamma)^{\frac{1}{2}}.$$

In particular, the measure $d\alpha$ is absolutely continuous with respect to both $d\beta$ and $d\gamma$.

Note that $|d\alpha|$ denotes the total variation measure corresponding to the measure $d\alpha$, the Lebesgue-Stieltjes signed measure induced by α. We use this theorem to prove an important inequality concerning the quadratic variation and bracket processes.

Theorem 25 (Kunita-Watanabe Inequality). *Let X and Y be two semimartingales, and let H and K be two measurable processes. Then one has a.s.*

$$\int_0^\infty |H_s||K_s||d[X,Y]_s| \leq (\int_0^\infty H_s^2 d[X,X]_s)^{\frac{1}{2}}(\int_0^\infty K_s^2 d[Y,Y]_s)^{\frac{1}{2}}.$$

Proof. By Theorem 24 we only need to show that there exists a null set N, such that for $\omega \notin N$, and (s,t) with $s \leq t$, we have

$$|\int_s^t d[X,Y]_u| \leq (\int_s^t d[X,X]_u)^{\frac{1}{2}}(\int_s^t d[Y,Y]_u)^{\frac{1}{2}}. \qquad (*)$$

Let N be the null set such that if $\omega \notin N$, then $0 \leq \int_s^t d[X+rY, X+rY]_u$, for every r, s, t; $s \leq t$, with r, s, t all rational numbers. Then

$$\begin{aligned}0 &\leq [X+rY, X+rY]_t - [X+rY, X+rY]_s \\ &= r^2([Y,Y]_t - [Y,Y]_s) + 2r([X,Y]_t - [X,Y]_s) + ([X,X]_t - [X,X]_s).\end{aligned}$$

The right side being positive for all rational r, it must be positive for all real r by continuity. Thus the discriminant of this quadratic equation in r must be non-negative, which gives us exactly the inequality $(*)$. Since we have, then, the inequality for all rational (s,t), it must hold for all real (s,t), by the right continuity of the paths of the processes. □

Corollary. Let X and Y be two semimartingales, and let H and K be two measurable processes. Then

$$E\{\int_0^\infty |H_s||K_s||d[X,Y]_s|\} \leq \|(\int_0^\infty H_s^2 d[X,X]_s)^{\frac{1}{2}}\|_{L^p} \|(\int_0^\infty K_s^2 d[Y,Y]_s)^{\frac{1}{2}}\|_{L^q}$$

if $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Apply Hölder's inequality to the Kunita-Watanabe inequality of Theorem 25. □

Since Theorem 25 and its corollary are path-by-path Lebesgue-Stieltjes results, we *do not have to assume* that the integrand processes H and K be adapted.

Since the process $[X,X]$ is non-decreasing with right continuous paths, and since $\Delta[X,X]_t = (\Delta X_t)^2$ for all $t \geq 0$ (*with the convention that* $X_{0-} = 0$), we can decompose $[X,X]$ path-by-path into its continuous part and its pure jump part.

Definition. For a semimartingale X, the process $[X,X]^c$ denotes the **path-by-path continuous part** of $[X,X]$.

We can then write

$$\begin{aligned}[X,X]_t &= [X,X]_t^c + X_0^2 + \sum_{0<s\leq t} (\Delta X_s)^2 \\ &= [X,X]_t^c + \sum_{0\leq s\leq t} (\Delta X_s)^2.\end{aligned}$$

Observe that $[X,X]_0^c = 0$. Analogously, $[X,Y]^c$ denotes the path-by-path continuous part of $[X,Y]$.

Comment. In Chap. IV, Sect. 7, we briefly discuss the *continuous local martingale part* of a semimartingale X satisfying an auxiliary hypothesis known as Hypothesis A. This (unique) continuous local martingale part of X is denoted X^c. It can be shown that a unique continuous local martingale part, X^c, exists for every semimartingale X.[4] (If X is an FV process, then $X^c \equiv 0$, as we shall see in Chap. III.) *It is always true that* $[X^c, X^c] = [X,X]^c$. Although we will have no need of this result in this book, this notation is often used in the literature.

We remark also that the conditional quadratic variation of a semimartingale X, denoted $\langle X,X\rangle$, is defined in Chap. III, Sect. 4. *It is also true that* $\langle X^c, X^c\rangle = [X^c, X^c] = [X,X]^c$, and these notations are used interchangeably in the literature. *If X is already continuous and $X_0 = 0$, then* $\langle X,X\rangle = [X,X] = [X,X]^c$.

[4] See, for example, Dellacherie-Meyer [46, page 355]. See also in this regard Exercise 6 of Chap. IV which proves the result for L^2 martingales.

Definition. A semimartingale X will be called **quadratic pure jump** if $[X,X]^c = 0$.

If X is quadratic pure jump, then $[X,X]_t = X_0^2 + \sum_{0<s\le t}(\Delta X_s)^2$. The Poisson process N is an obvious example of a quadratic pure jump semimartingale. From the definition or immediately from Theorem 22, we see that $[N,N]_t = N_t$. More generally it can be shown that if X is a Lévy process with a Lévy decomposition $X_t = B_t + Y_t$ as in Theorem 42 of Chap. I, where B is a Brownian motion and

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

(using the notation of Theorem 42 of Chap. I), then Y is a quadratic pure jump semimartingale. Last note that the trivial continuous process $X_t = t$ is quadratic pure jump since $[X,X]_t^c = [X,X]_t \equiv 0$.

The next theorem gives a simple criterion for a process X to be a quadratic pure jump semimartingale.

Theorem 26. *If X is adapted, càdlàg, with paths of finite variation on compacts, then X is a quadratic pure jump semimartingale.*

Proof. Without loss of generality we assume $X_0 = 0$. We have already seen that such an X is a semimartingale (Theorem 7), and that the stochastic integral with respect to X is nothing more than a path-by-path Lebesgue-Stieltjes integral (Theorem 17). The integration by parts formula for Lebesgue-Stieltjes differentials applied to X times itself yields $X^2 = \int X_- dX + \int X dX$, computed path-by-path. The semimartingale integration by parts formula (Corollary 2 of Theorem 22), on the other hand, yields $X^2 = 2\int X_- dX + [X,X]$. Moreover

$$\int X dX = \int (X_- + \Delta X) dX = \int X_- dX + \int \Delta X dX,$$

and

$$\int_0^t X_- dX_s + \int_0^t \Delta X_s dX_s = \int_0^t X_{s-} dX_s + \sum_{s\le t} (\Delta X_s)^2.$$

Equating the two formulas we deduce $[X,X]_t = \sum_{s\le t}(\Delta X_s)^2$, yielding the theorem. □

Note in particular that if X is adapted with *continuous* paths of finite variation, then $[X,X]_t = X_0^2$, all $t \ge 0$.

Theorem 27. *Let X be a local martingale with continuous paths that are not everywhere constant. Then $[X,X]$ is not the constant process X_0^2, and $X^2 - [X,X]$ is a continuous local martingale. Moreover if $[X,X]_t = 0$ for all t then $X_t = 0$ for all t.*

Proof. Note that a continuous local martingale is a semimartingale (Corollary 2 of Theorem 8). We have $X^2 - [X,X] = 2\int X_- dX$, and by the martingale preservation property (Theorem 20) we have that $2\int X_- dX$ is a local martingale. Moreover $\Delta 2\int X_- dX = 2(X_-)(\Delta X)$, and since X is continuous, $\Delta X = 0$, and thus $2\int X_- dX$ is a *continuous* local martingale, hence locally square integrable. Thus $X^2 - [X,X]$ is a locally square integrable local martingale.

By stopping, we can suppose X is a square integrable martingale. Assume further $X_0 = 0$. Next assume that $[X,X]$ actually were constant. Then $[X,X]_t = [X,X]_0 = X_0^2 = 0$, for all t. Since $X^2 - [X,X]$ is a local martingale, we conclude X^2 is a non-negative local martingale, with $X_0^2 = 0$. Thus $X_t^2 = 0$, all t. This is a contradiction. If X_0 is not identically 0, we set $\hat{X}_t = X_t - X_0$ and the result follows. □

The following corollary is of fundamental importance in the theory of martingales.

Corollary 1. Let X be a continuous local martingale, and $S \leq T \leq \infty$ be stopping times. If X has paths of finite variation on the stochastic interval (S,T), then X is constant on $[S,T]$. Moreover if $[X,X]$ is constant on $[S,T] \cap [0,\infty)$, then X is constant there too.

Proof. $M = X^T - X^S$ is also a continuous local martingale, and M has finite variation on compacts. Moreover $[M,M] = [X^T - X^S, X^T - X^S] = [X,X]^T - [X,X]^S$, and therefore $[M,M]$ is constant everywhere, hence by Theorem 27, M must be constant everywhere; thus X is constant on $[S,T]$. □

Observe that if $t > 0$ is arbitrary and we take $S = 0$ and $T = t$ in Corollary 1, then we can conclude that *a continuous local martingale with paths of finite variations on compacts is a.s. constant.* While non-trivial continuous local martingales must therefore always have paths of infinite variation on compacts, they are not the only such local martingales. For example, the Lévy process martingale Z^d of Theorem 41 of Chap. I will have paths of infinite variation if the Lévy measure ν has infinite mass in a neighborhood of the origin. Another example is Azéma's martingale, which is presented in detail in Sect. 8 of Chap. IV.

Corollary 2. Let X and Y be two locally square integrable local martingales. Then $[X,Y]$ is the unique adapted càdlàg process A with paths of finite variation on compacts satisfying the two properties:

(i) $XY - A$ is a local martingale; and
(ii) $\Delta A = \Delta X \Delta Y$, $A_0 = X_0 Y_0$.

Proof. Integration by parts yields

$$XY = \int X_- dY + \int Y_- dX + [X,Y],$$

but the martingale preservation property tells us that both stochastic integrals are local martingales. Thus $XY - [X,Y]$ is a local martingale. Property (ii) is simply an application of Theorem 23. Thus it remains to show uniqueness. Suppose A, B both satisfy properties (i) and (ii). Then $A - B = (XY - B) - (XY - A)$, the difference of two local martingales which is again a local martingale. Moreover,

$$\Delta(A - B) = \Delta A - \Delta B = \Delta X \Delta Y - \Delta X \Delta Y = 0.$$

Thus $A - B$ is a continuous local martingale, $A_0 - B_0 = 0$, and it has paths of finite variation on compacts. Corollary 1 yields $A_t - B_t - A_0 + B_0 = 0$ and we have uniqueness. □

Corollary 2 can be useful in determining the process $[X,Y]$. For example if X and Y are locally square integrable martingales without common jumps such that XY is a martingale, then $[X,Y] = X_0Y_0$. One can also easily verify (as a consequence of Theorem 23 and Corollaries 1 and 2 above, for example) that if X is a continuous square integrable martingale and Y is a square integrable martingale with paths of finite variation on compacts, then $[X,Y] = X_0Y_0$, and hence XY is a martingale. (An example would be X a Brownian motion and Y a compensated Poisson process.) Corollary 2 is true as well for X, Y local martingales, however we need Theorem 29 of Chap. III to prove the general result.

In Chap. III we show that any local martingale is a semimartingale (see the corollary of Theorem 26 of Chap. III), and therefore if M is a local martingale its quadratic variation $[M,M]_t$ always exists and is finite a.s. for every $t \geq 0$. We use this fact in the next corollary.

Corollary 3. Let M be a local martingale. Then M is a martingale with $E\{M_t^2\} < \infty$, all $t \geq 0$, if and only if $E\{[M,M]_t\} < \infty$, all $t \geq 0$. If $E\{[M,M]_t\} < \infty$, then $E\{M_t^2\} = E\{[M,M]_t\}$.

Proof. First assume that M is a martingale with $E\{M_t^2\} < \infty$ for all $t \geq 0$. Then M is clearly a locally square integrable martingale. Let $N_t = M_t^2 - [M,M]_t = 2\int_0^t M_{s-}dM_s$, which is a locally square integrable local martingale by Theorem 20. Let $(T^n)_{n\geq 1}$ be stopping times increasing to ∞ a.s. such that $N_t^{T^n}$ is a square integrable martingale. Then $E\{N_t^{T^n}\} = E\{N_0\} = 0$, all $t \geq 0$. Therefore

$$E\{M_{t\wedge T^n}^2\} = E\{[M,M]_{t\wedge T^n}\}.$$

Doob's maximal quadratic inequality gives $E\{(M_t^*)^2\} \leq 4E\{M_t^2\} < \infty$. Therefore by the Dominated Convergence Theorem

$$\begin{aligned} E\{M_t^2\} &= \lim_{n\to\infty} E\{M_{t\wedge T^n}^2\} \\ &= \lim_{n\to\infty} E\{[M,M]_{t\wedge T^n}\} \\ &= E\{[M,M]_t\}, \end{aligned}$$

where the last result is by the Monotone Convergence Theorem. In particular we have that $E\{[M,M]_t\} < \infty$.

For the converse, we now assume $E\{[M,M]_t\} < \infty$, all $t \geq 0$. Define stopping times by

$$T^n = \inf\{t > 0 : |M_t| > n\} \wedge n.$$

Then T^n increase to ∞ a.s. Furthermore $(M^{T^n})^* \leq n + |\Delta M_{T^n}| \leq n + [M,M]_n^{1/2}$, which is in L^2. By Theorem 51 of Chap. I, M^{T^n} is a uniformly integrable martingale for each n. Also we have that

$$E\{(M_t^{T^n})^2\} \leq E\{((M^{T^n})^*)^2\} < \infty,$$

for all $t \geq 0$. Therefore M^{T^n} satisfies the hypotheses of the first half of this theorem, and $E\{(M_t^{T^n})^2\} = E\{[M^{T^n}, M^{T^n}]_t\}$. Using Doob's inequality we have

$$\begin{aligned} E\{(M_{t\wedge T^n}^*)^2\} \leq 4E\{(M_t^{T^n})^2\} &= 4E\{[M^{T^n}, M^{T^n}]_t\} \\ &= 4E\{[M,M]_{T^n\wedge t}\} \\ &\leq 4E\{[M,M]_t\}. \end{aligned}$$

The Monotone Convergence Theorem next gives

$$\begin{aligned} E\{(M_t^*)^2\} &= \lim_{n\to\infty} E\{(M_{t\wedge T^n}^*)^2\} \\ &\leq 4E\{[M,M]_t\} < \infty. \end{aligned}$$

Therefore, again by Theorem 51 of Chap. I, we conclude that M is a martingale. The preceding gives $E\{M_t^2\} < \infty$. □

For emphasis we state as another corollary a special case of Corollary 3.

Corollary 4. If M is a local martingale and $E\{[M,M]_\infty\} < \infty$, then M is a square integrable martingale (that is $\sup_t E\{M_t^2\} = E\{M_\infty^2\} < \infty$). Moreover $E\{M_t^2\} = E\{[M,M]_t\}$ for all t, $0 \leq t \leq \infty$.

Example. Before continuing we consider again an example of a local martingale that exhibits many of the surprising pathologies of local martingales. Let B be a standard Brownian motion in $\mathbb{R}^3$ with $B_0 = (1,1,1)$. Let $M_t = \|B_t\|^{-1}$, where $\|x\|$ is standard Euclidean norm in $\mathbb{R}^3$. (We previously considered this example in Sect. 6 of Chap. I.) As noted in Chap. I, the process M is a continuous local martingale; hence it is a locally square integrable local martingale. Moreover $E\{M_t^2\} < \infty$ for all t. However instead of $t \mapsto E\{M_t^2\}$ being an increasing function as it would if M were a martingale, $\lim_{t\to\infty} E\{M_t^2\} = 0$. Moreover $E\{[M,M]_t\} \geq E\{[M,M]_0\} = 1$ since $[M,M]_t$ is increasing. Therefore we cannot have $E\{M_t^2\} = E\{[M,M]_t\}$ for all t. Indeed, by Corollary 3 and the preceding we see that we must have $E\{[M,M]_t\} = \infty$ for all $t > 0$. In conclusion, $M = \|B\|^{-1}$ is a continuous local martingale with $E\{M_t^2\} < \infty$ for all t which is both not a true martingale and for which $E\{M_t^2\} < \infty$ while $E\{[M,M]_t\} = \infty$ for all $t > 0$. (Also refer to Exercise 20 at the end of this chapter.)

Corollary 5. Let X be a continuous local martingale. Then X and $[X,X]$ have the same intervals of constancy a.s.

Proof. Let r be a positive rational, and define

$$T_r = \inf\{t \geq r : X_t \neq X_r\}.$$

Then $M = X^{T_r} - X^r$ is a local martingale which is constant. Hence $[M,M] = [X,X]^{T_r} - [X,X]^r$ is also constant. Since this is true for any rational r a.s., any interval of constancy of X is also one of $[X,X]$.

Since X is continuous, by stopping we can assume without loss of generality that X is a bounded martingale (and hence square integrable). For every positive, rational r we define

$$S_r = \inf\{t \geq r : [X,X]_t > [X,X]_r\}.$$

Then

$$E\{(X_{S_r} - X_r)^2\} = E\{X_{S_r}^2\} - E\{X_r^2\}$$

by Doob's Optional Sampling Theorem. Moreover

$$E\{X_{S_r}^2\} - E\{X_r^2\} = E\{[X,X]_{S_r} - [X,X]_r\} = 0,$$

by Corollary 3. Therefore $E\{(X_{S_r} - X_r)^2\} = 0$, and $X_{S_r} = X_r$ a.s. Moreover this implies $X_q = X_{S_q}$ a.s. on $\{S_q = S_r\}$ for each pair of rationals (r,q), and therefore we deduce that any interval of constancy of $[X,X]$ is also one of X. □

Note that the continuity of the local martingale X is essential in Corollary 5. Indeed, let N_t be a Poisson process, and let $M_t = N_t - t$. Then M is a martingale and $[M,M]_t = N_t$; clearly M has no intervals of constancy while N is constant except for jumps.

Theorem 28. *Let X be a quadratic pure jump semimartingale. Then for any semimartingale Y we have*

$$[X,Y]_t = X_0Y_0 + \sum_{0<s\leq t} \Delta X_s \Delta Y_s.$$

Proof. The Kunita-Watanabe inequality (Theorem 25) tells us $d[X,Y]_s$ is a.s. absolutely continuous with respect to $d[X,X]$ (path-by-path). Thus $[X,X]^c = 0$ implies $[X,Y]^c = 0$, and hence $[X,Y]$ is the sum of its jumps, and the result follows by Theorem 23. □

Theorem 29. *Let X and Y be two semimartingales, and let H, $K \in \mathbb{L}$. Then*

$$[H \cdot X, K \cdot Y]_t = \int_0^t H_s K_s d[X,Y]_s$$

and, in particular,

$$[H \cdot X, H \cdot X]_t = \int_0^t H_s^2 d[X, X]_s.$$

Proof. First assume (without loss of generality) that $X_0 = Y_0 = 0$. It suffices to establish the following result

$$[H \cdot X, Y]_t = \int_0^t H_s d[X, Y]_s, \tag{$*$}$$

and then apply it again, by the symmetry of the form $[\cdot, \cdot]$, and by the associativity of the stochastic integral (Theorem 19).

First suppose H is the indicator of a stochastic interval. That is, $H = 1_{[0,T]}$, where T is a stopping time. Establishing $(*)$ is equivalent in this case to showing $[X^T, Y] = [X, Y]^T$, a result that is an obvious consequence of Theorem 23, which approximates $[X, Y]$ by sums.

Next suppose $H = U1_{(S,T]}$, where S, T are stopping times, $S \leq T$ a.s., and $U \in \mathcal{F}_S$. Then $\int H_s dX_s = U(X^T - X^S)$, and by Theorem 23

$$\begin{aligned}
[H \cdot X, Y] &= U\{[X^T, Y] - [X^S, Y]\} \\
&= U\{[X, Y]^T - [X, Y]^S\} = \int H_s d[X, Y]_s.
\end{aligned}$$

The result now follows for $H \in \mathbf{S}$ by linearity. Finally, suppose $H \in \mathbb{L}$ and let H^n be a sequence in $\mathbf{S}$ converging in *ucp* to H. Let $Z^n = H^n \cdot X$, $Z = H \cdot X$. We know Z^n, Z are all semimartingales. We have $\int H_s^n d[X, Y]_s = [Z^n, Y]$, since $H^n \in \mathbf{S}$, and using integration by parts

$$\begin{aligned}
[Z^n, Y] &= YZ^n - \int Y_- dZ^n - \int Z_-^n dY \\
&= YZ^n - \int Y_- H^n dX - \int Z_-^n dY.
\end{aligned}$$

By the definition of the stochastic integral, we know $Z^n \to Z$ in *ucp*, and since $H^n \to H$ (*ucp*), letting $n \to \infty$ we have

$$\begin{aligned}
\lim_{n \to \infty} [Z^n, Y] &= YZ - \int Y_- H dX - \int Z_- dY \\
&= YZ - \int Y_- dZ - \int Z_- dY \\
&= [Z, Y],
\end{aligned}$$

again by integration by parts. Since $\lim_{n \to \infty} \int H_s^n d[X, Y]_s = \int H_s d[X, Y]_s$, we have $[Z, Y] = [H \cdot X, Y] = \int H_s d[X, Y]_s$, and the proof is complete. $\square$

Example. Let B_t be a standard Wiener process with $B_0 = 0$, (i.e., Brownian motion). $B_t^2 - t$ is a continuous martingale by Theorem 27 of Chap. I. Let $H \in \mathbb{L}$ be such that $E\{\int_0^t H_s^2 ds\} < \infty$, each $t \geq 0$. By Theorem 28 of Chap. I we have $[B,B]_t = t$, hence $[H \cdot B, H \cdot B]_t = \int_0^t H_s^2 ds$. By the martingale preservation property, $\int H_s dB_s$ is also a continuous local martingale, with $(H \cdot B)_0 = 0$. By Corollary 3 to Theorem 27

$$\begin{aligned} E\{(\int_0^t H_s dB_s)^2\} &= E\{[H \cdot B, H \cdot B]_t\} \\ &= E\{\int_0^t H_s^2 ds\}. \end{aligned}$$

It was this last equality,

$$E\{(\int_0^t H_s dB_s)^2\} = E\{\int_0^t H_s^2 ds\},$$

that was crucial in K. Itô's original treatment of a stochastic integral.

Theorem 30. *Let H be a càdlàg, adapted process, and let X, Y be two semimartingales. Let σ_n be a sequence of random partitions tending to the identity. Then*

$$\sum_i H_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n})$$

converges in ucp to $\int H_{s-} d[X,Y]_s$ ($H_{0-} = 0$). Here $\sigma_n = (0 \leq T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots \leq T_{k_n}^n)$.

Proof. By the definition of quadratic variation, $[X,Y] = XY - X_- \cdot Y - Y_- \cdot X$, where $X_- \cdot Y$ denotes the process $(\int_0^t X_{s-} dY_s)_{t \geq 0}$. By the associativity of the stochastic integral (Theorem 19)

$$\begin{aligned} H_- \cdot [X,Y] &= H_- \cdot (XY) - H_- \cdot (X_- \cdot Y) - H_- \cdot (Y_- \cdot X) \\ &= H_- \cdot (XY) - (H_- X_-) \cdot Y - (H_- Y_-) \cdot X \\ &= H_- \cdot (XY) - (HX)_- \cdot Y - (HY)_- \cdot X. \end{aligned}$$

By Theorem 21 the above is the limit of

$$\begin{aligned} &\sum_i H_{T_i^n}(X^{T_{i+1}^n} Y^{T_{i+1}^n} - X^{T_i^n} Y^{T_i^n}) - \sum_i H_{T_i^n} X_{T_i^n}(Y^{T_{i+1}^n} - Y^{T_i^n}) \\ &\qquad - \sum_i H_{T_i^n} Y_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n}) \\ &\quad = \sum_i H_{T_i^n}\{(X^{T_{i+1}^n} Y^{T_{i+1}^n} - X^{T_i^n} Y^{T_i^n}) - X^{T_i^n}(Y^{T_{i+1}^n} - Y^{T_i^n}) \\ &\qquad - Y^{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})\} \\ &\quad = \sum_i H_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n}). \end{aligned}$$

□

7 Itô's Formula (Change of Variables)

Let A be a process with continuous paths of finite variation on compacts. If $H \in \mathbb{L}$ we know by Theorem 17 that the stochastic integral $H \cdot A$ agrees a.s. with the path-by-path Lebesgue-Stieltjes integral $\int H_s dA_s$. In Sect. 7 of Chap. I (Theorem 54) we proved the change of variables formula for $f \in \mathcal{C}^1$, namely

$$f(A_t) - f(A_0) = \int_0^t f'(A_s)dA_s. \tag{$*$}$$

We also saw at the end of Sect. 4 of this chapter that for a standard Wiener process B with $B_0 = 0$,

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t. \tag{$**$}$$

Taking $f(x) = x^2/2$, the above formula is equivalent to

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds,$$

which does not agree with the Lebesgue-Stieltjes change of variables formula ($*$). In this section we will state and prove a change of variables formula valid for all semimartingales.

We first mention, however, that the change of variables formula for continuous Stieltjes integrals given in Theorem 54 of Chap. I has an extension to right continuous processes of finite variation on compacts. We state this result as a theorem but we do not prove it here because it is merely a special case of Theorem 32, thanks to Theorem 26.

Theorem 31 (Change of Variables). *Let V be an FV process with right continuous paths, and let f be such that f' exists and is continuous. Then $\big(f(V_t)\big)_{t\geq 0}$ is an FV process and*

$$f(V_t) - f(V_0) = \int_{0+}^t f'(V_{s-})dV_s + \sum_{0<s\leq t}\{f(V_s) - f(V_{s-}) - f'(V_{s-})\Delta V_s\}.$$

Recall that the notation $\int_{0+}^t = \int_{(0,t]}$ denotes the integral over the half open interval $(0,t]$. We wish to establish a formula analogous to the above, but for the stochastic integral; that is, when the process is a semimartingale. The formula is different in this case, as we can see by comparing equation ($*$) with equation ($**$); we must add an extra term!

Theorem 32 (Itô's Formula). *Let X be a semimartingale and let f be a $\mathcal{C}^2$ real function. Then $f(X)$ is again a semimartingale, and the following formula holds:*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s^c$$
$$+ \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\}.$$

Proof. Note that the jump part of the stochastic integral $\int f''(X_{s-})d[X,X]_s$ is given by $\sum_{s\leq t} f''(X_{s-})(\Delta X_s)^2$, and this is a convergent series. By adding and subtracting $1/2$ of this series, we can rewrite Itô's formula in the equivalent form

$$f(X_t)-f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s$$
$$+ \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\}$$

which is perhaps less obviously a generalization of the "classical" case, but notationally simpler to prove. The proof rests, of course, on Taylor's Theorem which says

$$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2}f''(x)(y-x)^2 + R(x,y)$$

where $|R(x,y)| \leq r(|y-x|)(y-x)^2$, such that $r : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function with $\lim_{u\downarrow 0} r(u) = 0$, which is valid for $f \in \mathcal{C}^2$ defined on a compact set. We now prove separately the continuous case and the general case.

Proof for the continuous case. We first restrict our attention to a continuous semimartingale X, since the proof is less complicated but nevertheless gives the basic idea. Without loss of generality we can take $X_0 = 0$. Define stopping times

$$R_m = \inf\{t : |X_t| \geq m\}.$$

Then the stopped process X^{R_m} is bounded by m, and if Itô's formula is valid for X^{R_m} for each m, it is valid for X as well. Therefore we assume that X takes its values in a compact set. We fix a $t > 0$, and let σ_n be a refining sequence of random partitions[5] of $[0,t]$ tending to the identity $[\sigma_n = (0 = T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n = t)]$. Then

$$f(X_t) - f(X_0) = \sum_{i=0}^{k_n} \left\{f\left(X_{T_{i+1}^n}\right) - f\left(X_{T_i^n}\right)\right\} \qquad (***)$$
$$= \sum_i f'\left(X_{T_i^n}\right)\left(X_{T_{i+1}^n} - X_{T_i^n}\right)$$
$$+ \frac{1}{2}\sum_i f''\left(X_{T_i^n}\right)\left(X_{T_{i+1}^n} - X_{T_i^n}\right)^2 + \sum_i R\left(X_{T_i^n}, X_{T_{i+1}^n}\right)$$

[5] Note that it would suffice for this proof to restrict attention to deterministic partitions.

The first sum converges in probability to the stochastic integral $\int_0^t f'(X_{s-})dX_s$ by Theorem 21; the second sum converges to $\frac{1}{2}\int_0^t f''(X_s)d[X,X]_s$ in probability by Theorem 30. It remains to consider the third sum $\sum_i R\left(X_{T_i^n}, X_{T_{i+1}^n}\right)$. But this sum is majorized, in absolute value, by

$$\sup_i r(|X_{T_{i+1}^n} - X_{T_i^n}|)\{\sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2\},$$

and since $\sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2$ converges in probability to $[X,X]_t$ (Theorem 22), the last term will tend to 0 if $\lim_{n\to\infty}\sup_i r(|X_{T_{i+1}^n} - X_{T_i^n}|) = 0$. However $s \mapsto X_s(\omega)$ is a continuous function on $[0,t]$, each fixed ω, and hence uniformly continuous. Since $\lim_{n\to\infty}\sup_i |T_{i+1}^n - T_i^n| = 0$ by hypothesis, we have the result. Thus, in the continuous case, $f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X,X_s]$, for each t, a.s. The continuity of the paths then permits us to remove the dependence of the null set on t, giving the complete result in the continuous case.

Proof for the general case. X is now given as a right continuous semimartingale. Once again we have a representation as in $(***)$, but we need a closer analysis. For any $t > 0$ we have $\sum_{0<s\le t}(\Delta X_s)^2 \le [X,X]_t < \infty$ a.s., hence $\sum_{0<s\le t}(\Delta X_s)^2$ is convergent. Given $\varepsilon > 0$ and $t > 0$, let $A = A(\varepsilon,t)$ be a set of jump times of X that has a.s. a finite number of times s, and let $B = B(\varepsilon,t)$ be such that $\sum_{s\in B}(\Delta X_s)^2 \le \varepsilon^2$, where A and B are disjoint and $A \cup B$ exhaust the jump times of X on $(0,t]$. We write

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_i \left\{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\right\} \\ &= \sum_{i,A}\left\{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\right\} + \sum_{i,B}\left\{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\right\} \end{aligned}$$

where $\sum_{i,A}$ denotes $\sum_i 1_{\{A\cap(T_i^n,T_{i+1}^n]\neq\emptyset\}}$. Then

$$\lim_n \sum_{i,A}\left\{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\right\} = \sum_{s\in A}\{f(X_s) - f(X_{s-})\},$$

and by Taylor's formula

$$\begin{aligned} &\sum_{i,B}\left\{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\right\} \\ &= \sum_i f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n}) + \frac{1}{2}\sum_i f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2 \\ &\quad - \sum_{i,A}\left\{f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n}) + \frac{1}{2}f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2\right\} \\ &\quad + \sum_{i,B} R(X_{T_i^n}, X_{T_{i+1}^n}). \end{aligned}$$

As in the continuous case, the first two sums on the right side above converge respectively to $\int_{0+}^{t} f'(X_{s-})dX_s$ and $\frac{1}{2}\int_{0+}^{t} f''(X_{s-})d[X,X]_s$. The third sum converges to

$$-\sum_{s\in A}\left\{f'(X_{s-})\Delta X_s + \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\right\}.$$

Assume temporarily that $|X_s| \leq k$, some constant k, all $s \leq t$. Then f'' is uniformly continuous, and using the right continuity of X we have

$$\limsup_n \sum_{i,B} R(X_{T_i^n}, X_{T_{i+1}^n}) \leq r(\varepsilon+)[X,X]_t,$$

where $r(\varepsilon+)$ is $\limsup_{\delta\downarrow\varepsilon} r(\delta)$. Next let ε tend to 0. Then $r(\varepsilon+)[X,X]_t$ tends to 0, and

$$\sum_{s\in A(\varepsilon,t)} \left\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\right\}$$

tends to the series in $(***)$, provided this series is absolutely convergent.

Let $V_k = \inf\{t > 0 : |X_t| \geq k\}$, with $X_0 = 0$. By first establishing $(***)$ for $X1_{[0,V_k)}$, which is a semimartingale since it is the product of two semimartingales (Corollary 2 of Theorem 22), it suffices to consider semimartingales taking their values in intervals of the form $[-k,k]$. For f restricted to $[-k,k]$ we have $|f(y) - f(x) - (y-x)f'(x)| \leq C(y-x)^2$. Then

$$\sum_{0<s\leq t} |f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s| \leq C \sum_{0<s\leq t} (\Delta X_s)^2 \leq C[X,X]_t < \infty,$$

and $\sum_{0<s\leq t} |f''(X_{s-})|(\Delta X_s)^2 \leq K\sum_{0<s\leq t}(\Delta X_s)^2 \leq K[X,X]_t < \infty$ a.s. Thus the series is absolutely convergent and this completes the proof. □

Corollary (Itô's Formula). Let X be a continuous semimartingale and let f be a C^2 real function. Then $f(X)$ is again a semimartingale and the following formula holds:

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_s)dX_s + \frac{1}{2}\int_{0+}^{t} f''(X_s)d[X,X]_s.$$

Theorem 32 has a multi-dimensional analog. We omit the proof.

Theorem 33. *Let $X = (X^1,\dots,X^n)$ be an n-tuple of semimartingales, and let $f : \mathbb{R}^n \to \mathbb{R}$ have continuous second order partial derivatives. Then $f(X)$ is a semimartingale and the following formula holds:*

$$
\begin{aligned}
& f(X_t) - f(X_0) \\
& = \sum_{i=1}^{n} \int_{0+}^{t} \frac{\partial f}{\partial x_i}(X_{s-}) dX_s^i + \frac{1}{2} \sum_{1 \le i,j \le n} \int_{0+}^{t} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s-}) d[X^i, X^j]_s^c \\
& \quad + \sum_{0<s\le t} \left\{ f(X_s) - f(X_{s-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s-}) \Delta X_s^i \right\}.
\end{aligned}
$$

The stochastic integral calculus, as revealed by Theorems 32 and 33, is different from the classical Lebesgue-Stieltjes calculus. By restricting the class of integrands to semimartingales made left continuous (instead of $\mathbb{L}$), *one can define a stochastic integral that obeys the traditional rules of the Lebesgue-Stieltjes calculus.*

Definition. Let X, Y be semimartingales. Define the **Fisk-Stratonovich integral** of Y with respect to X, denoted $\int_0^t Y_{s-} \circ dX_s$, by

$$
\int_0^t Y_{s-} \circ dX_s \equiv \int_0^t Y_{s-} dX_s + \frac{1}{2}[Y, X]_t^c.
$$

The Fisk-Stratonovich integral is often referred to as simply the **Stratonovich integral**. The notation "$\circ$" is called **Itô's circle**. Note that we have defined the Fisk-Stratonovich integral in terms of the semimartingale integral. With some work one can slightly enlarge the domain of the definition and we do so in Sect. 5 of Chap. V. In particular, Theorem 34 below is proved with the weaker hypothesis that $f \in C^2$ (Theorem 20 of Chap. V). We will write the **F-S integral** as an abbreviation for the Fisk-Stratonovich integral.

Theorem 34. *Let X be a semimartingale and let f be C^3. Then*

$$
f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-}) \circ dX_s + \sum_{0<s\le t} \{ f(X_s) - f(X_{s-}) - f'(X_{s-}) \Delta X_s \}.
$$

Proof. Note that f' is C^2, so that $f'(X)$ is a semimartingale by Theorem 32 and in the domain of the F-S integral. By Theorem 32 and the definition, it suffices to establish $\frac{1}{2}[f'(X), X]^c = \frac{1}{2}\int_0^t f''(X_{s-}) d[X, X]_s^c$. However

$$
\begin{aligned}
f'(X_t) - f'(X_0) &= \int_{0+}^{t} f''(X_{s-}) dX_s + \frac{1}{2} \int_{0+}^{t} f^{(3)}(X_{s-}) d[X, X]_s^c \\
& \quad + \sum_{0<s\le t} \{ f'(X_s) - f'(X_{s-}) - f''(X_{s-}) \Delta X_s \}.
\end{aligned}
$$

Thus

$$
[f'(X), X]^c = [f''(X_-) \cdot X, X]^c + [\frac{1}{2} f^{(3)}(X_-) \cdot [X, X], X]^c.
$$

The first term on the right side above is $\int_0^t f''(X_{s-}) d[X, X]_s^c$ by Theorem 29; the second term can easily be seen, as a consequence of Theorem 22 and the fact that $[X,X]$ has paths of finite variation, to be $\sum_{0<s\leq t} f^{(3)}(X_{s-})(\Delta X_s)^3$. That is, zero, and the theorem is proved. □

Note that if X is a semimartingale with *continuous paths*, then Theorem 34 reduces to the classical Riemann-Stieltjes formula $f(X_t)-f(X_0) = \int_0^t f'(X_s) \circ dX_s$. This is, of course, the main attraction of the Fisk-Stratonovich integral.

Corollary (Integration by Parts). Let X, Y be semimartingales, with at least one of X and Y continuous. Then

$$X_tY_t - X_0Y_0 = \int_0^t X_{s-} \circ dY_s + \int_0^t Y_{s-} \circ dX_s.$$

Proof. The standard integration by parts formula is

$$X_tY_t = \int_0^t X_{s-}dY_s + \int_0^t Y_{s-}dX_s + [X,Y]_t.$$

However $[X,Y]_t = [X,Y]_t^c + X_0Y_0$ if one of X or Y is continuous. Thus adding $\frac{1}{2}[X,Y]_t^c$ to each integral on the right side yields the result. □

One can extend the stochastic calculus to complex-valued semimartingales.

Theorem 35. *Let X, Y be continuous semimartingales, let $Z_t = X_t + iY_t$, and let f be analytic. Then*

$$f(Z_t) = f(Z_0) + \int_{0+}^t f'(Z_s)dZ_s + \frac{1}{2}\int_{0+}^t f''(Z_s)d[Z,Z]_s.$$

Proof. Using Itô's formula for the real and imaginary parts of f yields

$$\begin{aligned} f(Z_t) = f(Z_0) &+ \int_{0+}^t \frac{\partial f}{\partial x}(Z_s)dX_s + \int_{0+}^t \frac{\partial f}{\partial y}(Z_s)dY_s \\ &+ \frac{1}{2}\int_{0+}^t \frac{\partial^2 f}{\partial x^2}(Z_s)d[X,X]_s + \int_{0+}^t \frac{\partial^2 f}{\partial x \partial y}(Z_s)d[X,Y]_s \\ &+ \frac{1}{2}\int_{0+}^t \frac{\partial^2 f}{\partial y^2}(Z_s)d[Y,Y]_s. \end{aligned}$$

Since f is analytic by the Cauchy-Riemann equations $\frac{\partial f}{\partial x} = f'$ and $\frac{\partial f}{\partial y} = if'$. Differentiating again gives $\frac{\partial^2 f}{\partial x^2} = \frac{\partial f'}{\partial x} = f''$ and $\frac{\partial^2 f}{\partial x \partial y} = if''$ and $\frac{\partial^2 f}{\partial y^2} = -f''$. The result now follows by collecting terms and observing that $dZ_s = dX_s + idY_s$ and $[Z,Z] = [X+iY, X+iY] = [X,X] + 2i[X,Y] - [Y,Y]$. □

Theorem 35 also has a version for $Z_t = X_t + iY_t$, with both X and Y càdlàg semimartingales. The proof is almost the same as the proof of Theorem 35.

Theorem 36. *Let X, Y be càdlàg semimartingales, let $Z_t = X_t + iY_t$, and let f be analytic. Then*

$$f(Z_t) = f(Z_0) + \int_{0+}^{t} f'(Z_{s-})dZ_s + \frac{1}{2}\int_{0+}^{t} f''(Z_{s-})d[Z,Z]_s^c + \sum_{0<s\le t}\{f(Z_s) - f(Z_{s-}) - f'(Z_{s-})\Delta Z_s\}.$$

Observe that for a complex-valued semimartingale *the process $[Z, Z]$ is in general a complex-valued process.* For many applications, it is more appropriate to use the non-negative increasing process

$$[Z, \overline{Z}] = [X, X] + [Y, Y]$$

to play the role of the quadratic variation.

8 Applications of Itô's Formula

As an application of the change of variables formula, we investigate a simple, yet important and non-trivial, stochastic differential equation. We treat it, of course, in integral form.

Theorem 37. *Let X be a semimartingale, $X_0 = 0$. Then there exists a (unique) semimartingale Z that satisfies the equation $Z_t = 1 + \int_0^t Z_{s-}dX_s$. Z is given by*

$$Z_t = \exp\left\{X_t - \frac{1}{2}[X,X]_t\right\}\prod_{0<s\le t}(1+\Delta X_s)\exp\left\{-\Delta X_s + \frac{1}{2}(\Delta X_s)^2\right\}$$

where the infinite product converges.

Proof. We will not prove the uniqueness here since it is a trivial consequence of the general theory to be established in Chap. V. (For example, see Theorem 7 of Chap. V.) Note that the formula for Z_t is equivalent to the formula

$$Z_t = \exp\left\{X_t - \frac{1}{2}[X,X]_t^c\right\}\prod_{s\le t}(1+\Delta X_s)\exp\{-\Delta X_s\}.$$

Since $X_t - \frac{1}{2}[X,X]_t^c$ is a semimartingale, and e^x is C^2, we need only show that $\prod_{s\le t}(1+\Delta X_s)\exp\{-\Delta X_s\}$ is càdlàg, adapted, and of finite variation. It will then be a semimartingale too, and thus Z will be a semimartingale. The product is clearly càdlàg, adapted; it thus suffices to show the product converges and is of finite variation.

Since X has càdlàg paths, there are only a finite number of s such that $|\Delta X_s| \ge 1/2$ on each compact interval (fixed ω). Thus it suffices to show

$$V_t = \prod_{0<s\leq t} \left(1+\Delta X_s 1_{\{|\Delta X_s|<1/2\}}\right) \exp\left\{-\Delta X_s 1_{\{|\Delta X_s|<1/2\}}\right\}$$

converges and is of finite variation. Let $U_s = \Delta X_s 1_{\{|\Delta X_s|\leq 1/2\}}$. Then we have $\log V_t = \sum_{s\leq t}\{\log(1+U_s)-U_s\}$, which is an absolutely convergent series a.s., since $\sum_{0<s\leq t}(U_s)^2 \leq [X,X]_t < \infty$ a.s., because $|\log(1+x)-x| \leq x^2$ when $|x|<1/2$. Thus $\log(V_t)$ is a process with paths of finite variation, and hence so also is $\exp\{\log V_t\} = V_t$.

To show that Z is a solution, we set $K_t = X_t - \frac{1}{2}[X,X]_t^c$, and let $f(x,y) = ye^x$. Then $Z_t = f(K_t, S_t)$, where $S_t = \prod_{0\leq s\leq t}(1+\Delta X_s)\exp\{-\Delta X_s\}$. By the change of variables formula we have

$$\begin{aligned} Z_t - 1 &= \int_{0+}^t Z_{s-}dK_s + \int_{0+}^t e^{K_{s-}}dS_s + \frac{1}{2}\int_{0+}^t Z_{s-}d[K,K]_s^c \qquad (*) \\ &\quad + \sum_{0<s\leq t}\left(Z_s - Z_{s-} - Z_{s-}\Delta K_s - e^{K_{s-}}\Delta S_s\right) \\ &= \int_{0+}^t Z_{s-}dX_s - \frac{1}{2}\int_{0+}^t Z_{s-}d[X,X]_s^c + \int_{0+}^t e^{K_{s-}}dS_s \\ &\quad + \frac{1}{2}\int_{0+}^t Z_{s-}d[X,X]_s^c + \sum_{0<s\leq t}\left(Z_s - Z_{s-} - Z_{s-}\Delta K_s - e^{K_{s-}}\Delta S_s\right), \end{aligned}$$

since $[K,S]^c = [S,S]^c = 0$. Note that S is a pure jump process. Hence $\int_{0+}^t e^{K_{s-}}dS_s = \sum_{0<s\leq t} e^{K_{s-}}\Delta S_s$. Also $Z_s = Z_{s-}(1+\Delta X_s)$, and $Z_{s-}\Delta K_s = Z_{s-}\Delta X_s$, so the last sum on the right side of equation $(*)$ becomes

$$\sum_{0<s\leq t}\left(Z_{s-}(1+\Delta X_s) - Z_{s-} - Z_{s-}\Delta X_s - e^{K_{s-}}\Delta S_s\right) = \sum_{0<s\leq t} -e^{K_{s-}}\Delta S_s.$$

Thus equation $(*)$ simplifies due to cancellation to $Z_t - 1 = \int_0^t Z_{s-}dX_s$, and we have the result. □

Definition. For a semimartingale X, $X_0 = 0$, the **stochastic exponential** of X, written $\mathcal{E}(X)$, is the (unique) semimartingale Z that is a solution of $Z_t = 1 + \int_0^t Z_{s-}dX_s$.

The stochastic exponential is also known as the **Doléans-Dade exponential**. Theorem 37 gives a general formula for $\mathcal{E}(X)$. This formula simplifies considerably when X is continuous. Indeed, let X be a continuous semimartingale with $X_0 = 0$. Then

$$\mathcal{E}(X)_t = \exp\{X_t - \frac{1}{2}[X,X]_t\}.$$

An important special case is when the semimartingale X is a multiple λ of a standard Brownian motion $B = (B_t)_{t\geq 0}$. Since λB has no jumps we have

$$\mathcal{E}(\lambda B)_t = \exp\{\lambda B_t - \frac{\lambda^2}{2}[B,B]_t\} = \exp\{\lambda B_t - \frac{\lambda^2}{2}t\}.$$

Moreover, since $\mathcal{E}(\lambda B)_t = 1 + \lambda\int_0^t \mathcal{E}(\lambda B)_{s-}dB_s$ we see that $\mathcal{E}(\lambda B)_t = e^{\lambda B_t - \frac{\lambda^2}{2}t}$ is a continuous martingale. The process $\mathcal{E}(\lambda B)$ is sometimes referred to as *geometric Brownian motion*. Note that the previous theorem gives us $\mathcal{E}(X)$ in closed form. We also have the following pretty result.

Theorem 38. *Let X and Y be two semimartingales with $X_0 = Y_0 = 0$. Then $\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X,Y])$.*

Proof. Let $U_t = \mathcal{E}(X)_t$ and $V_t = \mathcal{E}(Y)_t$. Then the integration by parts formula gives that $U_tV_t - 1 = \int_{0+}^t U_{s-}dV_s + \int_{0+}^t V_{s-}dU_s + [U,V]_t$. Since U and V are exponentials, this is equivalent to

$$\int_{0+}^t U_{s-}V_{s-}dY_s + \int_{0+}^t U_{s-}V_{s-}dX_s + \int_{0+}^t U_{s-}V_{s-}d[X,Y]_s.$$

Letting $W_t = U_tV_t$, we deduce that $W_t = 1 + \int_0^t W_{s-}d(X + Y + [X,Y])_s$ so that $W = \mathcal{E}(X + Y + [X,Y])$, which was to be shown. □

Corollary. Let X be a continuous semimartingale, $X_0 = 0$. Then $\mathcal{E}(X)^{-1} = \mathcal{E}(-X + [X,X])$.

Proof. By Theorem 38,

$$\mathcal{E}(X)\mathcal{E}(-X + [X,X]) = \mathcal{E}(X + (-X + [X,X]) + [X,-X]),$$

since $[-X,[X,X]] = 0$. However $\mathcal{E}(0) = 1$, and we are done. □

In Sect. 9 of Chap. V we consider general linear equations. In particular, we obtain an explicit formula for the solution of the equation

$$X_t = H_t + \int_0^t X_s dZ_s,$$

where Z is a continuous semimartingale. We also consider more general inverses of stochastic exponentials. See, for example, Theorem 63 of Chap. V.

Another application of the change of variables theorem (and indeed of the stochastic exponential) is a proof of Lévy's characterization of Brownian motion in terms of its quadratic variation.

Theorem 39 (Lévy's Theorem). *A stochastic process $X = (X_t)_{t\geq 0}$ is a standard Brownian motion if and only if it is a continuous local martingale with $[X,X]_t = t$.*

Proof. We have already observed that a Brownian motion B is a continuous local martingale and that $[B,B]_t = t$ (see the remark following Theorem 22). Thus it remains to show sufficiency. Fix $u \in \mathbb{R}$ and set $F(x,t) = \exp\{iux + \frac{u^2}{2}t\}$. Let $Z_t = F(X_t,t) = \exp\{iuX_t + \left(\frac{u^2}{2}\right)t\}$. Since $F \in \mathcal{C}^2$ we can apply Itô's formula (Theorem 33) to obtain

$$\begin{aligned} Z_t &= 1 + iu\int_0^t Z_s dX_s + \frac{u^2}{2}\int_0^t Z_s ds - \frac{u^2}{2}\int_0^t Z_s d[X,X]_s \\ &= 1 + iu\int_0^t Z_s dX_s, \end{aligned}$$

which is the exponential equation. Since X is a continuous local martingale, we now have that Z is also one (complex-valued, of course) by the martingale preservation property. Moreover stopping Z at a fixed time t_0, Z^{t_0}, we have that Z^{t_0} is bounded and hence a martingale. It then follows for $0 \leq s < t$ that

$$E\{\exp\{iu(X_t - X_s)\}|\mathcal{F}_s\} = \exp\left\{-\frac{u^2}{2}(t-s)\right\}.$$

Since this holds for any $u \in \mathbb{R}$ we conclude that $X_t - X_s$ is independent of $\mathcal{F}_s$ and that it is normally distributed with mean zero and variance $(t-s)$. Therefore X is a Brownian motion. □

Observe that if M and N are two continuous martingales such that MN is a martingale, then $[M,N] = 0$ by Corollary 2 of Theorem 27. Therefore if $\mathbf{B}_t = (B_t^1, \ldots, B_t^n)$ is an n-dimensional standard Brownian motion, $B_t^i B_t^j$ is a martingale for $i \neq j$, and we have that

$$[B^i, B^j]_t = \begin{cases} t, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Theorem 39 then has a multi-dimensional version, which has an equally simple proof.

Theorem 40 (Lévy's Theorem: Multi-dimensional Version). *Let* $\mathbf{X} = (X^1, \ldots, X^n)$ *be continuous local martingales such that*

$$[X^i, X^j]_t = \begin{cases} t, & \textit{if } i = j, \\ 0, & \textit{if } i \neq j. \end{cases}$$

Then $\mathbf{X}$ *is a standard* n*-dimensional Brownian motion.*

As another application of Itô's formula, we exhibit the relationship between harmonic and subharmonic functions and martingales.

Theorem 41. *Let* $\mathbf{X} = (X^1, \ldots, X^n)$ *be an* n*-dimensional continuous local martingale with values in an open subset* D *of* $\mathbb{R}^n$*. Suppose that* $[X^i, X^j] = 0$ *if* $i \neq j$*, and* $[X^i, X^i] = A$*,* $1 \leq i \leq n$*. Let* $u : D \to \mathbb{R}$ *be harmonic (resp. subharmonic). Then* $u(\mathbf{X})$ *is a local martingale (resp. submartingale).*

Proof. By Itô's formula (Theorem 33) we have

$$u(\mathbf{X}_t) - u(\mathbf{X}_0) = \int_{0+}^{t} \nabla u(\mathbf{X}_s) \cdot d\mathbf{X}_s + \frac{1}{2}\int_{0+}^{t} \Delta u(\mathbf{X}_s) dA_s$$

where the "dot" denotes Euclidean inner product, ∇ denotes the gradient, and Δ denotes the Laplacian. If u is harmonic (subharmonic), then $\Delta u = 0 (\Delta u \geq 0)$ and the result follows. □

If $\mathbf{B}$ is a standard n-dimensional Brownian motion, then $\mathbf{B}$ satisfies the hypotheses of Theorem 41 with the process $A_t = t$. That $u(\mathbf{B}_t)$ is a submartingale (resp. supermartingale) when u is subharmonic (resp. superharmonic) is the motivation for the terminology submartingale and supermartingale. The relationship between stochastic calculus and potential theory suggested by Theorem 41 has proven fruitful (see, for example, Doob [56]).

Lévy's characterization of Brownian motion (Theorem 39) allows us to prove a useful *change of time* result.

Theorem 42. *Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale with $M_0 = 0$ and such that $\lim_{t\to\infty}[M,M]_t = \infty$ a.s. Let*

$$T_s = \inf\{t > 0 : [M,M]_t > s\}.$$

Define $\mathcal{G}_s = \mathcal{F}_{T_s}$ and $B_s = M_{T_s}$. Then $(B_s, \mathcal{G}_s)_{s\geq 0}$ is a standard Brownian motion. Moreover $([M,M]_t)_{t\geq 0}$ are stopping times for $(\mathcal{G}_s)_{s\geq 0}$ and

$$M_t = B_{[M,M]_t} \quad \text{a.s.} \quad 0 \leq t < \infty.$$

That is, M can be represented as a time change of a Brownian motion.

Proof. The $(T_s)_{s\geq 0}$ are stopping times by Theorem 3 of Chap. I. Each T_s is finite a.s. by the hypothesis that $\lim_{t\to\infty}[M,M]_t = \infty$ a.s. Therefore the σ-fields $\mathcal{G}_s = \mathcal{F}_{T_s}$ are well-defined. The filtration $(\mathcal{G}_s)_{s\geq 0}$ need not be right continuous, but one can take $\mathcal{H}_s = \mathcal{G}_{s+} = \mathcal{F}_{T_s+}$ to obtain one. Note further that $\{[M,M]_t \leq s\} = \{T_s \geq t\}$, hence $([M,M]_t)_{t\geq 0}$ are stopping times for the filtration $\mathcal{G} = (\mathcal{G}_s)_{s\geq 0}$.

By Corollary 3 of Theorem 27 we have $E\{M_{T_s}^2\} = E\{[M,M]_{T_s}\} = s < \infty$, since $[M,M]_{T_s} = s$ identically because $[M,M]$ is continuous. Thus the time changed process is square integrable. Moreover

$$E\{B_u - B_s | \mathcal{G}_s\} = E\{M_{T_u} - M_{T_s} | \mathcal{F}_{T_s}\} = 0$$

by the Optional Sampling Theorem. Also

$$\begin{aligned} E\{B_u^2 - B_s^2 | \mathcal{G}_s\} &= E\{(B_u - B_s)^2 | \mathcal{G}_s\} \\ &= E\{(M_{T_u} - M_{T_s})^2 | \mathcal{F}_{T_s}\} \\ &= E\{[M,M]_{T_u} - [M,M]_{T_s} | \mathcal{F}_{T_s}\} \\ &= u - s. \end{aligned}$$

Therefore $B_s^2 - s$ is a martingale, whence $[B,B]_s = s$ provided B has continuous paths, by Corollary 2 to Theorem 27.

We want to show that $B_s = M_{T_s}$ has continuous paths. However by Corollary 5 of Theorem 27 almost surely all intervals of constancy of $[M,M]$ are also intervals of constancy of M. It follows easily that B is continuous. It remains to show that $M_t = B_{[M,M]_t}$. Since $B_s = M_{T_s}$, we have that $B_{[M,M]_t} = M_{T_{[M,M]_t}}$, a.s. Since $(T_s)_{s\geq 0}$ is the right continuous inverse of $[M,M]$, we have that $T_{[M,M]_t} \geq t$, with equality holding if and only if t is a point of right increase of $[M,M]$. (If $(T_s)_{s\geq 0}$ were continuous, then we would always have that $T_{[M,M]_t} = t$.) However $T_{[M,M]_t} > t$ implies that $t \mapsto [M,M]_t$ is constant on the interval $(t, T_{[M,M]_t})$; thus by Corollary 5 of Theorem 27 we conclude M is constant on $(t, T_{[M,M]_t})$. Therefore $B_{[M,M]_t} = M_{T_{[M,M]_t}} = M_t$ a.s., and we are done. □

Another application of the change of variables formula is the determination of the distribution of *Lévy's stochastic area process*. Let $\mathbf{B}_t = (X_t, Y_t)$ be an $\mathbb{R}^2$-valued Brownian motion with $(X_0, Y_0) = (0,0)$. Then during the times s to $s+ds$ the chord from the origin to $\mathbf{B}$ sweeps out a triangular region of area $\frac{1}{2}R_s dN_s$, where

$$R_s = \sqrt{X_s^2 + Y_s^2}$$

and

$$dN_s = -\frac{Y_s}{R_s}dX_s + \frac{X_s}{R_s}dY_s.$$

Therefore the integral $A_t = \int_0^t R_s dN_s = \int_0^t (-Y_s dX_s + X_s dY_s)$ is equal to twice the area swept out from time 0 until time t. Paul Lévy found the characteristic function of A_t and therefore determined its distribution. Theorem 43 is known as **Lévy's stochastic area formula.**

Theorem 43. *Let $\mathbf{B}_t = (X_t, Y_t)$ be an $\mathbb{R}^2$-valued Brownian motion, $\mathbf{B}_0 = (0,0)$, $u \in \mathbb{R}$. Let $A_t = \int_0^t X_s dY_s - \int_0^t Y_s dX_s$. Then*

$$E\{e^{iuA_t}\} = \frac{1}{\cosh(ut)}, \quad 0 \leq t < \infty, -\infty < u < \infty.$$

Proof. Let $\alpha(t)$, $\beta(t)$ be C^1 functions, and set

$$V_t = iuA_t - \frac{\alpha(t)}{2}(X_t^2 + Y_t^2) + \beta(t).$$

Then

$$\begin{aligned} dV_t &= iudA_t - \frac{\alpha'(t)}{2}(X_t^2 + Y_t^2)dt \\ &\quad - \alpha(t)\{X_t dX_t + Y_t dY_t + dt\} + \beta'(t)dt \\ &= (-iuY_t - \alpha(t)X_t)dX_t + (iuX_t - \alpha(t)Y_t)dY_t \\ &\quad - \frac{1}{2}dt\{\alpha'(t)X_t^2 + \alpha'(t)Y_t^2 + 2\alpha(t) - 2\beta'(t)\}. \end{aligned}$$

Next observe that, from the above calculation,

$$\begin{aligned} d[V,V]_t &= (-iuY_t - \alpha(t)X_t)^2 dt + (iuX_t - \alpha(t)Y_t)^2 dt \\ &= (\alpha(t)^2 - u^2)(X_t^2 + Y_t^2)dt. \end{aligned}$$

Using the change of variables formula and the preceding calculations

$$\begin{aligned} de^{V_t} &= e^{V_t}(dV_t + \frac{1}{2}d[V,V]_t) \\ &= e^{V_t}(-iuY_t - \alpha(t)X_t)dX_t + e^{V_t}(iuX_t - \alpha(t)Y_t)dY_t \\ &\quad + \frac{1}{2}e^{V_t}dt\{(\alpha(t)^2 - u^2 - \alpha'(t))(X_t^2 + Y_t^2) + 2\beta'(t) - 2\alpha(t)\}. \end{aligned}$$

Therefore e^{V_t} is a local martingale provided

$$\begin{aligned} \alpha'(t) &= \alpha(t)^2 - u^2 \\ \beta'(t) &= \alpha(t). \end{aligned}$$

Next we fix $t_0 > 0$ and solve the above ordinary differential equations with $\alpha(t_0) = \beta(t_0) = 0$. The solution is

$$\begin{aligned} \alpha(t) &= u\ \tanh(u(t_0 - t)) \\ \beta(t) &= -\log\cosh(u(t_0 - t)), \end{aligned}$$

where tanh and cosh are hyperbolic tangent and hyperbolic cosine, respectively. Note that for $0 \le t \le t_0$,

$$|e^{V_t}| = \exp\{-\alpha(t)\left\{\frac{X_t^2 + Y_t^2}{2}\right\} + \beta(t)\} \le e^{\beta(t)} \le 1.$$

Thus e^{V_t}, $0 \le t \le t_0$ is bounded and is therefore a true martingale, not just a local martingale, by Theorem 51 of Chap. I. Therefore,

$$E\{e^{V_{t_0}}\} = E\{e^{V_0}\}.$$

However, $V_{t_0} = iuA_{t_0}$ since $\alpha(t_0) = \beta(t_0) = 0$. As $A_0 = X_0 = Y_0 = 0$, it follows that $V_0 = -\log\ \cosh(ut_0)$. We conclude that

$$\begin{aligned} E\{e^{iuA_{t_0}}\} &= \exp\{-\log\cosh(ut_0)\} \\ &= \frac{1}{\cosh(ut_0)}, \end{aligned}$$

and the proof is complete. □

There are of course other proofs of Lévy's stochastic area formula (e.g., Yor [246], or Lévy [150]). As a corollary to Theorem 43 we obtain the density for the distribution of A_t.

Corollary. Let $\mathbf{B}_t = (X_t, Y_t)$ be an $\mathbb{R}^2$-valued Brownian motion, $\mathbf{B}_0 = (0,0)$, and set $A_t = \int_0^t X_s dY_s - \int_0^t Y_s dX_s$. Then the density function for the distribution of A_t is

$$f_{A_t}(x) = \frac{1}{2t\cosh(\pi x/2t)}, \qquad -\infty < x < \infty.$$

Proof. By Theorem 43, the Fourier transform (or characteristic function) of A_t is $E\{e^{iuA_t}\} = \frac{1}{\cosh(ut)}$. Thus we need only to calculate $\frac{1}{2\pi}\int \frac{e^{-iux}}{\cosh(ut)}du$. The integrand is of the form $f(z) = \frac{P(z)}{Q(z)} = \frac{e^{-izx}}{\cosh(zt)}$. Since $\cosh(zt)$ has a pole at $z_0 = \frac{i\pi}{2t}$, we have

$$\mathrm{Res}(f, z_0) = \frac{P(z_0)}{Q'(z_0)} = \frac{e^{\pi x/2t}}{it}.$$

Next we integrate along the closed curve C_r traversed counter clockwise, and given by

$$C_r : \begin{cases} y = 0, & -r \le x \le r, & C_r^1 \\ x = r, & 0 \le y \le \frac{\pi}{t}, & C_r^2 \\ y = \frac{\pi}{t}, & r \ge x \ge -r, & C_r^3 \\ x = -r, & \frac{\pi}{t} \ge y \ge 0, & C_r^4 \end{cases}$$

as shown in Fig. 1 below.

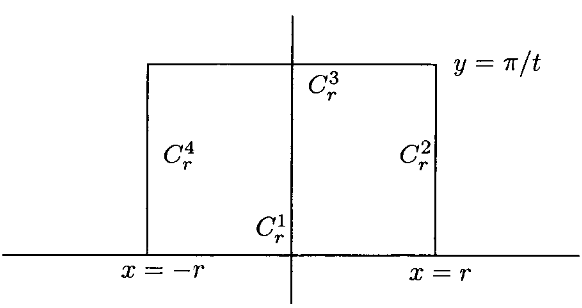

Fig. 1. The closed curve of integration $C_r = C_r^1 + C_r^2 + C_r^3 + C_r^4$.

Therefore $\int_{C_r} f(z)dz = 2\pi i\,\mathrm{Res}(f, z_0)$. Along C_r^3 the integral is

$$-\int_r^{-r} \frac{e^{-i(u+\frac{\pi i}{t})x}}{\cosh(ut)}du = e^{\frac{\pi x}{t}} \int_{-r}^{r} \frac{e^{-iux}}{\cosh(ut)}du.$$

The integrands on C_r^2 and C_r^4 are dominated by $2e^{-r}$, and therefore

$$\lim_{r\to\infty}\int_{C_r} f(z)dz = (1+e^{\pi x/t})\int_{-\infty}^{\infty}\frac{e^{-iux}}{\cosh(ut)}du$$
$$= 2\pi i \operatorname{Res}(f, z_0)$$
$$= 2\pi\frac{e^{\pi x/2t}}{t}.$$

Finally we can conclude

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-iux}}{\cosh(ut)}du = \frac{1}{2\pi}\left(\frac{2\pi e^{\frac{\pi x}{2t}}}{t(1+e^{\frac{\pi x}{t}})}\right)$$
$$= \frac{1}{t(e^{\frac{\pi x}{2t}}+e^{\frac{-\pi x}{2t}})}$$
$$= \frac{1}{2t\cosh(\frac{\pi x}{2t})}. \qquad \square$$

The stochastic area process A shares some of the properties of Brownian motion, as is seen by recalling that $A_t = \int_0^t R_s dN_s$, where N is a Brownian motion by Lévy's Theorem (Theorem 39), and N and R are independent (this must be proven, of course). For example A satisfies a *reflection principle*. If one changes the sign of the increments of A after a stopping time, the process obtained thereby has the same distribution as that of A. One can use this fact to show, for example, that if $S_t = \sup_{0\le s\le t} A_s$, then S_t has the same distribution as $|A_t|$, for $t>0$.

Bibliographic Notes

The definition of semimartingale and the treatment of stochastic integration as a Riemann-type limit of sums is in essence new. It has its origins in the fundamental theorem of Bichteler [13, 14], and Dellacherie [42]. The pedagogic approach used here was first suggested by Meyer [176], and it was then outlined by Dellacherie [42]. Dellacherie's outline was further expanded by Lenglart [145] and Protter [201, 202]. These ideas were also present in the work of M. Métivier and J. Pellaumail, and to a lesser extent Kussmaul, although this was not appreciated at the time by us. See for example [160], [162], and [138]. A similar idea was developed by Letta [148]. A recent treatment is in Bichteler [15].

We will not attempt to give a comprehensive history of stochastic integration here, but rather just a sketch. The important early work was that of Wiener [228, 229], and then, of course, Itô [93, 94, 95, 97, 98]. However, it has recently been discovered, by B. Bru and M. Yor, that the young French mathematician W. Doeblin[6] had essentially established Itô's formula in early

[6] Wolfgang Doeblin (Döblin) was born in 1915 in Berlin, and after acquiring French citizenship in 1936 was known as Vincent Dobbin. He was killed in June 1940.

1940, but was killed later that year in action in the Second World War, and thus his results were lost until recently [26, 27]. Doob stressed the martingale nature of the Itô integral in his book [55] and proposed a general martingale integral. Doob's proposed development depended on a decomposition theorem (the Doob-Meyer decomposition, Theorem 8 of Chap. III) which did not yet exist. Meyer proved this decomposition theorem in [163, 164], and commented that a theory of stochastic integration was now possible. This was begun by Courrège [36], and extended by Kunita and Watanabe [134], who revealed an elegant structure of square integrable martingales and established a general change of variables formula. Meyer [166, 167, 168, 169] extended Kunita and Watanabe's work, realizing that the restriction of integrands to predictable processes is essential. He also extended the integrals to local martingales, which had been introduced earlier by Itô and Watanabe [102]. Up to this point, stochastic integration was tied indirectly to Markov processes, by the assumption that the underlying filtration of σ-algebras be "quasi-left continuous." This hypothesis was removed by Doléans-Dade and Meyer [53], thereby making stochastic integration a purely martingale theory. It was also in this article that semimartingales were first proposed in the form we refer to as classical semimartingales in Chap. III.

A different theory of stochastic integration was developed independently by McShane [154, 155], which was close in spirit to the approach given here. However it was technically complicated and not very general. It was shown in Protter [200] (building on the work of Pop-Stojanovic [194]) that the theory of McShane could for practical purposes be viewed as a special case of the semimartingale theory.

The subject of stochastic integration essentially lay dormant for six years until Meyer [171] published a seminal "course" on stochastic integration. It was here that the importance of semimartingales was made clear, but it was not until the late 1970's that the theorem of Bichteler [13, 14], and Dellacherie [42] gave an *a posteriori* justification of semimartingales. The seemingly *ad hoc* definition of a semimartingale as a process having a decomposition into the sum of a local martingale and an FV process was shown to be the most general reasonable stochastic differential possible. (See also Kussmaul [138] in this regard, and the bibliographic notes in Chap. III.)

Most of the results of this chapter can be found in Meyer [171], though they are proven for classical semimartingales and hence of necessity the proofs are much more complicated. Theorem 4 (Stricker's Theorem) is (of course) due to Stricker [218]; see also Meyer [173]. Theorem 5 is due to Meyer [175]. There are many other methods of expanding a filtration and still preserving the semimartingale property. For further details, see Chap. VI on expansion of filtrations.

Theorem 14 is originally due to Lenglart [142]. Theorem 16 is known to be true only in the case of integrands in $\mathbb{L}$. The local behavior of the integral (Theorems 17 and 18) is due to Meyer [171] (see also McShane [154]). The a.s.

Kunita-Watanabe inequality, Theorem 25, is due to Meyer [171], while the expected version (the corollary to Theorem 25) is due to Kunita-Watanabe [134].

That continuous martingales have paths of infinite variation or are constant a.s. was first published by Fisk [73] (Corollary 1 of Theorem 27). The proof given here of Corollary 4 of Theorem 27 (that a continuous local martingale X and its quadratic variation $[X, X]$ have the same intervals of constancy) is due to Maisonneuve [151]. The proof of Itô's formula (Theorem 32) is by now classic; however we benefited from Föllmer's presentation of it [75]. A popular alternative proof which basically bootstraps up form the formula for integration by parts, can be found (for example) on pages 57-58 of J. Jacod and A. N. Shiryaev [110].

The Fisk-Stratonovich integral was developed independently by Fisk [73] and Stratonovich [217], and it was extended to general semimartingales by Meyer [171]. Theorem 35 is inspired by the work of Getoor and Sharpe [215].

The stochastic exponential of Theorem 37 is due to Doléans-Dade [49]. It has become extraordinarily important. See, for example, Jacod-Shiryaev [110]. The pretty formula of Theorem 38 is due to Yor [237]. Exponentials have of course a long history in analysis. For an insightful discussion of exponentials see Gill-Johansen [82]. That every continuous local martingale is the time change of a Brownian motion is originally due to Dubins-Schwarz [59] and Dambis [37]. The proof of Lévy's stochastic area formula (Theorem 43) is new and is due to Janson. See Janson-Wichura [113] for related results. The original result is in Lévy [150], and another proof can be found in Yor [246].

Exercises for Chapter II

Exercise 1. Let B be standard Brownian motion and let f be a function mapping $\mathbb{R} \to \mathbb{R}$ which is continuous except for one point where there is a jump discontinuity. Show that $X_t = f(B_t)$ cannot be a semimartingale.

Exercise 2. Let $Q << P$ (Q and P are both probability measures and Q is absolutely continuous with respect to P). Show that if $X_n \to X$ in P-probability, then $X_n \to X$ in Q-probability.

Exercise 3. Let (X, Y) be standard two dimensional Brownian motion. (That is, X and Y are independent one dimensional standard Brownian motions.) Let $0 < \alpha < 1$. Set

$$B_t = \alpha X_t + \sqrt{1 - \alpha^2} Y_t.$$

Show that B is a standard one dimensional Brownian motion. Compute $[X, B]$ and $[Y, B]$.

Exercise 4. Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be non-decreasing and continuous. Show that there exists a continuous martingale M such that $[M, M]_t = f(t)$.

Exercise 5. Let B be standard Brownian motion and let $H \in \mathbb{L}$ be such that $|H_t| = 1$, all t. Let $M_t = \int_0^t H_s dB_s$. Show that M is also a Brownian motion.

Exercise 6. Give an example of semimartingales X^n, X, Y such that $\lim_{n\to\infty} X_t^n = X_t$ a.s., all t, but $\lim_{n\to\infty}[X^n, Y]_t \neq [X, Y]_t$ a.s., for some t.

Exercise 7. Let Y, Z be semimartingales, and let H^n, $H \in \mathbb{L}$. Suppose $\lim_{n\to\infty} H^n = H$ in *ucp*. Let $X^n = \int H_s^n dY_s$ and $X = \int H_s dY_s$. Show that $\lim_{n\to\infty}[X^n, Z]_t = [X, Z]_t$ in *ucp*, all t.

Exercise 8. Suppose f and all f_n are C^2, and that f_n and f_n' converge uniformly on compacts to f and f', respectively. Let X, Y be continuous semimartingales and show that $\lim_{n\to\infty}[f_n(X), Y] = [f(X), Y]$.

Exercise 9. Let M be a continuous local martingale and let A be a continuous, adapted process with paths of finite variation on compact time sets, with $A_0 = 0$. Let $X = M + A$. Show that $[X, X] = [M, M]$ a.s.

Exercise 10. Let X be a semimartingale and let Q be another probability law, with $Q << P$. Let $[X, X]^P$ denote the quadratic variation of X considered as a P semimartingale. Show that $[X, X]^Q = [X, X]^P$, Q-a.s.

Exercise 11. Let B be standard Brownian motion, and let $T = \inf\{t > 0 : B_t = 1 \text{ or } B_t = -2\}$.

(a) Show that T is a stopping time.
(b) Let $M = B_{t\wedge T}$ and let $N = -M$. Show that M and N are continuous martingales and that $[M, M]_t = [N, N]_t = t \wedge T$. Show that nevertheless M and N have different distributions.

Exercise 12. Let X be a semimartinagle such that $\sum_{0<s\leq t} |\Delta X_s| < \infty$ a.s., each $t > 0$. Let $f \in C^2$ and show that $\sum_{0<s\leq t} |\Delta f(X_s)| < \infty$ a.s., each $t > 0$ as well.

Exercise 13. Let X be a semimartingale and suppose H^n and H are in $\mathbb{L}$. Further suppose that $H^n \to H$ in *ucp*. Show that $\int H_s^n dX_s \to \int H_s dX_s$ in *ucp*.

***Exercise 14.** Let X be a semimartingale, and let A be a non-negative, continuous, increasing process with $A_\infty = \infty$ a.s. Suppose $\lim_{t\to\infty} \int_0^t (1 + A_s)^{-1} dX_s$ exists and is finite a.s. Show that $\lim_{t\to\infty} \frac{X_t}{A_t} = 0$ a.s.

Exercise 15. Let M be a continuous local martingale and assume $[M, M]_\infty = \infty$ a.s. Show that $\lim_{t\to\infty} \frac{M_t}{[M,M]_t} = 0$ a.s.

Exercise 16. Let B be a standard Brownian motion and let H be adapted and continuous. Show that for fixed t,

$$\lim_{h\to 0} \frac{1}{B_{t+h} - B_t} \int_t^{t+h} H_s dB_s = H_t,$$

with convergence in probability.

Exercise 17. Let Y be a continuous, adapted process with $Y_0 = 0$ and with Y constant on $[1, \infty)$. Let

$$X_t^n = n \int_{t-\frac{1}{n}}^{t} Y_s ds 1_{\{t > \frac{1}{n}\}}.$$

Show that X^n is a semimartingale, each n, but that $\lim_{n\to\infty} X^n = Y$ need not be a semimartingale.

Exercise 18. (Continuation of Exercise 17.) Let B be a standard Brownian motion and let $X^n = n \int_{t-\frac{1}{n}}^{t} B_s ds 1_{\{t>\frac{1}{n}\}}$. Solve

$$dZ_t^n = Z_s^n dX_s^n, \qquad Z_0^n = 1,$$

and show that $\lim_{n\to\infty} X_t^n = B_t$ a.s., each t, but that $\lim_{n\to\infty} Z^n \neq Z$, where Z solves the equation $dZ_t = Z_t dB_t$.

Exercise 19. (Related to Exercises 17 and 18.) Let $A_t^n = \frac{1}{n}\sin(nt)$, $0 \le t \le \pi/2$. Show that A^n is a semimartingale for each n, and that $\int_0^{\frac{\pi}{2}} |dA_s^n| = 1$, each n (this means the total variation process, path-by-path, of A), but that

$$\lim_{n\to\infty} \sup_{0\le t\le \frac{\pi}{2}} |A_t^n| = 0.$$

***Exercise 20.** (This exercise refers to an example briefly considered in Sect. 6 of Chap. I on local martingles, and again in Sect. 6 of Chap. II.) Let B be a standard Brownian motion in $\mathbb{R}^3$, and let $u : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ be given by $u(y) = \|y\|^{-1}$. Assume $B_0 = x$ a.s., where $x \neq 0$, show the following.

(a) $M_t = u(B_t)$ is a local martingale.
(b) $\lim_{t\to\infty} E^x(u(B_t)) = 0$.
(c) $u(B_t) \in L^2$, each $t \ge 0$.
(d) $\sup_{t\ge 0} E^x(u(B_t)^2) < \infty$.

Conclude that $u(B)$ is not a martingale.

Exercise 21. Let A, C be two non-decreasing càdlàg processes with $A_0 = C_0 = 0$. Assume that both $A_\infty = \lim_{t\to\infty} A_t$ and $C_\infty = \lim_{t\to\infty} C_t$ are finite.

(a) Show that $A_\infty C_\infty = \int_0^\infty (A_\infty - A_s) dC_s + \int_0^\infty (C_\infty - C_{s-}) dA_s$.
(b) Deduce from part (a) the general formula

$$A_\infty C_\infty = \int_0^\infty A_s dC_s + \int_0^\infty C_{s-} dAs.$$

Note that the above formulae are not symmetric. See also Theorem 3 in Chap. VI.

Exercise 22. Let A be non-decreasing and continuous, and assume $A_\infty = \lim_{t\to\infty} A_t$ is finite, $A_0 = 0$. Show that for integer $p > 0$,

$$(A_\infty)^p = p! \int_0^\infty dA_{s_1} \int_{s_1}^\infty dA_{s_2} \cdots \int_{s_{p-1}}^\infty dA_{s_p}.$$

***Exercise 23 (expansion of filtrations).** Let B be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, and define a new filtration $\mathbb{G}$ by $\mathcal{G}_t^0 = \mathcal{F}_t \vee \sigma(B_1)$ and $\mathcal{G}_t = \bigcap_{u>t} \mathcal{G}_u^0$.

(a) Show that $E\{B_t - B_s | \mathcal{G}_s\} = \frac{t-s}{1-s}(B_1 - B_s)$ for $0 \le s \le t \le 1$.
(b) Show that

$$B_t = \beta_t + \int_0^{t\wedge 1} \frac{B_1 - B_s}{1-s} ds$$

where β is a $\mathbb{G}$-Brownian motion. (Note in particular that the process $(B_t - \int_0^t \frac{B_1 - B_s}{1-s} ds)_{0\le t<1}$ is a $\mathbb{G}$ martingale on $[0,1)$.)
(c) Show that

$$\int_0^1 \left|\frac{B_1 - B_s}{1-s}\right| ds < \infty \quad \text{a.s.}$$

and conclude that B is a $\mathbb{G}$ semimartingale.

This is a special case of Theorem 3 in Chap. VI.

Exercise 24. With the notation of Exercise 23, let $X_t = B_t - tB_1$, $0 \le t \le 1$.

(a) Show that

$$X_t = (1-t) \int_0^t \frac{1}{1-s} d\beta_s.$$

(b) Deduce from (a) that X satisfies the stochastic differential equation

$$X_t = \beta_t + \int_0^t -\frac{X_s}{1-s} ds.$$

(X is called a **Brownian bridge** from 0 to 0.)

***Exercise 25.** Let $(B_t, \mathcal{F}_t)_{t\ge 0}$ be a standard Brownian motion. Let $a, b > 0$ and $T = \inf\{t > 0 : B_t \in (-a, b)^c\}$.

(a) Show that

$$X_t = \exp\{\frac{1}{2}\theta^2 t\} \cos\left[\theta(B_t - \frac{b-a}{2})\right]$$

is a local martingale.
(b) Show that

$$X_T = \exp\{\frac{1}{2}\theta^2 T\} \cos(\frac{a+b}{2}\theta).$$

(c) For $0 \leq \theta < \frac{\pi}{a+b}$, show that X^T is a positive supermartingale, and deduce

$$\cos(\frac{a+b}{2}\theta)E\{\exp\{\frac{1}{2}\theta^2 T\}\} \leq \cos(\frac{a-b}{2}\theta).$$

(d) Use (c) to show that $X_T^* = \sup_{s \leq T} |X_s| \in L^1$, and conclude that X is a martingale.

(e) Conclude that

$$E\{\exp\{\frac{1}{2}\theta^2 T\}\} = \frac{\cos(\frac{ab}{2}\theta)}{\cos(\frac{a+b}{2}\theta)}.$$

Exercise 26. Let $(B_t, \mathcal{F}_t)_{t\geq 0}$ be standard Brownian motion, and let $T = \inf\{t > 0 : B_t \in (-d, d)^c\}$. Let $M = B^T$. Show that

(a) if $d < \frac{\pi}{2}$, then $E\{\exp\{\frac{1}{2}[M, M]_T\}\} < \infty$; but

(b) if $d = \frac{\pi}{2}$, then $E\{\exp\{\frac{1}{2}[M, M]_T\}\} = \infty$. (*Hint:* Use Exercise 25.)

Exercise 27. Let $(B_t, \mathcal{F}_t)_{t\geq 0}$ be standard Brownian motion, and let $X_t = e^{-\alpha t}(X_0 + \sigma \int_0^t e^{\alpha s} dB_s)$. Show that X is a solution to the stochastic differential equation

$$dX_t = -\alpha X_t dt + \sigma dB_t.$$

Exercise 28. Let B be a standard Brownian motion and let $\mathcal{E}(B)$ denote the stochastic exponential of B. Show that $\lim_{t\to\infty} \mathcal{E}(B)_t = 0$ a.s.

Exercise 29. Let X be a semimartingale. Show that

$$d(\frac{1}{\mathcal{E}(X)}) = \frac{-dX + d[X, X]}{\mathcal{E}(X)}.$$

Exercise 30. Let B be a standard Brownian motion.

(a) Show that M is a local martingale, where

$$M_t = \exp\{B_t - \frac{1}{2}t\}.$$

(b) Calculate $[M, M]_t$, and show that M is a martingale.

(c) Calculate $E\{e^{B_t}\}$.

The next eight problems involve a topic known as **changes of time**. For these problems, let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfy the usual hypotheses. A **change of time** $R = (R_t)_{t\geq 0}$ is a family of stopping times such that for every $\omega \in \Omega$, the function $R_\cdot(\omega)$ is non-decreasing, right continuous, $R_t < \infty$ a.s., and $R_0 = 0$. Let $\mathcal{G}_t = \mathcal{F}_{R_t}$. Change of time is discussed further in Sect. 3 of Chap. IV.

Exercise 31. Show that $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ satisfies the usual hypotheses.

Exercise 32. Show that if M is an $\mathbb{F}$ uniformly integrable martingale and $\overline{M}_t := M_{R_t}$, then $\overline{M}$ is a $\mathbb{G}$ martingale.

Exercise 33. If M is an $\mathbb{F}$ (right continuous) local martingale, show that $\overline{M}$ is a $\mathbb{G}$ semimartingale.

***Exercise 34.** Construct an example where M is an $\mathbb{F}$ local martingale, but $\overline{M}$ is not a $\mathbb{G}$ local martingale. (*Hint:* Let $(X_n)_{n\in\mathbb{N}}$ be an adapted process. It is a local martingale if and only if $|X_n|dP$ is a σ-finite measure on $\mathcal{F}_{n-1}$, and $E\{X_n|\mathcal{F}_{n-1}\} = X_{n-1}$, each $n \geq 1$. Find X_n where $|X_n|dP$ is not σ-finite on $\mathcal{F}_{n-2}$, any n, and let $R_n = 2n$.)

***Exercise 35.** Let R be a time change, with $s \mapsto R_s$ continuous, strictly increasing, $R_0 = 0$, and $R_t < \infty$, each $t \geq 0$. Show that for a continuous semimartingale X,

$$\int_0^{R_t} H_s dX_s = \int_0^t H_{R_u} dX_{R_u}$$

for bounded $H \in \mathbb{L}$.

***Exercise 36.** Let R and X be as in Exercise 35. No longer assume that $R_t < \infty$ a.s., each $t \geq 0$, but instead assume that X is a finite variation process. Let $A_t = \inf\{s > 0 : R_s > t\}$.

(a) Show that R strictly increasing implies that A is continuous.
(b) Show that R continuous implies that A is strictly increasing.
(c) Show that for general R, $R_{A_t} \geq t$, and if R is strictly increasing and continuous then $R_{A_t} = t$.
(d) Show that for bounded $H \in \mathbb{L}$ we have

$$\int_{R_0}^{R_t} H_s dX_s = \int_0^t 1_{(R_s<\infty)} H_{R_s} dX_{R_s} = \int_0^{t\wedge A_\infty} H_{R_s} dX_{R_s}.$$

(e) Show that for bounded $H \in \mathbb{L}$ we have

$$\int_{R_0}^{R_t} H_s dA_s = \int_0^{t\wedge A_\infty} H_{R_s} ds.$$

See in this regard *Lebesgue's change of time formula*, given in Theorem 45 of Chap. IV.

***Exercise 37.** Let R be a change of time and let $\mathbb{G}$ be the filtration given by $\mathcal{G}_t = \mathcal{F}_{R_t}$. Let $A_t = \inf\{s > 0 : R_s > t\}$. Show that $A = (A_t)_{t\geq 0}$ is a change of time for the filtration $\mathbb{G}$. Show also that if $t \to R_t$ is continuous a.s., $R_0 = 0$, and $R_\infty = \infty$, then $R_{A_t} = t$ a.s., $t \geq 0$.

***Exercise 38.** Let A, $\mathbb{G}$, be as in Exercise 37 and suppose that $R_{A_t} = t$ a.s., $t \geq 0$. Show that $\mathcal{G}_{A_t} \subset \mathcal{F}_t$, each $t \geq 0$.

***Exercise 39.** A function is **Hölder continuous** of order α if $|f(x)-f(y)| \leq K|x-y|^\alpha$. Show that the paths of a standard Brownian motion are a.s. nowhere locally Hölder continuous of order α for any $\alpha > 1/2$. (*Hint:* Use the fact that $\lim_{n\to\infty} \sum_{\pi^n[0,t]} (B_{t_{i+1}} - B_{t_i})^2 = t$.)

III

Semimartingales and Decomposable Processes

1 Introduction

In Chap. II we defined a semimartingale as a good integrator and we developed a theory of stochastic integration for integrands in $\mathbb{L}$, the space of adapted processes with left continuous, right-limited paths. Such a space of integrands suffices to establish a change of variables formula (or "Itô's formula"), and it also suffices for many applications, such as the study of stochastic differential equations. Nevertheless the space $\mathbb{L}$ is not general enough for the consideration of such important topics as local times and martingale representation theorems. We need a space of integrands analogous to measurable functions in the theory of Lebesgue integration. Thus defining an integral as a limit of sums—which requires a degree of smoothness on the sample paths—is inadequate. In this chapter we lay the groundwork necessary for an extension of our space of integrands, and the stochastic integral is then extended in Chap. IV.

Historically the stochastic integral was first proposed for Brownian motion, then for continuous martingales, then for square integrable martingales, and finally for processes which can be written as the sum of a locally square integrable local martingale and an adapted, càdlàg processes with paths of finite variation on compacts; that is, a decomposable process. Later Doléans-Dade and Meyer [53] showed that the local square integrability hypothesis could be removed, which led to the traditional definition of a semimartingale (what we call a *classical semimartingale*). More formally, let us recall two definitions from Chaps. I and II and then define classical semimartingales.

Definition. An adapted, càdlàg process A is a **finite variation process** (FV) if almost surely the paths of A are of finite variation on each compact interval of $[0, \infty)$. We write $\int_0^\infty |dA_s|$ or $|A|_\infty$ for the random variable which is the total variation of the paths of A.

Definition. An adapted, càdlàg process X is **decomposable** if there exist processes N, A such that

$$X_t = X_0 + N_t + A_t$$

with $N_0 = A_0 = 0$, N a locally square integrable local martingale, and A an FV process.

Definition. An adapted, càdlàg process Y is a **classical semimartingale** if there exist processes N, B with $N_0 = B_0 = 0$ such that

$$Y_t = Y_0 + N_t + B_t$$

where N is a local martingale and B is an FV process.

Clearly an FV process is decomposable, and both FV processes and decomposable processes are semimartingales (Theorems 7 and 9 of Chap. II). *The goal of this chapter is to show that a process X is a classical semimartingale if and only if it is a semimartingale.* To do this we have to develop a small amount of "the general theory of processes." The key result is Theorem 25 which states that any local martingale M can be written

$$M = N + A$$

where N is a local martingale with bounded jumps (and hence locally square integrable), and A is an FV process. An immediate consequence is that a classical semimartingale is decomposable and hence a semimartingale by Theorem 9 of Chap. II. The theorem of Bichteler and Dellacherie (Theorem 43) gives the converse: a semimartingale is decomposable.

We summarize the results of this chapter, that are important to our treatment, in Theorems 1 and 2 which follow.

Theorem 1. *Let X be an adapted, càdlàg process. The following are equivalent:*

(i) X is a semimartingale;
(ii) X is decomposable;
(iii) given $\beta > 0$, there exist M, A with $M_0 = A_0 = 0$, M a local martingale with jumps bounded by β, A an FV process, such that $X_t = X_0 + M_t + A_t$;
(iv) X is a classical semimartingale.

Definition. The **predictable σ-algebra** $\mathcal{P}$ on $\mathbb{R}_+ \times \Omega$ is the smallest σ-algebra making all processes in $\mathbb{L}$ measurable. We also let $\mathcal{P}$ (resp. $\mathbf{b}\mathcal{P}$) denote the processes (resp. bounded processes) that are predictably measurable.

The next definition is not used in this chapter, except in the Exercises, but it is natural to include it with the definition of the predictable σ-algebra.

Definition. The **optional σ-algebra** $\mathcal{O}$ on $\mathbb{R}_+ \times \Omega$ is the smallest σ-algebra making all càdlàg, adapted processes measurable. We also let $\mathcal{O}$ (resp. $\mathbf{b}\mathcal{O}$) denote the processes (resp. bounded processes) that are optional.

Theorem 2. *Let X be a semimartingale. If X has a decomposition $X_t = X_0 + M_t + A_t$ with M a local martingale and A a predictably measurable FV process, $M_0 = A_0 = 0$, then such a decomposition is unique.*

In Theorem 1, clearly (ii) or (iii) each imply (iv), and (iii) implies (ii), and (ii) implies (i). That (iv) implies (iii) is an immediate consequence of the Fundamental Theorem of Local Martingales (Theorem 25). While Theorem 25 (and Theorems 3 and 22) is quite deep, nevertheless the heart of Theorem 1 is the implication (i) implies (ii), essentially the theorem of K. Bichteler and C. Dellacherie, which itself uses the Doob-Meyer decomposition theorem, Rao's Theorem on quasimartingales, and the Girsanov-Meyer Theorem on changes of probability laws. Theorem 2 is essentially Theorem 30.

We have tried to present this succession of deep theorems in the most direct and elementary manner possible. In the first edition we were of the opinion that Meyer's original use of natural processes was simpler than the now universally accepted use of predictability. However, since the first edition, R. Bass has published an elementary proof of the key Doob-Meyer decomposition theorem which makes such an approach truly obsolete. We are pleased to use Bass' approach here; see [11].

2 The Classification of Stopping Times

We begin by defining three types of stopping times. The important ones are *predictable* times and *totally inaccessible* times.

Definition. A stopping time T is **predictable** if there exists a sequence of stopping times $(T_n)_{n\geq 1}$ such that T_n is increasing, $T_n < T$ on $\{T > 0\}$, all n, and $\lim_{n\to\infty} T_n = T$ a.s. Such a sequence (T_n) is said to **announce** T.

If X is a continuous, adapted process with $X_0 = 0$, and $T = \inf\{t : |X_t| \geq c\}$, for some $c > 0$, then T is predictable. Indeed, the sequence $T_n = \inf\{t : |X_t| \geq c - \frac{1}{n}\} \wedge n$ is an announcing sequence. Fixed times are also predictable.

Definition. A stopping time T is **accessible** if there exists a sequence $(T_k)_{k\geq 1}$ of predictable times such that

$$P(\bigcup_{k=1}^{\infty}\{\omega : T_k(\omega) = T(\omega) < \infty\}) = P(T < \infty).$$

Such a sequence $(T_k)_{k\geq 1}$ is said to **envelop** T.

Any stopping time that takes on a countable number of values is clearly accessible. The first jump time of a Poisson process is not an accessible stopping time (indeed, any jump time of a Lévy process is not accessible).

Definition. A stopping time T is **totally inaccessible** if for every predictable stopping time S,

$$P\{\omega : T(\omega) = S(\omega) < \infty\} = 0.$$

Let T be a stopping time and $\Lambda \in \mathcal{F}_T$. We define

$$T_\Lambda(\omega) = \begin{cases} T(\omega), & \text{if } \omega \in \Lambda, \\ \infty, & \text{if } \omega \notin \Lambda. \end{cases}$$

It is simple to check that since $\Lambda \in \mathcal{F}_T$, T_Λ is a stopping time. Note further that $T = \min(T_\Lambda, T_{\Lambda^c}) = T_\Lambda \wedge T_{\Lambda^c}$.

A simple but useful concept is that of the graph of a stopping time.

Definition. Let T be a stopping time. The **graph of the stopping time** T is the subset of $\mathbb{R}_+ \times \Omega$ given by $\{(t,\omega) :\ 0 \leq t = T(\omega) < \infty\}$; the graph of T is denoted by $[T]$.

Theorem 3. *Let T be a stopping time. There exist disjoint events A, B such that $A \cup B = \{T < \infty\}$ a.s., T_A is accessible and T_B is totally inaccessible, and $T = T_A \wedge T_B$ a.s. Such a decomposition is a.s. unique.*

Proof. If T is totally inaccessible there is nothing to show. So without loss of generality we assume it is not. We proceed with an inductive construction: Let $R_1 = T$ and take

$$\alpha_1 = \sup\{P(S = R_1 < \infty) \ :\ S \text{ is predictable}\}.$$

Choose S_1 predictable such that $P(S_1 = R_1 < \infty) > \frac{\alpha_1}{2}$ and set $V_1 = S_1$. Define $R_2 = R_{1\{V_1 \neq R_1\}}$. For the inductive step let

$$\alpha_i = \sup\{P(S = R_i < \infty) \ :\ S \text{ is predictable}\}$$

If $\alpha_i = 0$ we stop the induction. Otherwise, choose S_i predictable such that $P(S_i = R_i < \infty) > \frac{\alpha_i}{2}$ and set $V_i = S_{i\{S_i \text{ does not equal any of } V_j, 1 \leq j \leq i-1\}}$. Let $R_{i+1} = R_{i\{V_i \neq R_i\}}$. The graphs of the V_i are disjoint by construction (the finite valued parts, of course). Moreover the sets $\{R_{i+1} \neq R_i,\ R_i < \infty\}$ are disjoint. Thus $\sum \frac{\alpha_i}{2} \leq \sum P(R_{i+1} \neq R_i,\ R_i < \infty) \leq 1$, which implies that $\alpha_i \to 0$. The R_i form a non-decreasing sequence, so set $U = \lim_{i\to\infty} R_i$. If U is not totally inaccessible, there exists a predictable time W such that $P(W = U) > 0$ and hence $P(W = U) > \alpha_i$ for some i. This contradicts how we chose S_i at step i of the induction. Therefore U is totally inaccessible, and $B = \{U = T < \infty\}$. □

A beautiful application of Theorem 3 is Meyer's Theorem on the jumps of Markov processes, a special case of which we give here without proof. (We use the notation established in Chap. I, Sect. 5 and write $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{0 \leq t \leq \infty}$.)

Theorem 4 (Meyer's Theorem). *Let X be a (strong) Markov Feller process for the probability P^μ, where the distribution of X_0 is given by μ, and with its natural completed filtration $\mathbb{F}^\mu$. Let T be a stopping time with $P^\mu(T > 0) = 1$. Let $\Lambda = \{\omega : X_T(\omega) \neq X_{T-}(\omega)$ and $T(\omega) < \infty\}$. Then $T = T_\Lambda \wedge T_{\Lambda^c}$, where T_Λ is totally inaccessible and T_{Λ^c} is predictable.*

A consequence of Meyer's Theorem is that the jump times of a Poisson process (or more generally a Lévy process) are all totally inaccessible.

We will need a small refinement of the concept of a stopping time σ-algebra. This is particularly important when the stopping time is predictable.

Definition. Let T be a stopping time. The σ-algebra $\mathcal{F}_{T-}$ is the smallest σ-algebra containing $\mathcal{F}_0$ and all sets of the form $A \cap \{t < T\}$, $t > 0$ and $A \in \mathcal{F}_t$.

Observe that $\mathcal{F}_{T-} \subset \mathcal{F}_T$, and also the stopping time T is $\mathcal{F}_{T-}$ measurable. We also have the following elementary but intuitive result. We leave the proof to the reader.

Theorem 5. *Let T be a predictable stopping time and let $(T_n)_{n\geq 1}$ be an announcing sequence for T. Then $\mathcal{F}_{T-} = \sigma\{\bigcup_{n\geq 1} \mathcal{F}_{T_n}\} = \bigvee_{n\geq 1} \mathcal{F}_{T_n}$.*

3 The Doob-Meyer Decompositions

We begin with a definition. Let $\mathbb{N}$ denote the natural numbers.

Definition. An adapted, càdlàg process X is a **potential** if it is a nonnegative supermartingale such that $\lim_{t\to\infty} E\{X_t\} = 0$. A process $(X_n)_{n\in\mathbb{N}}$ is also called a **potential** if it is a nonnegative supermartingale for $\mathbb{N}$ and $\lim_{n\to\infty} E\{X_n\} = 0$.

Theorem 6 (Doob Decomposition). *A potential $(X_n)_{n\in\mathbb{N}}$ has a decomposition $X_n = M_n - A_n$, where $A_{n+1} \geq A_n$ a.s., $A_0 = 0$, $A_n \in \mathcal{F}_{n-1}$, and $M_n = E\{A_\infty|\mathcal{F}_n\}$. Such a decomposition is unique.*

Proof. Let $M_0 = X_0$ and $A_0 = 0$. Define $M_1 = M_0 + (X_1 - E\{X_1|\mathcal{F}_0\})$, and $A_1 = X_0 - E\{X_1|\mathcal{F}_0\}$. Define M_n, A_n inductively as follows:

$$M_n = M_{n-1} + (X_n - E\{X_n|\mathcal{F}_{n-1}\}),$$
$$A_n = A_{n-1} + (X_{n-1} - E\{X_n|\mathcal{F}_{n-1}\}).$$

Note that $E\{A_n\} = E\{X_0\} - E\{X_n\} \leq E\{X_0\} < \infty$, as is easily checked by induction. It is then simple to check that M_n and A_n so defined satisfy the hypotheses.

Next suppose $X_n = N_n - B_n$ is another such representation. Then $M_n - N_n = A_n - B_n$ and in particular $M_1 - N_1 = A_1 - B_1 \in \mathcal{F}_0$. Thus $M_1 - N_1 = E\{M_1 - N_1|\mathcal{F}_0\} = M_0 - N_0 = X_0 - X_0 = 0$, hence $M_1 = N_1$. Continuing inductively shows $M_n = N_n$, all n. □

We wish to extend Theorem 6 to continuous time supermartingales. Note the unusual measurability condition that $A_n \in \mathcal{F}_{n-1}$ which of course is stronger than simply being adapted. The continuous time analog is that the process A be *natural* or what turns out to be equivalent, *predictably measurable*.

Throughout this paragraph we assume given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses. Before we state the first decomposition theorem in continuous time, we establish a simple lemma.

Lemma. Let $(Y_n)_{n\geq 1}$ be a sequence of random variables converging to 0 in L^2. Then $\sup_t E\{Y_n|\mathcal{F}_t\} \to 0$ in L^2.

Proof. Let $M_t^n = E\{Y_n|\mathcal{F}_t\}$ which is of course a martingale for each fixed n. Using Doob's maximal quadratic inequality, $E\{\sup_t (M_t^n)^2\} \leq 4E\{(M_\infty^n)^2\} = 4E\{Y_n^2\} \to 0$. □

Definition. We will say a càdlàg supermartingale Z with $Z_0 = 0$ is of **Class D** if the collection $\{Z_T : T$ a finite valued stopping time$\}$ is uniformly integrable.

The name "Class D" was given by P. A. Meyer in 1963. Presumably he expected it to come to be known as "Doob Class" at some point, but it has stayed Class D for 40 years now, so we see no point in changing it. (There are no Classes A, B, and C.) We now come to our first version of the Doob-Meyer decomposition theorem. It is this theorem that gives the fundamental basis for the theory of stochastic integration.

Theorem 7 (Doob-Meyer Decomposition: Case of Totally Inaccessible Jumps). *Let Z be a càdlàg supermartingale with $Z_0 = 0$ of Class D, and such that all jumps of Z occur at totally inaccessible stopping times. Then there exists a unique, increasing, continuous, adapted process A with $A_0 = 0$ such that $M_t = Z_t + A_t$ is a uniformly integrable martingale.*

We first give the proof of uniqueness which is easy. For existence, we will first establish three lemmas.

Proof of uniqueness. Let $Z = M - A$ and $Z = N - C$ be two decompositions of Z. Subtraction yields $M - N = A - C$, which implies that $M - N$ is a continuous martingale with paths of finite variation. We know however by the Corollary of Theorem 27 of Chap. II that $M - N$ is then a constant martingale which implies $M_t - N_t = M_0 - N_0 = 0 - 0 = 0$ for all t. Thus $M = N$ and $A = C$. □

Lemma 1. Let $\mathbb{F}$ be a discrete time filtration and let C be a non-decreasing process with $C_0 = 0$, and $C_k \in \mathcal{F}_{k-1}$. Suppose there exists a constant $N > 0$ such that $E\{C_\infty - C_k|\mathcal{F}_k\} \leq N$ a.s. for all k. Then $E\{C_\infty^2\} \leq 2N^2$.

Proof of Lemma 1. First observe $E\{C_\infty\} = E\{E\{C_\infty - C_0|\mathcal{F}_0\}\} \leq N$. Letting $c_k = C_{k+1} - C_k \geq 0$, we obtain by rearranging terms:

$$C_\infty^2 = 2\sum_{k\geq 0}(C_\infty - C_k)c_k - \sum_{k\geq 0} c_k^2.$$

Thus it follows that

$$E\{C_\infty^2\} \leq 2E\{\sum_{k\geq 0} E\{C_\infty - C_k|\mathcal{F}_k\}c_k\} \leq 2NE\{\sum_{k\geq 0} c_k\} \leq 2NE\{C_\infty\}$$
$$\leq 2N^2. \qquad \square$$

Choose and fix a constant $\nu \in \mathbb{Z}_+$ and let $D_n = \{k2^{-n} : 0 \leq k2^{-n} \leq \nu\}$.

Lemma 2. Let T be a totally inaccessible stopping time. For $\delta > 0$, let $R(\delta) = \sup_{t\leq\nu} P(t \leq T \leq t+\delta|\mathcal{F}_t)$. Then $R(\delta) \to 0$ in probability as $\delta \to 0$.

Proof of Lemma 2. Let $a > 0$ and $S_n(\delta) = \inf_t\{t \in D_n : P(t \leq T \leq t + \delta|\mathcal{F}_t) > a\} \wedge \nu$. First we assume $S_n(\delta)$ is less than T. Since $S_n(\delta)$ is countably valued, it is accessible, and since T is totally inaccessible, $P(S_n(\delta) = T) = 0$. Suppose $\Gamma \subset \{T < t\}$, and also $\Gamma \in \mathcal{F}_t$. Then

$$E\{E\{1_{\{t\leq T\leq t+\delta\}}|\mathcal{F}_t\}1_\Gamma\} = E\{1_{\{t\leq T\leq t+\delta\}}1_\Gamma\} = 0.$$

Suppose now $P(T < S_n(\delta)) > 0$. Then for some $t \in D_n$, $P(T < t,\ S_n(\delta) = t) > 0$. Let $\Gamma = \{T < t,\ S_n(\delta) = t\}$. Then from the definition of $S_n(\delta)$ we obtain

$$E\{1_{(t\leq T\leq t+\delta)}|\mathcal{F}_t\}1_\Gamma > a1_\Gamma, \tag{$*$}$$

a contradiction. Thus we conclude $S_n(\delta) < T$ a.s.

Next we define a stopping time S as follows. Let $\bar{S}(\delta) = \inf_n S_n(\delta)$ and $S = \sup_n \bar{S}(\frac{1}{n})$. Fix n, so that $\bar{S}(\frac{1}{n}) = \inf_k S_k(\frac{1}{n})$. Thus on $\{S = T\}$, $\bar{S}(\frac{1}{n}) < S$. Hence since $S = \sup_n \bar{S}(\frac{1}{n})$, we have S is accessible on $\{S = T\}$, which implies that T is accessible on $\{S = T\}$, which in turn implies $P(S = T) = 0$. Consequently $E\{1_{\{S=T\}}|\mathcal{F}_{S-}\} = 0$ a.s.

Suppose now that the result is not true. That is suppose there exists an $\varepsilon > 0$ such that $P(R(\delta) > a) > \varepsilon$, for all δ tending to 0. Let $\beta > 0$ and $\delta < \beta$. For n sufficiently large

$$P(E\{1_{\{S_n(\delta)\leq T\leq S_n(\delta)+\delta\}}|\mathcal{F}_{S_n(\delta)}\} > a) \geq \varepsilon.$$

We also have $P(T = \bar{S}(\delta)+\delta) = 0$, since if not $\bar{S}(\delta)+\delta-\frac{1}{m}$ could announce part of T which would contradict that T is totally inaccessible. Thus $P(\{S_n(\delta) \leq T \leq S_n(\delta)+\delta\}\ \Delta\ \{\bar{S}(\delta) \leq T \leq \bar{S}(\delta)+\delta\}) \to 0$ as $n \to \infty$ where the symbol Δ denotes the symmetric difference set operation. Since the symmetric difference tends to 0, if $P(S_n(\delta) \leq T \leq S_n(\delta) + \delta|\mathcal{F}_{S_n(\delta)} \geq a) \geq \varepsilon$ for any $\delta > 0$, then we must also have

$$P(E\{1_{\{\bar{S}(\delta)\leq T\leq\bar{S}(\delta)+\delta\}}|\mathcal{F}_{\bar{S}(\delta)}\} > a) \geq \varepsilon,$$

since otherwise were it to tend to 0 along with the symmetric difference tending to 0, then we would have $(*) \to 0$, a contradiction. Thus we have $P(\bar{S}(\delta) \leq T \leq \bar{S}(\delta) + \beta|\mathcal{F}_{\bar{S}(\delta)}) > a$ on a set Λ, with $P(\Lambda) \geq \varepsilon$. From the definition of S and working on the set Λ, this implies $P(E\{1_{\{S\leq T\leq S+\beta\}}|\mathcal{F}_{S-}\} \geq a) \geq \varepsilon$. Let $\beta \to 0$ to obtain $P(T = S|\mathcal{F}_{S-}) \geq a$, and we have a contradiction. $\square$

Lemma 3. Let the hypotheses of Theorem 7 hold and also assume $|Z| \leq N$ where $N > 0$ is a constant, and further that the paths of Z are constant after a stopping time ν. Let $W(\delta) = \sup_{t \leq u \leq t+\delta} E\{Z_t - Z_u|\mathcal{F}_t\}$. Then $W(\delta) \to 0$ in L^2 as $\delta \to 0$.

Proof of Lemma 3. Since $|Z| \leq N$ we have $|W(\delta)| \leq 2N$. Thus it suffices to show $W(\delta) \to 0$ in probability. Let ε, $a > 0$ and $b = a\sqrt{\varepsilon}$. Define $Z_t^+ = \sum_{s \leq t} \Delta Z_s 1_{\{\Delta Z_s > b\}}$, $Z_t^- = \sum_{s \leq t} \Delta Z_s 1_{\{-\Delta Z_s > b\}}$, $Z_t^b = Z_t - (Z_t^+ + Z_t^-)$, and $W^b(\delta) = \sup_{t \leq u \leq t+\delta} |E\{Z_t^b - Z_u^b|\mathcal{F}_t\}|$, with $W^+(\delta)$ and $W^-(\delta)$ defined analogously. Then

$$W^b(\delta) \leq \sup_t E\{\sup_{r \leq s \leq r+\delta} |Z_r^b - Z_s^b||\mathcal{F}_t\},$$

which by Doob's inequality implies

$$P(W^b(\delta) > a) \leq \frac{1}{a^2} E\{\sup_{r \leq s \leq r+\delta} |Z_r^b - Z_s^b|^2\}.$$

However since Z^b is càdlàg and $|\Delta Z^b| \leq b$, Fatou's Lemma implies

$$\limsup_{\delta \to 0} \frac{1}{a^2} E\{\sup_{r \leq s \leq r+\delta} |Z_r^b - Z_s^b|^2\} \leq \frac{b^2}{a^2} = \varepsilon.$$

Also, again using that Z is càdlàg, for each fixed ω the number of jumps larger than b is finite. Define inductively

$$T_1 = \inf\{t > 0 : \Delta Z_t > b\},\ T_{i+1} = \inf\{t > T_i : \Delta Z_t > b\}.$$

Since $|Z| \leq N$ be hypothesis, $|\Delta Z_{T_i}| \leq 2N$. Choose k such that $P(T_k \leq \nu) < \varepsilon$. Then

$$\begin{aligned} P(W^+(\delta) > a) &\leq P(T_k \leq \nu) + \sum_{i=1}^k P(\sup_t E\{\Delta Z_{T_i} 1_{\{t \leq T_i \leq t+\delta\}}|\mathcal{F}_t\} > \frac{a}{k}) \\ &\leq \varepsilon + \sum_{i=1}^k P(\sup_t P(t \leq T_i \leq t+\delta|\mathcal{F}_t) > \frac{a}{2KN}). \end{aligned}$$

By the previous lemma we know that $R(\delta) \to 0$ in L^2, so by taking δ small enough, we get the above expression less than 2ε. The reasoning for W^- is analogous. We achieve $W(\delta) \leq W^b(\delta) + W^+(\delta) + W^-(\delta)$, which gives $W(\delta) \to 0$ in L^2. □

We return to the proof of Theorem 7.

Proof of existence. First suppose the jumps of Z are bounded by a constant c. Let $T_N = \inf\{t > 0 : |Z_t| \geq N - c\} \wedge N$, and $Z_t^N = Z_{t \wedge T_N}$. Then $|Z_t^N| \leq$

$|Z_{t-}^N| + c \le N$, and Z^N is constant after T_N. Thus we are now reduced to the case considered by the Lemma 2.

Fix n and let $\mathcal{F}_k^n = \mathcal{F}_{\frac{k}{2^n}}$. Define

$$a_k^n = E\{Z_{\frac{k-1}{2^n}} - Z_{\frac{k}{2^n}} | \mathcal{F}_{k-1}^n\}$$

All the $a_k^n \geq 0$ since Z is a supermartingale, and also $a_k^n \in \mathcal{F}_{k-1}^n$. Let $A_k^N = \sum_{j=1}^k a_j^n$. Then $L_k = Z_{\frac{k}{2^n}} + A_k^n$ is an $\mathcal{F}_k^n$ discrete time martingale. Define

$$B_t^n = A_k^n \quad \text{if} \quad \frac{k-1}{2^n} < t \leq \frac{k}{2^n}. \tag{$*$}$$

We want to show that $E\{\sup_t |B_t^n - B_t^m|^2\} \to 0$ as $n, m \to \infty$. Suppose $m \geq n$. Since B^m and B^n are constant on intervals of the form $(k2^{-m}, (k+1)2^{-m}]$, the sup of the difference occurs at some $k2^{-m}$.

Fix t and let $u = \inf\{s : s \in D_n \text{ and } s \geq t\}$. Observe that $E\{A_\infty^m - A_k^m | \mathcal{F}_k^m\} = E\{Z_t - Z_\infty | \mathcal{F}_t\}$ is bounded by $2N$. Also

$$\begin{aligned} E\{A_\infty^n - B_t^n | \mathcal{F}_t\} &= E\{E\{A_\infty^n - B_u^n | \mathcal{F}_u\} | \mathcal{F}_t\} \\ &= E\{E\{Z_u - Z_\infty | \mathcal{F}_u\} | \mathcal{F}_t\} = E\{Z_u - Z_\infty | \mathcal{F}_t\} \end{aligned}$$

which is bounded by $2N$. This implies

$$E\{B_\infty^m - B_\infty^n - (B_t^m - B_t^n) | \mathcal{F}_t\} = E\{Z_t - Z_u | \mathcal{F}_t\},$$

with the right side being nonnegative and bounded by $W(2^{-n})$, and of course also by $E\{W(2^{-n}) | \mathcal{F}_t\}$. Using Lemmas 1 and 2 we get that $E\{\sup_t |B_t^n - B_t^m|^2\} \to 0$ as $n, m \to \infty$. Therefore the sequence B_t^n is a Cauchy sequence and we denote its limits by A_t.

Next we show that the process A has continuous paths. We do this by analyzing the jumps of the approximating processes B^n. Note that for $t = k2^{-n}$,

$$\Delta B_t^n = E\{Z_{\frac{k-1}{2^n}} - Z_{\frac{k}{2^n}} | \mathcal{F}_{k-1}^n\}$$

which are bounded by $W(2^{-n})$. But then $\sup_t |\Delta B_t^n| \to 0$ in L^2. Thus there exists a subsequence (n_j) such that $\sup_t |\Delta B_t^{n_j}| \to 0$ a.s., and hence the limit A is continuous.

It remains to show that $Z_t + A_t$ is a uniformly integrable martingale. Both are square integrable, so we only need concern ourselves with the martingale property. Let $s, t \in D_n$, $s < t$, and $\Lambda \in \mathcal{F}_s$. Then $E\{(Z_t + A_t)1_\Lambda\} = E\{(Z_s + A_s)1_\Lambda\}$. The result follows by taking limits of $Z_t + B_t^n$.

At the beginning of the proof we made the simplifying assumption that Z had bounded jumps. We now treat the general case. Choose N and let $T = \inf\{t > 0 : |Z_t| \geq N\} \wedge N$. Then Z^T (that is, Z stopped at the time T) has at most one jump bigger than $2N$. By localization and the uniqueness of the decomposition, it suffices to prove the result for this case. Thus without loss of generality we assume Z has at most one jump greater than or equal to $2N$ in absolute value, that it occurs at the time T, and that Z is constant after the time T. We let

$$Z_t^+ = \sum_{s\leq t} \Delta Z_s 1_{\{\Delta Z_s \geq 2N\}} = \Delta Z_T 1_{\{\Delta Z_T \geq 2N\}} 1_{\{t\geq T\}}$$

$$Z_t^- = \sum_{s\leq t} \Delta Z_s 1_{\{\Delta Z_s \leq -2N\}} = \Delta Z_T 1_{\{\Delta Z_T \leq -2N\}} 1_{\{t\geq T\}}$$

Since $-Z^+$ and Z^- both have decreasing paths, they are both supermartingales. Suppose we can show the theorem is true for $-Z^+$ and Z^-. Let $-Z^+ = M^+ - A^+$ and $Z^- = M^- - A^-$ be the two decompositions, with A^+ and A^- both continuous. Then $\hat{Z} = Z + (-Z^+ + A^+) - (Z^- + A^-) = Z + M^+ - M^-$ is a supermartingale with jumps bounded by $2N$. Let $\hat{Z} = \hat{M} - \hat{A}$ be the unique decomposition of $\hat{Z}$ which we now know exists. Then $Z = \hat{Z} + M^+ - M^-$, and therefore $Z = (\hat{M} + M^+ - M^-) - \hat{A}$ is the desired (unique) Doob-Meyer decomposition.

Thus it remains to show that $-Z^+$ and Z^- both have Doob-Meyer decompositions. First observe that $|\Delta Z_T| \leq |Z_{T-}| + |Z_T| \leq N + |Z_T| \in L^1$, and hence $E\{|\Delta Z_T|\} < \infty$. Choose $a, \varepsilon > 0$. Next choose $R > N$ so large that $E\{|\Delta Z_T| 1_{\{|\Delta Z_T|\geq R\}}\} \leq \varepsilon a$. The cases for when the jump is positive and when the jump is negative being exactly the same except for a minus sign, we only treat the case where the jump is negative. Let $Z_t^R = \Delta Z_T 1_{\{t\geq T\}} 1_{\{-\Delta Z_T > R\}}$, and $Z_t^d = Z_t^- - Z_t^R$. We define B^{n-}, B^{nR}, B^{nd} analogously to the way we defined B^n in equation $(*)$ above.

We first show B_t^{n-} converges uniformly in t in probability. We have:

$$P(\sup_t |B_t^{n-} - B_t^{m-}| > a) \tag{$**$}$$
$$\leq P(\sup_t |B_t^{nd} - B_t^{md}| > \frac{a}{3}) + P(\sup_t |B_t^{nR}| > \frac{a}{3}) + P(\sup_t |B_t^{mR}| > \frac{a}{3})$$

The second term on the right side of $(**)$ above is less than 3ε:

$$\begin{aligned}P(\sup_t |B_t^{nR}| > \frac{a}{3}) &= P(B_\infty^{nR} > \frac{a}{3}) \leq \frac{3}{a} E\{B_\infty^{nR}\} \\ &\leq \frac{3}{a} E\{|Z_\infty^R|\} \leq \frac{3}{a} E\{|\Delta Z_T| 1_{\{|\Delta Z_T|>R\}}\} \\ &\leq 3\varepsilon.\end{aligned}$$

The third term on the right side of $(**)$ is shown to be less than 3ε similarly. Since $|Z^R|$ is bounded by R, the first term on the right side of $(**)$ can be made arbitrarily small by taking m and n large enough, analogous to what we did at the beginning of the proof for B^n.

Thus as we did before in this proof, we can conclude that B_t^{n+} converges uniformly in t in probability as $n \to \infty$, and we denote its limit process by A_t^+.

We prove continuity of A exactly as before. Finally by taking a subsequence n_j such that B^{n_j+} converges almost surely, we get

$$E\{A_\infty^+\} = E\{\lim_{n_j\to\infty} B_\infty^{n_j+}\} \leq \liminf_{n_j\to\infty} E\{B_\infty^{n_j+}\} = -E\{Z_\infty^+\}$$

by Fatou's Lemma. From this it easily follows that $Z^+ + A^+$ is a uniformly integrable martingale, and the proof is complete. □

While Theorem 7 is sufficient for most applications, the restriction to supermartingales having jumps only at totally inaccessible stopping times can be insufficient for some needs. When we move to the general case we no longer have that A is continuous in general (however see Exercises 24 and 26 for supplemental hypotheses that assure that the increasing process A of the Doob-Meyer decomposition is in fact continuous). Without the continuity we lose the uniqueness of the decomposition, since there exist many martingales of finite variation (for example, the compensated Poisson process) that we can add to the martingale term and subtract from the finite variation term of a given decomposition, to obtain a second, new decomposition. Instead of the continuity of A we add the condition that A be predictably measurable.

Theorem 8 (Doob-Meyer Decomposition: General Case). *Let Z be a càdlàg supermartingale with $Z_0 = 0$ of Class D. Then there exists a unique, increasing, predictable process A with $A_0 = 0$ such that $M_t = Z_t + A_t$ is a uniformly integrable martingale.*

Before proving Theorem 8 let us introduce the concept of a *natural process.* We introduce two definitions and prove the important properties of natural processes in the next three theorems.

Definition. An FV process A with $A_0 = 0$ is of **integrable variation** if the expected total variation is finite: $E\{\int_0^\infty |dA_s|\} < \infty$. A shorthand notation for this is $E\{|A|_\infty\} < \infty$. An FV process A is of **locally integrable variation** if there exists a sequence of stopping times $(T^n)_{n\geq 1}$ increasing to ∞ a.s. such that $E\{\int_0^{T^n} |dA_s|\} = E\{|A|_{T_n}\} < \infty$, for each n.

Definition. Let A be an (adapted) FV process, $A_0 = 0$, of integrable variation. Then A is a **natural process** if

$$E\{[M, A]_\infty\} = 0$$

for all bounded martingales M.

Here is the key theorem about natural processes. This use of natural processes was Meyer's original insight that allowed him to prove Doob's conjecture, which became the Doob-Meyer decomposition theorem.

Theorem 9. *Let A be an FV process, $A_0 = 0$, and $E\{|A|_\infty\} < \infty$. Then A is natural if and only if*

$$E\left\{\int_0^\infty M_{s-}dA_s\right\} = E\{M_\infty A_\infty\}$$

for any bounded martingale M.

Proof. By integration by parts we have

$$\int_0^\infty M_{s-}dA_s = M_\infty A_\infty - M_0 A_0 - \int_0^\infty A_{s-}dM_s - [M,A]_\infty.$$

Then $M_0 A_0 = 0$ and letting $N_t = \int_0^t A_{s-}dM_s$, we know that N is a local martingale (Theorem 20 of Chap. II). However using integration by parts we see that $E\{N^*_\infty\} < \infty$, hence N is a true martingale (Theorem 51 of Chap. I). Therefore $E\left\{\int_0^\infty A_{s-}dM_s\right\} = E\{N_\infty\} - E\{N_0\} = 0$, since N is a martingale, so that the equality holds if and only if $E\{[M,A]_\infty\} = 0$. □

Theorem 10. *Let A be an FV process of integrable variation which is natural. If A is a martingale then A is identically zero.*

Proof. Let T be a finite stopping time and let H be any bounded, nonnegative martingale. Then $E\{\int_0^T H_{s-}dA_s\} = 0$, as is easily seen by approximating sums and the Dominated Convergence Theorem, since $\int_0^T |dA_s| \in L^1$ and $E\{A_T\} = 0$. Using the naturality of A, $E\{H_T A_T\} = E\{\int_0^T H_{s-}dA_s\} = 0$, and letting $H_t = E\{1_{\{A_T>0\}}|\mathcal{F}_t\}$ then shows that $P(A_T > 0) = 0$. Since $E\{A_T\} = 0$, we conclude $A_T \equiv 0$ a.s., hence $A \equiv 0$. □

Theorem 11. *Let A be an FV process of integrable variation with $A_0 = 0$. If A is predictable, then A is natural.*

The proof of this theorem is quite intuitive and natural provided we accept a result from Chap. IV. (We do not need this theorem to prove the theorem we are using from Chap. IV.)

Proof. Let M be a bounded martingale. First assume A is bounded. Then the stochastic integral $\int_0^\infty A_s dM_s$ exists, and it is a martingale by Theorem 29 in Chap. IV combined with, for example, Corollary 3 of Theorem 27 of Chap. II. Therefore $E\{\int_0^\infty A_s dM_s\} = E\{A_0 M_0\} = 0$, since $A_0 = 0$. However $E\{\int_0^\infty A_{s-}dM_s\} = E\{A_{0-}M_0\} = 0$ as well, since $A_{0-} = 0$. Further,

$$\begin{aligned}\int_0^\infty A_s dM_s - \int_0^\infty A_{s-}dM_s &= \int_0^\infty (A_s - A_{s-})dM_s \\ &= \int_0^\infty \Delta A_s dM_s \\ &= \sum_{0<s<\infty} \Delta A_s \Delta M_s \\ &= [A,M]_\infty,\end{aligned}$$

since A is a quadratic pure jump semimartingale. Therefore

$$E\{[A,M]_\infty\} = E\{\int_0^\infty A_s dM_s - \int_0^\infty A_{s-} dM_s\} = 0.$$

Since M was an arbitrary bounded martingale, A is natural by definition. Finally we remove the assumption that A is bounded. Let $A^n = n \wedge (A \vee (-n))$. Then A^n is bounded and still predictable, hence it is natural. For a bounded martingale M

$$E\{[M,A]_\infty\} = \lim_n E\{[M,A^n]_\infty\} = 0,$$

by the Dominated Convergence Theorem. Therefore A is natural. □

The next theorem is now obvious and should be called a Corollary at best. Because of its importance, however, we give it the status of a theorem.

Theorem 12. *Let M be a local martingale with paths of finite variation on compact time sets. If M is predictably measurable, then M is constant. That is, $M_t = M_0$ for all t, almost surely.*

Proof. This theorem is a combination of Theorems 10 and 11. □

Proof of Theorem 8 (Doob-Meyer: General Case). We begin with the uniqueness. Suppose $Z = M - A$ and $Z = N - C$ are two decompositions. By subtraction we have $M - N = A - C$ is a martingale with paths of finite variation which is predictable. By Theorem 12 we have that $M - N$ is constant and since $M_0 - N_0 = 0$, it is identically zero. This gives the uniqueness.

The existence of the decomposition is harder. We begin by defining stopping times $T_{n,j}$, where $T_{n,j}$ is the j-th time $|\Delta Z_t|$ is in the half-open interval bounded by 2^{-n} and $2^{-(n-1)}$, where $-\infty < n < \infty$. We decompose $T_{n,j}$ into its accessible and totally inaccessible parts. Since we can cover the accessible part with a sequence of predictable times, we can thus separate each $T_{n,j}$ into a totally inaccessible time and a sequence of predictable times, with disjoint (stopping time) graphs. Therefore, by renumbering, we can obtain a sequence of stopping times $(S_i)_{i \geq 1}$ with disjoint graphs, each one of which is either predictable or totally inaccessible, which exhaust the jumps of Z (in the sense that the jump times of Z are contained in the union of the graphs of the S_i) and are such that for each i there exists a constant b_i such that $b_i < |\Delta Z_{S_i}| \leq 2b_i$.

We define $Z_0(t) = Z_t$ and inductively define $Z_{i+1}(t) = Z_i(t)$ and $A_i(t) = 0$ if S_i is totally inaccessible; whereas $A_i(t) = -E\{\Delta Z_{S_i} | \mathcal{F}_{S_i -}\} 1_{\{S_i \leq t\}}$ and then $Z_{i+1}(t) = Z_i(t) + A_i(t)$ in the case where S_i is predictable (note that $\Delta Z_{S_i} \in L^1$).

We will show that: each A_i is increasing; Z_i is a supermartingale for each i; and $E\{\sum_{j=1}^i A_j(\infty)\} \leq C$, for each i, where C is a constant not depending on i. We will prove these three properties by induction. We begin by showing that each A_i is increasing. This is trivial if S_i is totally inaccessible. Let then

S_i be a predictable time. Let S_i^n be an announcing sequence for S_i. Since each Z_i is a supermartingale, and using the Martingale Convergence Theorem, we have $E\{\Delta Z_i(S_i)|\mathcal{F}_{S_i-}\} = \lim_n E\{\Delta Z_i(S_i)|\mathcal{F}_{S_i^n}\}$. Now fix n:

$$\begin{aligned} E\{\Delta Z_i(S_i)|\mathcal{F}_{S_i^n}\} &= \lim_k E\{Z_i(S_i) - Z_i(S_i^k)|\mathcal{F}_{S_i^n}\} \\ &= \lim_k E\{E\{Z_i(S_i) - Z_i(S_i^k)|\mathcal{F}_{S_i^k}\}|\mathcal{F}_{S_i^n}\} \\ &\leq 0, \end{aligned}$$

and thus A_i is increasing. To see that Z_{i+1} is a supermartingale, it will suffice to show that whenever U_1 and U_2 are stopping times with $U_1 \leq U_2$ then $E\{Z_{i+1}(U_1)\} \geq E\{Z_{i+1}(U_2)\}$. Again letting S_i^n be an announcing sequence for S_i we have

$$\begin{aligned} E\{Z_i(U_1)\} - E\{Z_i(U_2)\} &= E\{Z_i(U_1)\} - E\{Z_i((U_1 \vee S_i^n) \wedge U_2)\} \\ &\quad + E\{Z_i((U_1 \vee S_i^n) \wedge U_2)\} - E\{Z_i((U_1 \vee S_i) \wedge U_2)\} \\ &\quad + E\{Z_i((U_1 \vee S_i) \wedge U_2)\} - E\{Z_i(U_2)\} \end{aligned}$$

where each of the summands on the right side above are nonnegative. Then let $n \to \infty$ to get $E\{Z_i(U_1)\} - E\{Z_i(U_2)\} \geq 0$. Finally we want to show that $E\{\sum_{j=1}^i A_j(\infty)\} \leq C$. It is easy to check (see Exercise 1 for example) that the minimum of predictable stopping times is predictable, hence we can order the predictable times so that $S_1 < S_2 < \cdots$ on the set where they are all finite. Let Λ_i be the collection of all j such that $j \leq i$ and such that the S_j are all predictable. Then

$$\begin{aligned} -E\{\sum_{\Lambda_i} \Delta Z_{S_j}\} &= \lim_k \sum E\{Z_{S_j^k} - Z_{S_j}\} \\ &\leq \lim_k \sum_{\Lambda_i} (E\{Z_{S_j^k \vee S_{j-1}}\} + E\{Z_{S_{j-1}} - Z_{S_j^k \vee S_{j-1}}\} \\ &\quad + E\{Z_{S_i} - Z_\infty\} + E\{Z_0 - Z_{S_i^k}\}) \\ &= E\{Z_0 - Z_\infty\}, \end{aligned}$$

which is bounded by a constant not depending on i.

To complete the proof, we note that because the processes A_i are increasing in t and the expectation of their sum is bounded independently of i, we have that $\sum_{i=1}^h A_i(t)$ converges uniformly in t a.s., as $h \to \infty$, and we call the limit $A_\infty(t)$. Fatou's Lemma gives us that $A_\infty(t)$ is integrable. Each $A_i(t)$ is easily seen to be predictable, and hence A_∞ is the almost sure limit of predictable measurable processes, so it too is predictable. Next set $Z_\infty(t) = Z_t + A_\infty(t) = \lim_i Z_i(t)$. Since each Z_i is a supermartingale, coupled with the uniform convergence in t of the partial sums of the A_i processes, we get that $Z_\infty(t)$ is a supermartingale. Moreover since each Z_i is càdlàg, again using the uniform convergence of the partial sums we obtain that Z_∞ is also càdlàg. Since the partial sums are uniformly bounded in expectation, $Z_\infty(t)$ is of Class D. Finally by our construction of the stopping times S_i, $E\{\Delta Z_\infty(T)|\mathcal{F}_{T-}\} = 0$ for all predictable times T. We can then show as before that $W(\delta) \to 0$ in probability as $\delta \to 0$ for Z_∞, and we obtain $Z_\infty(t) = M_t - \hat{A}(t)$. The process $A_t = A_\infty(t) + \hat{A}(t)$ is then the desired increasing predictable process. □

The next theorem can also be considered a Doob-Meyer decomposition theorem. It exchanges the uniform integrability for a weakening of the conclusion that M be a martingale to that of M being a *local martingale.*

Theorem 13 (Doob-Meyer Decomposition: Case Without Class D). *Let Z be a càdlàg supermartingale. Then Z has a decomposition $Z = Z_0 + M - A$ where M is a local martingale and A is an increasing process which is predictable, and $M_0 = A_0 = 0$. Such a decomposition is unique. Moreover if* $\lim_{t\to\infty} E\{Z_t\} > -\infty$, *then* $E\{A_\infty\} < \infty$.

Proof. First consider uniqueness. Let $Z = Z_0 + M - A$ and $Z = Z_0 + N - C$ be two decompositions. Then $M - N = A - C$ by subtraction. Hence $A - C$ is a local martingale. Let M^{T^n} be a uniformly integrable martingale. Then

$$E\{Z_0 - Z_{t\wedge T^n}\} = E\{-M_t^{T^n}\} + E\{A_t^{T^n}\} = E\{A_t^{T^n}\}$$

and therefore $E\{A_t^{T^n}\} \leq E\{Z_0 - Z_t\}$, using Theorem 17 of Chap. I. Letting n tend to ∞ yields $E\{A_t\} \leq E\{Z_0 - Z_t\}$. Thus A is integrable on $[0, t_0]$, each t_0, as is C. Therefore $A - C$ is of locally integrable variation, predictable, and a local martingale. Since $A_0 - C_0 = 0$, by localization and Theorems 10 and 11, $A - C = 0$. That is, $A = C$, and hence $M = N$ as well and we have uniqueness.

Next we turn to existence. Let $T^m = \inf\{t : |Z_t| \geq m\} \wedge m$. Then T^m increase to ∞ a.s. and since they are bounded stopping times $Z_{T^m} \in L^1$ each m (Theorem 17 of Chap. I). Moreover the stopped process Z^{T^m} is dominated by the integrable random variable $\max(|Z_{T^m}|, m)$. Hence if $X = Z^{T^m}$ for fixed m, then $\mathcal{H} = \{X_T; T \text{ a stopping time}\}$ is uniformly integrable. Let us implicitly stop Z at the stopping time T^m and for $n > 0$ define $Y_t^n = Z_t - E\{Z_n|\mathcal{F}_t\}$, with $Y_t^n = Y_{t\wedge n}^n = 0$ when $t \geq n$. Then Y^n is a positive supermartingale of Class D and hence $Y_t^n = Y_0^n + M_t^n - A_t^n$. Letting $N_t^n = E\{Z_n|\mathcal{F}_t\}$, a martingale, we have on $[0, n]$ that

$$Z_t = Z_0 + M_t^n + N_t^n - A_t^n.$$

To conclude, therefore, it suffices to show that $A_t^m = A_t^n$ on $[0, n]$, for $m \geq n$. This is a consequence of the uniqueness already established. The uniqueness also allows us to remove the assumption that Z is stopped at the time T^m.

Finally, note that since $E\{A_t\} \leq E\{Z_0 - Z_t\}$, and since A is increasing, we have by the Monotone Convergence Theorem

$$E\{A_\infty\} = \lim_{t\to\infty} E\{A_t\} \le \lim_{t\to\infty} E\{Z_0 - Z_t\}$$

which is finite if $\lim_{t\to\infty} E\{Z_t\} > -\infty$. □

4 Quasimartingales

Let X be a càdlàg, adapted process defined on $[0,\infty]$.[1]

Definition. A finite tuple of points $\tau = (t_0, t_1, \dots, t_{n+1})$ such that $0 = t_0 < t_1 < \dots < t_{n+1} = \infty$ is a **partition** of $[0,\infty]$.

Definition. Suppose that τ is a partition of $[0,\infty]$ and that $X_{t_i} \in L^1$, each $t_i \in \tau$. Define

$$C(X,\tau) = \sum_{i=0}^{n} |E\{X_{t_i} - X_{t_{i+1}} | \mathcal{F}_{t_i}\}|.$$

The **variation of X along τ** is defined to be

$$\mathrm{Var}_\tau(X) = E\{C(X,\tau)\}.$$

The **variation of X** is defined to be

$$\mathrm{Var}(X) = \sup_\tau \mathrm{Var}_\tau(X),$$

where the supremum is taken over all such partitions.

Definition. An adapted, càdlàg process X is a **quasimartingale** on $[0,\infty]$ if $E\{|X_t|\} < \infty$, for each t, and if $\mathrm{Var}(X) < \infty$.

Before stating the next theorem we recall the following notational convention. If X is a random variable, then

$$X^+ = \max(X,0),$$
$$X^- = -\min(X,0).$$

Also recall that by convention if X is defined only on $[0,\infty)$, we set $X_\infty = 0$.

Theorem 14. *Let X be a process indexed by $[0,\infty)$. Then X is a quasimartingale if and only if X has a decomposition $X = Y - Z$ where Y and Z are each positive right continuous supermartingales.*

[1] It is convenient when discussing quasimartingales to include ∞ in the index set, thus making it homeomorphic to $[0,t]$ for $0 < t \le \infty$. If a process X is defined only on $[0,\infty)$ we extend it to $[0,\infty]$ by setting $X_\infty = 0$.

Proof. For given $s \geq 0$, let $\sum(s)$ denote the set of finite subdivisions of $[s, \infty]$. For each $\tau \in \sum(s)$, set

$$Y_s^\tau = E\{C(X,\tau)^+|\mathcal{F}_s\} \text{ and } Z_s^\tau = E\{C(X,\tau)^-|\mathcal{F}_s\}$$

where $C(X,\tau)^+$ denotes $\sum_{t_i\in\tau} E\{X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}\}^+$, and analogously for $C(X,\tau)^-$. Also let $\prec$ denote the ordering of set containment. Suppose $\sigma, \tau \in \sum(s)$ with $\sigma \prec \tau$. We claim $Y_s^\sigma \leq Y_s^\tau$ a.s. To see this let $\sigma = (t_0, \ldots, t_n)$. It suffices to consider what happens upon adding a subdivision point t before t_0, after t_n, or between t_i and t_{i+1}. The first two situations being clear, let us consider the third. Set

$$\begin{aligned} A &= E\{X_{t_i} - X_t|\mathcal{F}_{t_i}\};\ B = E\{X_t - X_{t_{i+1}}|\mathcal{F}_t\}; \\ C &= E\{X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}\}; \end{aligned}$$

then $C = A + E\{B|\mathcal{F}_{t_i}\}$, hence

$$\begin{aligned} C^+ &\leq A^+ + E\{B|\mathcal{F}_{t_i}\}^+ \\ &\leq A^+ + E\{B^+|\mathcal{F}_{t_i}\}, \end{aligned}$$

by Jensen's inequality. Therefore

$$E\{C^+|\mathcal{F}_s\} \leq E\{A^+|\mathcal{F}_s\} + E\{B^+|\mathcal{F}_s\}$$

and we conclude $Y_s^\sigma \leq Y_s^\tau$. Since $E\{Y_s^\tau\}$ is bounded by $\mathrm{Var}(X)$, taking limits in L^1 along the directed ordered set $\sum(s)$ we define

$$\hat{Y}_s = \lim_\tau Y_s^\tau,$$

and we can define $\hat{Z}_s$ analogously. Taking a subdivision with $t_0 = s$ and $t_{n+1} = \infty$, we see $Y_s^\tau - Z_s^\tau = E\{C^+ - C^-|\mathcal{F}_s\} = X_s$, and we deduce that $\hat{Y}_s - \hat{Z}_s = X_s$. Moreover if $s < t$ it is easily checked that $\hat{Y}_s \geq E\{\hat{Y}_t|\mathcal{F}_s\}$ and $\hat{Z}_s \geq E\{\hat{Z}_t|\mathcal{F}_s\}$. Define the right continuous processes $Y_t \equiv \hat{Y}_{t+}$, $Z_t \equiv \hat{Z}_{t+}$, with the right limits taken through the rationals. Then Y and Z are positive supermartingales and $Y_s - Z_s = X_s$.

For the converse, suppose $X = Y - Z$, where Y and Z are each positive supermartingales. Then for a partition τ of $[0,t]$

$$\begin{aligned} &E\{\sum_{t_i\in\tau} |E\{X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}\}|\} \\ &\leq E\{\sum_{t_i\in\tau} E\{Y_{t_i} - Y_{t_{i+1}}|\mathcal{F}_{t_i}\}\} + E\{\sum_{t_i\in\tau} E\{Z_{t_i} - Z_{t_{i+1}}|\mathcal{F}_{t_i}\}\} \\ &= \sum_{t_i\in\tau} (E\{Y_{t_i}\} - E\{Y_{t_{i+1}}\} + E\{Z_{t_i}\} - E\{Z_{t_{i+1}}\}) \\ &= E\{Y_0\} + E\{Z_0\} - (E\{Y_t\} + E\{Z_t\}). \end{aligned}$$

Thus X is a quasimartingale on $[0,t]$, each $t > 0$. □

Theorem 15 (Rao's Theorem). *A quasimartingale X has a unique decomposition $X = M + A$, where M is a local martingale and A is a predictable process with paths of locally integrable variation and $A_0 = 0$.*

Proof. This theorem is a combination of Theorems 13 and 14. □

5 Compensators

Let A be a process of locally integrable variation, hence *a fortiori* an FV process. A is then locally a quasimartingale, and hence by Rao's Theorem (Theorem 15), there exists a *unique* decomposition

$$A = M + \widetilde{A}$$

where $\widetilde{A}$ is a predictable FV process. In other words, there exists a unique, predictable FV process $\widetilde{A}$ such that $A - \widetilde{A}$ is a local martingale.

Definition. Let A be an FV process with $A_0 = 0$, with locally integrable total variation. The unique FV predictable process $\widetilde{A}$ such that $A - \widetilde{A}$ is a local martingale is called the **compensator** of A.

The most common examples are when A is an increasing process locally of integrable variation. When A is an increasing process it is of course a submartingale, and thus by the Doob-Meyer Theorem we know that *its compensator $\widetilde{A}$ is also increasing.* We also make the obvious observation that $E\{A_t\} = E\{\widetilde{A}_t\}$ for all t, $0 \le t \le \infty$.

Theorem 16. *Let A be an increasing process of integrable variation, and let $H \in \mathbb{L}$ be such that $E\{\int_0^t H_s dA_s\} < \infty$. Then,*

$$E\{\int_0^t H_s dA_s\} = E\{\int_0^t H_s d\widetilde{A}_s\}.$$

Proof. Since $A - \widetilde{A}$ is a martingale, so also is $\int H_s d(A_s - \widetilde{A}_s)$, and hence it has constant expectation equal to 0. □

In Chap. IV we develop stochastic integration for integrands which are predictable, and Theorem 16 extends with $H \in \mathbb{L}$ replaced with H predictable.

One of the simplest examples to consider as an illustration is that of the Poisson process $N = (N_t)_{t \ge 0}$ with parameter λ. Recall that $N_t - \lambda t$ is a martingale. Since the process $A_t = \lambda t$ is continuous and obviously adapted, it is predictable (natural). Therefore $\widetilde{N}_t = \lambda t$, $t \ge 0$. A natural extension of the Poisson process case is that of counting processes without explosions. We begin however with a counting process that has only one jump. Let η be the counting process $\eta_t = 1_{\{t \ge \tau\}}$, where τ is a nonnegative random variable. Let $\mathbb{F}$ be the minimal filtration making τ a stopping time.

Theorem 17. *Let $P(\tau = 0) = 0$ and $P(\tau > t) > 0$, each $t > 0$. Then the $\mathbb{F}$ compensator A of η, where $\eta_t = 1_{\{t \geq \tau\}}$, is given by*

$$A_t = \int_0^{\tau \wedge t} \frac{1}{1 - F(u-)} dF(u),$$

where F is the cumulative distribution function of τ. If τ has a diffuse distribution (that is, if F is continuous), then A is continuous and $A_t = -\ln(1 - F(\tau \wedge t))$.

Before proving Theorem 17 we formalize an elementary result as a lemma.

Lemma. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. In addition, suppose τ is a positive $\mathcal{F}$ measurable random variable and $\mathcal{F}_t^{00} = \sigma\{\tau \wedge t\}$, where $\mathcal{F}_t^0 = \mathcal{F}_t^{00} \vee \mathcal{N}$, $\mathcal{N}$ are the null sets of $\mathcal{F}$, and $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u^0$. Then $\mathbb{F}$ so constructed is the **smallest filtration making τ a stopping time**. Let $Y \in L^1(\mathcal{F})$. Then

$$E\{Y|\mathcal{F}_t\} = E\{Y|\tau\}1_{\{t \geq \tau\}} + 1_{\{\tau > t\}} \frac{E\{1_{\{\tau > t\}} Y\}}{P(\tau > t)}.$$

Proof. By the hypotheses on τ the σ-algebra $\mathcal{F}_t^0$ is equal to the Borel σ-algebra on $[0, t]$ together with the indivisible atom (t, ∞). Observe that $\mathcal{F}_t^0 = \sigma\{\tau \wedge u;\ u \leq t\}$, and the result follows easily. □

Proof of Theorem 17. Fix $t_0 > 0$ and let π_n be a sequence of partitions of $[0, t_0]$ with $\lim_{n \to \infty} \text{mesh}(\pi_n) = 0$. Define $A_t^n = \sum_{\pi^n} E\{\eta_{t_{i+1}} - \eta_{t_i} | \mathcal{F}_{t_i}\}$ for $0 \leq t \leq t_0$. Then

$$E\{\eta_{t_{i+1}} | \mathcal{F}_{t_i}\} = \eta_{t_{i+1}} 1_{\{\tau \leq t_i\}} + \frac{E\{1_{\{\tau > t_i\}} \eta_{t_{i+1}}\}}{P(\tau > t_i)} 1_{\{\tau > t_i\}}$$

by the lemma preceding this proof. The first term on the right above is $\eta_{t_{i+1}} 1_{\{t_i \geq \tau\}} = 1_{\{t_{i+1} \geq \tau\}} 1_{\{t_i \geq \tau\}} = 1_{\{t_i \geq \tau\}}$. Furthermore, $1_{\{\tau > t_i\}} \eta_{t_{i+1}} = 1_{\{\tau > t_i\}} 1_{\{t_{i+1} \geq \tau\}} = 1_{\{t_i < \tau \leq t_{i+1}\}}$. Hence,

$$E\{\eta_{t_{i+1}} | \mathcal{F}_{t_i}\} = 1_{\{t_i \geq \tau\}} + \frac{P(t_i < \tau \leq t_{i+1})}{P(\tau > t_i)} 1_{\{\tau > t_i\}}.$$

Therefore,

$$\begin{aligned} E\{\eta_{t_{i+1}} - \eta_{t_i} | \mathcal{F}_{t_i}\} &= 1_{\{t_i \geq \tau\}} + \frac{F(t_{i+1}) - F(t_i)}{1 - F(t_i)} 1_{\{\tau > t_i\}} - 1_{\{t_i \geq \tau\}} \\ &= \frac{1}{1 - F(t_i)} (F(t_{i+1}) - F(t_i)) 1_{\{\tau > t_i\}}. \end{aligned}$$

These simple calculations yield

$$A_t^n = \sum_{\pi^n} E\{\eta_{t_{i+1}} - \eta_{t_i} | \mathcal{F}_{t_i}\} = \sum_{\pi^n} \frac{F(t_{i+1}) - F(t_i)}{1 - F(t_i)} 1_{\{\tau > t_i\}}$$

which are Riemann sums converging to $\int_0^{t\wedge\tau}(1 - F(u-))^{-1}dF(u)$. The remaining claims of the theorem are obvious. □

Corollary 1. If F in Theorem 17 is absolutely continuous, then so also is A.

Corollary 2. If τ in Theorem 17 has an exponential distribution, then $A_t = \tau \wedge t$.

Proof. Since in this case F is continuous and a distribution function, we can represent it as $F(x) = 1 - e^{-\phi(x)}$. We have that $A_t = -\ln(1 - F(\tau \wedge t))$ and equivalently we can write $A_t = \phi(\tau \wedge t)$. In the case of this corollary, the function $\phi(x) = x$, and the result follows. □

Remark. Theorem 17 gives an explicit formula for the compensator of a process $\eta_t = 1_{\{t \geq \tau\}}$ when one uses the minimal filtration making τ a stopping time. A perhaps more interesting question is the following: suppose that one has a non-trivial filtration $\mathbb{F}$ satisfying the usual hypotheses, and a strictly positive random variable L that is not a stopping time. Can one expand the filtration to make a larger filtration $\mathbb{G}$ that renders L a $\mathbb{G}$ stopping time, and then calculate the $\mathbb{G}$ compensator of $\mu_t = 1_{\{t \geq L\}}$? This has a positive answer with the right hypotheses. To see how to do this, let $X_t = 1_{[0,L)}(t)$, which is not an $\mathbb{F}$ adapted process. Let Z denote the optional projection of X onto $\mathbb{F}$. We prove in Chap. VI that Z is a supermartingale. Let $Z = M - A$ be its Doob-Meyer decomposition with A predictable. (Note that $M_t = E\{A_\infty | \mathcal{F}_t\}$.) Using the techniques presented in Chap. VI, one can show that if L is "the end of an $\mathbb{F}$ predictable set," then the $\mathbb{G}$ compensator of $\mu_t = 1_{\{t \geq L\}}$ is A_t, which is an $\mathbb{F}$ predictable process! By the end of a predictable set, we mean that there exists an $\mathbb{F}$ predictable set $\Lambda \subset [0, \infty] \times \Omega$ such that $L(\omega) = \sup\{t \leq \infty : (t, \omega) \in \Lambda\}$. In Chap. VI we discuss times that are the ends of optional sets, known as "honest times," and this is of course less restrictive than being the end of a predictable set, since any predictable set is also optional.[2] See [115] for this result described here and related ones.

Theorem 17 treats the case of a point process with one jump. A special case of interest is counting processes. We have the following.

Theorem 18. *Let N be a counting process without explosions, adapted to a filtration $\mathbb{G}$ satisfying the usual hypotheses. Then the compensator of N, call it A, always exists.*

Proof. Since N has non-decreasing paths, to ensure the existence of a compensator we need only to show that N is locally of integrable variation. Let T_n be the time of the n-th jump of N. Since N has no explosions, the times

[2] See Exercise 4.

T_n increase to ∞ a.s. Moreover $|N_{t\wedge T_n}| = N_{t\wedge T_n} \leq n$ and thus N is locally of bounded variation, hence certainly locally of integrable variation. $\square$

Let $N_t = \sum_{i\geq 1} 1_{\{t\geq T_i\}}$ be a counting process without explosions and let $S_i = T_i - T_{i-1}$ be its interarrival times, where we take $T_0 = 0$. Let $\mathcal{F}_t^0 = \sigma\{N_s;\ s \leq t\}$, and let $\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N}$, the filtration completed by the $\mathcal{F}$-null sets, as usual. (Note that this filtration is already right continuous.) Let us define the cumulative distribution functions for the interarrival times slightly informally as follows (note that we are using the unusual format of $P(S > x)$ rather than the customary $P(S \leq x)$):

$$F_1(t) = P(S_1 > t)$$
$$F_i(s_1, \ldots, s_{i-1}; t) = P(S_i > t | S_1 = s_1, \ldots, S_{i-1} = s_{i-1})$$

where the s_j are in $[0, \infty]$, $j \geq 1$. (If one of the s_j takes the value $+\infty$, then the corresponding cumulative distribution function is concentrated at $+\infty$.) Define

$$\phi_i(s_1, \ldots, s_{i-1};\ t) = \int_0^t \frac{-1}{F_i(s_1, \ldots, s_{i-1};\ s-)} dF_i(s_1, \ldots, s_{i-1};\ s). \qquad (*)$$

We now have the tools to describe the compensator of N for its minimal completed filtration. We omit the proof, since it is simply a more notationally cumbersome version of the proof of Theorem 17.

Theorem 19. *Let N be a counting process without explosions and let $\mathbb{F}$ be its minimal completed filtration. Then the compensator A of N is given by*

$$A_t = \phi_1(S_1) + \phi_2(S_1; S_2) + \cdots + \phi_{i-1}(S_1, \ldots, S_{i-2};\ S_{i-1}) + \phi_i(S_1, \ldots, S_{i-1};\ t - T_i)$$

on the event $\{T_i < t \leq T_{i+1}\}$.

Corollary. Let N be a counting process without explosions and independent interarrival times. Then the functions ϕ_i defined in equation $(*)$ have the simplified form

$$\phi_i(s_1, \ldots, s_{i-1};\ t) = \int_0^t \frac{-1}{F_i(u-)} dF_i(u),$$

and the compensator A is given by

$$A_t = \sum_{i\geq 1} [\sum_{j=1}^{i-1} \phi_j(S_j) + \phi_i(t - T_i)] 1_{\{T_i \leq t < T_{i+1}\}}$$

Example (hazard rates and censored data). In applications, an intuitive interpretation often used is that of hazard rates. Let T be a positive random variable with a continuous density for its distribution function F, given by f. The hazard rate λ is defined, when it exists, to be

$$\lambda(t) = \lim_{h \to 0} \frac{1}{h} P(t \leq T < t + h | T \geq t).$$

The intuition is that this is the probability the event will happen in the next infinitesimal unit of time, given that it has not yet happened. Another way of viewing this is that $P(t \leq T < t + h | T \geq t) = \lambda h + o(h)$. By Theorem 17 we have that the compensator of $N_t = 1_{\{t \geq T\}}$ is $A_t = \int_0^{t \wedge T} \lambda(s) ds$. Or in language closer to stochastic integration,

$$N_t - \int_0^{t \wedge T} \lambda(s) ds \text{ is a local martingale.}$$

In medical trials and other applications, one is faced with situations where one is studying random arrivals, but some of the data cannot be seen (for example when patients disappear unexpectedly from clinical trials). This is known as arrivals with *censored data.* A simple example is as follows. Let T be a random time with a density and let U be another time, with an arbitrary distribution. T is considered to be the arrival time and U is the censoring time. Let $X = T \wedge U$, $N_t = 1_{\{t \geq X\}} 1_{\{U \geq T\}}$, and $N_t^U = 1_{\{t \geq X\}} 1_{\{T > U\}}$ with $\mathcal{F}_t = \sigma\{N_u, N_u^U;\ u \leq t\}$. Let λ be the hazard rate function for T. λ is known as the *net hazard rate.* We define $\lambda^{\#}$, known as the *crude hazard rate*, by

$$\lambda^{\#} = \lim_{h \to 0} \frac{1}{h} P(t \leq T < t + h | T \geq t, U \geq t)$$

when the limit exists. We then have that the compensator of N is given by $A_t = \int_0^t 1_{\{X \geq u\}} \lambda^{\#}(u) du = \int_0^{t \wedge X} \lambda^{\#}(u) du = \int_0^{t \wedge T \wedge U} \lambda^{\#}(u) du$, or equivalently

$$N_t - \int_0^t 1_{\{X \geq u\}} \lambda^{\#}(u) du \text{ is a local martingale.}$$

If we impose the condition that $\lambda = \lambda^{\#}$, which can be intuitively interpreted as

$$P(s \leq T < s + ds | T \geq s) = P(s \leq T < s + ds | T \geq s, U \geq s),$$

then we have the satisfying result that

$$N_t - \int_0^t 1_{\{X \geq u\}} \lambda(u) du \text{ is a local martingale.}$$

As we saw in Chap. II, processes of fundamental importance to the theory of stochastic integration are the quadratic variation processes $[X, X] = ([X, X]_t)_{t \geq 0}$, where X is a semimartingale.

Definition. Let X be a semimartingale such that its quadratic variation process $[X, X]$ is locally integrable. Then the **conditional quadratic variation** of X, denoted $\langle X, X \rangle = (\langle X, X \rangle_t)_{t \geq 0}$, *exists and it is defined to be the compensator of* $[X, X]$. That is $\langle X, X \rangle = \widetilde{[X, X]}$.

If X is a continuous semimartingale then $[X,X]$ is also continuous and hence already predictable; *thus* $[X,X] = \langle X,X\rangle$ *when* X *is continuous.* In particular for a standard Brownian motion B, $[B,B]_t = \langle B,B\rangle_t = t$, all $t \geq 0$. The conditional quadratic variation is also known in the literature by its notation. It is sometimes called the *sharp bracket*, the *angle bracket*, or the *oblique bracket*. It has properties analogous to that of the quadratic variation processes. For example, if X and Y are two semimartingales such that $\langle X,X\rangle$, $\langle Y,Y\rangle$, and $\langle X+Y,X+Y\rangle$ all exist, then $\langle X,Y\rangle$ exists and can be defined by **polarization**

$$\langle X,Y\rangle = \frac{1}{2}(\langle X+Y,X+Y\rangle - \langle X,X\rangle - \langle Y,Y\rangle).$$

However $\langle X,Y\rangle$ can be defined independently as the compensator of $[X,Y]$ provided of course that $[X,Y]$ is locally of integrable variation. In other words, there exist stopping times $(T^n)_{n\geq 1}$ increasing to ∞ a.s. such that $E\{\int_0^{T^n} |d[X,Y]_s|\} < \infty$ for each n. Also, $\langle X,X\rangle$ is a non-decreasing process by the preceding discussion, since $[X,X]$ is non-decreasing. The conditional quadratic variation is inconvenient since unlike the quadratic variation it doesn't always exist. Moreover while $[X,X]$, $[X,Y]$, and $[Y,Y]$ all remain invariant with a change to an equivalent probability measure, *the sharp brackets in general change with a change to an equivalent probability measure and may even no longer exist.* Although the angle bracket is ubiquitous in the literature it is sometimes unnecessary as one can often use the quadratic variation instead, and indeed *whenever possible we use the quadratic variation rather than the conditional quadratic variation* $\langle X,X\rangle$ *of a semimartingale* X *in this book.* Nevertheless the process $\langle X,X\rangle$ occurs naturally in extensions of Girsanov's theorem for example, and it has become indispensable in many areas of advanced analysis in the theory of stochastic processes.

We end this section with several useful observations that we formalize as theorems and a corollary. Note that the second theorem below is a refinement of the first (and see also Exercise 25).

Theorem 20. *Let* A *be an increasing process of locally integrable variation, all of whose jumps occur at totally inaccessible stopping times. Then its compensator* $\widetilde{A}$ *is continuous.*

Theorem 21. *Let* A *be an increasing process of locally integrable variation, and let* T *be a jump time of* A *which is totally inaccessible. Then its compensator* $\widetilde{A}$ *is continuous at* T.

Proof. Both theorems are simple consequences of Theorem 7. □

Corollary. Let A be an increasing predictable process of locally integrable variation and let T be a stopping time. If $P(A_T \neq A_{T-}) > 0$ then T is not totally inaccessible.

Proof. Suppose T were totally inaccessible. Let $\widetilde{A}$ be the compensator of A. Then $\widetilde{A}$ is continuous at T. But since A is already predictable, $A = \widetilde{A}$, and we have a contradiction by Theorem 21. □

Theorem 22. *Let T be a totally inaccessible stopping time. There exists a martingale M with paths of finite variation and with exactly one jump, of size one, occurring at time T (that is, $M_T \neq M_{T-}$ on $\{T < \infty\}$).*

Proof. Define

$$U_t = 1_{\{t \geq T\}}.$$

Then U is an increasing, bounded process of integrable variation, and we let $A = \widetilde{U}$ be the compensator of U. A is continuous by Theorem 20, and $M = U - A$ is the required martingale. □

6 The Fundamental Theorem of Local Martingales

We begin with two preliminary results.

Theorem 23 (Le Jan's Theorem). *Let T be a stopping time and let H be an integrable random variable such that $E\{H|\mathcal{F}_{T-}\} = 0$ on $\{T < \infty\}$. Then the right continuous martingale $H_t = E\{H|\mathcal{F}_t\}$ is zero on $[0, T) = \{(t, \omega) : 0 \leq t < T(\omega)\}$.*

Proof. Since the martingale $(H_t)_{t \geq 0}$ is right continuous it suffices to show that $H_t 1_{\{t<T\}} = 0$ almost surely for all t. Let $\Lambda \in \mathcal{F}_t$. Then $\Lambda \cap \{t < T\}$ belongs both to $\mathcal{F}_{T-}$ and also to $\mathcal{F}_t$. Hence

$$0 = E\{H 1_{\Lambda \cap \{t<T\}}\} = E\{H_t 1_{\Lambda \cap \{t<T\}}\}.$$

Since $\Lambda \in \mathcal{F}_t$ is arbitrary, this implies that $H_t 1_{\{t<T\}} = 0$ a.s. If H is $\mathcal{F}_T$ measurable then $H 1_{\{T \leq t\}} \in \mathcal{F}_t$, and $E\{H|\mathcal{F}_t\} = H$ on $\{T \leq t\}$ and the theorem is proved. □

Theorem 24. *Let T be a predictable stopping time and let A be increasing, predictable, and locally of integrable variation. Then A_T and ΔA_T are each $\mathcal{F}_{T-}$ measurable.*

Proof. Let S_n be a sequence of stopping times announcing T. Then $A_{S_n} \in \mathcal{F}_{S_n} \subset \mathcal{F}_{T-}$ for each n. Since $A_{T-} = \lim_{n \to \infty} A_{S_n}$ we have that $A_{T-} \in \mathcal{F}_{T-}$. To show that $A_T \in \mathcal{F}_{T-}$ it suffices to show $E\{A_T H\} = 0$ for every bounded random variable H such that $E\{H|\mathcal{F}_{T-}\} = 0$. Let H be such a random variable and let $H_t = E\{H|\mathcal{F}_t\}$, the right continuous martingale with terminal value H. Then $H_t(\omega) = 0$ for $0 \leq t < T(\omega)$ by Le Jan's Theorem (Theorem 23). This implies that the left continuous version of H, H_{t-}, equals 0 on $[0, T]$. Then since the hypotheses imply that A is natural, we have

$$E\{HA_T\} = E\{\int_0^\infty H_{s-} 1_{\{s \leq T\}} dA_s\} + E\{[H, A]_T\} = 0 + 0 = 0. \qquad \square$$

Theorem 25 (Fundamental Theorem of Local Martingales). *Let M be a local martingale and let $\beta > 0$. Then there exist local martingales N, D such that D is an FV process, the jumps of N are bounded by 2β, and $M = N + D$.*

Proof. We first set

$$C_t = \sum_{0<s\le t} |\Delta M_s| 1_{\{|\Delta M_s|\ge\beta\}},$$

and we want to show C is locally integrable. By stopping *we can assume without loss that M is a uniformly integrable martingale and that $M_0 = 0$.* For each ω, $C_t(\omega)$ is a finite sum. Define

$$R_n = \inf\{t : C_t \ge n \text{ or } |M_t| \ge n\}.$$

Then $|\Delta M_{R_n}| \le |M_{R_n}| + |M_{R_n-}| \le |M_{R_n}| + n$. Moreover $C_{R_n} \le C_{R_n-} + |\Delta M_{R_n}| \le |M_{R_n}| + 2n$, which is in L^1. Since R_n increases to ∞, we have C is locally integrable. By stopping we further assume that $E\{C_\infty\} < \infty$.

Next we define

$$A_t = \sum_{0<s\le t} \Delta M_s 1_{\{|\Delta M_s|\ge\beta\}},$$

and since C is assumed integrable we deduce that A is a quasimartingale. By Theorems 13 and 14 we know that A decomposes as follows:

$$A = L + B^1_t - B^2_t,$$

where L is a local martingale and B^1, B^2 are non-decreasing, nonnegative, predictable processes, such that $E\{B^1_\infty\} < \infty$, $E\{B^2_\infty\} < \infty$. Let

$$\widetilde{A}_t = B^1_t - B^2_t.$$

Then $|A_t - \widetilde{A}_t| \le C_\infty + B^1_\infty + B^2_\infty \in L^1$, and hence the martingales $A - \widetilde{A}$ and

$$N = M - (A - \widetilde{A})$$

are uniformly integrable. As usual, we write $\widetilde{A}$ for the compensator of A. Then $D = A - \widetilde{A}$.

It remains only to show that the jumps of N are bounded by 2β. To this end, let T be a stopping time. If $\Delta\widetilde{A}_T = 0$ a.s., then

$$\begin{aligned} \Delta N_T &= \Delta M_T - \Delta A_T \\ &= \Delta M_t 1_{\{|\Delta M_T|\le\beta\}}. \end{aligned}$$

This implies $|\Delta N_T| \le \beta$, and we are done by Theorem 7 of Chap. I.

We therefore assume if $\Lambda = \{|\widetilde{\Delta} A_T| > 0\}$, then $P(\Lambda) > 0$. Let $R = T_\Lambda$. By Theorem 3 we can decompose R as $R = R_A \wedge R_B$, where R_A is accessible and R_B is totally inaccessible. Then $P(R_B < \infty) = 0$ by Theorem 7, whence

$R = R_A$. Let $(T_k)_{k\geq 1}$ be a sequence of predictable times enveloping R. It will suffice to show $|\Delta N_{T_k}| \leq 2\beta$, for each T_k. Thus without loss of generality we can take $R = T$ to be predictable. By convention, we set $\Delta N_T = 0$ on $\{T = \infty\}$. Since we know that $\Delta \widetilde{A}_T \in \mathcal{F}_{T-}$ by Theorem 24, we have

$$\begin{aligned}\Delta N_T &= \Delta N_T - E\{\Delta N_T|\mathcal{F}_{T-}\} \\ &= \Delta(M - A)_T + \Delta \widetilde{A}_T - E\{\Delta(M - A)_T|\mathcal{F}_{T-}\} - E\{\Delta \widetilde{A}_T|\mathcal{F}_{T-}\} \\ &= \Delta(M - A)_T - E\{\Delta(M - A)_T|\mathcal{F}_{T-}\}.\end{aligned}$$

Since $|\Delta(M - A)_T| \leq \beta$, the result follows. □

If the jumps of a local martingale M with $M_0 = 0$ are bounded by a constant β, then M itself is locally bounded. Let $T_n = \inf\{t : |M_t| \geq n\}$. Then $|M_{t\wedge T_n}| \leq n + \beta$. Therefore M is *a fortiori* locally square integrable. We thus have a corollary.

Corollary. A local martingale is decomposable.

Of course if all local martingales were locally square integrable, they would then be trivially decomposable. The next example shows that there are martingales that are not locally square integrable (a more complex example is published in Doléans-Dade [50]).

Example. Let $(\Omega, \mathcal{F}, P)$ be complete probability space and let X be a random variable such that $X \in L^1$, but $X \notin L^2$. Define the filtration

$$\mathcal{F}_t^0 = \begin{cases} \{\emptyset, \Omega\}, & 0 \leq t < 1, \\ \mathcal{F}, & t \geq 1, \end{cases}$$

where $\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N}$, with $\mathcal{N}$ all the P-null sets of $\mathcal{F}$. Let $M_t = E\{X|\mathcal{F}_t\}$, the right continuous version. Then M is not a locally square integrable martingale.

The next example shows another way in which local martingales differ from martingales. A local martingale need not remain a local martingale under a shrinkage of the filtration. They do, however, remain semimartingales and thus they still have an interpretation as a differential.

Example. Let Y be a symmetric random variable with a continuous distribution and such that $E\{|Y|\} = \infty$. Let $X_t = Y 1_{\{t\geq 1\}}$, and define

$$\mathcal{G}_t^0 = \begin{cases} \sigma\{|Y|\}, & 0 \leq t < 1, \\ \sigma\{Y\}, & t \geq 1, \end{cases}$$

where $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ is the completed filtration. Define stopping times T^n by

$$T^n = \begin{cases} 0, & \text{if } |Y| \geq n, \\ \infty, & \text{otherwise.} \end{cases}$$

Then T^n reduce X and show that it is a local martingale. However X is not a local martingale relative to its completed minimal filtration. Note that X is still a semimartingale however.

The full power of Theorem 25 will become apparent in Sect. 7.

7 Classical Semimartingales

We have seen that a decomposable process is a semimartingale (Theorem 9 of Chap. II). We can now show that a classical semimartingale is indeed a semimartingale as well.

Theorem 26. *A classical semimartingale is a semimartingale.*

Proof. Let X be a classical semimartingale. Then $X_t = M_t + A_t$ where M is a local martingale and A is an FV process. The process A is a semimartingale by Theorem 7 of Chap. II, and M is decomposable by the corollary of Theorem 25, hence also a semimartingale (Theorem 9 of Chap. II). Since semimartingales form a vector space (Theorem 1 of Chap. II) we conclude X is a semimartingale. □

Corollary. A càdlàg local martingale is a semimartingale.

Proof. A local martingale is a classical semimartingale. □

Theorem 27. *A càdlàg quasimartingale is a semimartingale.*

Proof. By Theorem 15 a quasimartingale is a classical semimartingale. Hence it is a semimartingale by Theorem 26. □

Theorem 28. *A càdlàg supermartingale is a semimartingale.*

Proof. Since a local semimartingale is a semimartingale (corollary to Theorem 6 of Chap. II), it suffices to show that for a supermartingale X, the stopped process X^t is a semimartingale. However for a partition τ of $[0,t]$,

$$\begin{aligned} E\{\sum_{t_i\in\tau} |E\{X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}\}|\} &= E\{\sum_{t_i\in\tau} E\{X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}\}\} \\ &= \sum_{t_i\in\tau}(E\{X_{t_i}\} - E\{X_{t_{i+1}}\}) \\ &= E\{X_0\} - E\{X_t\}. \end{aligned}$$

Therefore X^t is a quasimartingale, hence a semimartingale by Theorem 27. □

Corollary. A submartingale is a semimartingale.

We saw in Chap. II that if X is a locally square integrable local martingale and $H \in \mathbb{L}$, then the stochastic integral $H \cdot X$ is also a locally square integrable local martingale (Theorem 20 of Chap. II). Because of the corollary of Theorem 25 we can now improve this result.

Theorem 29. *Let M be a local martingale and let $H \in \mathbb{L}$. Then the stochastic integral $H \cdot M$ is again a local martingale.*

Proof. A local martingale is a semimartingale by the corollary of Theorem 25 and Theorem 9 of Chap. II; thus $H \cdot M$ is defined. By the Fundamental Theorem of Local Martingales (Theorem 25) for $\beta > 0$ we can write $M = N + A$ where N, A are local martingales, the jumps of N are bounded by β, and A has paths of finite variation on compacts. Since N has bounded jumps, by stopping we can assume N is bounded. Define T by

$$T = \inf\{t > 0 : \int_0^t |dA_s| > m\}.$$

Then $E\{\int_0^{t \wedge T} |dA_s|\} \leq m + \beta + E\{|\Delta M_T|\} < \infty$, and thus by stopping A can be assumed to be of integrable variation. Also by replacing H by $H^S 1_{\{S>0\}}$ for an appropriate stopping time S we can assume without loss of generality that H is bounded, since H is left continuous. We also assume without loss that $M_0 - N_0 = A_0 = 0$. We know $H \cdot N$ is a local martingale by Theorem 20 of Chap. II, thus we need show only that $H \cdot A$ is a local martingale.

Let σ_n be a sequence of random partitions of $[0, t]$ tending to the identity. Then $\sum H_{T_i^n}(A^{T_i^{n+1}} - A^{T_i^n})$ tends to $(H \cdot A)_t$ in *ucp*, where σ_n is the sequence $0 = T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots$. Let (n_k) be a subsequence such that the sums converge uniformly a.s. on $[0, t]$. Then

$$\begin{aligned} E\left\{\int_0^t H_u dA_u | \mathcal{F}_s\right\} &= E\left\{\lim_{n_k} \sum_i H_{T_i^{n_k}}(A_t^{T_{i+1}^{n_k}} - A_t^{T_i^{n_k}}) | \mathcal{F}_s\right\} \\ &= \lim_{n_k} E\left\{\sum_i H_{T_i^{n_k}}(A_t^{T_{i+1}^{n_k}} - A_t^{T_i^{n_k}}) | \mathcal{F}_s\right\} \\ &= \lim_{n_k} \sum_i H_{T_i^{n_k}}(A_s^{T_{i+1}^{n_k}} - A_s^{T_i^{n_k}}) \end{aligned}$$

by Lebesgue's Dominated Convergence Theorem. Since the last limit above equals $(H \cdot A)_s$, we conclude that $H \cdot A$ is indeed a local martingale. □

A note of caution is in order here. Theorem 29 does not extend completely to processes that are not in $\mathbb{L}$ but are only predictably measurable, as we will see in Emery's example of a stochastic integral behaving badly on page 176 in Chap. IV.

Let X be a classical semimartingale, and let $X_t = X_0 + M_t + A_t$ be a decomposition where $M_0 = A_0 = 0$, M is a local martingale, and A is an FV

process. Then if the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ supports a Poisson process N, we can write

$$X_t = X_0 + \{M_t + N_t - t\} + \{A_t - N_t + t\}$$

as another decomposition of X. In other words, the decomposition of a classical semimartingale need not be unique. This problem can often be solved by choosing a certain canonical decomposition which is unique.

Definition. Let X be a semimartingale. If X has a decomposition $X_t = X_0 + M_t + A_t$ with $M_0 = A_0 = 0$, M a local martingale, A an FV process, and with A *predictable*, then X is said to be a **special semimartingale**.

To simplify notation we henceforth assume $X_0 = 0$.

Theorem 30. *If X is a special semimartingale, then its decomposition $X = M + A$ with A predictable is unique.*

Proof. Let $X = N+B$ be another such decomposition. Then $M-N = B-A$, hence $B - A$ is an FV process which is a local martingale. Moreover, $B - A$ is predictable, and hence constant by Theorem 12. Since $B_0 - A_0 = 0$, we conclude $B = A$. □

Definition. If X is a special semimartingale, then the unique decomposition $X = M + A$ with $M_0 = X_0$ and $A_0 = 0$ and A predictable is called the **canonical decomposition**.

Theorem 15 shows that any quasimartingale is special. A useful sufficient condition for a semimartingale X to be special is that X be a classical semimartingale, or equivalently decomposable, *and also have bounded jumps*.

Theorem 31. *Let X be a classical semimartingale with bounded jumps. Then X is a special semimartingale.*

Proof. Let $X_t = X_0 + M_t + A_t$ be a decomposition of X with $M_0 = A_0 = 0$, M a local martingale, and A an FV process. By Theorem 25 we can then also write

$$X_t = X_0 + N_t + B_t$$

where N is a local martingale with bounded jumps and B is an FV process. Since X and N each have bounded jumps, so also does B. Consequently, it is locally a quasimartingale and therefore decomposes

$$B = L + \widetilde{B}$$

where L is a local martingale and $\widetilde{B}$ is a predictable FV process (Theorem 15). Therefore

$$X_t = X_0 + \{N_t + L_t\} + \widetilde{B}_t$$

is the canonical decomposition of X and hence X is special. □

Corollary. Let X be a classical semimartingale with continuous paths. Then X is special and in its canonical decomposition

$$X_t = X_0 + M_t + A_t,$$

the local martingale M and the FV process A have continuous paths.

Proof. X is continuous hence trivially has bounded jumps, so it is special by Theorem 31. Since X is continuous we must have

$$\Delta M_T = -\Delta A_T$$

for any stopping time T ($\Delta A_T = 0$ by convention on $\{T = \infty\}$). Suppose A jumps at a stopping time T. By Theorem 3, $T = T_A \wedge T_B$, where T_A is accessible and T_B is totally inaccessible. By Theorem 21 it follows that $P(|\Delta A_{T_B}| > 0) = 0$. Hence without loss of generality we can assume T is accessible. It then suffices to consider T predictable since countably many predictable times cover the stopping time T. Let S_n be a sequence of stopping times announcing T. Since A is predictable, we know by Theorem 24 that ΔA_T is $\mathcal{F}_{T-}$ measurable. Therefore ΔM_T is also $\mathcal{F}_{T-}$ measurable. Stop M so that it is a uniformly integrable martingale. Then

$$\begin{aligned}\Delta M_T &= E\{\Delta M_T | \mathcal{F}_{T-}\}\\ &= 0,\end{aligned}$$

and M, and hence A, are continuous, using Theorem 7 of Chap. I. □

Theorem 31 can be strengthened, as the next two theorems show. The criteria given in these theorems are quite useful.

Theorem 32. *Let X be a semimartingale. X is special if and only if the process $J_t = \sup_{s \le t} |\Delta X_s|$ is locally integrable.*

Theorem 33. *Let X be a semimartingale. X is special if and only if the process $X_t^* = \sup_{s \le t} |X_s|$ is locally integrable.*

Before proving the theorems, we need a preliminary result, which is interesting in its own right.

Theorem 34. *Let M be a local martingale and let $M_t^* = \sup_{s \le t} |M_s|$. Then the increasing process M^* is locally integrable.*

Proof. Without loss of generality assume $M_0 = 0$. Let T_n be a sequence of stopping times increasing to ∞ such that M^{T_n} is a uniformly integrable martingale for each n. Since we can replace T_n with $T_n \wedge n$ if necessary, without loss of generality we can assume that the stopping times T_n are each bounded. Set $S_n = T_n \wedge \inf_t\{|M_t| \ge n\}$. Since M^{T_n} is uniformly integrable, and $S_n \le T_n$, we have that M_{S_n} is integrable. But we also have that $M_{S_n}^* \le n \vee |M_{S_n}|$ which is in L^1. Since S_n increases to ∞ a.s., the proof is complete. □

Proof of Theorems 32 and 33. Since X is special the process A of its canonical decomposition $X = M + A$ is of locally integrable variation. Writing $X_t^* \leq M_t^* + \int_0^t |dA_s|$, and since M_t^* is locally integrable by Theorem 34 and since A is of locally integrable variation, we have that X^* is locally integrable. Further, note that $|\Delta X_t| = |X_t - X_{t-}| \leq 2|X_t^*|$, hence $J_t \leq X_t^*$ and we have that J_t is also locally integrable.

For the converse, it will suffice to show that there exists a decomposition $X = M + A$ where A is of locally integrable variation, since then we can take the compensator of A and obtain a canonical decomposition. To this end note that $|\Delta A_s| \leq |\Delta M_s| + |\Delta X_s| \leq 2M_s^* + J_s$, which is locally integrable by the hypotheses of Theorem 32 together with Theorem 34. Since $J_t \leq X_t^*$ we have that $|\Delta A_s|$ is locally integrable by the hypotheses of Theorem 33 together with Theorem 34 as well. To complete the proof, let $T_n = \inf\{t > 0 : \int_0^t |dA_s| \geq n\}$, and note that $\int_0^{T_n} |dA_s| = \int_0^{T_n-} |dA_s| + |\Delta A_{T_n}| \leq n + \sup_{s \leq T_n} |\Delta A_s|$, which is in L^1 as soon as $\sup_{s \leq T_n} |\Delta A_s|$ is. However this last term is locally in L^1 as we have seen, and the proof is complete with one more sequence of stopping times R_n putting $\sup_{s \leq R_n} |\Delta A_s|$ in L^1 and then constructing a new, single sequence by taking minimums from the two sequences. □

A class of examples of special semimartingales is Lévy processes with bounded jumps. By Theorem 31 we need only show that a Lévy process Z with bounded jumps is a classical semimartingale. This is merely Theorem 40 of Chap. I.

For examples of semimartingales which are *not* special, we need only to construct semimartingales with finite variation terms that are not locally integrable. For example, let X be a compound Poisson process with non-integrable, nonnegative jumps. Or, more generally, let X be a Lévy process with Lévy measure ν such that $\int_1^\infty x\nu(dx) = \infty$ and/or $\int_{-\infty}^{-1} |x|\nu(dx) = \infty$.

8 Girsanov's Theorem

Let X be a semimartingale on a space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses. We saw in Chap. II that if Q is another probability law on $(\Omega, \mathcal{F})$, and $Q \ll P$, then X is a Q semimartingale as well. If X is a classical semimartingale (or equivalently, if X is decomposable) and has a decomposition $X = M + A$, M a local martingale and A an FV process, then it is often useful to be able to calculate the analogous decomposition $X = N + B$, if it exists, under Q. This is rather tricky unless we make a simplifying assumption which usually holds in practice: that both $Q \ll P$ *and* $P \ll Q$.

Definition. Two probability laws P, Q on $(\Omega, \mathcal{F})$ are said to be **equivalent** if $P \ll Q$ and $Q \ll P$. (Recall that $P \ll Q$ denotes that P is *absolutely continuous* with respect to Q.) We write $Q \sim P$ to denote equivalence.

If $Q \ll P$, then there exists a random variable Z in $L^1(dP)$ such that $\frac{dQ}{dP} = Z$ and $E_P\{Z\} = 1$, where E_P denotes expectation with respect to the law P. We let

$$Z_t = E_P\left\{\frac{dQ}{dP}|\mathcal{F}_t\right\}$$

be the right continuous version. Then Z is a uniformly integrable martingale and hence a semimartingale (by the corollary of Theorem 26). Note that if Q is equivalent to P, then $\frac{dP}{dQ} \in L^1(dQ)$ and $\frac{dP}{dQ} = \left(\frac{dQ}{dP}\right)^{-1}$.

We begin with a simple lemma, the easy proof of which we leave as an exercise for the reader.

Lemma. Let $Q \sim P$, and $Z_t = E_P\left\{\frac{dQ}{dP}|\mathcal{F}_t\right\}$. An adapted, càdlàg process M is a Q local martingale if and only if MZ is a P local martingale.

Proof. See Exercise 20. □

Theorem 35 (Girsanov-Meyer Theorem). *Let P and Q be equivalent. Let X be a classical semimartingale under P with decomposition $X = M + A$. Then X is also a classical semimartingale under Q and has a decomposition $X = L + C$, where*

$$L_t = M_t - \int_0^t \frac{1}{Z_s} d[Z, M]_s$$

is a Q local martingale, and $C = X - L$ is a Q FV process.

Proof. Recall that by Theorem 2 of Chap. II it is trivial that X is a Q semimartingale. We need to show it is a classical semimartingale, with the above decomposition being valid.

Since M and Z are P local martingales, they are semimartingales (corollary of Theorem 26) and

$$\int Z_- dM + \int M_- dZ$$

is a local martingale as well (Theorem 29). Using integration by parts we have

$$ZM - [Z, M] = \int Z_- dM + \int M_- dZ \tag{$*$}$$

is also a P local martingale. Since Z is a version of $E_P\{\frac{dQ}{dP}|\mathcal{F}_t\}$, we have $\frac{1}{Z}$ is a càdlàg version of $E_Q\{\frac{dP}{dQ}|\mathcal{F}_t\}$; therefore $\frac{1}{Z}$ is a Q semimartingale. Since $P \ll Q$, it is also a P semimartingale. Multiplying equation $(*)$ by $\frac{1}{Z}$ we have

$$M - \left(\frac{1}{Z}\right)[Z, M] = \frac{1}{Z}\left(\int Z_- dM + \int M_- dZ\right). \tag{$**$}$$

Note that if we were to multiply the right side of $(**)$ by Z we would obtain a P local martingale. We conclude by the lemma preceding this theorem that

$M - (\frac{1}{Z})[Z,M]$ is a Q local martingale. We next use integration by parts (under Q):

$$\left(\frac{1}{Z}\right)[Z,M] = \int \frac{1}{Z_-}d[Z,M] + \int [Z,M]_- d\left(\frac{1}{Z}\right) + [[Z,M],\frac{1}{Z}].$$

Let $N = \int [Z,M]_- d\left(\frac{1}{Z}\right)$. Since $\frac{1}{Z}$ is a Q local martingale, so also is N (Theorem 29). The above becomes:

$$\begin{aligned}\left(\frac{1}{Z_t}\right)[Z,M]_t &= \int_0^t \frac{1}{Z_{s-}}d[Z,M]_s + N_t + [[Z,M],\frac{1}{Z}]_t \\ &= \int_0^t \frac{1}{Z_{s-}}d[Z,M]_s + N_t + \sum_{0<s\le t} \Delta\left(\frac{1}{Z_s}\right)\Delta[Z,M]_s \\ &= \int_0^t \frac{1}{Z_s}d[Z,M]_s + N_t.\end{aligned}$$

Adding our two Q local martingales yields

$$\begin{aligned}N + M - \left(\frac{1}{Z}\right)[Z,M] &= N + M - \int \frac{1}{Z}d[Z,M] - N \\ &= M - \int \frac{1}{Z}d[Z,M]\end{aligned}$$

which, being the sum of two Q local martingales, is itself one. This establishes the theorem. □

The Girsanov-Meyer Theorem has another version using the sharp bracket. It transforms a P local martingale into a Q special semimartingale and gives its canonical decomposition. However an extra integrability hypothesis is needed for the theorem to be true. Note that in the predictable version the integrand $\frac{1}{Z_s}$ is replaced by its left continuous version $\frac{1}{Z_{s-}}$.

Theorem 36 (Girsanov-Meyer Theorem: Predictable Version). *Let X be a P local martingale with $X_0 = 0$. Let Q be another probability equivalent to P and let $Z_t = E\{\frac{dQ}{dP}|\mathcal{F}_t\}$. If $\langle X,Z\rangle$ exists for the P probability, then the canonical Q decomposition of X is*

$$X_t = (X_t - \int_0^t \frac{1}{Z_{s-}}d\langle X,Z\rangle_s) + \int_0^t \frac{1}{Z_{s-}}d\langle X,Z\rangle_s.$$

Proof. Using the same ideas as in the proof of Theorem 35, we want to find a predictable finite variation term A such that $Y = Z(X-A)$ is a P local martingale. Using integration by parts yields

$$\begin{aligned}Z_tX_t &= \int_0^t Z_{s-}dX_s + \int_0^t X_{s-}dZ_s + [Z,X] \\ &= N_t + [Z,X]_t - \langle X,Z\rangle_t + \langle X,Z\rangle_t \\ &= M_t + \langle X,Z\rangle_t\end{aligned}$$

where both $N = \int_0^t Z_{s-}dX_s + \int_0^t X_{s-}dZ_s$ and $M = N + [Z,X] - \langle X,Z\rangle$ are P local martingales. Therefore $ZX - \langle X,Z\rangle$ is a P local martingale.

Let $A = \langle X,Z\rangle$, which of course is a predictable FV process. Moreover we know that A does not jump at totally inaccessible times, and that we can cover its jumps with predictable times. This implies that A is locally bounded. Let $T_n = \inf\{t > 0 : |A_t| > n\}$, and let A^{T_n-} be defined by $A^{T_n-} = \int 1_{[0,T_n)}(s)dA_s$. From now on we assume $A = A^{T_n-}$. Then A is still an FV process and is still predictable, $|A| \le n$, and the stopping times T_n increase to ∞ a.s. Integration by parts gives us also

$$\frac{1}{Z_t}A_t = \int_0^t \frac{1}{Z_{s-}}dA_s + \int_0^t A_{s-}d(\frac{1}{Z_s}) + [A, \frac{1}{Z}]_t.$$

If we show that $[A, \frac{1}{Z}]$ is also a Q local martingale, we will be done. It suffices to show that $E_Q\{[A, \frac{1}{Z}]_T\} = 0$ for every stopping time T. Since A is FV and predictable, by stopping at $T-$ if necessary, we can assume the jumps of A are bounded. And since $\frac{1}{Z}$ is a Q local martingale, its supremum process is locally integrable (Theorem 34). This is enough to give us that $[A, \frac{1}{Z}]$ is Q locally of integrable variation. By stopping we assume it is integrable. We then have that A is Q natural, so if T is a stopping time, we replace A with A^T and we still have a natural process, and thus we have $E_Q\{[A, \frac{1}{Z}]_T\} = 0$. Since T was arbitrary, this implies that $[A, \frac{1}{Z}]$ is a martingale. Of course, the process is implicitly stopped at several stopping times, so what we have actually shown is that $[A, \frac{1}{Z}]$ is a Q local martingale.

We now have that

$$\frac{1}{Z_t}A_t = Q\text{-local martingale} + \int_0^t \frac{1}{Z_{s-}}dA_s$$

which in turn gives us, recalling that $A = \langle X,Z\rangle$,

$$\begin{aligned} X_t &= \frac{1}{Z_t}(Z_tX_t) = \frac{1}{Z_t}(M_t + \langle X,Z\rangle_t) \\ &= \frac{1}{Z_t}M_t + Q\text{-local martingale} + \int_0^t \frac{1}{Z_{s-}}d\langle X,Z\rangle_s. \end{aligned}$$

Since $Z(\frac{1}{Z}M) = M = P$ local martingale, we have that $\frac{1}{Z}M$ is also a Q local martingale. Thus finally $X_t - \int_0^t \frac{1}{Z_{s-}}d\langle X,Z\rangle_s$ is a Q local martingale, and the proof is done. □

Let us next consider the interesting case where Q is absolutely continuous with respect to P, but the two probability measures are not equivalent. (That is, *we no longer assume that P is also absolutely continuous with respect to Q*.) We begin with a simple result, and we will assume $\mathcal{F}_0 = \{\emptyset, \Omega\}$ a.s.

Theorem 37. *Let X be a P local martingale with $X_0 = 0$. Let Q be another probability absolutely continuous with respect to P, and let $Z_t = E\{\frac{dQ}{dP}|\mathcal{F}_t\}$.*

Assume that $\langle X, Z\rangle$ exists for P. Then $A_t = \int_0^t \frac{1}{Z_{s-}} d\langle X, Z\rangle_s$ exists a.s. for the probability Q, and $X_t - \int_0^t \frac{1}{Z_{s-}} d\langle X, Z\rangle_s$ is a Q local martingale.

Proof. $Z_0 = E\{Z\} = 1$, so if we let $R_n = \inf\{t > 0 : Z_t \leq 1/n\}$, then R_n increase to ∞, Q-a.s., and the process $\frac{1}{Z_{t-}}$ is bounded on $[0, R_n]$. By Theorem 36, $X_t^{R_n} - \int_0^t \frac{1}{Z_{s-}} d\langle X, Z\rangle_s^{R_n}$ is a Q local martingale, each n. Since a local, local martingale is a local martingale, we are done. □

We now turn to the general case, where we no longer assume the existence of $\langle X, Z\rangle$, calculated with P. (As before, we take X to be a P local martingale.) We begin by defining a key stopping time: $R = \inf\{t > 0 : Z_t = 0,\ Z_{t-} > 0\}$. Note that $Q(R < \infty) = 0$, but it is entirely possible that $P(R < \infty) > 0$. We further define $U_t = \Delta X_R 1_{\{t \geq R\}}$. Then U is an FV process, and moreover U is locally integrable (dP). Let T_n increase to ∞ and be such that X^{T_n} is a uniformly integrable martingale. Then

$$|\Delta X_R^{T_n}| = |X_{T_n \wedge R} - X_{(T_n \wedge R)-}| \leq |X_{T_n \wedge R}| + |X_{(T_n \wedge R)-}| \in L^1.$$

Thus U has a compensator $\widetilde{U}$, and of course $\widetilde{U}$ is predictable and $U - \widetilde{U}$ is a P local martingale.

Theorem 38 (Lenglart-Girsanov Theorem). *Let X be a P local martingale with $X_0 = 0$. Let Q be a probability absolutely continuous with respect to P, and let $Z_t = E_P\{\frac{dQ}{dP} | \mathcal{F}_t\}$, $R = \inf\{t > 0 : Z_t = 0,\ Z_{t-} > 0\}$, and $U_t = \Delta X_R 1_{\{t \geq R\}}$. Then*

$$X_t - \int_0^t \frac{1}{Z_s} d[X, Z]_s + \widetilde{U}_t$$

is a Q local martingale.

Proof. Let $R_n = \inf\{t > 0 : Z_t \leq \frac{1}{n}\}$. (Recall that $Z_0 = 1$, and also note that it is possible that $R_n = R$.) Then both X^{R_n} and Z^{R_n} are P local martingales. Also note that $A_t^{R_n} = \int_0^t \frac{1}{Z_s^{R_n}} 1_{\{Z_s^{R_n} > 0\}} d[X^{R_n}, Z^{R_n}]_s$, $\widetilde{U}^{R_n}$, and $Y^{R_n} = X^{R_n} - A^{R_n} + \widetilde{U}^{R_n}$ are all P-well-defined. We can define

$$A_t = \int_0^t \frac{1}{Z_s} 1_{\{Z_s > 0\}} d[X, Z]_s$$

on $[0, R)$, since $d[X^{R_n}, Z^{R_n}]_s$ does not charge (R, ∞), and $\frac{1}{Z_s^{R_n}} 1_{\{Z_s^{R_n} > 0\}} = 0$ at R. Thus we need only to show Y^{R_n} is a Q local martingale for each fixed n, which is the same as showing that $Z^{R_n} Y^{R_n}$ is a P local martingale. Let us assume all these processes are stopped at R_n to simplify notation. We have

$$ZY = ZX - ZA + Z\widetilde{U}.$$

Hence,

$$\begin{aligned} d(ZX) &= Z_- dX + X_- dZ + d[Z,X] \\ &= \text{ local martingale } + d[Z,X] \\ d(AZ) &= A_- dZ + Z dA \\ &= \text{ local martingale } + Z dA \\ &= \text{ local martingale } + 1_{\{Z>0\}} d[X,Z] \\ d(Z\widetilde{U}) &= Z_- d\widetilde{U} + \widetilde{U} dZ \\ &= \text{ local martingale } + Z_- d\widetilde{U} \\ &= \text{ local martingale } + Z_- dU \end{aligned}$$

where the last equality uses that $U - \widetilde{U}$ is a local martingale (dP). Summarizing we have

$$\begin{aligned} ZY &= ZX - ZA + Z\widetilde{U} \\ &= \text{ local martingale } + [Z,X] - \text{ local martingale} \\ &\quad + \int 1_{\{Z>0\}} d[X,Z]) + \text{ local martingale } + \int Z_- dU \end{aligned}$$

which we want to be a local martingale under dP. This will certainly be the case if

$$d[Z,X] - 1_{\{Z>0\}} d[X,Z] + Z_- dU = 0. \tag{$*$}$$

However $(*)$ equals

$$\Delta Z_R \Delta X_R 1_{\{t \wedge R_n \geq R\}} + Z_{R-} \Delta X_R 1_{\{t \wedge R_n \geq R\}}.$$

But $\Delta Z_R = Z_R - Z_{R-} = 0 - Z_{R-} = -Z_{R-}$, and this implies that equation $(*)$ is indeed zero, and thus the Lenglart-Girsanov Theorem holds. □

Corollary. Let X be a continuous local martingale under P. Let Q be absolutely continuous with respect to P. Then $\langle X, Z\rangle = [Z,X] = [Z,X]^c$ exists, and

$$X_t - \int_0^t \frac{1}{Z_{s-}} d[Z,X]_s^c = X_t - \int_0^t \alpha_s d[X,X]_s$$

which is a Q local martingale.

Proof. By the Kunita-Watanabe inequality we have

$$\int_0^t \frac{1}{Z_s} d[Z,X]_s^c \leq \left(\int_0^t \left(\frac{1}{Z_s}\right)^2 d[Z,Z]_s\right)^{\frac{1}{2}} \left(\int_0^t d[X,X]_s\right)^{\frac{1}{2}},$$

which shows that it is absolutely continuous with respect to $d[X,X]_s$ a.s., whence the result. □

We remark that if Z is the solution of a stochastic exponential equation of the form $dZ_s = Z_{s-}H_s dX_s$ (which it often is), then $\alpha_s = H_s$.

Example. A problem that arises often in *mathematical finance theory* is that one has a semimartingale $S = M + A$ defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses, and one wants to find an equivalent probability measure Q such that under Q the semimartingale X is a local martingale, or better, a martingale. In essence this amounts to finding a probability measure that "removes the drift." To be concrete, let us suppose S is the solution of a stochastic differential equation[3]

$$dS_s = h(s, S_s)dB_s + b(s; S_r; r \leq s)ds,$$

where B is a standard Wiener process (Brownian motion) under P. Let us postulate the existence of a Q and let $Z = \frac{dQ}{dP}$ and $Z_t = E\{Z|\mathcal{F}_t\}$, which is clearly a càdlàg martingale. By Girsanov's Theorem

$$\int_0^t h(s, S_s)dB_s - \int_0^t \frac{1}{Z_s} d[Z, \int_0^\cdot h(r, S_r)dB_r]_s$$

is a Q local martingale. We want to find the martingale Z. In Chap. IV we will study martingale representation and show in particular that every local martingale on a Brownian space is a stochastic integral with respect to Brownian motion. Thus we can write $Z_t = 1 + \int_0^t J_s dB_s$ for some predictable process J. If we assume Z is well behaved enough to define $H_s = \frac{J_s}{Z_s}$, then we have $Z_t = 1 + \int_0^t H_s Z_s dB_s$, which gives us a linear stochastic differential equation to solve for Z. Thus if we let $N_t = \int_0^t H_s dB_s$, we get that $Z_t = \mathcal{E}(N)_t$.[4] It remains to determine H. We do this by observing from our previous Girsanov calculation that

$$\int_0^t h(s, S_s)dB_s - \int_0^t \frac{1}{Z_s} Z_s H_s h(s, S_s)ds = \int_0^t h(s, S_s)dB_s - \int_0^t H_s h(s, S_s)ds$$

is a Q local martingale. We then choose $H_s = \dfrac{-b(s; S_r; r \leq s)}{h(s, S_s)}$, which yields

$$S_t = \int_0^t h(s, S_s)dB_s + \int_0^t b(s; S_r; r \leq s)ds$$

is a local martingale under Q. Letting $M_t = B_t + \displaystyle\int_0^t \frac{b(s; S_r; r \leq s)}{h(s, S_s)} ds$ denote this Q local martingale, we get that $[M, M]_t = [B, B]_t = t$, and by Lévy's Theorem M is a Q-Brownian motion. Finally, under Q we have that S satisfies the stochastic differential equation

$$dS_t = h(t, S_t)dM_t.$$

---

[3] Stochastic differential equations are introduced and studied in some detail in Chap. V.

[4] The stochastic exponential $\mathcal{E}$ is defined on page 85.

There is one problem with the preceding example: we do not know *a priori* whether our solution Z is the Radon-Nikodym density for simply a measure Q, or whether Q is an actual *bona fide* probability measure. This is a constant problem. Put more formally, we wish to address this problem:

Let M be a local martingale. When is $\mathcal{E}(M)$ a martingale?

The only known general conditions that solve this problem are Kazamaki's criterion and Novikov's criterion.[5] Moreover, these criteria apply only to local martingales with continuous paths. Novikov's is a little less powerful than Kazamaki's, but it is much easier to check in practice. Since Novikov's criterion follows easily from Kazamaki's, we present both criteria here. Note that if M is a continuous local martingale, then of course $\mathcal{E}(M)$ is also a continuous local martingale. Even if, however, it is a uniformly integrable local martingale, it still need not be a martingale; we need a stronger condition. As an example, one can take $u(x) = \|x\|^{-1}$ and $N_t = u(B_t)$ where B is standard three dimensional Brownian motion. Then N is a uniformly integrable local martingale but not a martingale.

Nevertheless, whenever M is a continuous local martingale, then $\mathcal{E}(M)$ is a positive supermartingale, as is the case with any nonnegative local martingale. Since $\mathcal{E}(M)_0 = 1$, and since $\mathcal{E}(M)$ is a positive supermartingale, we have $E\{\mathcal{E}(M)_t\} \leq 1$, all t. (See the lemma below.) It is easy to see that Z *is a true martingale if one also has* $E\{\mathcal{E}(M)_t\} = 1$, *for all* t. We begin with some preliminary results.

Lemma. Let M be a continuous local martingale with $M_0 = 0$. Then $E\{\mathcal{E}(M)_t\} \leq 1$ for all $t \geq 0$.

Proof. Recall that $\mathcal{E}(M)_0 = 1$. Since M is a local martingale, $\mathcal{E}(M)$ is a nonnegative local martingale. Let T_n be a sequence of stopping times reducing $\mathcal{E}(M)$. Then $E\{\mathcal{E}(M)_{t\wedge T_n}\} = 1$, and using Fatou's Lemma,

$$E\{\mathcal{E}(M)_t\} = E\{\liminf_{n\to\infty} \mathcal{E}(M)_{t\wedge T_n}\} \leq \liminf_{n\to\infty} E\{\mathcal{E}(M)_{t\wedge T_n}\} = 1. \qquad \square$$

Theorem 39. *Let M be a continuous local martingale. Then*

$$E\{e^{\frac{1}{2}M_t}\} \leq E\{e^{\frac{1}{2}[M,M]_t}\}^{1/2}.$$

Proof.

$$\begin{aligned}(\mathcal{E}(M))^{\frac{1}{2}} &= (e^{M_t - \frac{1}{2}[M,M]_t})^{\frac{1}{2}} \\ &= e^{\frac{1}{2}M_t}(e^{-\frac{1}{2}[M,M]_t})^{\frac{1}{2}}\end{aligned}$$

which implies that

[5] There also exist partial results when the local martingale is no longer continuous but only has càdlàg paths. See Exercise 14 of Chap. V for an example of these results.

$$e^{\frac{1}{2}M_t} = (\mathcal{E}(M)_t)^{\frac{1}{2}}(e^{\frac{1}{2}[M,M]_t})^{\frac{1}{2}}$$

and this together with the Cauchy-Schwarz inequality and the fact that $E\{\mathcal{E}(M)_t\} \leq 1$ gives the result. □

Lemma. Let M be a continuous local martingale. Let $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Taking the supremum below over all bounded stopping times, assume that

$$\sup_T E\{e^{(\frac{\sqrt{p}}{2\sqrt{p}-1})M_T}\} < \infty.$$

Then $\mathcal{E}(M)$ is an L^q bounded martingale.

Proof. Let $1 < p < \infty$ and $r = \frac{\sqrt{p}+1}{\sqrt{p}-1}$. Then $s = \frac{\sqrt{p}+1}{2}$ and $\frac{1}{r} + \frac{1}{s} = 1$. Also we note that $(q - \sqrt{\frac{q}{r}})s = \frac{\sqrt{p}}{2(\sqrt{p}-1)}$ which we use in the last equality of the proof. We have

$$\begin{aligned}\mathcal{E}(M) &= e^{qM - \frac{q}{2}[M,M]} \\ &= e^{\sqrt{\frac{q}{r}}M - \frac{q}{2}[M,M]}e^{(q-\sqrt{\frac{q}{r}})M}\end{aligned}$$

We now apply Hölder's inequality for a stopping time S:

$$\begin{aligned}E\{\mathcal{E}(M)^q_S\} &= E\{e^{\sqrt{qr}M_S - \frac{qr}{2}[M,M]_S}\}^{\frac{1}{r}}E\{e^{(q-\sqrt{\frac{q}{r}})M_S}\}^{\frac{1}{s}} \\ &= (E\{\mathcal{E}(\sqrt{qr}M)_S\})^{\frac{1}{r}}(E\{e^{(\frac{\sqrt{p}}{2(\sqrt{p}-1)})M_S}\})^{\frac{1}{s}}\end{aligned}$$

Recalling that $E\{\mathcal{E}(\sqrt{qr}M)_S\} \leq 1$, we have the result. □

Theorem 40 (Kazamaki's Criterion). *Let M be a continuous local martingale. Suppose $\sup_T E\{e^{(\frac{1}{2}M_T)}\} < \infty$, where the supremum is taken over all bounded stopping times. Then $\mathcal{E}(M)$ is a uniformly integrable martingale.*

Proof. Let $0 < a < 1$, and $p > 1$ be such that $\frac{\sqrt{p}}{(\sqrt{p}-1)} < \frac{1}{a}$. Our hypothesis combined with the preceding lemma imply that $\mathcal{E}(aM)$ is an L^q bounded martingale, where $\frac{1}{p} + \frac{1}{q} = 1$, which in turn implies it is a uniformly integrable martingale. However

$$\begin{aligned}\mathcal{E}(aM) &= e^{aM - \frac{a^2}{2}[M,M]} = e^{a^2M - \frac{a^2}{2}[M,M]}e^{a(1-a)M} \\ &= \mathcal{E}(M)^{a^2}e^{a(1-a)M},\end{aligned}$$

and using Hölder's inequality with a^{-2} and $(1-a^2)^{-1}$ yields (where the 1 on the left side comes from the uniform integrability):

$$\begin{aligned}1 &= E\{\mathcal{E}(aM)_\infty\} \leq E\{(\mathcal{E}(M)_\infty)\}^{a^2}E\{(e^{a(1-a)M_\infty})^{\frac{1}{1-a^2}}\}^{1-a^2} \\ &= E\{\mathcal{E}(M)_\infty\}^{a^2}E\{e^{\frac{a}{1+a}M_\infty}\}^{1-a^2} \\ &\leq E\{\mathcal{E}(M)_\infty\}^{a^2}E\{e^{\frac{1}{2}M_\infty}\}^{2a(1-a)}\end{aligned}$$

Now let a increase to 1 and the second term on the right side of the last inequality above converges to 1 since $2a(1-a) \to 0$. Thus $1 \leq E\{\mathcal{E}(M)_\infty\}$, and since we know that it is always true that $1 \geq E\{\mathcal{E}(M)_\infty\}$, we are done. □

As a corollary we get the very useful Novikov's criterion. Because of its importance, we call it a theorem.

Theorem 41 (Novikov's Criterion). *Let M be a continuous local martingale, and suppose that*

$$E\{e^{\frac{1}{2}[M,M]_\infty}\} < \infty.$$

Then $\mathcal{E}(M)$ is a uniformly integrable martingale.

Proof. By Theorem 39 we have $E\{e^{\frac{1}{2}M_T}\} \leq E\{e^{\frac{1}{2}[M,M]_T}\}^{\frac{1}{2}}$, and we need only to apply Kazamaki's criterion (Theorem 40). □

We remark that it can be shown that $1/2$ is the best possible constant in Novikov's criterion, even though Kazamaki's criterion is slightly stronger. Note that in the case of the example treated earlier, we have $[N,N]_t = \int_0^t H_s^2 ds$ where $H = \frac{b(s;S_r;r\leq s)}{h(s,S_s)}$. Clearly whether or not $\frac{1}{2}[N,N]_t$ has a finite exponential moment will depend on our choice of b and of h. A standard simple solution is to choose b and h so that b/h is bounded.

Partial results exist that give an analogue of Novikov's criterion in the general (càdlàg) case. See for example Exercise 14 of Chap. V, and the sources [156] and [157].

Example. This example is taken from *statistical communication theory*. Suppose we are receiving a signal corrupted by noise, and we wish to determine if there is indeed a signal, or if we are just receiving noise (e.g., we could be searching for signs of intelligent life in our galaxy with a radio telescope). Let $x(t)$ be the received signal, $\xi(t)$ the noise, and $s(t)$ the actual (transmitted) signal. Then

$$x(t) = s(t) + \xi(t).$$

A frequent assumption is that the noise is "white." A **white noise** is usually described as a second order wide sense stationary process with a constant spectral density function (that is, $E\{\xi_t\} = 0$, and $V(\tau) = E\{X_t X_{t+\tau}\} = \delta_\tau S_0$, where δ_τ is the Dirac delta function at τ, and S_0 is the constant spectral density; one then has $S(\nu) = \int_{-\infty}^{\infty} e^{i2\pi\nu\tau} V(\tau) d\tau$). Such a process does not exist in a rigorous mathematical sense. Indeed it can be interpreted as the derivative of the Wiener process, in a generalized function sense. (See Arnold [2, page 53], for example.) This suggests that we consider the integrated version of our signal-noise equation:

$$\begin{aligned} X_t &= \int_0^t s(u)du + W_t \\ &= S_t + W_t \end{aligned}$$

where W is a standard Wiener process and where S_t is thought of as "$\int_0^t x(s)ds$," the cumulative received signal.

The key step in our analysis is the following consequence of the Girsanov-Meyer Theorem.

Theorem 42. *Let W be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, P)$, and let $H \in \mathbb{L}$ be bounded. Let*

$$X_t = \int_0^t H_s ds + W_t$$

and define Q by $\frac{dQ}{dP} = \exp\{\int_0^T -H_s dW_s - \frac{1}{2}\int_0^T H_s^2 ds\}$, for some $T > 0$. Then under Q, X is a standard Brownian motion for $0 \le t \le T$.

Proof. Let $Z_T = \exp\{\int_0^T -H_s dW_s - \frac{1}{2}\int_0^T H_s^2 ds\}$. Then if $Z_t = E\{Z_T|\mathcal{F}_t\}$ we know by Theorem 37 of Chap. II that Z satisfies the equation

$$Z_t = 1 - \int_0^t Z_{s-} H_s dW_s.$$

By the Girsanov-Meyer Theorem (Theorem 35), we know that

$$N_t = W_t - \int_0^t \frac{1}{Z_s} d[Z, W]_s$$

is a Q local martingale. However

$$\begin{aligned}[Z, W]_t = [-Z_- H \cdot W, W]_t &= \int_0^t -Z_s H_s d[W, W]_s \\ &= -\int_0^t Z_s H_s ds,\end{aligned}$$

since $[W, W]_t = t$ for Brownian motion. Therefore

$$\begin{aligned}N_t &= W_t - \int_0^t -\frac{1}{Z_s} Z_s H_s ds \\ &= W_t + \int_0^t H_s ds \\ &= X_t,\end{aligned}$$

hence X is a Q local martingale. Since $(\int_0^t H_s ds)_{t \ge 0}$ is a continuous FV process we have that $[X, X]_t = [W, W]_t = t$, and by Lévy's Theorem (Theorem 39 of Chap. II) we conclude that X is a standard Brownian motion. □

Corollary. Let W be a standard Brownian motion and $H \in \mathbb{L}$ be bounded. Then the law of

$$X_t = \int_0^t H_s ds + W_t,$$

$0 \le t \le T < \infty$, is equivalent to Wiener measure.

Proof. Let $\mathcal{C}[0,T]$ be the space of continuous functions on $[0,T]$ with values in $\mathbb{R}$ (such a space is called a *path space*). If $W = (W_t)_{0\leq t\leq T}$ is a standard Brownian motion, it induces a measure μ_W on $\mathcal{C}[0,T]$:

$$\mu_W(A) = P\{\omega : t \mapsto W_t(\omega) \in A\}.$$

Let μ_X be the analogous measure induced by X. Then by Theorem 42 we have $\mu_X \sim \mu_W$ and further we have

$$\frac{d\mu_W}{d\mu_X} = \exp\left\{\int_0^T -H_s dW_s - \frac{1}{2}\int_0^T H_s^2 ds\right\}. \qquad \square$$

Remark. We have not tried for maximum generality here. For example the hypothesis that H be bounded can be weakened. It is also desirable to weaken the restriction that $H \in \mathbb{L}$. Indeed we only needed that hypothesis to be able to form the integral $\int_0^t H_s dW_s$. This is one example to indicate why we need a space of integrands more general than $\mathbb{L}$.

We are now in a position to consider the problem posed earlier: is there a signal corrupted by noise, or is there just noise (that is, does $s(t) = 0$ a.e., a.s.)? In terms of hypothesis testing, let $\mathbf{H}_0$ denote the null hypothesis, $\mathbf{H}_1$ the alternative. We have:

$$\mathbf{H}_0 : X_T = W_T$$
$$\mathbf{H}_1 : X_T = \int_0^T H_s ds + W_T.$$

We then have

$$\frac{d\mu_W}{d\mu_X} = \exp\left(\int_0^T -H_s dW_s - \frac{1}{2}\int_0^T H_s^2 ds\right),$$

by the preceding corollary. This leads to a likelihood ratio test:

$$\text{if} \quad \frac{d\mu_W}{d\mu_X}(\omega) \leq \lambda, \quad \text{reject } \mathbf{H}_0,$$
$$\text{if} \quad \frac{d\mu_W}{d\mu_X}(\omega) > \lambda, \quad \text{fail to reject } \mathbf{H}_0,$$

where the threshold level λ is chosen so that the fixed Type I error is achieved.

To indicate another use of the Girsanov-Meyer Theorem let us consider stochastic differential equations. Since stochastic differential equations[6] are treated systematically in Chap. V we are free here to restrict our attention to a simple but illustrative situation. Let W be a standard Brownian motion on a space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses. Let $f_i(\omega, s, x)$ be functions satisfying $(i = 1, 2)$:

[6] Stochastic "differential" equations have meaning only if they are interpreted as stochastic integral equations.

(i) $|f_i(\omega, s, x) - f_i(\omega, s, y)| \leq K|x - y|$ for fixed (ω, s);
(ii) $f_i(\cdot, s, x) \in \mathcal{F}_s$ for fixed (s, x);
(iii) $f_i(\omega, \cdot, x)$ is left continuous with right limits for fixed (ω, x).

By a Picard-type iteration procedure one can show there exists a unique solution (with continuous paths) of

$$X_t = X_0 + \int_0^t f_1(\cdot, s, X_s)dW_s + \int_0^t f_2(\cdot, s, X_s)ds. \tag{$***$}$$

The Girsanov-Meyer Theorem allows us to establish the existence of solutions of analogous equations where the Lipschitz hypothesis on the "drift coefficient" f_2 is removed. Indeed if X is the solution of $(***)$, let γ be any bounded, measurable function such that $\gamma(\omega, s, X_s) \in \mathbb{L}$. Define

$$g(\omega, s, x) = f_2(\omega, s, x) + f_1(\omega, s, x)\gamma(s, \omega, x).$$

We will see that we can find a solution of

$$Y_t = Y_0 + \int_0^t f_1(\cdot, s, Y_s)dB_s + \int_0^t g(\cdot, s, Y_s)ds$$

provided we choose a new Brownian motion B appropriately.

We define a new probability law Q by

$$\frac{dQ}{dP} = \exp\left(\int_0^T \gamma(s, X_s)dW_s - \frac{1}{2}\int_0^T \gamma(s, X_s)^2 ds\right).$$

By Theorem 42 we have that

$$B_t = W_t - \int_0^t \gamma(s, X_s)ds$$

is a standard Brownian motion under Q. We then have that the solution X of $(***)$ also satisfies

$$\begin{aligned} X_t &= X_0 + \int_0^t f_1(\cdot, s, X_s)dB_s + \int_0^t (f_2 + f_1\gamma)(\cdot, s, X_s)ds \\ &= X_0 + \int_0^t f_1(\cdot, s, X_s)dB_s + \int_0^t g(\cdot, s, X_s)ds, \end{aligned}$$

which is a solution of a stochastic differential equation driven by a Brownian motion, under the law Q.

9 The Bichteler-Dellacherie Theorem

In Sect. 7 we saw that a classical semimartingale is a semimartingale. In this section we will show the converse.

Theorem 43 (Bichteler-Dellacherie Theorem). *An adapted, càdlàg process X is a semimartingale if and only if it is a classical semimartingale. That is, X is a semimartingale if and only if it can be written $X = M + A$, where M is a local martingale and A is an FV process.*

Proof. The sufficiency is exactly Theorem 26. We therefore establish here only the necessity.

Since X is càdlàg, the process $J_t = \sum_{0<s\leq t} \Delta X_s 1_{\{|\Delta X_s| \geq 1\}}$ has paths of finite variation on compacts; hence it is an FV process. Let $Y = X - J$. Then Y has bounded jumps. If we show that X is a classical semimartingale on $[0, t_0]$, some t_0, then $Y = X - J$ is one as well; moreover Y is special by Theorem 31. Let $Y = Y_0 + \overline{M} + \overline{A}$ be the canonical decomposition of Y. Suppose $t_1 > t_0$ and X is also shown to be a classical semimartingale on $[0, t_1]$. Then let $Y = Y_0 + \overline{N} + \overline{B}$ be the canonical decomposition of Y on $[0, t_1]$. We have that, for $t \leq t_0$,

$$\overline{M} - \overline{N} = \overline{B} - \overline{A}$$

by subtraction. Thus $\overline{B} - \overline{A}$ is a predictable, FV process which is a local martingale; hence by Theorem 12 we have $\overline{B} = \overline{A}$. Thus if we take t_n increasing to ∞ and show that X is a classical semimartingale on $[0, t_n]$, each n, we have that X is a classical semimartingale on $[0, \infty)$.

We now choose $u_0 > 0$, and by the above it suffices to show X is a classical semimartingale on $[0, u_0]$. Thus it is no loss to assume X is a total semimartingale on $[0, u_0]$. We will show that X is a quasimartingale, under an equivalent probability Q. Rao's Theorem (Theorem 15) shows that X is a classical semimartingale under Q, and the Girsanov-Meyer Theorem (Theorem 35) then shows that X is a classical semimartingale under P.

Let us take $H \in \mathbf{S}$ of the special form:

$$H_t = \sum_{i=0}^{n-1} H_i 1_{(T_i, T_{i+1}]} \tag{$*$}$$

where $0 = T_0 \leq T_1 \leq \cdots \leq T_{n-1} < T_n = u_0$. In this case the mapping I_X is given by

$$I_X(H) = (H \cdot X)_{u_0} = H_0(X_{T_1} - X_0) + \cdots + H_{n-1}(X_{u_0} - X_{T_{n-1}}).$$

The mapping $I_X : \mathbf{S}_u \to L^0$ is continuous, where L^0 is endowed with the topology of convergence in probability, by the hypothesis that X is a total semimartingale. Let

$$\mathcal{B} = \{H \in \mathbf{S} : H \text{ has a representation } (*) \text{ and } |H| \leq 1\}.$$

Let $\beta = I_X(\mathcal{B})$, the image of $\mathcal{B}$ under I_X. It will now suffice to find a probability Q equivalent to P such that $X_t \in L^1(dQ)$, $0 \leq t \leq u_0$ and such that $\sup_{U \in \beta} E_Q(U) = c < \infty$. The reason this suffices is that if we take, for a given $0 = t_0 < t_1 < \cdots < t_n = u_0$, the random variables

$H_0 = \text{sign}(E_Q\{X_{t_1} - X_0|\mathcal{F}_0\})$, $H_1 = \text{sign}(E_Q\{X_{t_2} - X_{t_1}|\mathcal{F}_{t_1}\}), \ldots$, we have that for this $H \in \mathcal{B}$,

$$E_Q(I_X(H)) = E_Q\{|E_Q\{X_{t_0} - X_{t_1}|\mathcal{F}_0\}| + \cdots + |E_Q\{X_{t_{n-1}} - X_{t_n}|\mathcal{F}_{t_{n-1}}\}|\}.$$

Since this partition τ was arbitrary, we have $\text{Var}(X) = \sup_\tau \text{Var}_\tau(X) \leq \sup_{U\in\beta} E_Q(U) = c < \infty$, and so X is a Q quasimartingale.

Lemma 1. $\lim_{c\to\infty} \sup_{Y\in\beta} P(|Y| > c) = 0$.

Proof of Lemma 1. Suppose $\lim_{c\to\infty} \sup_{Y\in\beta} P(|Y| > c) > 0$. Then there exists a sequence c_n tending to ∞, $Y_n \in \beta$, and $a > 0$ such that $P(|Y_n| > c_n) \geq a$, all n. This is equivalent to

$$P\left(\frac{|Y_n|}{c_n} > 1\right) \geq a > 0. \tag{$*$}$$

Since $Y_n \in \beta$, there exists $H^n \in \mathcal{B}$ such that $I_X(H^n) = Y_n$. Then $I_X(\frac{1}{c_n}H^n) = \frac{1}{c_n}I_X(H^n) = \frac{1}{c_n}Y_n \in \beta$, if $c_n \geq 1$. But $\frac{1}{c_n}H^n$ tends to 0 uniformly a.s. which implies that $I_X(\frac{1}{c_n}H^n) = \frac{1}{c_n}Y_n$ tends to 0 in probability. This contradicts $(*)$. □

Lemma 2. There exists a law Q equivalent to P such that $X_t \in L^1(dQ)$, $0 \leq t \leq u_0$.

Proof of Lemma 2. Let $Y = \sup_{0\leq t\leq u_0} |X_t|$. Since X has càdlàg paths, $Y < \infty$ a.s. Moreover if D is a countable dense subset of $[0, u_0]$, then $Y = \sup_{t\in D} |X_t|$. Hence Y is a random variable. Let $A_m = \{m \leq Y < m+1\}$, and set $Z = \sum_{m=0}^{\infty} 2^{-m} 1_{A_m}$. Then Z is bounded, strictly positive, and $YZ \in L^1(dP)$. Define Q by setting $\frac{dQ}{dP} = \frac{1}{E_P(Z)}Z$. Then $E_Q\{|X_t|\} \leq E_Q\{Y\} = E_P\{YZ\}/E_P\{Z\} < \infty$. Hence $E_Q\{|X_t|\} < \infty$, $0 \leq t \leq u_0$. □

Observe that $\beta \subset L^1(dQ)$ for the law Q constructed in Lemma 2. Lemma 3 below implies that X is an R quasimartingale for $R \sim Q \sim P$. Hence by Rao's Theorem (Theorem 15) it is a classical R semimartingale, and by the Girsanov-Meyer Theorem (Theorem 35) it is a classical semimartingale for the equivalent law P as well. Thus Lemma 3 below will complete the proof of Theorem 43. We follow Yan [233].

Lemma 3. Let β be a convex subset of $L^1(dQ)$, $0 \in \beta$, that is bounded in probability. That is, for any $\varepsilon > 0$ there exists a $c > 0$ such that $Q(\zeta > c) \leq \varepsilon$, for any $\zeta \in \beta$. Then there exists a probability R equivalent to Q, with a bounded density, such that $\sup_{U\in\beta} E_R(U) < \infty$.

Proof of Lemma 3 and end of proof of Theorem 43. To begin, note that the hypotheses imply that $\beta \subset L^1(dR)$. What we must show is that there exists a bounded random variable Z, such that $P(Z > 0) = 1$, and such that $\sup_{\zeta\in\beta} E_Q(Z\zeta) < \infty$.

Let $A \in \mathcal{F}$ such that $Q(A) > 0$. Then there exists a constant d such that $Q(\zeta > d) \le Q(A)/2$, for all $\zeta \in \beta$, by assumption. Using this constant d, let $c = 2d$, and we have that $0 \le c1_A \notin \beta$, and moreover if B_+ denotes all bounded, positive r.v., then $c1_A$ is not in the $L^1(dQ)$ closure of $\beta - B_+$, denoted $\overline{\beta - B_+}$. That is, $c1_A \notin \overline{\beta - B_+}$. Since the dual of L^1 is L^∞, and $\beta - B_+$ is convex, by a version of the Hahn-Banach Theorem (see, e.g., Treves [223, page 190]) there exists a bounded random variable Y such that

$$\sup_{\zeta \in \beta, \eta \in B_+} E_Q\{Y(\zeta - \eta)\} < cE_Q\{Y1_A\}. \tag{$*$}$$

Replacing η by $a1_{\{Y<0\}}$ and letting a tend to ∞ shows that $Y \ge 0$ a.s., since otherwise the expectation on the left side above would get arbitrarily large. Next suppose $\eta = 0$. Then the inequality above gives

$$\sup_{\zeta \in \beta} E_Q\{Y\zeta\} \le cE_Q\{Y1_A\} < +\infty.$$

Now set $\mathcal{H} = \{Y \in B_+ : \sup_{\zeta \in \beta} E_Q\{Y\zeta\} < \infty\}$. Since $0 \in B_+$, we know $\mathcal{H}$ is not empty. Let $\mathcal{A} = \{$all sets of the form $\{Z = 0\}, Z \in \mathcal{H}\}$. We wish to show that there exists a $Z \in \mathcal{H}$ such that $Q(Z = 0) = \inf_{A \in \mathcal{A}} Q(A)$. Suppose, then, that Z_n is a sequence of elements of $\mathcal{H}$. Let $c_n = \sup_{\zeta \in \beta} E\{Z_n\zeta\}$ and $d_n = \|Z_n\|_{L^\infty}$. (Since $0 \in \beta$, we have $c_n \ge 0$). Choose b_n such that $\sum b_n c_n < \infty$ and $\sum b_n d_n < \infty$, and set $Z = \sum b_n Z_n$. Then clearly $Z \in \mathcal{H}$. Moreover, $\{Z = 0\} = \bigcap_n \{Z_n = 0\}$. Thus $\mathcal{A}$ is stable under countable intersections, and so there exists a Z such that $Q(Z = 0) = \inf_{A \in \mathcal{A}} Q(A)$.

We now wish to show $Z > 0$ a.s. Suppose not. That is, suppose $Q(Z = 0) > 0$. Let Y satisfy ($*$) (we have seen that there exists such a Y and that it hence is in $\mathcal{H}$). Further we take for our set A in ($*$) the set $A = \{Z = 0\}$, for which we are assuming $Q(A) > 0$. Since $0 \in \beta$ and $0 \in B_+$, we have from Lemma 2 that

$$0 < E\{Y1_A\} = E\{Y1_{\{Z=0\}}\}.$$

Since each of Y and Z are in $\mathcal{H}$, their sum is in $\mathcal{H}$ as well. But then the above implies

$$Q\{Y + Z = 0\} = Q\{Z = 0\} - Q(\{Z = 0\} \cap \{Y > 0\}) < Q(Z = 0).$$

This, then, is a contradiction, since $Q(Z = 0)$ is minimal for $Z \in \mathcal{H}$. Therefore we conclude $Z > 0$ a.s., and since $Z \in B_+$, it is bounded as well, and Lemma 3 is proved; thus also, Theorem 43 is proved. □

We state again, for emphasis, that Theorems 26 and 43 together allow us to conclude that semimartingales (as we have defined them) and classical semimartingales are the same.

Bibliographic Notes

The material of Chap. III comprises a large part of the core of the "general theory of processes" as presented, for example in Dellacherie [41], or alternatively Dellacherie and Meyer [45, 46, 44]. We have tried once again to keep the proofs non-technical, but instead of relying on the concept of a natural process as we did in the first edition, we have used the approach of R. Bass [11], which uses the P. A. Meyer classification of stopping times and Doob's quadratic inequality to prove the Doob-Meyer Theorem (Theorem 13), the key result of the whole theory.

The Doob decomposition is from Doob [55], and the Doob-Meyer decomposition (Theorem 13) is originally due to Meyer [163, 164]. The theory of quasimartingales was developed by Fisk [73], Orey [187], K. M. Rao [207], Stricker [218], and Métivier-Pellaumail [159].

The treatment of compensators is new to this edition. The simple example of the compensator of a process with one jump of size one (Theorem 17), dates back to 1970 with the now classic paper of Dellacherie [40]. The case of many jumps (Theorem 19) is due to C. S. Chou and P. A. Meyer [30], and can be found in many texts on point processes, such as [24] or [139]. The example of hazard rates comes from Fleming and Harrington [74].

The Fundamental Theorem of Local Martingales is due to J. A. Yan and appears in an article of Meyer [172]; it was also proved independently by Doléans-Dade [51]. Le Jan's Theorem is from [112].

The notion of special semimartingales and canonical decompositions is due to Meyer [171]; see also Yoeurp [234]. The Girsanov-Meyer theorems (Theorems 35 and 36) trace their origin to the 1954 work of Maruyama [152], followed by Girsanov [83], who considered the Brownian case only. The two versions presented here are due to Meyer [171], and the cases where the two measures are not equivalent (Theorems 37 and 38) are due to Lenglart [142]. The example from finance theory (starting on page 137) is inspired by [203]. Kazamaki's criterion was published originally in 1977 [124]; see also [125], whereas Novikov's condition dates to 1972 [184].

The Bichteler-Dellacherie Theorem (Theorem 43) is due independently to Bichteler [13, 14] and Dellacherie [42]. It was proved in the late 1970's, but the first time it appeared in print was when J. Jacod included it in his 1979 tome [103]. Many people have made contributions to this theorem, which had at least some of its origins in the works of Métivier-Pellaumail, Mokobodzki, Nikishin, Letta, and Lenglart. Our treatment was inspired by Meyer [176] and by Yan [233].

Exercises for Chapter III

Exercise 1. Show that the maximum and the minimum of a finite number of predictable stopping times is still a predictable stopping time.

Exercise 2. Let S be a totally inaccessible stopping time, and let $T = S+1$. Show that T is a predictable stopping time.

Exercise 3. Let S, T be predictable stopping times. Let $\Lambda = \{S = T\}$. Show that S_Λ is a predictable stopping time.

Exercise 4. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses. Show that the predictable σ-algebra (on $\mathbb{R}_+ \times \Omega$) is contained in the optional σ-algebra. (*Hint:* Show that a càdlàg, adapted process can be approximated by processes in $\mathbb{L}$.)

Exercise 5. Let S, T be stopping times with $S \leq T$. Show that $(S,T] = \{(t,\omega) : S(\omega) \leq t \leq T(\omega)\}$ is a predictable set. Show further that $[S,T)$ and (S,T) are optional sets. Last, show that if T is predictable, then (S,T) is a predictable set, and if both S and T are predictable, then $[S,T)$ is a predictable set.

Exercise 6. Let T be a predictable stopping time, and let $(S_n)_{n\geq 1}$ be a sequence of stopping times announcing T. Show that $\mathcal{F}_{T-} = \bigvee_n \mathcal{F}_{S_n}$.

Exercise 7. A filtration $\mathbb{F}$ is called **quasi left continuous** if for every *predictable* stopping time T one has $\mathcal{F}_T = \mathcal{F}_{T-}$.[7] Show that if $\mathbb{F}$ is a quasi left continuous filtration, then whenever one has a non-decreasing sequence of stopping times $(S_n)_{n\geq 1}$, with $\lim_{n\to\infty} S_n = T$, it follows that $\mathcal{F}_T = \bigvee_n \mathcal{F}_{S_n}$. (*Note:* One can have a quasi left continuous filtration such that there exists a stopping time T with $\mathcal{F}_{T-} \neq \mathcal{F}_T$.)

Exercise 8. Show that the natural completed filtration of a Lévy process is a quasi left continuous filtration.

***Exercise 9.** Let X be a semimartingale with $E\{[X,X]_\infty\} < \infty$ and suppose X has a decomposition $X = X_0 + M + A$, with A predictable. Show that $E\{[A,A]_\infty\} \leq E\{[X,X]_\infty\}$. (*Hint*: Recall that if T is a predictable stopping time, then

$$\Delta A_T 1_{\{T<\infty\}} = E\{\Delta X_T 1_{\{T<\infty\}} | \mathcal{F}_{T-}\},$$

and also that the set $\{(t,\omega) : \Delta A_t(\omega) \neq 0\}$ is a countable disjoint union of sets of the form $\{(t,\omega) : T_i(\omega) = t\}$, where the stopping times $(T_i)_{i\geq 1}$ are predictable.)

Exercise 10. Let X and Y be semimartingales with the processes $[X,X]$ and $[Y,Y]$ locally integrable. Prove the following "sharp bracket" version of the Kunita-Watanabe inequality, for jointly measurable processes H, K:

$$\int_0^t |H_s K_s| |d\langle X,Y\rangle_s| \leq \Big(\int_0^t H_s^2 d\langle X,X\rangle_s\Big)^{\frac{1}{2}} \Big(\int_0^t K_s^2 d\langle Y,Y\rangle_s\Big)^{\frac{1}{2}} \quad \text{a.s.}$$

(*Warning*: Recall that $\langle X,Y\rangle$ can exist as the compensator of $[X,Y]$ even when $\langle X,X\rangle$ and $\langle Y,Y\rangle$ do not exist.)

---

[7] A quasi left continuous filtration is formally defined in Chap. IV on page 189.

Exercise 11. Show that if A is a predictable finite variation process, then it is of locally integrable variation. That is, show that there exists a sequence of stopping times $(T_n)_{n\geq 1}$ such that $E\{\int_0^{T_n} |dA_s|\} < \infty$.

Exercise 12. Let N be a counting process with its minimal completed filtration, with independent interarrival times. Show that the compensator A of N has absolutely continuous paths if and only if the cumulative distribution functions of the interarrival times are absolutely continuous.

Exercise 13. Let T be an exponential random variable and $N_t = 1_{\{t\geq T\}}$. Let $\mathcal{F}_0 = \sigma\{T\}$ and suppose the filtration $\mathbb{F}$ is constant: $\mathcal{F}_t = \mathcal{F}_0$ for all $t \geq 0$. Show that the compensator A of N is $A_t = 1_{\{t\geq T\}}$. (This illustrates the importance of the filtration when calculating the compensator.)

Exercise 14. Let N be a compound Poisson process, $N_t = \sum_{i\geq 1} U_i 1_{\{t\geq T_i\}}$ where the times T_i are the arrival times of a standard Poisson process with parameter λ and the U_i are i.i.d. and independent of the arrival times. Suppose $E\{|U_i|\} < \infty$ and $E\{U_i\} = \mu$. Show that the compensator A of N for the natural filtration is given by $A_t = \lambda\mu t$.

Exercise 15. Show that if T is exponential with parameter λ, its hazard rate is constant and equal to λ. Show also that if R is Weibull with parameters (α, β), then its hazard rate is $\lambda(t) = \alpha\beta^\alpha t^{\alpha-1}$.

Exercise 16. Let T be exponential with parameter λ and have joint distribution with U given by $P(T > t, U > s) = \exp\{-\lambda t - \mu s - \theta ts\}$ for $t \geq 0$, $s \geq 0$, where λ, μ, and θ are all positive constants and also $\theta \leq \lambda\mu$. Show that the crude hazard rate of (T, U) is given by $\lambda^{\#}(t) = \lambda + \theta t$.

***Exercise 17.** Let M be a martingale on a filtered space where the filtration is quasi left continuous. Show that $\langle M, M\rangle$ is continuous. (*Hint:* See the discussion on quasi left continuous filtrations on page 189 of Chap. IV.)

Exercise 18. Let X be a semimartingale such that the process $D_t = \sup_{s\leq t} |\Delta X_s|$ is locally integrable. Show that every decomposition of X, $X = X_0 + M + C$, has that the process C is of locally integrable variation.

Exercise 19. Suppose that X is **locally special**, that is, there exists a sequence $(T_n)_{n\geq 1}$ of stopping times increasing to ∞ a.s. such that $X^{T_n} = M^n + A^n$, with A^n a predictable finite variation process, for each n. Show that X is special. (That is, *a semimartingale which is locally special is also special.*)

Exercise 20. Prove the Lemma on page 132. Let $Q \sim P$, and $Z_t = E_P\left\{\frac{dQ}{dP}|\mathcal{F}_t\right\}$. Show that an adapted, càdlàg process M is a Q local martingale if and only if MZ is a P local martingale.

***Exercise 21.** Let $\mathbb{F} \subset \mathbb{G}$ be filtrations satisfying the usual hypotheses, and let X be a $\mathbb{G}$ quasimartingale. Suppose that $Y_t = E\{X_t|\mathcal{F}_t\}$, and that Y can be taken càdlàg (this can be proved to be true, although it is a little hard to prove). Show that Y is also a quasimartingale for the filtration $\mathbb{F}$. (*Hint*: Use Rao's Theorem.)

***Exercise 22.** Suppose that A is a predictable finite variation process with $E\{[A,A]_t^{\frac{1}{2}}\} < \infty$, and that M is a bounded martingale. Show that $[A,M]$ is a uniformly integrable martingale.

Exercise 23. Let T be a strictly positive random variable, and let $\mathcal{F}_t = \sigma\{T \wedge s; s \leq t\}$. Show that $(\mathcal{F}_t)_{t\geq 0}$ is the smallest filtration making T a stopping time.

Exercise 24. Let Z be a càdlàg supermartingale of Class D with $Z_0 = 0$ and suppose for all predictable stopping times T one has $E\{\Delta Z_T|\mathcal{F}_{T-}\} = 0$, a.s. Show that if $Z = M - A$ is the unique Doob-Meyer decomposition of Z, then A has continuous paths almost surely.

Exercise 25. Let A be an increasing process of integrable variation, and let T be a predictable jump time of A such that $E\{\Delta A_T|\mathcal{F}_{T-}\} = 0$. Then its compensator $\widetilde{A}$ is continuous at T. (This exercise complements Theorem 21.)

***Exercise 26.** A supermartingale Z is said to be **regular** if whenever a sequence of stopping times $(T_n)_{n\geq 1}$ increases to T, then $\lim_{n\to\infty} E\{Z_{T_n}\} = E\{Z_T\}$. Let Z be a càdlàg supermartingale of Class D with Doob-Meyer decomposition $Z = M - A$. Show that A is continuous if and only if Z is regular.

***Exercise 27 (approximation of the compensator by Laplacians).** Let Z be a càdlàg positive supermartingale of Class D with $\lim_{t\to\infty} E\{Z_t\} = 0$. (Such a supermartingale is called a **potential**.) Let $Z = M - A$ be its Doob-Meyer decomposition and assume further that A is continuous. Define

$$A_t^h = \frac{1}{h}\int_0^t (Z_s - E\{Z_{s+h}|\mathcal{F}_s\})ds$$

Show that for any stopping time T, $\lim_{h\to 0} A_T^h = A_T$ with convergence in L^1.

***Exercise 28.** Let Z be a càdlàg positive supermartingale of Class D with $\lim_{t\to\infty} E\{Z_t\} = 0$. Let $Z = M - A$ be its Doob-Meyer decomposition. Let A^h be as given in Exercise 27. Show that for any stopping time T, $\lim_{h\to 0} A_T^h = A_T$, but in this case the convergence is weak for L^1; that is, the convergence is in the topology $\sigma(L^1, L^\infty)$.[8]

Exercise 29 (discrete Laplacian approximations). Let Z be a càdlàg positive supermartingale of Class D with $\lim_{t\to\infty} E\{Z_t\} = 0$. Let $Z = M - A$ be its Doob-Meyer decomposition and assume further that A is continuous. Define

---
[8] X^n converges to X in $\sigma(L^1, L^\infty)$ if X^n, X are in L^1 and for any a.s. bounded random variable Y, $E(X^nY) \to E(XY)$.

$$A_\infty^n = \sum_{k=0}^{\infty} E\{Z_{\frac{k}{2^n}} - Z_{\frac{k+1}{2^n}} | \mathcal{F}_{\frac{k}{2^n}}\}$$

Show that $\lim_{n\to\infty} A_\infty^n = A_\infty$ with convergence in L^1.

Exercise 30. Use Meyer's Theorem (Theorem 4) to show that if X is a strong (Markov) Feller process for its natural completed filtration $\mathbb{F}^\mu$, and if X has continuous paths, then the filtration $\mathbb{F}^\mu$ has no totally inaccessible stopping times. (This implies that the natural filtration of Brownian motion does not have any totally inaccessible stopping times.)

***Exercise 31.** Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be the standard Brownian space. Show that the optional σ-algebra and the predictable σ-algebra coincide. (*Hint:* Use Meyer's Theorem (Theorem 4) and Exercise 30.)

***Exercise 32.** Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses. Let X be a (not necessarily adapted) càdlàg stochastic process such that for $\lambda > 0$, $E\{\int_0^\infty e^{-\lambda t}|X_t|dt\} < \infty$. Let $R_\lambda(X_t) = E\{\int_0^\infty e^{-\lambda s} X_{t+s} ds | \mathcal{F}_t\}$, the right continuous version. Show that

$$M_\lambda(t) = R_\lambda(X_t) - R_\lambda(X_0) + \int_0^t X_s - \lambda R_\lambda(X_s) ds$$

is an $\mathbb{F}$ martingale.

***Exercise 33 (Knight's compensator calculation method).** Let X be a càdlàg semimartingale. In the framework of Exercise 32 suppose the limits below exist both pathwise a.s. and are in L^1, and are of finite variation in finite time intervals:

$$A_t = \lim_{\lambda\to\infty} \lambda \int_0^t (X_u - \lambda R_\lambda X_u) du, \quad 0 < t.$$

Show that X is a special semimartingale, and A is the predictable term in its semimartingale decomposition.

IV

General Stochastic Integration and Local Times

1 Introduction

We defined a semimartingale as a "good integrator" in Chap. II, and this led naturally to defining the stochastic integral as a limit of sums. To express an integral as a limit of sums requires some path smoothness of the integrands and we limited our attention to processes in $\mathbb{L}$, the space of adapted processes with paths that are left continuous and have right limits. The space $\mathbb{L}$ is sufficient to prove Itô's formula, the Girsanov-Meyer Theorem, and it also suffices in some applications such as stochastic differential equations. But other uses, such as martingale representation theory or local times, require a larger space of integrands.

In this chapter we define stochastic integration for predictable processes. Our extension from Chap. II is very roughly analogous to how the Lebesgue integral extends the Riemann integral. We first define stochastic integration for bounded, predictable processes and a subclass of semimartingales known as $\mathcal{H}^2$. We then extend the definition to arbitrary semimartingales and to locally bounded predictable integrands.

We also treat the issue of when a stochastic integral with respect to a martingale or a local martingale is still a local martingale, which is not always the case. In this respect we treat the subject of sigma martingales, which has recently been shown to be important for the theory of mathematical finance.

2 Stochastic Integration for Predictable Integrands

In this section, we will weaken the restriction that an integrand H must be in $\mathbb{L}$. We will show our definition of stochastic integrals can be extended to a class of *predictably measurable* integrands.

Throughout this section X will denote a semimartingale such that $X_0 = 0$. This is a convenience involving no loss of generality. If Y is any semimartingale we can set $\hat{Y}_t = Y_t - Y_0$, and if we have defined stochastic integrals for semimartingales that are zero at 0, we can next define

$$\int_0^t H_s dY_s \equiv \int_0^t H_s d\hat{Y}_s + H_0 Y_0.$$

When $Y_0 \neq 0$, recall that we write $\int_{0+}^t H_s dY_s$ to denote integration on $(0, t]$, and $\int_0^t H_s dY_s$ denotes integration on the closed interval $[0, t]$.

We recall for convenience the definition of the predictable σ-algebra, already defined in Chap. III.

Definition. The **predictable σ-algebra** $\mathcal{P}$ on $\mathbb{R}_+ \times \Omega$ is the smallest σ-algebra making all processes in $\mathbb{L}$ measurable. That is, $\mathcal{P} = \sigma\{H : H \in \mathbb{L}\}$. We let $\mathbf{b}\mathcal{P}$ denote bounded processes that are $\mathcal{P}$ measurable.

Let $X = M + A$ be a decomposition of a semimartingale X, with $X_0 = M_0 = A_0 = 0$. Here M is a local martingale and A is an FV process (such a decomposition exists by the Bichteler-Dellacherie Theorem (Theorem 43 of Chap. III)). We will first consider special semimartingales. Recall that a semimartingale X is called *special* if it has a decomposition

$$X = \overline{N} + \overline{A}$$

where $\overline{N}$ is a local martingale and $\overline{A}$ is a predictable FV process. This decomposition is unique by Theorem 30 in Chap. III and it is called the *canonical decomposition*.

Definition. Let X be a special semimartingale with canonical decomposition $X = \overline{N} + \overline{A}$. The $\mathcal{H}^2$ **norm** of X is defined to be

$$\|X\|_{\mathcal{H}^2} = \|[\overline{N}, \overline{N}]_\infty^{1/2}\|_{L^2} + \|\int_0^\infty |d\overline{A}_s|\|_{L^2}.$$

The **space of semimartingales** $\mathcal{H}^2$ consists of all special semimartingales with finite $\mathcal{H}^2$ norm.

In Chap. V we define an equivalent norm which we denote $\|\cdot\|_{\underline{\underline{H}}^2}$.

Theorem 1. *The space of $\mathcal{H}^2$ semimartingales is a Banach space.*

Proof. The space is clearly a normed linear space and it is easy to check that $\|\cdot\|_{\mathcal{H}^2}$ is a norm (recall that $E\{\overline{N}_\infty^2\} = E\{[\overline{N}, \overline{N}]_\infty\}$, and therefore $\|X\|_{\mathcal{H}^2} = 0$ implies that $E\{\overline{N}_\infty^2\} = 0$ which implies, since $\overline{N}$ is a martingale, that $\overline{N} \equiv 0$).

To show completeness we treat the terms $\overline{N}$ and $\overline{A}$ separately. Consider first $\overline{N}$. Since $E\{\overline{N}_\infty^2\} = \|[\overline{N}, \overline{N}]_\infty^{1/2}\|_{L^2}^2$, it suffices to show that the space of L^2 martingales is complete. However an L^2 martingale M can be identified with $M_\infty \in L^2$, and thus the space is complete since L^2 is complete.

Next suppose (A^n) is a Cauchy sequence of predictable FV processes in $\|\cdot\|_2$ where $\|A\|_p = \|\int_0^\infty |dA_s|\|_{L^p}$, $p \geq 1$. To show (A^n) converges it suffices to show a subsequence converges. Therefore without loss of generality we can assume $\sum_n \|A^n\|_2 < \infty$.

Then $\sum A^n$ converges in $\|\cdot\|_1$ to a limit A. Moreover

$$\lim_{m\to\infty} \sum_{n\geq m} \int_0^\infty |dA_s^n| = 0$$

in L^1 and is dominated in L^2 by $\sum_n \int_0^\infty |dA_s^n|$. Therefore $\sum A^n$ converges to the limit A in $\|\cdot\|_2$ as well, and there is a subsequence converging almost surely. To see that the limit A is predictable, note that since each term in the sequence $(A^n)_{n\geq 1}$ is predictable, the limit A is the limit of predictably measurable processes and hence also predictable. □

For convenience we recall here the definition of $\mathbb{L}$.

Definition. $\mathbb{L}$ (resp. $\mathbf{b}\mathbb{L}$) denotes the space of adapted processes with càglàd[1] (resp. bounded, càglàd) paths.

We first establish a useful technical lemma.

Lemma. Let A be a predictable FV process, and let H be in $\mathbb{L}$ such that $E\{\int_0^\infty |H_s||dA_s|\} < \infty$. Then the FV process $(\int_0^t H_s dA_s)_{t\geq 0}$ is also predictable.

Proof. We need only to write the integral $\int_0^t H_s dA_s$ as the limit of Riemann sums, each one of which is predictable, and which converge in *ucp* to $\int_0^t H_s dA_s$, showing that it too is predictable. □

The results that follow will enable us to extend the class of stochastic integrands from $\mathbf{b}\mathbb{L}$ to $\mathbf{b}\mathcal{P}$, with $X \in \mathcal{H}^2$ (and $X_0 = 0$). First we observe that if $H \in \mathbf{b}\mathbb{L}$ and $X \in \mathcal{H}^2$, then the stochastic integral $H \cdot X \in \mathcal{H}^2$. Also if $X = \overline{N} + \overline{A}$ is the canonical decomposition of X, then $H \cdot \overline{N} + H \cdot \overline{A}$ is the canonical decomposition of $H \cdot X$ by the preceding lemma. Moreover,

$$\|H \cdot X\|_{\mathcal{H}^2} = \|(\int_0^\infty H_s^2 d[\overline{N},\overline{N}]_s)^{\frac{1}{2}}\|_{L^2} + \|\int_0^\infty |H_s||d\overline{A}_s|\|_{L^2}.$$

The key idea in extending our integral is to notice that $[\overline{N},\overline{N}]$ and $\overline{A}$ are FV processes, and therefore ω-by-ω the integrals $\int_0^t H_s^2(\omega)d[\overline{N},\overline{N}]_s(\omega)$ and $\int_0^t |H_s||d\overline{A}_s|$ make sense for any $H \in \mathbf{b}\mathcal{P}$ and not just $H \in \mathbb{L}$.

Definition. Let $X \in \mathcal{H}^2$ with $X = \overline{N} + \overline{A}$ its canonical decomposition, and let H, $J \in \mathbf{b}\mathcal{P}$. We define $d_X(H, J)$ by

$$d_X(H,J) \equiv \|(\int_0^\infty (H_s - J_s)^2 d[\overline{N},\overline{N}]_s)^{1/2}\|_{L^2} + \|\int_0^\infty |H_s - J_s||d\overline{A}_s|\|_{L^2}.$$

[1] "càglàd" is the French acronym for *left continuous with right limits.*

Theorem 2. *For $X \in \mathcal{H}^2$ the space* $\mathbf{b}\mathbb{L}$ *is dense in* $\mathbf{b}\mathcal{P}$ *under* $d_X(\cdot,\cdot)$.

Proof. We use the Monotone Class Theorem. Define

$$\mathcal{A} = \{H \in \mathbf{b}\mathcal{P} : \text{ for any } \varepsilon > 0, \text{ there exists } J \in \mathbf{b}\mathbb{L} \text{ such that } d_X(H,J) < \varepsilon\}.$$

Trivially $\mathcal{A}$ contains $\mathbf{b}\mathbb{L}$. If $H^n \in \mathcal{A}$ and H^n increases to H with H bounded, then $H \in \mathbf{b}\mathcal{P}$, and by the Dominated Convergence Theorem if $\delta > 0$ then for some $N(\delta)$, $n > N(\delta)$ implies $d_X(H,H^n) < \delta$. Since each $H^n \in \mathcal{A}$, we choose $n_0 > N(\delta)$ and there exists $J \in \mathbf{b}\mathbb{L}$ such that $d_X(J,H^{n_0}) < \delta$. Therefore given $\varepsilon > 0$, by taking $\delta = \varepsilon/2$ we can find $J \in \mathbf{b}\mathbb{L}$ such that $d_X(J,H) < \varepsilon$, and therefore $H \in \mathcal{A}$. An application of the Monotone Class Theorem yields the result. □

Theorem 3. *Let $X \in \mathcal{H}^2$ and $H^n \in \mathbf{b}\mathbb{L}$ such that H^n is Cauchy under d_X. Then $H^n \cdot X$ is Cauchy in $\mathcal{H}^2$.*

Proof. Since $\|H^n \cdot X - H^m \cdot X\|_{\mathcal{H}^2} = d_X(H^n,H^m)$, the theorem is immediate. □

Theorem 4. *Let $X \in \mathcal{H}^2$ and $H \in \mathbf{b}\mathcal{P}$. Suppose $H^n \in \mathbf{b}\mathbb{L}$ and $J^m \in \mathbf{b}\mathbb{L}$ are two sequences such that $\lim_n d_X(H^n,H) = \lim_m d_X(J^m,H) = 0$. Then $H^n \cdot X$ and $J^m \cdot X$ tend to the same limit in $\mathcal{H}^2$.*

Proof. Let $Y = \lim_{n\to\infty} H^n \cdot X$ and $Z = \lim_{m\to\infty} J^m \cdot X$, where the limits are taken in $\mathcal{H}^2$. For $\varepsilon > 0$, by taking n and m large enough we have

$$\begin{aligned}
\|Y - Z\|_{\mathcal{H}^2} &\le \|Y - H^n \cdot X\|_{\mathcal{H}^2} + \|H^n \cdot X - J^m \cdot X\|_{\mathcal{H}^2} + \|J^m \cdot X - Z\|_{\mathcal{H}^2} \\
&\le 2\varepsilon + \|H^n \cdot X - J^m \cdot X\|_{\mathcal{H}^2} \\
&\le 2\varepsilon + d_X(H^n, J^m) \\
&\le 2\varepsilon + d_X(H^n, H) + d_X(H, J^m) \\
&\le 4\varepsilon,
\end{aligned}$$

and the result follows. □

We are now in a position to define the stochastic integral for $H \in \mathbf{b}\mathcal{P}$ (and $X \in \mathcal{H}^2$).

Definition. Let X be a semimartingale in $\mathcal{H}^2$ and let $H \in \mathbf{b}\mathcal{P}$. Let $H^n \in \mathbf{b}\mathbb{L}$ be such that $\lim_{n\to\infty} d_X(H^n,H) = 0$. The **stochastic integral** $H \cdot X$ is the (unique) semimartingale $Y \in \mathcal{H}^2$ such that $\lim_{n\to\infty} H^n \cdot X = Y$ in $\mathcal{H}^2$. We write $H \cdot X = (\int_0^t H_s dX_s)_{t\ge 0}$.

We have defined our stochastic integral for predictable integrands and semimartingales in $\mathcal{H}^2$ as limits of our (previously defined) stochastic integrals. In order to investigate the properties of this more general integral, we need to have approximations converging uniformly. The next theorem and its corollary give us this.

Theorem 5. *Let X be a semimartingale in $\mathcal{H}^2$. Then*

$$E\{(\sup_t |X_t|)^2\} \leq 8\|X\|_{\mathcal{H}^2}^2.$$

Proof. For a process H, let $H^* = \sup_t |H_t|$. Let $X = \overline{N} + \overline{A}$ be the canonical decomposition of X. Then

$$X^* \leq \overline{N}^* + \int_0^\infty |d\overline{A}_s|.$$

Doob's maximal quadratic inequality (Theorem 20 of Chap. I) yields

$$E\{(\overline{N}^*)^2\} \leq 4E\{\overline{N}_\infty^2\} = 4E\{[\overline{N},\overline{N}]_\infty\},$$

and using $(a+b)^2 \leq 2a^2 + 2b^2$ we have

$$\begin{aligned} E\{(X^*)^2\} &\leq 2E\{(\overline{N}^*)^2\} + 2E\{(\int_0^\infty |d\overline{A}_s|)^2\} \\ &\leq 8E\{[\overline{N},\overline{N}]_\infty\} + 2\|\int_0^\infty |d\overline{A}_s|\|_{L^2}^2 \\ &\leq 8\|X\|_{\mathcal{H}^2}^2. \end{aligned}$$ □

Corollary. Let (X^n) be a sequence of semimartingales converging to X in $\mathcal{H}^2$. Then there exists a subsequence (n_k) such that $\lim_{n_k\to\infty}(X^{n_k} - X)^* = 0$ a.s.

Proof. By Theorem 5 we know that $(X^n - X)^* = \sup_t |X_t^n - X_t|$ converges to 0 in L^2. Therefore there exists a subsequence converging a.s. □

We next investigate some of the properties of this generalized stochastic integral. Almost all of the properties established in Chap. II (Sect. 5) still hold.[2]

Theorem 6. *Let X, $Y \in \mathcal{H}^2$ and H, $K \in \mathrm{b}\mathcal{P}$. Then*

$$(H+K)\cdot X = H\cdot X + K\cdot X,$$

and

$$H\cdot(X+Y) = H\cdot X + H\cdot Y.$$

Proof. One need only check that it is possible to take a sequence $H^n \in \mathrm{b}\mathbb{L}$ that approximates H in both d_X and d_Y. □

Theorem 7. *Let T be a stopping time. Then $(H\cdot X)^T = H1_{[0,T]}\cdot X = H\cdot(X^T)$.*

[2] Indeed, it is an open question whether or not Theorem 16 of Chap. II extends to integrands in $\mathrm{b}\mathcal{P}$. See the discussion at the end of this section.

Proof. Note that $1_{[0,T]} \in \mathrm{b}\mathbb{L}$, so $H1_{[0,T]} \in \mathrm{b}\mathcal{P}$. Also, X^T is clearly still in $\mathcal{H}^2$. Since we know this result is true for $H \in \mathrm{b}\mathbb{L}$ (Theorem 12 of Chap. II), the result follows by uniform approximation, using the corollary of Theorem 5. □

Theorem 8. *The jump process* $(\Delta(H \cdot X)_s)_{s\geq 0}$ *is indistinguishable from* $(H_s(\Delta X_s))_{s\geq 0}$.

Proof. Recall that for a process J, $\Delta J_t = J_t - J_{t-}$, the jump of J at time t. (Note that $H \cdot X$ and X are càdlàg semimartingales, so Theorem 8 makes sense.) By Theorem 13 of Chap. II we know the result is true for $H \in \mathrm{b}\mathbb{L}$. Let $H \in \mathrm{b}\mathcal{P}$, and let $H^n \in \mathrm{b}\mathbb{L}$ such that $\lim_{n\to\infty} d_X(H^n, H) = 0$. By the corollary of Theorem 5, there exists a subsequence (n_k) such that

$$\lim_{n_k\to\infty} (H^{n_k} \cdot X - H \cdot X)^* = 0 \quad \text{a.s.}$$

This implies that, considered as processes,

$$\lim_{n_k\to\infty} \Delta(H^{n_k} \cdot X) = \Delta(H \cdot X),$$

outside of an evanescent set.[3] Since each $H^{n_k} \in \mathrm{b}\mathbb{L}$, we have $\Delta(H^{n_k} \cdot X) = H^{n_k}(\Delta X)$, outside of another evanescent set. Combining these, we have

$$\begin{aligned}\lim_{n_k\to\infty} H^{n_k}(\Delta X)1_{\{\Delta X\neq 0\}} &= \lim_{n_k\to\infty} \Delta(H^{n_k} \cdot X)1_{\{\Delta X\neq 0\}}\\ &= \Delta(H \cdot X)1_{\{\Delta X\neq 0\}},\end{aligned}$$

and therefore

$$\lim_{n_k\to\infty} H^{n_k}1_{\{\Delta X\neq 0\}} = \frac{\Delta(H\cdot X)}{\Delta X}1_{\{\Delta X\neq 0\}}.$$

In particular, the above implies that $\lim_{n_k\to\infty} H_t^{n_k}(\omega)$ exists for all (t,ω) in $\{\Delta X \neq 0\}$, a.s. We next form

$$\Lambda = \{\omega : \text{there exists } t>0 \text{ such that } \lim_{n_k\to\infty} H_t^{n_k}(\omega) \neq H_t(\omega) \text{ and } \Delta X_t(\omega)\neq 0\}.$$

Suppose $P(\Lambda) > 0$. Then

$$\begin{aligned}d_X(H^{n_k}, H) \geq {} & \|1_\Lambda\{\int_0^\infty (H_s^{n_k} - H_s)^2 d(\sum_{0<u\leq s} (\Delta\overline{N}_u)^2)\}\|_{L^2}\\ &+ \|1_\Lambda\{\int_0^\infty |H_s^{n_k} - H_s| d(\sum_{0<u\leq s} |\Delta\overline{A}_u|)\}\|_{L^2},\end{aligned} \tag{$*$}$$

and if $\Delta X_s \neq 0$, then $|\Delta\overline{N}_s| + |\Delta\overline{A}_s| > 0$. The left side of $(*)$ tends to 0 as $n_k \to \infty$, and the right side of $(*)$ does not. Therefore $P(\Lambda) = 0$, and we conclude $\Delta(H \cdot X) = \lim_{n_k\to\infty} H^{n_k}\Delta X = H\Delta X$. □

[3] A set $\Lambda \subset \mathbb{R}_+ \times \Omega$ is *evanescent* if 1_Λ is a process that is indistinguishable from the zero process.

Corollary. Let $X \in \mathcal{H}^2$, $H \in \mathbf{b}\mathcal{P}$, and T a finite stopping time. Then

$$H \cdot (X^{T-}) = (H \cdot X)^{T-}.$$

Proof. By Theorem 8, $(H \cdot X)^{T-} = (H \cdot X)^T - H_T \Delta X_T 1_{\{t \geq T\}}$. On the other hand, $X^{T-} = X^T - \Delta X_T 1_{\{t \geq T\}}$. Let $A_t = \Delta X_T 1_{\{t \geq T\}}$. By the bilinearity (Theorem 6), $H \cdot (X^{T-}) = H \cdot (X^T) - H \cdot A$. Since $H \cdot (X^T) = (H \cdot X)^T$ by Theorem 7, and $H \cdot A = H_T \Delta X_T 1_{\{t \geq T\}}$, the result follows. □

The next three theorems all involve the same simple proofs. The result is known to be true for processes in $\mathbf{b}\mathbb{L}$; let $(H^n) \in \mathbf{b}\mathbb{L}$ approximate $H \in \mathbf{b}\mathcal{P}$ in $d_X(\cdot, \cdot)$, and by the corollary of Theorem 5 let n_k be a subsequence such that

$$\lim_{n_k \to \infty} (H^{n_k} \cdot X - H \cdot X)^* = 0 \quad \text{a.s.}$$

Then use the uniform convergence to obtain the desired result. We state these theorems, therefore, without proofs.

Theorem 9. *Let $X \in \mathcal{H}^2$ have paths of finite variation on compacts, and $H \in \mathbf{b}\mathcal{P}$. Then $H \cdot X$ agrees with a path-by-path Lebesgue-Stieltjes integral.*

Theorem 10 (Associativity). *Let $X \in \mathcal{H}^2$ and $H, K \in \mathbf{b}\mathcal{P}$. Then $K \cdot X \in \mathcal{H}^2$ and $H \cdot (K \cdot X) = (HK) \cdot X$.*

Theorem 11. *Let $X \in \mathcal{H}^2$ be a (square integrable) martingale, and $H \in \mathbf{b}\mathcal{P}$. Then $H \cdot X$ is a square integrable martingale.*

Theorem 12. *Let $X, Y \in \mathcal{H}^2$ and $H, K \in \mathbf{b}\mathcal{P}$. Then*

$$[H \cdot X, K \cdot Y]_t = \int_0^t H_s K_s d[X, Y]_s, \quad (t \geq 0),$$

and in particular

$$[H \cdot X, H \cdot X]_t = \int_0^t H_s^2 d[X, X]_s, \quad (t \geq 0).$$

Proof. As in the proof of Theorem 29 of Chap. II, it suffices to show

$$[H \cdot X, Y]_t = \int_0^t H_s d[X, Y]_s.$$

Let $(H^n) \in \mathbf{b}\mathbb{L}$ such that $\lim_{n \to \infty} d_X(H^n, H) = 0$. Let $T^m = \inf\{t > 0 : |Y_t| > m\}$. Then (T^m) are stopping times increasing to ∞ a.s. and $|Y_-^{T^m}| \leq m$.[4] Since it suffices to show the result holds on $[0, T^m)$, each m, we can assume without loss of generality that Y_- is in $\mathbf{b}\mathbb{L}$. Moreover, the Dominated Convergence Theorem gives $\lim_{n \to \infty} d_X(H^n Y_-, HY_-) = 0$. By Theorem 29 of Chap. II, we have

[4] Recall that Y_- denotes the left continuous version of Y.

$$[H^n \cdot X, Y]_t = \int_0^t H_s^n d[X,Y]_s \qquad (t \geq 0),$$

and again by dominated convergence

$$\lim_{n\to\infty} [H^n \cdot X, Y] = \int_0^t H_s d[X,Y]_s \qquad (t \geq 0).$$

It remains only to show $\lim_{n\to\infty}[H^n \cdot X, Y] = [H \cdot X, Y]$. Let $Z^n = H^n \cdot X$, and let n_k be a subsequence such that $\lim_{n_k\to\infty}(Z^{n_k} - Z)^* = 0$ a.s., where $Z = H \cdot X$ (by the corollary to Theorem 5). Integration by parts yields

$$\begin{aligned}
[Z^{n_k}, Y] &= Z^{n_k} Y - (Y_-) \cdot Z^{n_k} - (Z_-^{n_k}) \cdot Y \\
&= Z^{n_k} Y - (Y_- H^{n_k}) \cdot X - (Z_-^{n_k}) \cdot Y,
\end{aligned}$$

where we have used associativity (Theorem 10). We take limits so that

$$\begin{aligned}
\lim_{n_k\to\infty} [Z^{n_k}, Y] &= ZY - Y_- \cdot (H \cdot X) - Z_- \cdot Y \\
&= ZY - Y_- \cdot (Z) - Z_- \cdot Y \\
&= [Z, Y] = [H \cdot X, Y].
\end{aligned}$$

□

At this point the reader may wonder how to calculate in practice a canonical decomposition of a semimartingale X in order to verify that $X \in \mathcal{H}^2$. Fortunately Theorem 13 will show that $\mathcal{H}^2$ is merely a mathematical convenience.

Lemma. Let A be an FV process with $A_0 = 0$ and $\int_0^\infty |dA_s| \in L^2$. Then $A \in \mathcal{H}^2$. Moreover $\|A\|_{\mathcal{H}^2} \leq 6 \| \int_0^\infty |dA_s| \|_{L^2}$.

Proof. If we can prove the result for A increasing then the general result will follow by decomposing $A = A^+ - A^-$. Therefore we assume without loss of generality that A is increasing. Hence as we noted in Sect. 5 of Chap. III, the compensator $\widetilde{A}$ of A is also increasing and $E\{\widetilde{A}_\infty\} = E\{A_\infty\} < \infty$.

Let M be a martingale bounded by a constant k. Since $A - \widetilde{A}$ is a local martingale, Corollary 2 to Theorem 27 of Chap. II shows that

$$L = M(A - \widetilde{A}) - [M, A - \widetilde{A}]$$

is a local martingale. Moreover

$$\begin{aligned}
\sup_t |L_t| &\leq k(A_\infty + \widetilde{A}_\infty) + 2k \sum_s |\Delta(A - \widetilde{A})_s| \\
&\leq 3k(A_\infty + \widetilde{A}_\infty) \in L^1.
\end{aligned}$$

Therefore L is a uniformly integrable martingale (Theorem 51 of Chap. I) and $E\{L_\infty\} = E\{L_0\} = 0$. Hence

$$\begin{aligned}E\{M_\infty(A-\widetilde{A})_\infty\} &= E\{[M,A-\widetilde{A}]_\infty\}\\ &= E\{[M,A]_\infty\} - E\{[M,\widetilde{A}]_\infty\}\\ &= E\{[M,A]_\infty\},\end{aligned}$$

because $\widetilde{A}$ is natural. By the Kunita-Watanabe inequality (the corollary to Theorem 25 of Chap. II)

$$\begin{aligned}E\{|[M,A]_\infty|\} &\le (E\{[M,M]_\infty\}E\{[A,A]_\infty\})^{1/2}\\ &\le \frac{1}{2}E\{[M,M]_\infty\} + \frac{1}{2}E\{[A,A]_\infty\},\end{aligned}$$

where the second inequality uses $2ab \le a^2 + b^2$. However

$$E\{[M,M]_\infty\} = E\{M_\infty^2\}$$

(Corollary 4 of Theorem 27 of Chap. II) and also $[A,A]_\infty \le A_\infty^2$ a.s. Therefore

$$E\{M_\infty(A-\widetilde{A})_\infty\} \le \frac{1}{2}E\{M_\infty^2\} + \frac{1}{2}E\{A_\infty^2\}.$$

Since M is an arbitrary bounded martingale we are free to choose

$$M_\infty = (A-\widetilde{A})_\infty 1_{\{|(A-\widetilde{A})_\infty| \le n\}},$$

and we obtain

$$\frac{1}{2}E\{(A-\widetilde{A})_\infty^2 1_{\{|(A-\widetilde{A})_\infty| \le n\}}\} \le \frac{1}{2}E\{A_\infty^2\},$$

and using the Monotone Convergence Theorem we conclude

$$E\{(A-\widetilde{A})_\infty^2\} \le E\{A_\infty^2\}.$$

Consequently

$$E\{\widetilde{A}_\infty^2\} \le 2E\{A_\infty^2\} + 2E\{(A-\widetilde{A})_\infty^2\} \le 4E\{A_\infty^2\} < \infty,$$

and $A - \widetilde{A}$ is a square integrable martingale, and

$$\|A\|_{\mathcal{H}^2} = \|(A-\widetilde{A})_\infty\|_{L^2} + \|\widetilde{A}_\infty\|_{L^2} \le 3\|A_\infty\|_{L^2}$$

for A increasing. □

Remarks. The constant 6 can be improved to $1+\sqrt{8} \le 4$ by not decomposing A into A^+ and A^-. This lemma can also be proved using the Burkholder-Gundy inequalities (see Meyer [171, page 347]).

In Chap. V we use an alternative norm for semimartingales which we denote $\|\cdot\|_{\underline{\underline{H}}^p}$, $1 \le p < \infty$. The preceding lemma shows that the norms $\|\cdot\|_{\mathcal{H}^2}$ and $\|\cdot\|_{\underline{\underline{H}}^2}$ are equivalent.

The restrictions of integrands to $\mathbf{b}\mathcal{P}$ and semimartingales to $\mathcal{H}^2$ are mathematically convenient but not necessary. A standard method of relaxing such hypothesis is to consider cases where they hold *locally*. Recall from Sect. 6 of Chap. I that a **property π is said to hold locally** for a process X if there exists a sequence of stopping times $(T^n)_{n\ge 0}$ such that $0 = T^0 \le T^1 \le T^2 \le \cdots \le T^n \le \cdots$ and $\lim_{n\to\infty} T^n = \infty$ a.s., and such that $X^{T^n} 1_{\{T^n>0\}}$ has property π for each n. Since we are assuming our semimartingales X satisfy $X_0 = 0$, we could as well require only that X^{T^n} has property π for each n. A related condition is that a property hold *prelocally*.

Definition. A **property π is said to hold prelocally** for a process X with $X_0 = 0$ if there exists a sequence of stopping times $(T^n)_{n\ge 1}$ increasing to ∞ a.s. such that X^{T^n-} has property π for each $n \ge 1$.

Recall that $X^{T-} = X_t 1_{\{0\le t<T\}} + X_{T-} 1_{\{t\ge T\}}$. The next theorem shows that the restriction of semimartingales to $\mathcal{H}^2$ is not really a restriction at all.

Theorem 13. *Let X be a semimartingale, $X_0 = 0$. Then X is prelocally in $\mathcal{H}^2$. That is, there exists a non-decreasing sequence of stopping times (T^n), $\lim_{n\to\infty} T^n = \infty$ a.s., such that $X^{T^n-} \in \mathcal{H}^2$ for each n.*

Proof. Recall that $X^{T^n-} = X_t 1_{[0,T^n)} + X_{T^n-} 1_{[T^n,\infty)}$. By the Bichteler-Dellacherie Theorem (Theorem 43 of Chap. III) we can write $X = M + A$, where M is a local martingale and A is an FV process. By Theorem 25 of Chap. III we can further take M to have bounded jumps. Let β be the bound for the jumps of M. We define

$$T^n = \inf\{t > 0 : [M,M]_t > n \;\text{ or }\; \int_0^t |dA_s| > n\}$$

and let $Y = X^{T^n-}$. Then Y has bounded jumps and hence it is a special semimartingale (Theorem 31 of Chap. III). Moreover

$$Y = X^{T^n-} = M^{T^n} + A^{T^n-} - (\Delta M_{T^n}) 1_{[T^n,\infty)},$$

or

$$Y = L + C,$$

where $L = M^{T^n}$ and $C = A^{T^n-} - (\Delta M_{T^n}) 1_{[T^n,\infty)}$. Then $[L,L] \le n + \beta^2$, so L is a martingale in $\mathcal{H}^2$ (Corollary 4 to Theorem 27 of Chap. II), and also

$$\int_0^\infty |dC_s| \le n + |\Delta M_{T^n}| \le 2n + \beta,$$

hence $C \in \mathcal{H}^2$ by the lemma. Therefore $X^{T^n-} = L + C \in \mathcal{H}^2$. $\square$

We are now in a position to define the stochastic integral *for an arbitrary semimartingale*, as well as for predictable processes which need not be bounded.

Let X be a semimartingale in $\mathcal{H}^2$. To define a stochastic integral for predictable processes H which are not necessarily bounded (written $H \in \mathcal{P}$), we approximate them with $H^n \in \mathbf{b}\mathcal{P}$.

Definition. Let $X \in \mathcal{H}^2$ with canonical decomposition $X = \overline{N} + \overline{A}$. We say $H \in \mathcal{P}$ is $(\mathcal{H}^2, X)$ **integrable** if

$$E\{\int_0^\infty H_s^2 d[\overline{N}, \overline{N}]_s\} + E\{(\int_0^\infty |H_s||d\overline{A}_s|)^2\} < \infty.$$

Theorem 14. *Let X be a semimartingale and let $H \in \mathcal{P}$ be $(\mathcal{H}^2, X)$ integrable. Let $H^n = H1_{\{|H| \le n\}} \in \mathbf{b}\mathcal{P}$. Then $H^n \cdot X$ is a Cauchy sequence in $\mathcal{H}^2$.*

Proof. Since $H^n \in \mathbf{b}\mathcal{P}$, each n, the stochastic integrals $H^n \cdot X$ are defined. Note also that $\lim_{n\to\infty} H^n = H$ and that $|H^n| \le |H|$, each n. Then

$$\begin{aligned}\|H^n \cdot X - H^m \cdot X\|_{\mathcal{H}^2} &= d_X(H^n, H^m) \\ &= \|(\int_0^\infty (H_s^n - H_s^m)^2 d[\overline{N}, \overline{N}]_s)^{1/2}\|_{L^2} + \|\int_0^\infty |H_s^n - H_s^m||d\overline{A}_s|\|_{L^2},\end{aligned}$$

and the result follows by two applications of the Dominated Convergence Theorem. □

Definition. Let X be a semimartingale in $\mathcal{H}^2$, and let $H \in \mathcal{P}$ be $(\mathcal{H}^2, X)$ integrable. The **stochastic integral** $H \cdot X$ is defined to be $\lim_{n\to\infty} H^n \cdot X$, with convergence in $\mathcal{H}^2$, where $H^n = H1_{\{|H| \le n\}}$.

Note that $H \cdot X$ in the preceding definition exists by Theorem 14. We can "localize" the above theorem by allowing both more general $H \in \mathcal{P}$ and arbitrary semimartingales with the next definition.

Definition. Let X be a semimartingale and $H \in \mathcal{P}$. The stochastic integral $H \cdot X$ is said to **exist** if there exists a sequence of stopping times T^n increasing to ∞ a.s. such that $X^{T^n-} \in \mathcal{H}^2$, each $n \ge 1$, and such that H is $(\mathcal{H}^2, X^{T^n-})$ integrable for each n. In this case we say H is X **integrable**, written $H \in L(X)$, and we define the **stochastic integral** by

$$H \cdot X = H \cdot (X^{T^n-}), \qquad \text{on } [0, T^n),$$

each n.

Note that if $m > n$ then

$$H^k \cdot (X^{T^m-})^{T^n-} = H^k \cdot (X^{T^m \wedge T^n-}) = H^k \cdot (X^{T^n-}),$$

where $H^k = H1_{\{|H| \le k\}}$, by the corollary of Theorem 8. Hence taking limits we have $H \cdot (X^{T^m-})^{T^n-} = H \cdot (X^{T^n-})$, and the stochastic integral is well-defined for $H \in L(X)$. Moreover let R^ℓ be another sequence of stopping times such that $X^{R^\ell-} \in \mathcal{H}^2$ and such that H is $(\mathcal{H}^2, X^{R^\ell-})$ integrable, for each ℓ. Again using the corollary of Theorem 8 combined with taking limits we see that

$$H \cdot (X^{R^\ell-}) = H \cdot (X^{T^n-})$$

on $[0, R^\ell \wedge T^n)$, each $\ell \ge 1$ and $n \ge 1$. Thus in this sense the definition of the stochastic integral *does not depend on the particular sequence of stopping times.*

If $H \in \mathbf{b}\mathcal{P}$ (i.e., H is bounded), then $H \in L(X)$ for all semimartingales X, since every semimartingale is prelocally in $\mathcal{H}^2$ by Theorem 13.

Definition. A process H is said to be **locally bounded** if there exists a sequence of stopping times $(S^m)_{m\ge 1}$ increasing to ∞ a.s. such that for each $m \ge 1$, $(H_{t\wedge S^m} 1_{\{S^m>0\}})_{t\ge 0}$ is bounded.

Note that any process in $\mathbb{L}$ is locally bounded. The next example is sufficiently important that we state it as a theorem.

Theorem 15. *Let X be a semimartingale and let $H \in \mathcal{P}$ be locally bounded. Then $H \in L(X)$. That is, the stochastic integral $H \cdot X$ exists.*

Proof. Let $(S^m)_{m\ge1}$, $(T^n)_{n\ge1}$ be two sequences of stopping times, each increasing to ∞ a.s., such that $H^{S^m} 1_{\{S^m>0\}}$ is bounded for each m, and $X^{T^n-} \in \mathcal{H}^2$ for each n. Define $R^n = \min(S^n, T^n)$. Then $H = H^{R^n} 1_{\{R^n>0\}}$ on $(0, R^n)$ and hence it is bounded there. Since X^{R^n-} charges only $(0, R^n)$, we have that H is $(\mathcal{H}^2, X^{R^n-})$ integrable for each $n \ge 1$. Therefore using the sequence R^n which increases to ∞ a.s., we are done. □

We now turn our attention to the properties of this more general integral. Many of the properties are simple extensions of earlier theorems and we omit their proofs. Note that trivially the stochastic integral $H \cdot X$, for $H \in L(X)$, is also a semimartingale.

Theorem 16. *Let X be a semimartingale and let $H, J \in L(X)$. Then $\alpha H + \beta J \in L(X)$ and $(\alpha H + \beta J) \cdot X = \alpha H \cdot X + \beta J \cdot X$. That is, $L(X)$ is a linear space.*

Proof. Let (R^m) and (T^n) be sequences of stopping times such that H is $(\mathcal{H}^2, X^{R^m-})$ integrable, each m, and J is $(\mathcal{H}^2, X^{T^n-})$ integrable, each n. Taking $S^n = R^n \wedge T^n$, it is easy to check that $\alpha H + \beta J$ is $(\mathcal{H}^2, X^{S^n-})$ integrable for each n. □

Theorem 17. *Let X, Y be semimartingales and suppose $H \in L(X)$ and $H \in L(Y)$. Then $H \in L(X + Y)$ and $H \cdot (X + Y) = H \cdot X + H \cdot Y$.*

Theorem 18. *Let X be a semimartingale and $H \in L(X)$. The jump process $(\Delta(H \cdot X)_s)_{s \geq 0}$ is indistinguishable from $(H_s(\Delta X_s))_{s \geq 0}$.*

Theorem 19. *Let T be a stopping time, X a semimartingale, and $H \in L(X)$. Then*

$$(H \cdot X)^T = H1_{[0,T]} \cdot X = H \cdot (X^T).$$

Moreover, letting $\infty-$ equal ∞, we have moreover

$$(H \cdot X)^{T-} = H \cdot (X^{T-}).$$

Theorem 20. . *Let X be a semimartingale with paths of finite variation on compacts. Let $H \in L(X)$ be such that the Stieltjes integral $\int_0^t |H_s||dX_s|$ exists a.s., each $t \geq 0$. Then the stochastic integral $H \cdot X$ agrees with a path-by-path Stieltjes integral.*

Theorem 21 (Associativity). *Let X be a semimartingale with $K \in L(X)$. Then $H \in L(K \cdot X)$ if and only if $HK \in L(X)$, in which case $H \cdot (K \cdot X) = (HK) \cdot X$.*

Theorem 22. *Let X, Y be semimartingales and let $H \in L(X)$, $K \in L(Y)$. Then*

$$[H \cdot X, K \cdot Y]_t = \int_0^t H_s K_s d[X,Y]_s \quad (t \geq 0).$$

Note that in Theorem 22 since $H \cdot X$ and $H \cdot Y$ are semimartingales, the quadratic covariation exists and the content of the theorem is the formula. Indeed, Theorem 22 gives a necessary condition for H to be in $L(X)$, namely that $\int_0^t H_s^2 d[X,X]_s$ exists and is finite for all $t \geq 0$. The next theorem (Theorem 23) is a special case of Theorem 25, but we include it because of the simplicity of its proof.

Theorem 23. *Let X be a semimartingale, let $H \in L(X)$, and suppose Q is another probability with $Q \ll P$. If $H_Q \cdot X$ exists, it is Q indistinguishable from $H_P \cdot X$.*

Proof. $H_Q \cdot X$ denotes the stochastic integral computed under Q. By Theorem 14 of Chap. II, we know that $H_Q \cdot X = H_P \cdot X$ for $H \in \mathbb{L}$, and therefore if $X \in \mathcal{H}^2$ for both P and Q, they are equal for $H \in b\mathcal{P}$ by the corollary of Theorem 5. Let $(R^\ell)_{\ell \geq 1}$, $(T^n)_{n \geq 1}$ be two sequences of stopping times increasing to ∞ a.s. such that H is $(\mathcal{H}^2, X^{R^\ell -})$ integrable under Q, and H is $(\mathcal{H}^2, X^{T^n -})$ integrable under P, each ℓ and n. Let $S^m = R^m \wedge T^m$, so that H is $(\mathcal{H}^2, X^{S^m -})$ integrable under both P and Q. Then $H \cdot X = \lim_{n \to \infty} H^n \cdot X$ on $[0, S^m)$ in both $d_X(P)$ and $d_X(Q)$, where $H^n = H1_{\{|H| \leq n\}} \in b\mathcal{P}$. Since $H_P^n \cdot X = H_Q^n \cdot X$, each n, the limits are also equal. □

Much more than Theorem 23 is indeed true, as we will see in Theorem 25, which contains Theorem 23 as a special case. We need several preliminary results.

Lemma. Let $X \in \mathcal{H}^2$ and $X = \overline{N} + \overline{A}$ be its canonical decomposition. Then

$$E\{[\overline{N}, \overline{N}]_\infty\} \leq E\{[X, X]_\infty\}.$$

Proof. First observe that

$$[X, X] = [\overline{N}, \overline{N}] + 2[\overline{N}, \overline{A}] + [\overline{A}, \overline{A}].$$

It suffices to show $E\{[\overline{N}, \overline{A}]_\infty\} = 0$, since then

$$E\{[\overline{N}, \overline{N}]_\infty\} = E\{[X, X]_\infty - [\overline{A}, \overline{A}]_\infty\},$$

and the result follows since $[\overline{A}, \overline{A}]_\infty \geq 0$. Note that

$$E\{|[\overline{N}, \overline{A}]_\infty|\} \leq (E\{[\overline{N}, \overline{N}]_\infty\})^{1/2}(E\{[\overline{A}, \overline{A}]_\infty\})^{1/2} < \infty,$$

by the Kunita-Watanabe inequalities. Also $E\{[M, \overline{A}]_\infty\} = 0$ for all bounded martingales because $\overline{A}$ is natural. Since bounded martingales are dense in the space of L^2 martingales, there exists a sequence $(M^n)_{n \geq 1}$ of bounded martingales such that $\lim_{n \to \infty} E\{[M^n - \overline{N}, M^n - \overline{N}]_\infty\} = 0$. Again using the Kunita-Watanabe inequalities we have

$$E\{|[\overline{N} - M^n, \overline{A}]_\infty|\} \leq (E\{[\overline{N} - M^n, \overline{N} - M^n]_\infty\})^{1/2}(E\{[\overline{A}, \overline{A}]_\infty\})^{1/2}$$

and therefore $\lim_{n \to \infty} E\{[M^n, \overline{A}]_\infty\} = E\{[\overline{N}, \overline{A}]_\infty\}$. Since $E\{[M^n, \overline{A}]_\infty\} = 0$, each n, it follows that $E\{[\overline{N}, \overline{A}]_\infty\} = 0$. □

Note that in the preceding proof we established the useful equality $E\{[X, X]_t\} = E\{[\overline{N}, \overline{N}]_t\} + E\{[\overline{A}, \overline{A}]_t\}$ for a semimartingale $X \in \mathcal{H}^2$ with canonical decomposition $X = \overline{N} + \overline{A}$.

Theorem 24. *For X a semimartingale in $\mathcal{H}^2$,*

$$\|X\|_{\mathcal{H}^2} \leq \sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2} + 2\|[X, X]^{1/2}_\infty\|_{L^2} \leq 5\|X\|_{\mathcal{H}^2}.$$

Proof. By Theorem 5 for $|H| \leq 1$

$$\|(H \cdot X)^*_\infty\|_{L^2} \leq \sqrt{8}\|H \cdot X\|_{\mathcal{H}^2} \leq \sqrt{8}\|X\|_{\mathcal{H}^2}.$$

Since

$$\begin{aligned} 2\|[X, X]^{1/2}_\infty\|_{L^2} &\leq 2\|[M, M]^{1/2}_\infty\|_{L^2} + 2\left\|\int_0^\infty |dA_s|\right\|_{L^2} \\ &= 2\|X\|_{\mathcal{H}^2}, \end{aligned}$$

where $X = M + A$ is the canonical decomposition of X, we have the right inequality.

For the left inequality we have $\|[M, M]^{1/2}_\infty\|_{L^2} \leq \|[X, X]^{1/2}_\infty\|_{L^2}$ by the lemma preceding this theorem. Moreover if $|H| \leq 1$, then

$$\begin{aligned}\|(H \cdot A)_t\|_{L^2} &\leq \|(H \cdot X)_t\|_{L^2} + \|(H \cdot M)_t\|_{L^2} \\ &\leq \|(H \cdot X)^*_\infty\|_{L^2} + \|[M, M]^{1/2}_\infty\|_{L^2}.\end{aligned}$$

Next take $H = \frac{dA}{|dA|}$; this exists as a predictable process since A is predictable, and therefore A predictable implies that $|A|$, the total variation process, is also predictable. Consequently we have

$$\begin{aligned}\|X\|_{\mathcal{H}^2} &= \|[M, M]^{1/2}_\infty\|_{L^2} + \|\int_0^\infty |dA_s|\|_{L^2} \\ &= \|[M, M]^{1/2}_\infty\|_{L^2} + \|(H \cdot A)_\infty\|_{L^2} \\ &\leq \|[M, M]^{1/2}_\infty\|_{L^2} + \|(H \cdot X)^*_\infty\|_{L^2} + \|[M, M]^{1/2}_\infty\|_{L^2} \\ &\leq \|(H \cdot X)^*_\infty\|_{L^2} + 2\|[X, X]^{1/2}_\infty\|_{L^2}.\end{aligned}$$

□

We present as a corollary to Theorem 24, the equivalence of the two pseudonorms $\sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2}$ and $\|X\|_{\mathcal{H}^2}$. We will not have need of this corollary in this book,[5] but it is a pretty and useful result nevertheless. It is originally due to Yor [241], and it is actually true for all p, $1 \leq p < \infty$. (See also Dellacherie-Meyer [46, pages 303–305].)

Corollary. For a semimartingale X (with $X_0 = 0$),

$$\|X\|_{\mathcal{H}^2} \leq 3 \sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2} \leq 9\|X\|_{\mathcal{H}^2},$$

and in particular $\sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2} < \infty$ if and only if $\|X\|_{\mathcal{H}^2} < \infty$.

Proof. By Theorem 5 if $\|X\|_{\mathcal{H}^2} < \infty$ and $|H| \leq 1$ we have

$$\|(H \cdot X)^*_\infty\|_{L^2} \leq \sqrt{8}\|H \cdot X\|_{\mathcal{H}^2} \leq \sqrt{8}\|X\|_{\mathcal{H}^2}.$$

Thus we need to show only the left inequality. By Theorem 24 it will suffice to show that

$$\|[X, X]^{1/2}_\infty\|_{L^2} \leq \sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2},$$

for a semimartingale X with $X_0 = 0$.

To this end fix a $t > 0$ and let $0 = T_0 \leq T_1 \leq \cdots \leq T_n = t$ be a random partition of $[0, t]$. Choose $\varepsilon_1, \ldots, \varepsilon_n$ non-random and each equal to 1 or -1 and let $H = \sum_{i=1}^n \varepsilon_i 1_{(T_{i-1}, T_i]}$. Then H is a simple predictable process and

$$(H \cdot X)_\infty = \sum_{i=1}^n \varepsilon_i (X_{T_i} - X_{T_{i-1}}).$$

Let $\alpha = \sup_{|H| \leq 1} \|(H \cdot X)^*_\infty\|_{L^2}$. We then have

[5] However the corollary does give some insight into the relationship between Theorems 12 and 14 in Chap. V.

$$\alpha^2 \geq E\{(H \cdot X)_\infty^2\} = \sum_{i,j=1}^{n} \varepsilon_i \varepsilon_j E\{(X_{T_i} - X_{T_{i-1}})(X_{T_j} - X_{T_{j-1}})\}.$$

If we next average over all sequences $\varepsilon_1, \ldots, \varepsilon_n$ taking values in the space $\{\pm 1\}^n$, we deduce

$$\alpha^2 \geq \sum_{i=1}^{n} E\{(X_{T_i} - X_{T_{i-1}})^2\} = E\{\sum_{i=1}^{n} (X_{T_i} - X_{T_{i-1}})^2\}.$$

Next let $\sigma_m = \{T_i^m\}$ be a sequence of random partitions of $[0, t]$ tending to the identity. Then $\sum_i (X_{T_i^m} - X_{T_{i-1}^m})^2$ converges in probability to $[X, X]_t$. Let $\{m_k\}$ be a subsequence so that $\sum_i (X_{T_i^{m_k}} - X_{T_{i-1}^{m_k}})^2$ converges to $[X, X]_t$ a.s. Finally by Fatou's Lemma we have

$$E\{[X, X]_t\} \leq \liminf_{m_k \to \infty} E\{\sum_i (X_{T_i^{m_k}} - X_{T_{i-1}^{m_k}})^2\} \leq \alpha^2.$$

Letting t tend to ∞ we conclude that

$$E\{[X, X]_\infty\} \leq \alpha^2 = \sup_{|H| \leq 1} \|(H \cdot X)_\infty^*\|_{L^2}^2.$$

It remains to show that if $\sup_{|H| \leq 1} \|(H \cdot X)_\infty^*\|_{L^2} < \infty$, then $X \in \mathcal{H}^2$. We will show the contrapositive. If $X \notin \mathcal{H}^2$, then $\sup_{|H| \leq 1} \|(H \cdot X)_\infty^*\|_{L^2} = \infty$. Indeed, let T^n be stopping times increasing to ∞ such that $X^{T^n -}$ is in $\mathcal{H}^2$ for each n (cf., Theorem 13). Then

$$\|X^{T^n -}\|_{\mathcal{H}^2} \leq 3 \sup_{|H| \leq 1} \|(H \cdot X^{T^n -})_\infty^*\|_{L^2} \leq 9 \|X^{T^n -}\|_{\mathcal{H}^2}.$$

Letting n tend to ∞ gives the result. □

Before proving Theorem 25 we need two technical lemmas.

Lemma 1. Let A be a non-negative increasing FV process and let Z be a positive uniformly integrable martingale. Let T be a stopping time such that $A = A^{T-}$ (that is, $A_\infty = A_{T-}$) and let k be a constant such that $Z \leq k$ on $[0, T)$. Then

$$E\{A_\infty Z_\infty\} \leq k E\{A_\infty\}.$$

Proof. Since $A_{0-} = Z_{0-} = 0$, by integration by parts

$$\begin{aligned} A_t Z_t &= \int_0^t A_{s-} dZ_s + \int_0^t Z_{s-} dA_s + [A, Z]_t \\ &= \int_0^t A_{s-} dZ_s + \int_0^t Z_s dA_s \end{aligned}$$

where the second integral in the preceding is a path-by-path Stieltjes integral. Let R^n be stopping times increasing to ∞ a.s. that reduce the local martingale $(\int_0^t A_{s-} dZ_s)_{t\geq 0}$. Since dA_s charges only $[0,T)$ we have

$$\int_0^t Z_s dA_s \leq \int_0^t k dA_s \leq kA_\infty$$

for every $t \geq 0$. Therefore

$$\begin{aligned} E\{(AZ)_{R^n}\} &= E\{\int_0^{R^n} A_{s-} dZ_s\} + E\{\int_0^{R^n} Z_s dA_s\} \\ &\leq 0 + E\{kA_\infty\}. \end{aligned}$$

The result follows by Fatou's Lemma. □

Lemma 2. Let X be a semimartingale with $X_0 = 0$, let Q be another probability with $Q \ll P$, and let $Z_t = E_P\{\frac{dQ}{dP}|\mathcal{F}_t\}$. If T is a stopping time such that $Z_t \leq k$ on $[0,T)$ for a constant k, then

$$\|X^{T-}\|_{\mathcal{H}^2(Q)} \leq 5\sqrt{k}\|X^{T-}\|_{\mathcal{H}^2(P)}.$$

Proof. By Theorem 24 we have

$$\begin{aligned} \|X^{T-}\|_{\mathcal{H}^2(Q)} &\leq \sup_{|H|\leq 1} E_Q\{((H \cdot X^{T-})^*)^2\}^{\frac{1}{2}} + 2E_Q\{[X^{T-}, X^{T-}]^{\frac{1}{2}}\} \\ &\leq \sup_{|H|\leq 1} E_P\left\{\frac{dQ}{dP}((H \cdot X^{T-})^*)^2\right\}^{\frac{1}{2}} + 2E_P\left\{\frac{dQ}{dP}[X^{T-}, X^{T-}]^{\frac{1}{2}}\right\} \\ &\leq \sqrt{k} \sup_{|H|\leq 1} E_P\{((H \cdot X^{T-})^*)^2\}^{\frac{1}{2}} + 2\sqrt{k}E_P\{[X^{T-}, X^{T-}]^{\frac{1}{2}}\}. \end{aligned}$$

where we have used Lemma 1 on both terms to obtain the last inequality above. The result follows by the right inequality of Theorem 24. □

Note in particular that an important consequence of Lemma 2 is that if $Q \ll P$ with $\frac{dQ}{dP}$ bounded, then $X \in \mathcal{H}^2(P)$ implies that $X \in \mathcal{H}^2(Q)$ as well, with the estimate $\|X\|_{\mathcal{H}^2(Q)} \leq 5\sqrt{k}\|X\|_{\mathcal{H}^2(P)}$, where k is the bound for $\frac{dQ}{dP}$. Note further that this result (without the estimate) is obvious if one uses the equivalent pseudonorm given by the corollary to Theorem 24, since

$$\begin{aligned} \sup_{|H|\leq 1} E_Q\{((H \cdot X)^*_\infty)^2\} &= \sup_{|H|\leq 1} E_P\{\frac{dQ}{dP}((H \cdot X)^*_\infty)^2\} \\ &\leq k \sup_{|H|\leq 1} E_P\{((H \cdot X)^*_\infty)^2\}, \end{aligned}$$

where again k is the bound for $\frac{dQ}{dP}$.

Theorem 25. *Let X be a semimartingale and $H \in L(X)$. If $Q \ll P$, then $H \in L(X)$ under Q as well, and $H_Q \cdot X = H_P \cdot X$, Q-a.s.*

Proof. Let T^n be a sequence of stopping times increasing to ∞ a.s. such that H is $(X^{T^n-}, \mathcal{H}^2)$ integrable under P, each $n \geq 1$. Let $Z_t = E_P\{\frac{dQ}{dP}|\mathcal{F}_t\}$, the càdlàg version. Define $S^n = \inf\{t > 0 : |Z_t| > n\}$ and set $R^n = S^n \wedge T^n$. Then $X^{R^n-} \in \mathcal{H}^2(P) \cap \mathcal{H}^2(Q)$ by Lemma 2, and H is $(\mathcal{H}^2, X^{R^n-})$ integrable under P. We need to show H is $(\mathcal{H}^2, X^{R^n-})$ integrable under Q, which will in turn imply that $H \in L(X)$ under Q. Let $X^{R^n-} = N + C$ be the canonical decomposition under Q. Let $H^m = H1_{\{|H| \leq m\}}$. Then

$$\begin{aligned}
(E_Q\{\int (H_s^m)^2 d[N,N]_s\})^{1/2} &+ \| \int |H_s^m||dC_s| \|_{L^2(Q)} \\
&= \|H^m \cdot X^{R^n-}\|_{\mathcal{H}^2(Q)} \leq 5\sqrt{n}\|H^m \cdot X^{R^n-}\|_{\mathcal{H}^2(P)} \\
&\leq 5\sqrt{n}\|H \cdot X^{R^n-}\|_{\mathcal{H}^2(P)} < \infty,
\end{aligned}$$

and then by monotone convergence we see that H is $(\mathcal{H}^2, X^{R^n-})$ integrable under Q. Thus $H \in L(X)$ under Q, and it follows that $H_Q \cdot X = H_P \cdot X$, Q-a.s. □

Theorem 25 can be used to extend Theorem 20 in a way analogous to the extension of Theorem 17 by Theorem 18 in Chap. II.

Theorem 26. *Let X, $\overline{X}$ be two semimartingales, and let $H \in L(X)$, $\overline{H} \in L(\overline{X})$. Let $A = \{\omega : H_\cdot(\omega) = \overline{H}_\cdot(\omega)$ and $X_\cdot(\omega) = \overline{X}_\cdot(\omega)\}$, and let $B = \{\omega : t \mapsto X_t(\omega)$ is of finite variation on compacts$\}$. Then $H \cdot X = \overline{H} \cdot \overline{X}$ on A, and $H \cdot X$ is equal to a path-by-path Lebesgue-Stieltjes integral on B.*

Proof. Without loss of generality assume $P(A) > 0$. Define Q by $Q(\Lambda) = P(\Lambda|A)$. Then $Q \ll P$ and therefore $H \in L(X)$, $\overline{H} \in L(\overline{X})$ under Q as well as under P by Theorem 25. However under Q the processes H and $\overline{H}$ as well as X and $\overline{X}$ are indistinguishable. Thus $H_Q \cdot X = \overline{H}_Q \cdot \overline{X}$ and hence $H \cdot X = \overline{H} \cdot \overline{X}$ P-a.s. on A by Theorem 25, since $Q \ll P$.

The second assertion has an analogous proof (see the proof of Theorem 18 of Chap. II). □

Note that one can use stopping times to localize the result of Theorem 26. The proof of the following corollary is analogous to the proof of the corollary of Theorem 18 of Chap. II.

Corollary. With the notation of Theorem 26, let S, T be two stopping times with $S < T$. Define

$$\begin{aligned}
C &= \{\omega : H_t(\omega) = \overline{H}_t(\omega); X_t(\omega) = \overline{X}_t(\omega); S(\omega) < t \leq T(\omega)\}, \\
D &= \{\omega : t \mapsto X_t(\omega) \text{ is of finite variation on } S(\omega) < t < T(\omega)\}.
\end{aligned}$$

Then $H \cdot X^T - H \cdot X^S = \overline{H} \cdot \overline{X}^T - \overline{H} \cdot \overline{X}^S$ on C and $H \cdot X^T - H \cdot X^S$ equals a path-by-path Lebesgue-Stieltjes integral on D.

Theorem 27. *Let P_k be a sequence of probabilities such that X is a P_k semimartingale for each k. Let $R = \sum_{k=1}^{\infty} \lambda_k P_k$ where $\lambda_k \geq 0$, each k, and $\sum_{k=1}^{\infty} \lambda_k = 1$. Let $H \in L(X)$ under R. Then $H \in L(X)$ under P_k and $H_R \cdot X = H_{P_k} \cdot X$, P_k-a.s., for all k such that $\lambda_k > 0$.*

Proof. If $\lambda_k > 0$ then $P_k \ll R$. Moreover since $P_k(\Lambda) \leq \frac{1}{\lambda_k} R(\Lambda)$, it follows that $H \in L(X)$ under P_k. The result then follows by Theorem 25. □

We now turn to the relationship of stochastic integration to martingales and local martingales. In Theorem 11 we saw that if M is a square integrable martingale and $H \in \mathrm{b}\mathcal{P}$, then $H \cdot M$ is also a square integrable martingale. When M is *locally* square integrable we have a simple sufficient condition for H to be in $L(M)$.

Lemma. Let M be a square integrable martingale and let $H \in \mathcal{P}$ be such that $E\{\int_0^{\infty} H_s^2 d[M,M]_s\} < \infty$. Then $H \cdot M$ is a square integrable martingale.

Proof. If $H^k \in \mathrm{b}\mathcal{P}$, then $H^k \cdot M$ is a square integrable martingale by Theorem 11. Taking $H^k = H 1_{\{|H| \leq k\}}$, and since H is $(\mathcal{H}^2, M)$ integrable, by Theorem 14 $H^k \cdot M$ converges in $\mathcal{H}^2$ to $H \cdot M$ which is hence a square integrable martingale. □

Theorem 28. *Let M be a locally square integrable local martingale, and let $H \in \mathcal{P}$. The stochastic integral $H \cdot M$ exists (i.e., $H \in L(M)$) and is a locally square integrable local martingale if there exists a sequence of stopping times $(T^n)_{n \geq 1}$ increasing to ∞ a.s. such that $E\{\int_0^{T^n} H_s^2 d[M,M]_s\} < \infty$.*

Proof. We assume that M is a square integrable martingale stopped at the time T^n. The result follows by applying the lemma. □

Theorem 29. *Let M be a local martingale, and let $H \in \mathcal{P}$ be locally bounded. Then the stochastic integral $H \cdot M$ is a local martingale.*

Proof. By stopping we may, as in the proof of Theorem 29 of Chap. III, assume that H is bounded, M is uniformly integrable, and that $M = N + A$ where N is a bounded martingale and A is of integrable variation. We know that there exists R^k increasing to ∞ a.s. such that $M^{R^k-} \in \mathcal{H}^2$, and since $H \in \mathrm{b}\mathcal{P}$ there exist processes $H^\ell \in \mathrm{b}\mathbb{L}$ such that $\|H^\ell \cdot M^{R^k-} - H \cdot M^{R^k-}\|_{\mathcal{H}^2}$ tends to zero. In particular, $H^\ell \cdot M^{R^k-}$ tends to $H \cdot M^{R^k-}$ in *ucp*. Therefore we can take H^ℓ such that $H^\ell \cdot M^{R^k-}$ tends to $H \cdot M$ in *ucp*, with $H^\ell \in \mathrm{b}\mathbb{L}$. Finally without loss of generality we assume $H^\ell \cdot M$ converges to $H \cdot M$ in *ucp*. Since $H^\ell \cdot M = H^\ell \cdot N + H^\ell \cdot A$ and $H^\ell \cdot N$ converges to $H \cdot N$ in *ucp*, we deduce $H^\ell \cdot A$ converges to $H \cdot A$ in *ucp* as well. Let $0 \leq s < t$ and assume $Y \in \mathrm{b}\mathcal{F}_s$. Therefore, since A is of integrable total variation, and since $H^\ell \cdot A$ is a martingale for $H^\ell \in \mathrm{b}\mathbb{L}$ (see Theorem 29 of Chap. III), we have

$$
\begin{aligned}
E\{Y\int_{s+}^{t} H_u dA_u\} = E\{Y\int_{s+}^{t} \lim_{\ell\to\infty} H_u^\ell dA_u\} &= E\{Y \lim_{\ell\to\infty}\int_{s+}^{t} H_u^\ell dA_u\} \\
&= \lim_{\ell\to\infty} E\{Y\int_{s+}^{t} H_u^\ell dA_u\} \\
&= 0,
\end{aligned}
$$

where we have used Lebesgue's Dominated Convergence Theorem both for the Stieltjes integral ω-by-ω (taking a subsequence if necessary to have a.s. convergence) and for the expectation. We conclude that $(\int_0^t H_s dA_s)_{t\geq 0}$ is a martingale, hence $H\cdot M = H\cdot N + H\cdot A$ is also a martingale under the assumptions made; therefore it is a local martingale. □

In the proof of Theorem 29 the hypothesis that $H\in\mathcal{P}$ was locally bounded was used to imply that if $M = N + A$ with N having locally bounded jumps and A an FV local martingale, then the two processes

$$\int_0^t H_s^2 d[N,N]_s \quad \text{and} \quad \int_0^t |H_s||dA_s|$$

are locally integrable. Thus one could weaken the hypothesis that H is locally bounded, but it would lead to an awkward statement. *The general result, that M a local martingale and $H\in L(M)$ implies that $H\cdot M$ is a local martingale, is not true!*

See Emery's example, which precedes Theorem 34 for a stochastic integral with respect to an $\mathcal{H}^2$ martingale which is not even a local martingale! Emery's counterexample has lead to the development of what are now known as sigma martingales, a class of processes which are not local martingales, but can be thought to be "morally local martingales." This is treated in Section 9 of this chapter.

Corollary. Let M be a local martingale, $M_0 = 0$, and let T be a predictable stopping time. Then M^{T-} is a local martingale.

Proof. The notation M^{T-} means

$$M_t^{T-} = M_t 1_{\{t<T\}} + M_{T-}1_{\{t\geq T\}},$$

where $M_{0-} = M_0 = 0$. Let $(S^n)_{n\geq 1}$ be a sequence of stopping times announcing T. On $\{T>0\}$,

$$\lim_{n\to\infty} 1_{[0,S^n]} = 1_{[0,T)},$$

and since $1_{[0,S^n]}$ is a left continuous, adapted process it is predictable; hence $1_{[0,T)}$ is predictable. But $\int_0^t 1_{[0,T)}(s)dM_s = M_t^{T-}$ by Theorem 18, and hence it is a local martingale by Theorem 29. □

Note that if M is a local martingale and T is an arbitrary stopping time, it is not true in general that M^{T-} is still a local martingale.

Since a continuous local martingale is locally square integrable, the theory is particularly nice.

Theorem 30. *Let M be a continuous local martingale and let $H \in \mathcal{P}$ be such that $\int_0^t H_s^2 d[M,M]_s < \infty$ a.s., each $t \geq 0$. Then the stochastic integral $H \cdot M$ exists (i.e., $H \in L(M)$) and it is a continuous local martingale.*

Proof. Since M is continuous, we can take

$$R^k = \inf\{t > 0 : |M_t| > k\}.$$

Then $|M_{t \wedge R^k}| \leq k$ and therefore M is locally bounded, hence locally square integrable. Also M continuous implies $[M,M]$ is continuous, whence if

$$T^k = \inf\{t > 0 : \int_0^t H_s^2 d[M,M]_s > k\},$$

we see that $(\int_0^t H_s^2 d[M,M]_s)_{t \geq 0}$ is also locally bounded. Then $H \cdot M$ is a locally square integrable local martingale by Theorem 28. The stochastic integral $H \cdot M$ is continuous because $\Delta(H \cdot M) = H(\Delta M)$ and $\Delta M = 0$ by hypothesis. □

In the classical case where the continuous local martingale M equals B, a standard Brownian motion, Theorem 30 yields that if $H \in \mathcal{P}$ and $\int_0^t H_s^2 ds < \infty$ a.s., each $t \geq 0$, then the stochastic integral $(H \cdot B_t)_{t \geq 0} = (\int_0^t H_s dB_s)_{t \geq 0}$ exists, since $[B,B]_t = t$.

Corollary. Let X be a continuous semimartingale with (unique) decomposition $X = M + A$. Let $H \in \mathcal{P}$ be such that

$$\int_0^t H_s^2 d[M,M]_s + \int_0^t |H_s||dA_s| < \infty \quad \text{a.s.}$$

each $t \geq 0$. Then the stochastic integral $(H \cdot X)_t = \int_0^t H_s dX_s$ exists and it is continuous.

Proof. By the corollary of Theorem 31 of Chap. III, we know that M and A have continuous paths. The integral $H \cdot M$ exists by Theorem 30. Since $H \cdot A$ exists as a Stieltjes integral, it is easy to check that $H \in L(A)$, since A is continuous, and the result follows from Theorem 20. □

In the preceding corollary the semimartingale X is continuous, hence $[X,X] = [M,M]$ and the hypothesis can be written equivalently as

$$\int_0^t H_s^2 d[X,X]_s + \int_0^t |H_s||dA_s| < \infty \quad \text{a.s.}$$

each $t \geq 0$.

We end our treatment of martingales with a special case that yields a particularly simple condition for H to be in $L(M)$.

Theorem 31. *Let M be a local martingale with jumps bounded by a constant β. Let $H \in \mathcal{P}$ be such that $\int_0^t H_s^2 d[M,M]_s < \infty$ a.s., $t \geq 0$, and $E\{H_T^2\} < \infty$ for any bounded stopping time T. Then the stochastic integral $(\int_0^t H_s dM_s)_{t\geq 0}$ exists and it is a local martingale.*

Proof. Let $R^n = \inf\{t > 0 : \int_0^t H_s^2 d[M,M]_s > n\}$, and let $T^n = \min(R^n, n)$. Then T^n are bounded stopping times increasing to ∞ a.s. Note that

$$\begin{aligned} E\{\int_0^{T^n} H_s^2 d[M,M]_s\} &\leq n + E\{H_{T^n}^2(\Delta M_{T^n})^2\} \\ &\leq n + \beta^2 E\{H_{T^n}^2\} < \infty, \end{aligned}$$

and the result follows from Theorem 28. □

The next theorem is, of course, an especially important theorem, the *Dominated Convergence Theorem for stochastic integrals.*

Theorem 32 (Dominated Convergence Theorem). *Let X be a semimartingale, and let $H^m \in \mathcal{P}$ be a sequence converging a.s. to a limit H. If there exists a process $G \in L(X)$ such that $|H^m| \leq G$, all m, then H^m, H are in $L(X)$ and $H^m \cdot X$ converges to $H \cdot X$ in ucp.*

Proof. First note that if $|J| \leq G$ with $J \in \mathcal{P}$, then $J \in L(X)$. Indeed, let $(T^n)_{n\geq 1}$ increase to ∞ a.s. such that G is $(\mathcal{H}^2, X^{T^n-})$ integrable for each n. Then clearly

$$\begin{aligned} &E\{\int_0^\infty J_s^2 d[\overline{N},\overline{N}]_s\} + E\{(\int_0^\infty |J_s||d\overline{A}_s|)^2\} \\ &\qquad \leq E\{\int_0^\infty G_s^2 d[\overline{N},\overline{N}]_s\} + E\{(\int_0^\infty |G_s||d\overline{A}_s|)^2\} < \infty, \end{aligned}$$

and thus J is $(\mathcal{H}^2, X^{T^n-})$ integrable for each n. (Here $\overline{N} + \overline{A}$ is the canonical decomposition of X^{T^n-}.)

To show convergence in *ucp*, it suffices to show uniform convergence in probability on intervals of the form $[0, t_0]$ for t_0 fixed. Let $\varepsilon > 0$ be given, and choose n such that $P(T^n < t_0) < \varepsilon$, where $X^{T^n-} \in \mathcal{H}^2$ and G is $(\mathcal{H}^2, X^{T^n-})$ integrable. Let $X^{T^n-} = \overline{N} + \overline{A}$, the canonical decomposition. Then

$$\begin{aligned} &E\{\sup_{t\leq t_0}|H^m \cdot X^{T^n-} - H \cdot X^{T^n-}|^2\} \\ &\qquad \leq 2E\{\sup_{t\leq t_0}|(H^m - H)\cdot \overline{N}|^2\} + 2E\{(\int_0^{t_0} |H_s^m - H_s||d\overline{A}_s|)^2\}. \end{aligned}$$

The second term tends to zero by Lebesgue's Dominated Convergence Theorem. Since $|H^m - H| \leq 2G$, the integral $(H^m - H)\cdot \overline{N}$ is a square integrable martingale (by the lemma preceding Theorem 28). Therefore using Doob's maximal quadratic inequality, we have

$$E\{\sup_{t\le t_0}|(H^m-H)\cdot\overline{N}|^2\}\le 4E\{|(H^m-H)\cdot\overline{N}_{t_0}|^2\}$$
$$=4E\{\int_0^{t_0}(H^m_s-H_s)^2d[\overline{N},\overline{N}]_s\},$$

and again this tends to zero by the Dominated Convergence Theorem. Since convergence in L^2 implies convergence in probability, we conclude for $\delta>0$,

$$\begin{aligned}&\limsup_{m\to\infty}P\{\sup_{t\le t_0}|H^m\cdot X_t-H\cdot X_t|>\delta\}\\&\quad\le\limsup_{m\to\infty}P\{\sup_{t\le t_0}|((H^m-H)\cdot X^{T^n-})_t|>\delta\}+P(T^n<t_0)\\&\quad\le\varepsilon,\end{aligned}$$

and since ε is arbitrary, the limit is zero. □

We use the Dominated Convergence Theorem to prove a seemingly innocuous result. Generalizations, however, are delicate as we indicate following the proof.

Theorem 33. *Let $\mathbb{F}=(\mathcal{F}_t)_{t\ge0}$ and $\mathbb{G}=(\mathcal{G}_t)_{t\ge0}$ be two filtrations satisfying the usual hypotheses and suppose $\mathcal{F}_t\subset\mathcal{G}_t$, each $t\ge0$, and that X is a semimartingale for both $\mathbb{F}$ and $\mathbb{G}$. Let H be locally bounded and predictable for $\mathbb{F}$. Then the stochastic integrals $H_{\mathbb{F}}\cdot X$ and $H_{\mathbb{G}}\cdot X$ both exist, and they are equal.*[6]

Proof. It is trivial that H is locally bounded and predictable for $(\mathcal{G}_t)_{t\ge0}$ as well. By stopping, we can assume without loss of generality that H is bounded. Let

$$\mathcal{H}=\{\text{all bounded, }\mathcal{F}\text{ predictable }H\text{ such that }H_{\mathbb{F}}\cdot X=H_{\mathbb{G}}\cdot X\}.$$

Then $\mathcal{H}$ is clearly a monotone vector space, and $\mathcal{H}$ contains the multiplicative class $\mathbf{b}\mathbb{L}$ by Theorem 16 of Chap. II. Thus using Theorem 32 and the Monotone Class Theorem we are done. □

It is surprising that the assumption that H be locally bounded is important. Indeed, Jeulin [114, pages 46, 47] has exhibited an example which shows that Theorem 33 is false in general.

Theorem 33 is not an exact generalization of Theorem 16 of Chap. II. Indeed, suppose $\mathbb{F}$ and $\mathbb{G}$ are two arbitrary filtrations such that X is a semimartingale for both $\mathbb{F}$ and $\mathbb{G}$, and H is bounded and predictable for both of them. If $\mathcal{I}_t=\mathcal{F}_t\cap\mathcal{G}_t$, then X is still an $(\mathcal{I}_t)_{t\ge0}$ semimartingale by Stricker's Theorem, but it is not true in general that H is $(\mathcal{I}_t)_{t\ge0}$ predictable. It is an open question as to whether or not $H_{\mathcal{F}}\cdot X=H_{\mathcal{G}}\cdot X$ in this situation. For a partial result, see Zheng [248].

---

6 $H_{\mathbb{F}}\cdot X$ and $H_{\mathbb{G}}\cdot X$ denote the stochastic integrals computed with the filtrations $(\mathcal{F}_t)_{t\ge0}$ and $(\mathcal{G}_t)_{t\ge0}$, respectively.

Example (Emery's example of a stochastic integral behaving badly). The following simple example is due to M. Emery, and it has given rise to the study of sigma martingales, whose need in mathematical finance has become apparent. Let $X = (X_t)_{t\geq 0}$ be a stochastic process given by the following description. Let T be an exponential random variable with parameter $\lambda = 1$, let U be an independent random variable such that $P\{U = 1\} = P\{U = -1\} = 1/2$, and set $X = U1_{\{t\geq T\}}$. Then X together with its minimal filtration satisfying the usual hypotheses is a martingale in $\mathcal{H}^2$. That is, X is a stopped compound Poisson process with mean zero and is an L^2 martingale. Let $H_t = \frac{1}{t}1_{\{t>0\}}$. Therefore H is a deterministic integrand, continuous on $(0, \infty)$, and hence predictable. Consequently the path-by-path Lebesgue-Stieltjes integral $Z_t = \int_0^t H_s dX_s$ exists a.s. However $H \cdot X$ is not locally in $\mathcal{H}^p$ for any $p \geq 1$. (However since it is still a semimartingale[7], it is prelocally in $\mathcal{H}^2$.) Moreover, even though X is an L^2 martingale and H is a predictable integrand, *the stochastic integral $H \cdot X$ is not a local martingale* because $E\{|Z_S|\} = \infty$ for every stopping time S such that $P(S > 0) > 0$.

The next theorem is useful, since it allows one to work in the convenient space $\mathcal{H}^2$ through a change of measure. It is due originally to Bichteler and Dellacherie, and the proof here is due to Lenglart. We remark that the use of the exponent 2 is not important, and that the theorem is true for $\mathcal{H}^p$ for any $p \geq 1$.

Theorem 34. *Let X be a semimartingale on a filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses. Then there exists a probability Q which is equivalent to P such that under Q, X is a semimartingale in $\mathcal{H}^2$. Moreover, $\frac{dQ}{dP}$ can be taken to be bounded.*

Before we begin the proof we establish a useful lemma.

Lemma. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and let X_n be a sequence of a.s. finite valued random variables. There exists another probability Q, equivalent to P and with a bounded density, such that every X_n is in $L^2(dQ)$.

Proof. Assume without loss that each of the X_n is positive. For a single random variable X we take $\Lambda_k = \{k \leq X < k+1\}$ and $Y = \sum_{k\geq 1} 2^{-k}1_{\Lambda_k}$. Then Y is bounded and

$$E_Q\{X^2\} = E\{X^2Y\} \leq \sum_{k\geq 1} \frac{(k+1)^2}{2^k} < \infty.$$

For the general case, choose constants α_n such that $P(X_n > \alpha_n) \leq 2^{-n}$. The Borel-Cantelli Lemma implies that a.s. $X_n \leq \alpha_n$ for all n sufficiently large. Next choose constants c_n such that $\sum_{n\geq 1} c_n\alpha_n < \infty$. Let Y_n be the bounded

[7] $H \cdot X$ is still a semimartingale since it is of finite variation a.s. on compact time sets.

density chosen for X_n individually as done in the first part of the proof, and take $Y = \sum_{n\geq 1} c_n Y_n$ so that we have the result. □

Proof of Theorem 34. Using the lemma we make a first change of measure making all of the random variables $[X, X]_n$ integrable. Recall that $[X, X]$ is invariant under a change to an equivalent probability measure. By abusing notation, we will still denote this new measure by P. This implies that if $J_t = \sup_{s\leq t} |\Delta X_s|$, then $J_t^2 \leq [X, X]_t$ and thus J is locally in L^2. Hence it is also *a fortiori* locally in L^1, and thus X is special. We write its canonical decomposition $X = M + A$, with $X_0 = A_0$. We now make a second change to an equivalent law, this time called Q, again using the lemma, such that $\int_0^n |dA_s| \in L^2(dQ)$ for each n. Then X is special under Q as well, since each $[X, X]_n$ remains in L^1 for Q. Let $X = N + C$ denote its Q decomposition. We apply Girsanov's Theorem to get

$$C_t = A_t + \int_0^t \frac{1}{Z_{s-}} d\langle M, Z\rangle_s$$

where $Z_t = E_P\{Z|\mathcal{F}_t\}$ with $Z = \frac{dQ}{dP}$. For convenience we write this as $C_t = A_t + V_t$. We now need to establish that $E_Q\{(\int_0^t |dV_s|)^2\} = E_Q\{|V|_t^2\} < \infty$. Note that we have $E_Q\{V_t^2\} < \infty$, since $V_t = M_t - A_t$, and $A_t \in L^2(dQ)$. It is easy to check that $E_P\{[M, M]_n\} \leq 4E_P\{[X, X]_n\} < \infty$ by an argument similar to the proof of the lemma preceding Theorem 24. Finally $L^2(dP) \subset L^2(dQ)$. Since V is predictable, there exists a predictable process H taking values in $\{-1, 1\}$ such that $\int_0^t |dV_s| = \int_0^t H_s dV_s$. Setting $S = H \cdot X$, we have $[S, S] = [X, X]$, and the random variables $[S, S]_n$ are then integrable (dP). S has a decomposition $S = L + D$ where $L = H \cdot M$ and $D = H \cdot A$, hence under the law Q we have $L = H \cdot N + H \cdot V$, always using that the stochastic integral is invariant under a change to an equivalent probability measure. We deduce that $E_Q\{(\int_0^t H_s dV_s)^2\} < \infty$. Since $E_Q\{(\int_0^t H_s dV_s)^2\} = E_Q\{(\int_0^t |dV_s|)^2\}$, coupled with $|H \cdot A|_t = |A|_t$, the result follows. □

Corollary (Lenglart's Inclusion Theorem). Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfy the usual hypotheses, and suppose Q is an equivalent probability measure with a bounded density with respect to P. Then $\mathcal{H}^2(P) \subset \mathcal{H}^2(Q)$. That is, $\mathcal{H}^2$ with respect to P is contained in the space of $\mathcal{H}^2$ semimartingales with respect to Q.

Proof. Let X be an $\mathcal{H}^2$ semimartingale with respect to P. As we saw in the proof of Theorem 34, since Q has a bounded density, we obtain that $[X, X]$ is in $L^1(dQ)$, and hence X is special for Q. We then obtain a canonical decomposition of X and show both terms are in $\mathcal{H}^2$ analogously to how we did it in the proof of Theorem 34. □

3 Martingale Representation

In this section we will be concerned with martingales, rather than semimartingales. The question of martingale representation is the following. Given a collection $\mathcal{A}$ of martingales (or local martingales), when can *all* martingales (or all local martingales) be represented as stochastic integrals with respect to processes in $\mathcal{A}$? This question is surprisingly important in applications, and it is particularly interesting (in finance theory, for example) when $\mathcal{A}$ consists of just one element.

Throughout this section we assume as given an underlying complete, filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the usual hypotheses. We begin by considering only L^2 martingales. Later we indicate how to extend these results to locally square integrable local martingales.

Definition. The space $\mathbf{M}^2$ of L^2 **martingales** is all martingales M such that $\sup_t E\{M_t^2\} < \infty$, and $M_0 = 0$ a.s. Notice that if $M \in \mathbf{M}^2$, then $\lim_{t\to\infty} E\{M_t^2\} = E\{M_\infty^2\} < \infty$, and $M_t = E\{M_\infty|\mathcal{F}_t\}$. Thus each $M \in \mathbf{M}^2$ can be identified with its terminal value M_∞. We can endow $\mathbf{M}^2$ with a norm

$$\|M\| = E\{M_\infty^2\}^{1/2} = E\{[M,M]_\infty\}^{1/2},$$

and also with an inner product

$$(M,N) = E\{M_\infty N_\infty\},$$

for $M, N \in \mathbf{M}^2$. It is evident that $\mathbf{M}^2$ is a Hilbert space and that its dual space is also $\mathbf{M}^2$.

If $E\{M_t^2\} < \infty$ for each t, we call M a **square integrable martingale.** If in addition $\sup_t E\{M_t^2\} = E\{M_\infty^2\} < \infty$, then we call M an L^2 martingale.

The next definition is a key idea in the theory of martingale representation. It differs slightly from the customary definition because we are assuming all martingales are zero at time $t = 0$. If we did not have this hypothesis, we would have to add the condition that for any event $\Lambda \in \mathcal{F}_0$, any martingale M in the subspace F, then $M 1_\Lambda \in F$.

Definition. A closed subspace F of $\mathbf{M}^2$ is called a **stable subspace** if it is *stable under stopping* (that is, if $M \in F$ and if T is a stopping time, then $M^T \in F$).[8]

Theorem 35. *Let F be a closed subspace of $\mathbf{M}^2$. Then the following are equivalent.*

(a) F is closed under the following operation. For $M \in F$, $(M - M^t)1_\Lambda \in F$ for $\Lambda \in \mathcal{F}_t$, any $t \geq 0$.
(b) F is a stable subspace.

[8] Recall that $M_t^T = M_{t\wedge T}$, and $M_0 = 0$.

(c) If $M \in F$ and H is bounded, predictable, then $(\int_0^t H_s dM_s)_{t\geq 0} = H \cdot M \in F$.

(d) If $M \in F$ and H is predictable with $E\{\int_0^\infty H_s^2 d[M,M]_s\} < \infty$, then $H \cdot M \in F$.

Proof. Property (d) implies (c), and it is simple that (c) implies (b). To get (b) implies (a), let $T = t_\Lambda$, where

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

Then $T = t_\Lambda$ is a stopping time when $\Lambda \in \mathcal{F}_t$, and $(M - M^t)1_\Lambda = M - M^T$; since F is assumed stable, both M and M^T are in F. It remains to show only that (a) implies (d). Note that if H is simple predictable of the special form

$$H_t = 1_{\Lambda_0} 1_{\{0\}} + \sum_{i=1}^n 1_{\Lambda_i} 1_{(t_i, t_{i+1}]}$$

with $\Lambda_i \in \mathcal{F}_{t_i}$, $0 \leq i \leq n$, $0 = t_0 \leq t_1 \leq \cdots \leq t_{n+1} < \infty$, then $H \cdot M \in F$ whenever $M \in F$. Linear combinations of such processes are dense in $\mathbf{b}\mathbb{L}$ which in turn is dense in $\mathbf{b}\mathcal{P}$ under $d_M(\cdot,\cdot)$ by Theorem 2. But then $\mathbf{b}\mathcal{P}$ is dense in the space of predictable processes $\mathcal{H}$ such that $E\{\int_0^\infty H_s^2 d[M,M]_s\} < \infty$, as is easily seen (cf., Theorem 14). Therefore (a) implies (d) and the theorem is proved. □

Since the arbitrary intersection of closed, stable subspaces is still closed and stable, we can make the following definition.

Definition. Let $\mathcal{A}$ be a subset of $\mathbf{M}^2$. The **stable subspace generated** by $\mathcal{A}$, denoted $\mathcal{S}(\mathcal{A})$, is the intersection of all closed, stable subspaces containing $\mathcal{A}$.

As already noted on page 178, we can identify a martingale $M \in \mathbf{M}^2$ with its terminal value $M_\infty \in L^2$. Therefore another martingale $N \in \mathbf{M}^2$ is **(weakly) orthogonal** to M if $E\{N_\infty M_\infty\} = 0$. There is however another, stronger notion of orthogonality for martingales in $\mathbf{M}^2$.

Definition. Two martingales N, $M \in \mathbf{M}^2$ are said to be **strongly orthogonal** if their product $L = NM$ is a (uniformly integrable) martingale.

Note that if N, $M \in \mathbf{M}^2$ are strongly orthogonal, then NM being a (uniformly integrable) martingale implies that $[N, M]$ is also a local martingale by Corollary 2 of Theorem 27 of Chap. II. It is a uniformly integrable martingale by the Kunita-Watanabe inequality (Theorem 25 of Chap. II). Thus $M, N \in \mathbf{M}^2$ *are strongly orthogonal if and only if* $[M,N]$ *is a uniformly integrable martingale.* If N and M are strongly orthogonal then $E\{N_\infty M_\infty\} = E\{L_\infty\} = E\{L_0\} = 0$, so strong orthogonality implies orthogonality. The converse is not true however. For example let $M \in \mathbf{M}^2$, and let $Y \in \mathcal{F}_0$, independent of M, with $P(Y = 1) = P(Y = -1) = \frac{1}{2}$. Let $N_t = YM_t$, $t \geq 0$. Then $N \in \mathbf{M}^2$ and

$$E\{N_\infty M_\infty\} = E\{YM_\infty^2\} = E\{Y\}E\{M_\infty^2\} = 0,$$

so M and N are orthogonal. However $MN = YM^2$ is not a martingale (unless $M = 0$) because $E\{YM_t^2|\mathcal{F}_0\} = YE\{M_t^2|\mathcal{F}_0\} \neq 0 = YM_0^2$.

Definition. For a subset $\mathcal{A}$ of $\mathbf{M}^2$ we let $\mathcal{A}^\perp$ (resp. $\mathcal{A}^\times$) denote the set of all elements of $\mathbf{M}^2$ orthogonal (resp. strongly orthogonal) to each element of $\mathcal{A}$.

Lemma 1. If $\mathcal{A}$ is any subset of $\mathbf{M}^2$, then $\mathcal{A}^\times$ is (closed and) stable.

Proof. Let M^n be a sequence of elements of $\mathcal{A}^\times$ converging to M, and let $N \in \mathcal{A}$. Then $M^n N$ is a martingale for each n and $\mathcal{A}^\times$ will be shown to be closed if MN is also one, or equivalently that $[M, N]$ is a martingale. However

$$\begin{aligned} E\{|[M^n, N] - [M, N]_t|\} &= E\{|[M^n - M, N]_t|\} \\ &\leq (E\{[M^n - M, M^n - M]_t\})^{1/2}(E\{[N, N]_t\})^{1/2} \end{aligned}$$

by the Kunita-Watanabe inequalities. It follows that $[M^n, N]_t$ converges to $[M, N]_t$ in L^1, and therefore $[M, N]$ is a martingale, and $\mathcal{A}^\times$ is closed. Also $\mathcal{A}^\times$ is stable because $M \in \mathcal{A}^\times$, $N \in \mathcal{A}$ implies $[M^T, N] = [M, N]^T$ is a martingale and thus M^T is strongly orthogonal to N. □

Lemma 2. Let N, M be in $\mathbf{M}^2$. Then the following are equivalent.

(a) M and N are strongly orthogonal.
(b) $\mathcal{S}(M)$ and N are strongly orthogonal.
(c) $\mathcal{S}(M)$ and $\mathcal{S}(N)$ are strongly orthogonal.
(d) $\mathcal{S}(M)$ and N are weakly orthogonal.
(e) $\mathcal{S}(M)$ and $\mathcal{S}(N)$ are weakly orthogonal.

Proof. If M and N are strongly orthogonal, let $\mathcal{A} = \{N\}$ and then $M \in \mathcal{A}^\times$. Since $\mathcal{A}^\times$ is a closed stable subspace by Lemma 1, $\mathcal{S}(M) \subset \{N\}^\times$. Therefore (b) holds and hence (a) implies (b). The same argument yields that (b) implies (c). That (c) implies (e) which implies (d) is obvious. It remains to show that (d) implies (a).

Suppose N is weakly orthogonal to $\mathcal{S}(M)$. It suffices to show that $[N, M]$ is a martingale. By Theorem 21 of Chap. I it suffices to show $E\{[N, M]_T\} = 0$ for any stopping T. However $E\{[N, M]_T\} = E\{[N, M^T]_\infty\} = 0$, since N is orthogonal to M^T which is in $\mathcal{S}(M)$. □

Theorem 36. *Let $M^1, \dots, M^n \in \mathbf{M}^2$, and suppose M^i, M^j are strongly orthogonal for $i \neq j$. Then $\mathcal{S}(M^1, \dots, M^n)$ consists of the set of stochastic integrals*

$$H^1 \cdot M^1 + \cdots + H^n \cdot M^n = \sum_{i=1}^n H^i \cdot M^i,$$

where H^i is predictable and

$$E\{\int_0^\infty (H_s^i)^2 d[M^i, M^i]_s\} < \infty, \qquad 1 \le i \le n.$$

Proof. Let $\mathcal{I}$ denote the space of processes $\sum_{i=1}^n H^i \cdot M^i$, where H^i satisfy the hypotheses of the theorem. By Theorem 35 any closed, stable subspace must contain $\mathcal{I}$. It is simple to check that $\mathcal{I}$ is stable, so we need to show only that $\mathcal{I}$ is closed. Let

$$L_M^2 = \{H \in \mathcal{P} : E\{\int_0^\infty H_s^2 d[M, M]_s\} < \infty\}.$$

Then the mapping $(H^1, H^2, \ldots, H^n) \to \sum_{i=1}^n H^i \cdot M^i$ is an isometry from $L_{M^1}^2 \oplus \cdots \oplus L_{M^n}^2$ into $\mathbf{M}^2$. Since it is a Hilbert space isometry its image $\mathcal{I}$ is complete, and therefore closed. $\square$

Theorem 37. *Let $\mathcal{A}$ be a subset of $\mathbf{M}^2$ which is stable. Then $\mathcal{A}^\perp$ is a stable subspace, and if $M \in \mathcal{A}^\perp$ then M is strongly orthogonal to $\mathcal{A}$. That is, $\mathcal{A}^\perp = \mathcal{A}^\times$, and $\mathcal{S}(\mathcal{A}) = \mathcal{A}^{\perp\perp} = \mathcal{A}^{\times\perp} = \mathcal{A}^{\times\times}$.*

Proof. We first show that $\mathcal{A}^\perp = \mathcal{A}^\times$. Let $M \in \mathcal{A}$ and $N \in \mathcal{A}^\perp$. Since N is orthogonal to $\mathcal{S}(M)$, by Lemma 2, N and M are strongly orthogonal. Therefore $\mathcal{A}^\perp \subset \mathcal{A}^\times$. However clearly $\mathcal{A}^\times \subset \mathcal{A}^\perp$, whence $\mathcal{A}^\times = \mathcal{A}^\perp$, and thus $\mathcal{A}^\perp = \mathcal{A}^\times$ is a stable subspace by Lemma 1.

By the above applied to $\mathcal{A}^\perp$, we have that $(\mathcal{A}^\perp)^\perp = (\mathcal{A}^\perp)^\times$.

It remains to show that $\mathcal{S}(\mathcal{A}) = \mathcal{A}^{\perp\perp}$. Since $\mathcal{A}^{\perp\perp} = \overline{\mathcal{A}}$, the closure of $\mathcal{A}$ in $\mathbf{M}^2$, it suffices to show that $\overline{\mathcal{A}}$ is stable. However it is simple to check that condition (a) of Theorem 35 is satisfied for $\overline{\mathcal{A}}$, since it already is satisfied for $\mathcal{A}$, and we conclude $\overline{\mathcal{A}}$ is a stable subspace. $\square$

Corollary 1. Let $\mathcal{A}$ be a stable subspace of $\mathbf{M}^2$. Then each $M \in \mathbf{M}^2$ has a unique decomposition $M = A + B$, with $A \in \mathcal{A}$ and $B \in \mathcal{A}^\times$.

Proof. $\mathcal{A}$ is a closed subspace of $\mathbf{M}^2$, so each $M \in \mathbf{M}^2$ has a unique decomposition into $M = A + B$ with $A \in \mathcal{A}$ and $B \in \mathcal{A}^\perp$. However $\mathcal{A}^\perp = \mathcal{A}^\times$ by Theorem 37. $\square$

Corollary 2. Let M, $N \in \mathbf{M}^2$, and let L be the projection of N onto $\mathcal{S}(M)$, the stable subspace generated by M. Then there exists a predictable process H such that $L = H \cdot M$.

Proof. We know that such an L exists by Corollary 1. Since $\{M\}$ consists of just one element we can apply Theorem 36 to obtain the result. $\square$

Definition. Let $\mathcal{A}$ be finite set of martingales in $\mathbf{M}^2$. We say that $\mathcal{A}$ has the **(predictable) representation property** if $\mathcal{I} = \mathbf{M}^2$, where

$$\mathcal{I} = \{X : X = \sum_{i=1}^{n} H^i \cdot M^i, M^i \in \mathcal{A}\},$$

each H^i predictable such that

$$E\{\int_0^\infty (H_s^i)^2 d[M^i, M^i]_s\} < \infty, \quad 1 \le i \le n.$$

Corollary 3. Let $\mathcal{A} = \{M^1, \ldots, M^n\} \subset \mathbf{M}^2$, and suppose M^i, M^j are strongly orthogonal for $i \neq j$. Suppose further that if $N \in \mathbf{M}^2$, $N \perp \mathcal{A}$ in the strong sense implies that $N = 0$. Then $\mathcal{A}$ has the predictable representation property.

Proof. By Theorem 36 we have $\mathcal{S}(\mathcal{A}) = \mathcal{I}$. The hypotheses imply that $\mathcal{S}(\mathcal{A})^\perp = \{0\}$, hence $\mathcal{S}(\mathcal{A}) = \mathbf{M}^2$. □

Stable subspaces and predictable representation can be considered from an alternative perspective. Up to this point we have assumed as given and fixed an underlying space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, P)$, and a set of martingales $\mathcal{A}$ in $\mathbf{M}^2$. We will see that the property that $\mathcal{S}(\mathcal{A}) = \mathbf{M}^2$, intimately related to predictable representation (cf., Theorem 36), is actually a property of the probability measure P, considered as one element among the collection of probability measures that make L^2 $(\mathcal{F}_t)_{t\ge0}$ martingales of all the elements of $\mathcal{A}$.

Our first observation is that since the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge0}$ is assumed to be P complete, it is reasonable to consider only probability measures that are absolutely continuous with respect to P.

Definition. Let $\mathcal{A} \subset \mathbf{M}^2$. The set of $\mathbf{M}^2$ **martingale measures** for $\mathcal{A}$, denoted $\mathcal{M}^2(\mathcal{A})$, is the set of all probability measures Q defined on $\bigvee_{0\le t<\infty} \mathcal{F}_t$ such that

(i) $Q \ll P$,
(ii) $Q = P$ on $\mathcal{F}_0$,
(iii) if $X \in \mathcal{A}$ then $X \in \mathbf{M}^2(Q)$,

where $\mathbf{M}^2(Q)$ denotes all $(Q, \mathbb{F})$ L^2 martingales.[9]

Lemma. The set $\mathcal{M}^2$ is a convex set.

[9] Note that while $\mathbb{F}$ satisfies the usual hypotheses under P, $\mathcal{F}_0$ does not contain all of the Q-null sets, and thus $\mathbb{F}$ does not completely satisfy the usual hypotheses under Q. This does not create any problems, however.

Proof. Let Q and $R \in \mathcal{M}^2(\mathcal{A})$, and let $S = \lambda Q + (1-\lambda)R$, $0 < \lambda < 1$. Then for $X \in \mathcal{M}^2(\mathcal{A})$,

$$\sup_t E_S\{M_t^2\} = \sup_t [\lambda E_Q\{M_t^2\} + (1-\lambda)E_R\{M_t^2\}] < \infty,$$

since Q, $R \in \mathcal{M}^2(\mathcal{A})$. Also if $H \in \mathbf{b}\mathcal{F}_s$, $s < t$, then

$$\begin{aligned} E_S\{M_t H\} &= \lambda E_Q\{M_t H\} + (1-\lambda)E_R\{M_t H\} \\ &= \lambda E_Q\{M_s H\} + (1-\lambda)E_R\{M_s H\} \\ &= E_S\{M_s H\}, \end{aligned}$$

and $S \in \mathcal{M}^2(\mathcal{A})$. □

Definition. A measure $Q \in \mathcal{M}^2(\mathcal{A})$ is an **extremal point** of $\mathcal{M}^2(\mathcal{A})$ if whenever $Q = \lambda R + (1-\lambda)S$ with R, $S \in \mathcal{M}^2(\mathcal{A})$, $R \neq S$, $0 \leq \lambda \leq 1$, then $\lambda = 0$ or 1.

Theorem 38. *Let $\mathcal{A} \subset \mathbf{M}^2$. If $\mathcal{S}(\mathcal{A}) = \mathbf{M}^2$ then P is an extremal point of $\mathcal{M}^2(\mathcal{A})$.*

Proof. Suppose P is not extremal. We will show that $\mathcal{S}(\mathcal{A}) \neq \mathbf{M}^2$. Since P is not extremal, there exist $Q, R \in \mathcal{M}^2(\mathcal{A})$, $Q \neq R$, such that $P = \lambda Q + (1-\lambda)R$, $0 < \lambda < 1$. Let

$$L_\infty = \frac{dQ}{dP},$$

and let $L_t = E\{\frac{dQ}{dP}|\mathcal{F}_t\}$. Then $1 = \frac{dP}{dP} = \lambda L_\infty + (1-\lambda)\frac{dR}{dP} \geq \lambda L_\infty$ a.s., so that $L_\infty \leq \frac{1}{\lambda}$ a.s. Therefore L is a bounded martingale with $L_0 = 1$ (since $Q = P$ on $\mathcal{F}_0$), and thus $L - L_0$ is a nonconstant martingale in $\mathbf{M}^2(P)$. However, if $X \in \mathcal{A}$ and $H \in \mathbf{b}\mathcal{F}_s$, then X is a Q martingale and for $s < t$,

$$\begin{aligned} E_P\{X_t L_t H\} = E_P\{X_t L_\infty H\} = E_Q\{X_t H\} &= E_Q\{X_s H\} \\ &= E_P\{X_s L_\infty H\} \\ &= E_P\{X_s L_s H\}, \end{aligned}$$

and XL is a P martingale. Therefore $X(L - L_0)$ is a P martingale, and $L - L_0 \in \mathbf{M}^2$ and it is strongly orthogonal to $\mathcal{A}$. By Theorem 37 we cannot have $\mathcal{S}(\mathcal{A}) = \mathbf{M}^2$. □

Theorem 39. *Let $\mathcal{A} \subset \mathbf{M}^2$. If P is an extremal point of $\mathcal{M}^2(\mathcal{A})$, then every bounded P martingale strongly orthogonal to $\mathcal{A}$ is null.*

Proof. Let L be a bounded nonconstant martingale strongly orthogonal to $\mathcal{A}$. Let c be a bound for $|L|$, and set

$$dQ = (1 - \frac{L_\infty}{2c})dP \quad \text{and} \quad dR = (1 + \frac{L_\infty}{2c})dP.$$

We have Q, $R \in \mathcal{M}^2(\mathcal{A})$, and $P = \frac{1}{2}Q + \frac{1}{2}R$ is a decomposition that shows that P is not extremal which is a contradiction. □

Theorem 40. *Let $\mathcal{A} = \{M^1, \dots, M^n\} \subset \mathbf{M}^2$, with M^i continuous and M^i, M^j strongly orthogonal for $i \neq j$. Suppose P is an extremal point of $\mathcal{M}^2(\mathcal{A})$. Then*

(a) every stopping time is accessible;
(b) every bounded martingale is continuous;
(c) every uniformly integrable martingale is continuous; and
(d) $\mathcal{A}$ has the predictable representation property.

Proof. (a) Suppose T is a totally inaccessible stopping time and $P(T < \infty) > 0$. By Theorem 22 of Chap. III, there exists a martingale M with $\Delta M_T = 1_{\{T<\infty\}}$ and M continuous elsewhere. Moreover, the martingale M can be taken of finite variation with $M_0 = 0$. Therefore, since each M^i is continuous, $[M, M^i] = 0$. Moreover taking (for example)

$$R^n = \inf\{t > 0 : |M_t| > n\},$$

$|M^{R^n}| \le n + 1$ and thus M is locally bounded. Indeed, we then have M^{R^n} are bounded martingales strongly orthogonal to M^i. By Theorem 39 we have $M^{R^n} = 0$ for each n. Since $\lim_{n\to\infty} R^n = \infty$ a.s., we conclude $M = 0$, a contradiction.

(b) Let M be a bounded martingale which is not continuous, and assume $M_0 = 0$. Let $T^\varepsilon = \inf\{t > 0 : |\Delta M_t| > \varepsilon\}$. Then there exists $\varepsilon > 0$ such that for $T = T^\varepsilon$, $P\{|\Delta M_T| > 0\} > 0$. By part (a) the stopping time T is accessible, hence without loss we may assume that T is predictable. Therefore M^{T-} is a bounded martingale by the corollary to Theorem 29, whence $N = M^T - M^{T-} = \Delta M_T 1_{\{t \ge T\}}$ is also a bounded martingale. However N is a finite variation bounded martingale, hence $[N, M^i] = 0$ each i. That is, N is a bounded martingale strongly orthogonal to $\mathcal{A}$. Hence $N = M^T - M^{T-} = 0$ by Theorem 39, and we conclude that M is continuous.

(c) Let M be a uniformly integrable martingale closed by M_∞. Define

$$M_t^n = E\{M_\infty 1_{\{|M_\infty| \le n\}} | \mathcal{F}_t\}.$$

Then M^n are bounded martingales and therefore continuous by part (b). However

$$P\{\sup_t |M_t^n - M_t| > \varepsilon\} \le \frac{1}{\varepsilon} E\{|M_\infty^n - M_\infty|\}$$

by an inequality of Doob[10], and the right side tends to 0 as n tends to ∞. Therefore there exists a subsequence (n_k) such that $\lim_{k\to\infty} M_t^{n_k} = M_t$ a.s., uniformly in t. Thus M is continuous.

(d) By Corollary 3 of Theorem 37 it suffices to show that if $N \in \mathcal{A}^\times$ then $N = 0$. Suppose $N \in \mathcal{A}^\times$. Then N is continuous by (c). Therefore N is locally bounded, and hence by stopping, N must be 0 by Theorem 39. □

[10] See, for example, Breiman [23, page 88].

The next theorem allows us to consider subspaces generated by countably infinite collections of martingales.

Theorem 41. *Let $M \in \mathbf{M}^2$, $Y^n \in \mathbf{M}^2$, $n \geq 1$, and suppose Y^n_∞ converges to Y_∞ in L^2, and that there exists a sequence $H^n \in L(M)$ such that $Y^n_t = \int_0^t H^n_s dM_s$, $n \geq 1$. Then there exists a predictable process $H \in L(M)$ such that $Y_t = \int_0^t H_s dM_s$.*

Proof. If Y^n_∞ converges to Y_∞ in L^2, then Y^n converges to Y in $\mathbf{M}^2$. By Theorem 36 we have that $\mathcal{S}(M) = \mathcal{I}(M)$, the stochastic integrals with respect to M. Moreover $Y^n \in \mathcal{S}(M)$, each n. Therefore Y is in the closure of $\mathcal{S}(M)$; but $\mathcal{S}(M)$ is closed, so $Y \in \mathcal{S}(M) = \mathcal{I}(M)$, and the theorem is proved. □

Theorem 42. *Let $\mathcal{A} = \{M^1, M^2, \dots, M^n, \dots\}$, with $M^i \in \mathbf{M}^2$, and suppose there exist disjoint predictable sets Λ^i such that $1_{\Lambda^i} d[M^i, M^i] = d[M^i, M^i]$, $i \geq 1$. Let $A_t = \sum_{i=1}^\infty \int_0^t 1_{\Lambda_i}(s) d[M^i, M^i]_s$. Suppose that*

(a) $E\{A_\infty\} < \infty$; and
(b) for $\mathcal{F}^i \subset \mathcal{F}_\infty$ such that for any $X^i \in \mathbf{b}\mathcal{F}^i$, we have $X^i_t = E\{X^i|\mathcal{F}_t\} = \int_0^t H^i_s dM^i_s$, $t \geq 0$, for some predictable process H^i.

Then $M = \sum_{i=1}^\infty M^i$ exists and is in $\mathbf{M}^2$, and for any $Y \in \mathbf{b}\bigvee_i \mathcal{F}^i$, if $Y_t = E\{Y|\mathcal{F}_t\}$, we have that $Y_t = \int_0^t H_s dM_s$, for the martingale $M = \sum_{i=1}^\infty M^i$ and for some $H \in L(M)$.

Proof. Let $N^n = \sum_{i=1}^n M^i$. Then $[N^n, N^n]_t = \sum_{i=1}^n \int_0^t 1_{\Lambda_i}(s) d[M^i, M^i]_s$, hence $E\{(N^n_\infty)^2\} = E\{[N^n, N^n]_\infty\} \leq E\{A_\infty\}$, and N^n is Cauchy in $\mathbf{M}^2$ with limit equal to M. By hypothesis we have that if $X^i \in \mathbf{b}\mathcal{F}^i$ then

$$\begin{aligned} E\{X^i|\mathcal{F}_t\} = X^i_t = \int_0^t H^i_s dM^i_s &= \int_0^t 1_{\Lambda^i} H^i_s dM^i_s \\ &= \int_0^t 1_{\Lambda^i} H^i_s d(M^i_s + \sum_{i \neq j} M^j_s) \\ &= \int_0^t 1_{\Lambda^i} H^i_s dM_s. \end{aligned}$$

Therefore if $i \neq j$ we have that

$$\begin{aligned} [X^i, X^j] &= \int_0^t 1_{\Lambda^i} H^i_s 1_{\Lambda^j} H^j_s d[M, M]_s \\ &= \int_0^t 1_{\Lambda^i \cap \Lambda^j} H^i_s H^j_s d[M, M]_s \\ &= 0, \end{aligned}$$

since $\Lambda^i \cap \Lambda^j = \emptyset$, by hypothesis. However using integration by parts we have

$$\begin{aligned}
X_t^i X_t^j &= \int_0^t X_{s-}^i dX_s^j + \int_0^t X_{s-}^j dX_s^i + [X^i, X^j]_t \\
&= \int_0^t X_{s-}^i dX_s^j + \int_0^t X_{s-}^j dX_s^i \\
&= \int_0^t X_{s-}^i 1_{A^j} H_s^j dM_s + \int_0^t X_{s-}^j 1_{A^i} H_s^i dM_s \\
&= \int_0^t H_s^{i,j} dM_s,
\end{aligned}$$

where $H^{i,j}$ is defined in the obvious way. By iteration we have predictable representation for all finite products $\prod_{i\leq n} X^i$. The Monotone Class Theorem together with Theorem 41 then yields the result. □

We conclude this section by applying these results to a very important special case, namely n-dimensional Brownian motion. This example shows how these results can be easily extended to locally square integrable local martingales.

Theorem 43. *Let $X = (X^1, \ldots, X^n)$ be an n-dimensional Brownian motion and let $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq\infty}$ denote its completed natural filtration. Then every locally square integrable local martingale M for $\mathbb{F}$ has a representation*

$$M_t = M_0 + \sum_{i=1}^n \int_0^t H_s^i dX_s^i,$$

where H^i is (predictable, and) in $L(X^i)$.

Proof. Fix t_0, $0 < t < t_0$, and assume X is stopped at t_0. Then letting $\mathcal{A} = \{X^1, \ldots, X^n\}$, we have that $\mathcal{A} \subset \mathbf{M}^2$. Let $\mathcal{M}^2(\mathcal{A})$ be all probability measures Q such that $Q \ll P$, and under Q all of $\mathcal{A} \subset \mathbf{M}^2(Q)$.[11] Since $[X^i, X^j]$ is the same process under Q as under P we have by Lévy's Theorem (Theorem 40 of Chap. II) that $X = (X^1, \ldots, X^n)$ is a Q-Brownian motion as well. Therefore, for bounded Borel functions $f_1, \ldots, f_n$, we have that for $t_1 < t_2 \cdots < t_n \leq t_0$, $f_1(X_{t_1}) \prod_{i=1}^{n-1} f_i(X_{t_{i+1}} - X_{t_i})$ have the same expectation for Q as they do for P. Thus $P = Q$ on $\mathcal{F}_{t_0}$, and we conclude that $\mathcal{M}^2(\mathcal{A})$ is the singleton $\{P\}$. The probability law P is then trivially extremal for $\mathcal{M}^2(\mathcal{A})$, and therefore by Theorem 40 we know that $\mathcal{A}$ has the predictable representation property.

Suppose M is locally square integrable and $M_t = M_{t\wedge t_0}$. Let $(T^k)_{k\geq 0}$ be a sequence of stopping times increasing to ∞ a.s., $T_0 = 0$, and such that $M^{T^k} \in \mathbf{M}^2$, each $k \geq 0$. Then by the preceding there exist processes $H^{i,k}$ such that

$$M^{T^k} = \sum_{i=1}^n H^{i,k} \cdot X^i, \quad \text{each } k \geq 0.$$

[11] Recall that $\mathbf{M}^2(Q)$ denotes all L^2 Q martingales.

By defining $H^i = \sum_{k=1}^{\infty} H^{i,k} 1_{(T^{k-1},T^k]}$, we have that $H^i \in L(X^i)$ and also

$$M = \sum_{i=1}^{n} H^i \cdot X^i.$$

Next let $(t_\ell)_{\ell \geq 0}$ be fixed times increasing to ∞ with $t_0 = 0$. For each t_ℓ we know there exist $H^{i,\ell}$ such that for a given locally square integrable local martingale N

$$N_{t \wedge t_\ell} = \sum_{i=1}^{n} (H^{i,\ell} \cdot X^i)_{t \wedge t_\ell}.$$

By defining $H^i = \sum_{\ell=1}^{\infty} H^{i,\ell} 1_{(t_{\ell-1},t_\ell]}$ one easily checks that $H^i \in L(X^i)$ and that

$$N_t = \sum_{i=1}^{n} H^i \cdot X^i. \qquad \square$$

Corollary 1. As in Theorem 43, let $\mathbb{F}$ be the completed natural filtration of an n-dimensional Brownian motion. Then every local martingale M for $\mathbb{F}$ is continuous.

Proof. In the proof of Theorem 43 we saw that the underlying probability law P is extremal for $\mathcal{A} = \{X^1, \ldots, X^n\}$. Therefore by Theorem 40(c), every uniformly integrable martingale is continuous. The corollary follows by stopping. $\square$

Corollary 2. Let $X = (X^1, \ldots, X^n)$ be an n-dimensional Brownian motion and let $\mathbb{F}$ be its completed natural filtration. Then every local martingale M for $\mathbb{F}$ has a representation

$$M_t = M_0 + \sum_{i=1}^{n} \int_0^t H_s^i dX_s^i$$

where H^i are predictable.

Proof. By Corollary 1 any local martingale M is continuous, hence it is locally square integrable. It remains only to apply Theorem 43. $\square$

Corollary 3. Let $X = (X^1, \ldots, X^n)$ be an n-dimensional Brownian motion and let $\mathbb{F}$ be its completed natural filtration. Let $Z \in \mathcal{F}_\infty$ be in L^1. Then there exist H^i predictable in $L(X^i)$ with $\int_0^\infty (H_s^i)^2 ds < \infty$ a.s. such that

$$Z = E\{Z\} + \sum_{i=1}^{n} \int_0^\infty H_s^i dX_s^i.$$

Proof. Let $Z_t = E\{Z|\mathcal{F}_t\}$, taking the càdlàg (and hence continuous) version. By Corollary 2 we have

$$Z_t = Z_0 + \sum_{i=1}^{n} \int_0^t H_s^i dX_s^i.$$

By Theorem 42 of Chap. II we have that $Z_t = B_{[Z,Z]_t}$ for some Brownian motion B, $0 < t < \infty$. Letting t tend to ∞ shows that $Z_\infty = B_{\lim_{t\to\infty}[Z,Z]_t}$, which shows that $[Z,Z]_\infty < \infty$ a.s. However,

$$\sum_{i=1}^{n} \int_0^\infty (H_s^i)^2 ds = \lim_{t\to\infty} \sum_{i=1}^{n} \int_0^t (H_s^i)^2 ds = \lim_{t\to\infty} [Z,Z]_t = [Z,Z]_\infty < \infty \quad \text{a.s.}$$

Finally, take $t = \infty$, observe that $Z_\infty = Z$ and that $\mathcal{F}_0$ is a.s. trivial. Hence Z_0 is constant and therefore $Z_0 = E\{Z\}$. □

Corollary 4. Let $X = (X^1, \dots, X^n)$ be an n-dimensional Brownian motion and let $\mathbb{F}$ be its completed natural filtration. Let $Z \in L^1(\mathcal{F}_\infty)$ and $Z > 0$ a.s. Then there exist J^i predictable with $\int_0^\infty (J_s^i)^2 ds < \infty$ a.s. such that

$$Z = E\{Z\} \exp\left\{ \sum_{i=1}^{n} \int_0^\infty J_s^i dX_s^i - \sum_{i=1}^{n} \frac{1}{2} \int_0^\infty (J_s^i)^2 ds \right\}.$$

Proof. By Corollary 3 there exist predictable H^i such that if $Z_t = E\{Z|\mathcal{F}_t\}$, then

$$Z_t = E\{Z\} + \sum_{i=1}^{n} \int_0^t H_s^i dX_s^i.$$

Therefore

$$\log(Z_t) = \log(Z_0) + \sum_{i=1}^{n} \int_0^t \frac{1}{Z_s} H_s^i dX_s^i - \frac{1}{2} \sum_{i=1}^{n} \int_0^t \frac{1}{Z_s^2} (H_s^i)^2 ds.$$

(Note that since $(Z_t)_{t\geq 0}$ is continuous and never 0, $\frac{1}{Z}$ is locally bounded.) The proof is completed by setting $J_s^i = \frac{1}{Z_s} H_s^i$ and taking exponentials of both sides. □

Corollary 5. Let $\mathbb{F}$ be the completed natural filtration of an n-dimensional Brownian motion. If T is a totally inaccessible stopping time, then $T = \infty$ a.s.

Proof. This is merely Theorem 40 (a). □

We end this section with a quite general martingale representation result, which while often true, nevertheless requires a countable number of martingales to have the representation. We then apply it to give a condition for compensators to have absolutely continuous paths. But first we state a definition from measure theory that is not widely known.

Definition. A measurable space $(\Omega, \mathcal{F})$ is said to be **separable** if there exists a countable family of functions on (or subsets of) Ω which generate the σ-algebra $\mathcal{F}$.

Theorem 44. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered complete probability space satisfying the usual hypotheses. Assume also that the σ-algebra $\mathcal{F} = \mathcal{F}_\infty$ and that it is separable. Then there exists a countable sequence of martingales $(M_1, M_2, \ldots)$ in L^2 such that they are orthogonal, $\sum_{i \geq 1} E\{[M^i, M^i]_\infty\} < \infty$, and such that if N is any L^2 martingale, then there exists a sequence of predictable processes $(H^1, H^2, \ldots)$ such that $N_t = \sum_{i \geq 1} \int_0^t H_s^i dM_s^i$. That is, the martingales $\{M^i, i \geq 1\}$ have martingale representation.*

Proof. Since $\mathcal{F}$ is separable, there exists a countable basis of L^2, call it $(M^0, M^1, M^2, \ldots)$, where M^0 is constant and $E\{M^i\} = 0$ for $i \neq 0$. One can take such a basis to be orthonormal in the usual way. That is, $E\{M^i M^j\} = 0$ for $i \neq j$, and $E\{(M^i)^2\} = 1$ for all i. If one then multiplies M^i by 2^{-i} and defines the martingales $M_t^i = E\{M^i | \mathcal{F}_t\}$, we have the construction. The space of stochastic integrals with respect to all of the M^i where the sum converges in L^2 is easily seen to be a stable subspace. Next let N be any martingale in L^2 with $N_0 = 0$. Then we have

$$N_t = \sum_{i \geq 1} \int_0^t H_s^i dM_s^i + L_t$$

where L is orthogonal to the stable subspace generated by the countable collection of martingales $(M^i)_{i \geq 1}$. This implies that L_∞ is orthogonal to all of the random variables $M_\infty^i = M^i$. Since these random variables form a basis of L^2, this implies that L must be constant. Since $L_0 = 0$ we have that L is identically zero. □

Before we apply Theorem 44 to compensators, we need to recall and formalize two definitions, and prove Lebesgue's change of time formula. The first definition we have already seen in Exercise 7 of Chap. III, and the second one we saw in Theorem 42 of Chap. II as well as Exercises 31 through 36 of that chapter.

Definition. A (right continuous) filtration $\mathbb{F}$ on a filtered probability space satisfying the usual hypotheses is called **quasi left continuous** if for every *predictable* stopping time T one has $\mathcal{F}_T = \mathcal{F}_{T-}$.

Note that if T is a predictable stopping time, and if T_n is an announcing sequence of stopping times for T, and if $X \in L^1$, then $\lim_{n\to\infty} E\{X | \mathcal{F}_{T_n}\} = E\{X | \mathcal{F}_{T-}\}$. If M is a uniformly integrable martingale, then

$$E\{M_T | \mathcal{F}_{T-}\} = \lim_{n\to\infty} E\{M_T | \mathcal{F}_{T_n}\} = \lim_{n\to\infty} M_{T_n} = M_{T-}.$$

Therefore $E\{\Delta M_T | \mathcal{F}_{T-}\} = 0$ a.s. and if ΔM_T is $\mathcal{F}_{T-}$ measurable, it must be 0 almost surely. So if $\mathbb{F}$ is quasi left continuous, no martingales can jump at predictable times. Since the accessible part of any stopping time can be covered by a countable sequence of predictable times, we conclude that *if* $\mathbb{F}$ *is quasi left continuous then martingales jump only at totally inaccessible stopping times.*

Definition. Let $A = (A_t)_{t\geq 0}$ be an adapted, right continuous increasing process, which need not always be finite-valued. The **change of time** (also known as a **time change**) associated to A is the process

$$\tau_t = \inf\{s > 0 : A_s > t\}.$$

Some observations are in order. We have that $t \mapsto \tau_t$ is non-decreasing and hence τ_{t-} exists. Also, since $\{A_t > s\} = \bigcup_{\varepsilon>0}\{A_t > s+\varepsilon\}$, we have that $t \mapsto \tau_t$ is right continuous. It is continuous if A has strictly increasing paths. Moreover $A_{\tau_t} \geq A_{\tau_{t-}} \geq t$, and $A_{(\tau_{t-})-} \leq A_{(\tau_t)-} \leq t$. (Here we use the convention $\tau_{0-} = 0$.) Finally note that $\{\tau_{s-} \leq t\} = \{s \leq A_t\}$, which implies that τ_{t-} is a stopping time, and since $\tau_t = \lim_{\varepsilon\to 0}\tau_{(t+\varepsilon)-}$ we conclude that τ_t is also a stopping time.

Theorem 45 (Lebesgue's Change of Time Formula). *Let a be a positive, finite, right continuous, increasing function on $[0,\infty)$. Let c denote its right continuous inverse (change of time). Let f be a positive Borel function on $[0,\infty)$. If G is any positive, finite, right continuous function on $[0,\infty)$ with $G(0-)=0$, then*

$$\int_0^\infty f(s)dG(a(s)) = \int_0^\infty f(c(s-))1_{\{c(s-)<\infty\}}dG(s),$$

and in particular

$$\int_0^\infty f(s)da(s) = \int_0^\infty f(c(s))1_{\{c(s)<\infty\}}ds.$$

Here the integrals are taken over the set $[0,\infty)$.

Proof. First consider f of the form $f(s) = 1_{[0,u]}(s)$, for $0 \leq u < \infty$. The left side of the equation then reduces to $G(a(u))$. For the right side note that $f(c(s-))1_{\{c(s-)<\infty\}} = 1_{\{c(s-)\leq u\}} = 1_{\{s\leq a(u)\}}$ and it follows that the right side is also equal to $G(a(u))$. By subtraction and linearity we also have the result for f of the form $1_{(u,v]}$. The Monotone Class Theorem gives the result for any f positive with compact support, and approximations then gives the general case. If we take $G(s) = s$ we get $\int_0^\infty f(s)da(s) = \int_0^\infty f(c(s-))1_{\{c(s-)<\infty\}}ds$, and since c only jumps at most countably often because it is increasing, and since ds does not charge countable sets, we have $\int_0^\infty f(c(s-))1_{\{c(s-)<\infty\}}ds = \int_0^\infty f(c(s))1_{\{c(s)<\infty\}}ds$ and the second statement is proved, whence the theorem. □

Corollary. Let a be a positive, finite, continuous, strictly increasing function on $[0, \infty)$. Let c denote its continuous inverse (change of time). Let f be a positive Borel function on $[0, \infty)$. Then

$$\int_0^{c(t)} f(s)da(s) = \int_0^t f(c(s))ds.$$

Proof. It suffices to rewrite the left side of the equation as

$$\int_0^\infty 1_{[0,c(t)]}(s)f(s)da(s)$$

and observe that c is also continuous and strictly increasing, which implies $1_{[0,c(t)]}(c(s)) = 1_{[0,t]}(s)$. □

Our goal, stated informally in words, is if a filtration is quasi left continuous, then modulo a change of time, all compensators of adapted counting processes with totally inaccessible jumps have paths which are absolutely continuous. This is achieved in Theorem 47. We begin with two definitions.

Definition. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses, and let $\mathcal{S}$ denote the space of all square integrable martingales with continuous paths a.s. It is easy to see that $\mathcal{S}$ is a stable subspace.[12] For an arbitrary square integrable martingale M, let M^c denote the orthogonal projection of M onto $\mathcal{S}$. Then M^c is called the **continuous martingale part** of M. If we write $M = M^c + (M - M^c)$ as its orthogonal decomposition, then we have $M^d = (M - M^c)$ where M^d is called the **purely discontinuous part** of the martingale M.

Note that if a local martingale is locally square integrable, we can extend the definition of continuous part and purely discontinuous part trivially, by stopping. Also note that the term "purely discontinuous" is misleading: it is *not* a description of a path property of a martingale, but rather simply refers to the property of being orthogonal to the stable subspace of continuous martingales. See Exercises 6 and 7 in this regard.

Definition. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses. We call it an **absolutely continuous space** if for any purely discontinuous locally square integrable martingale M, $d\langle M, M\rangle_t \ll dt$.

Theorem 46. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be an absolutely continuous space. Then the compensators for all adapted counting processes with totally inaccessible jump times and without explosions, are absolutely continuous.*

Proof. Let N be an adapted counting process and let $\widetilde{N}$ be its compensator, so that $X = N - \widetilde{N}$ is a locally square integrable local martingale. Since $\widetilde{N}$ is continuous we have $[X, X]_t = \sum_{s \leq t}(\Delta N_s)^2 = N_t$, since $(\Delta N_s)^2 = \Delta N_s = 1$

[12] See Exercise 6.

when N jumps at s. Therefore $\widetilde{N} = \langle X, X\rangle$, and since $\langle X, X\rangle$ has absolutely continuous paths by hypothesis, we are done. □

Remark (Yan Zeng). With a little more work, one can show the following: Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfy the usual hypotheses. If it is an absolutely continuous space, then all jump times of square integrable martingales are totally inaccessible. Moreover it is an absolutely continuous space if and only if for every totally inaccessible stopping time T, if $N_t = 1_{\{t\geq T\}}$, then $d\widetilde{N}_t \ll dt$ a.s.

If we take the uninteresting counting process $N_t = \sum_{i\geq 1} 1_{\{t\geq i\}}$ which jumps at constant (and hence predictable) times, then its compensator is simply $\widetilde{N}_t = N_t$, and the martingale $X = N - \widetilde{N}$ is identically zero. To avoid these trivialities we assume that the process $\widetilde{N}$ is continuous. Note however that $\widetilde{N}$ is not *a priori* absolutely continuous. This is to explain our hypotheses in the next theorem:

Theorem 47 (Absolutely Continuous Compensators). *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses. Assume that the σ-algebra $\mathcal{F} = \mathcal{F}_\infty$ and that it is separable.*[13] *Assume further that $\mathbb{F}$ is quasi left continuous. To prevent trivialities assume also that it is not an absolutely continuous space. Then there exists a change of time τ_t and a new filtration $\mathbb{G}$ given by $\mathcal{G}_t = \mathcal{F}_{\tau_t}$, such that if N is a $\mathbb{G}$ adapted counting process with totally inaccessible jump times and without explosions, then $d\widetilde{N}_t \ll dt$.*

Proof. Let $(M^i)_{i\geq 1}$ be the countable sequence constructed in Theorem 44. Define the strictly increasing process $A_t = t + \sum_{i\geq 1}[M^i, M^i]_t$, and note that $E\{A_t\} < \infty$ for $0 \leq t < \infty$. Let C denote the compensator of A. One can check that $C_t = t + \sum_{i\geq 1}\langle M^i, M^i\rangle_t$. Then C is both strictly increasing and continuous, since $\mathbb{F}$ is quasi left continuous. Let

$$\tau_t = \inf\{s > 0 : C_s > t\} \quad \text{and then} \quad C_t = \inf\{s > 0 : \tau_s > t\},$$

where τ is the time change of the theorem statement.

Let N be a $\mathbb{G}$ adapted counting process and let $X = N - \widetilde{N}$. Assume, by stopping if necessary, that $X_\infty \in L^2$. We write $\langle X, X\rangle^{\mathbb{G}}$ to denote the sharp bracket taken using the $\mathbb{G}$ filtration. Note that $L_t = N_{C_t}$ is also a counting process, and since τ continuous implies that $\mathcal{G}_{C_t} \subset \mathcal{F}_t$ (see Exercise 38 of Chap. II), we conclude L is adapted to $\mathbb{F}$. We then have $[L, L]_t = \sum_{s\leq t}(\Delta L_s)^2 = L_t$, hence $\widetilde{L}_t = \langle L, L\rangle_t^{\mathbb{F}}$. But also $[X, X]_t = N_t$, and we have

$$[X, X]_t - \langle X, X\rangle_t^{\mathbb{G}} = N_t - \langle X, X\rangle_t^{\mathbb{G}}.$$

Hence

$$N_{C_t} - \langle X, X\rangle_{C_t}^{\mathbb{G}} = L_t - \langle X, X\rangle_{C_t}^{\mathbb{G}},$$

[13] A separable σ-algebra is defined on page 189.

and by the uniqueness of the compensator of L we have $\langle X,X\rangle^{\mathbb{G}}_{C_t} = \langle L,L\rangle^{\mathbb{F}}_t$. On the other hand, by Theorem 44 we know that $L_t - \widetilde{L}_t = \sum_{i\geq 1}\int_0^t H^i_s dM^i_s$ for some predictable processes (H^i), and thus

$$\langle L-\widetilde{L},L-\widetilde{L}\rangle^{\mathbb{F}}_t = \sum_{i\geq 1}\int_0^t (H^i_s)^2 d\langle M^i,M^i\rangle^{\mathbb{F}}_s = \sum_{i\geq 1}\int_0^t (H^i_s)^2 c^i_s dC_s$$
$$= \int_0^t J_s dC_s$$

for some (predictable) nonnegative processes c^i whose existence is assured because $d\langle M^i,M^i\rangle_t \ll dC_t$ a.s. The process $J_s = \sum_{i\geq 1}(H^i_s)^2 c^i_s$, by Fubini's Theorem. Next, note that since $\langle L-\widetilde{L},L-\widetilde{L}\rangle^{\mathbb{F}}_t = \langle X,X\rangle^{\mathbb{G}}_{C_t}$, we have $\langle X,X\rangle^{\mathbb{G}}_t = \langle L-\widetilde{L},L-\widetilde{L}\rangle^{\mathbb{F}}_{\tau_t} = \int_0^{\tau_t} J_s dC_s$, which by the corollary of Lebesgue's change of time formula, implies $\langle X,X\rangle^{\mathbb{G}}_t = \int_0^t J_{\tau_s} ds$. □

4 Martingale Duality and the Jacod-Yor Theorem on Martingale Representation

We have already defined the space $\mathcal{H}^2$ for semimartingales; we use the same definition for martingales, which in fact historically came before the definition for semimartingales. We make the definition for all p, $p \geq 1$.

Definition. Let M be a local martingale with càdlàg paths. We define the $\mathcal{H}^p$ norm of M to be $\|M\|_p = E\{[M,M]^{p/2}_\infty\}^{1/p}$. If $\|M\|_{\mathcal{H}^p} < \infty$, we say that M is in $\mathcal{H}^p$. Finally we call $\mathcal{H}^p$ the space of all càdlàg martingales with finite $\mathcal{H}^p$ norm.

It is simple to verify that $\|\cdot\|_{\mathcal{H}^p}$ is a norm if we identify martingales that are almost surely equal. Indeed, the only property that is not obvious is the subadditivity. For the case $p = 1$ for example, we have that $[M+N,M+N]^{1/2} \leq [M,M]^{1/2} + [N,N]^{1/2}$ since by the Kunita-Watanabe inequality $[M,N] \leq [M,M]^{1/2}[N,N]^{1/2}$, which gives the subadditivity. It is also simple to verify that $\mathcal{H}^p$ is a linear space. We state without proof the important martingale inequalities known as the Burkholder-Davis-Gundy inequalities, and often referred to by the acronym BDG. (For a proof one can consult many different sources, including [146] and [44].)

Theorem 48 (Burkholder-Davis-Gundy Inequalities). *Let M be a martingale with càdlàg paths and let $p \geq 1$ be fixed. Let $M^*_t = \sup_{s\leq t}|M_s|$. Then there exist constants c_p and C_p such that for any such M*

$$E\{[M,M]^{\frac{p}{2}}_t\}^{\frac{1}{p}} \leq c_p E\{(M^*_t)^p\}^{\frac{1}{p}} \leq C_p E\{[M,M]^{\frac{p}{2}}_t\}^{\frac{1}{p}}$$

for all t, $0 \leq t \leq \infty$. The constants are universal: they do not depend on the choice of M.

We have the following simple results concerning $\mathcal{H}^1$.

Theorem 49. *$\mathcal{H}^2 \subset \mathcal{H}^1$ and local martingales of integrable variation are a subset of $\mathcal{H}^1$.*

Proof. First note that if M is a càdlàg martingale, then $E\{\sqrt{[M,M]_\infty}\} \leq (E\{[M,M]_\infty\})^{1/2}$ which gives the first statement. For the second, if M has integrable variation then $[M,M]_\infty = \sum_s (\Delta M_s)^2 \leq (\sum_s |\Delta M_s|)^2$, whence $\sqrt{[M,M]_\infty} \leq \int_0^\infty |dM_s|$, and taking expectations, we have the result. □

Theorem 50. *$\mathcal{H}^2$ is dense in $\mathcal{H}^1$ and bounded martingales are dense in $\mathcal{H}^1$.*

Proof. Let M be a local martingale in $\mathcal{H}^1$. By the Fundamental Theorem of Local Martingales (Theorem 25 of Chap. III) we have $M = N + U$ where N is a local martingale with jumps bounded by 1 and U has paths of locally integrable variation. If $(T_n)_{n\geq 1}$ is a sequence of stopping times increasing to ∞ a.s., then $\|M^{T_n}\|_{\mathcal{H}^1} \leq \|M\|_{\mathcal{H}^1}$, and also M^{T_n} converges to M in $\mathcal{H}^1$. Choose T_n to be the first time $[U,U]_t$ has total variation larger than n, and also $|[N,N]^{T_n}| \leq n-1$; that is, $T_n = \inf\{t \geq 0 : \int_0^t [U,U]_t > n, \text{ or } [N,N]_t > n-1\}$ with T_n increasing to ∞. Then $[N,N]^{T_n} \leq n$ and thus N^{T_n} is in $\mathcal{H}^1$ for each n. Therefore we turn out attention to U^{T_n}, and since $[U,U]^{T_n} \leq n + |\Delta U_{T_n}|$, we need consider only the one jump of U at time T_n : ΔU_{T_n}. Let ξ denote ΔU_{T_n}. Letting $\xi_k = \xi 1_{\{|\xi|\leq k\}}$, since ξ_k is bounded we can compensate $\xi_k 1_{\{t\geq T_n\}}$, and call the resulting martingale V_k^n. The compensator of $\xi_k 1_{\{t\geq T_n\}}$ is in $\mathcal{H}^2$ because ξ_k is bounded. We have then that V_n^k is in $\mathcal{H}^2$ and converges to V_n as $k \to \infty$ in total variation norm, and hence also in $\mathcal{H}^1$. This proves that $\mathcal{H}^2$ is dense in $\mathcal{H}^1$, and since bounded martingales are dense in $\mathcal{H}^2$, they too are dense in $\mathcal{H}^1$. □

Theorem 51. *Let M be a local martingale. Then M is locally in $\mathcal{H}^1$.*

Proof. By the Fundamental Theorem of Local Martingales we know that $M = N + U$, where N has jumps bounded by a constant β and U is locally of integrable variation. By stopping, we thus assume that N is bounded and U has paths of integrable variation. The result then follows by the previous theorem (Theorem 49). □

One can further show, by identifying $\mathcal{H}^p$ with L^p through the terminal value of the martingale, that $\mathcal{H}^p$ is a Banach space (in particular it is complete) for each $p \geq 1$. One can further show that the dual space of continuous linear functional on $\mathcal{H}^p$ is $\mathcal{H}^q$, where $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$. This continues nicely the analogy with L^p *except for the case* $p = 1$. It can be shown (see [47]) that the dual of $\mathcal{H}^1$ *is not* $\mathcal{H}^\infty$. It turns out that a better analogy than L^p is that of Hardy spaces in complex analysis, where C. Fefferman showed that the dual of $\mathcal{H}^1$ can be identified with the space of functions of bounded mean oscillation, known by its acronym as BMO. With this in mind we define the space of BMO martingales.

Definition. Let M be a local martingale. M is said to be in ***BMO*** if M is in $\mathcal{H}^2$ and if there exists a constant c such that for any stopping time T we have

$$E\{(M_\infty - M_{T-})^2 | \mathcal{F}_T\} \leq c^2 \text{ a.s.},$$

where $M_{0-} = 0$ by convention. The smallest such c is defined to be the BMO norm of M, and it is written $\|M\|_{BMO}$. If the constant c does not exist, or if M is not in $\mathcal{H}^2$, then we set $\|M\|_{BMO} = \infty$.

Note that in the above definition $E\{M_\infty^2 | \mathcal{F}_0\} \leq c^2$ (with the convention that $M_{0-} = 0$) and therefore $\|M\|_{\mathcal{H}^2} \leq \|M\|_{BMO}$. Note in particular that $\|M\|_{BMO} = 0$ implies that $M = 0$.

Let T be a stopping time and $\Lambda \in \mathcal{F}_T$. Replacing T with T_Λ shows that the above definition is equivalent to the statement

$$E\{(M_\infty - M_{T-})^2\} \leq c^2 P(T < \infty)$$

for every stopping time T. This in turn gives us an equivalent description of the BMO norm:

$$\|M\|_{BMO} = \sup_T \sqrt{\frac{E\{(M_\infty - M_{T-})^2\}}{P(T < \infty)}},$$

where the supremum is taken over all stopping times T. Note that this second characterization gives that $\|\cdot\|_{BMO}$ is a semi-norm, since it is the supremum of quadratic semi-norms. An elementary property of BMO martingales is that $M \in BMO$ if and only if all of the jumps of M are uniformly bounded. Thus trivially, continuous L^2 martingales and bounded martingales are in BMO.

The next inequality is quite powerful.

Theorem 52 (Fefferman's Inequality). *Let M and N be two local martingales. Then there exists a constant c such that*

$$E\{\int_0^\infty |d[M,N]_s|\} \leq c\|M\|_{\mathcal{H}^1} \|N\|_{BMO}.$$

Fefferman's inequality is a special case of the following more general result.

Theorem 53 (Strengthened Fefferman Inequality). *There exists a constant c such that for all local martingales M and N, and U an optional process,*

$$E\{\int_0^\infty |U_s||d[M,N]_s|\} \leq cE\{(\int_0^\infty U_s^2 d[M,M]_s)^{1/2}\}\|N\|_{BMO}.$$

Proof. Let $C_t = \int_0^t U_s^2 d[M,M]_s$ and define H and K by

$$H_t^2 = \frac{U_t^2}{\sqrt{C_t} + \sqrt{C_{t-}1_{\{t>0\}}}} 1_{\{C_t>0\}}, \quad K_t^2 = \sqrt{C_t}.$$

Using integration by parts yields

$$H_t^2 d[M,M]_t = 1_{\{C_t>0\}} \frac{dC_t}{\sqrt{C_t} + \sqrt{C_{t-}1_{\{t>0\}}}} = 1_{\{t>0\}} d\sqrt{C_t}.$$

From the definitions of H and K we have

$$H_t^2 K_t^2 \geq \frac{1}{2} U_t^2 1_{\{C_t>0\}}.$$

The Kunita-Watanabe inequality implies

$$\begin{aligned}\int_0^\infty |U_s| 1_{\{C_s=0\}} |d[M,N]_s| &\leq (\int_0^\infty U_s^2 1_{\{C_s=0\}} d[M,M]_s)^{1/2} ([N,N]_\infty)^{1/2} \\ &= (\int_0^\infty 1_{\{C_s=0\}} dC_s)^{1/2} ([N,N]_\infty)^{1/2} = 0 \text{ a.s.,}\end{aligned}$$

and since $|d[M,N]_s|$ is absolutely continuous with respect to $d[M,M]_s$ as a consequence of the Kunita-Watanabe inequality, we have

$$\begin{aligned}\frac{1}{\sqrt{2}} E\{\int_0^\infty |U_s||d[M,N]_s|\} &= \frac{1}{\sqrt{2}} E\{\int_0^\infty |1_{\{C_s>0\}} U_s||d[M,N]_s|\} \\ &\leq E\{\int_0^\infty |H_s K_s||d[M,N]_s|\} \\ &\leq \sqrt{E\{\int_0^\infty H_s^2 d[M,M]_s\}} \sqrt{E\{\int_0^\infty K_s^2 d[N,N]_s\}}.\end{aligned}$$

But

$$\begin{aligned}E\{\int_0^\infty H_s^2 d[M,M]_s\} &\leq E\{\int_0^\infty d\sqrt{C_s}\} = E\{\sqrt{C_\infty}\} \\ &= E\{(\int_0^\infty U_s^2 d[M,M]_s)^{1/2}\}\end{aligned}$$

and

$$\begin{aligned}E\{\int_0^\infty K_s^2 d[N,N]_s\} &= E\{\int_0^\infty ([N,N]_\infty - [N,N]_{s-}) dK_s^2\} \\ &= E\{\int_0^\infty (E\{[N,N]_\infty | \mathcal{F}_s\} - [N,N]_{s-}) dK_s^2\}.\end{aligned}$$

But $E\{E\{[N,N]_\infty|\mathcal{F}_s\} - [N,N]_{s-}\}$ is bounded by $\|N\|_{BMO}^2$ on $(0,\infty)$, hence we have that

$$\begin{aligned}E\{\int_0^\infty K_s^2 d[N,N]_s\} &= E\{\int_0^\infty (E\{[N,N]_\infty|\mathcal{F}_s\} - [N,N]_{s-}) dK_s^2\} \\ &\leq \|N\|_{BMO}^2 E\{\sqrt{C_\infty}\},\end{aligned}$$

and the result follows. □

Remark. The constant c in Theorems 52 and 53 can be taken to be $\sqrt{2}$, as can be seen from an analysis of the preceding proof.

Theorem 54. *Let $N \in \mathcal{H}^2$. Then N is in BMO if and only if there is a constant $c > 0$ such that for all $M \in \mathcal{H}^2$,*

$$|E\{[M,N]_\infty\}| \leq c\,\|M\|_{\mathcal{H}^1}.$$

Moreover $\|N\|_{BMO} \leq \sqrt{6}\,c$.

Proof. If N is in BMO, then we can take $c = \sqrt{2}\,\|N\|_{BMO}$ from Fefferman's inequality (Theorem 52) and the remark following the proof of Theorem 53.

Now suppose that M is in $\mathcal{H}^2$ and that $|E\{[M,N]_\infty\}| \leq c\|M\|_{\mathcal{H}^1}$; we want to show that N is in BMO. We do this by first showing that $|N_0| \leq c$ a.s., and then showing that N has bounded jumps.

Let $\Lambda = \{|N_0| > c\}$. Suppose $P(\Lambda) > 0$. Let $\xi = \frac{\operatorname{sign}(N_0)}{P(\Lambda)} 1_\Lambda$. Then $E\{|\xi|\} = 1$, and if we define the trivial martingale $M_t = \xi$ for all $t \geq 0$, then $M \in \mathcal{H}^2$ and $\|M\|_{\mathcal{H}^1} = E\{|\xi|\} = 1$, whence

$$|E\{[M,N]_\infty\}| = E\{M_0N_0\} = E\{\frac{|N_0|1_{\{|N_0|>c\}}}{P(|N_0| > c)}\} > c = c\|M\|_{\mathcal{H}^1}.$$

This of course is a contradiction and we conclude $|N_0| \leq c$ a.s.

We next show $|\Delta N| \leq 2c$. Since every stopping time T can be decomposed into its accessible and totally inaccessible parts, and since each accessible time can be covered by a countable collection of predictable times with disjoint graphs, we can assume without loss of generality that T is either totally inaccessible or predictable. We further assume $P(T > 0) = 1$. Suppose then that $P(|\Delta N_T| > 2c) > 0$, and set

$$\xi = \frac{\operatorname{sign}(\Delta N_T)}{P(|\Delta N_T| > 2c)} 1_{\{|\Delta N_T|>2c\}}.$$

Let M be the martingale consisting of $\xi 1_{\{t \geq T\}}$ minus its compensator. Then M is in $\mathcal{H}^2$ and has at most one jump, which occurs at T. The jump is given by

$$\Delta M_T = \begin{cases} \xi - E\{\xi|\mathcal{F}_{T-}\}, & T \text{ is predictable,} \\ \xi & T \text{ is totally inaccessible.} \end{cases}$$

Note that we also have $\|M\|_{\mathcal{H}^1}$ is less than the expected total variation of M, which in turn is less than $2E\{|\xi|\} = 2$. If T is totally inaccessible then

$$\begin{aligned} E\{[M,N]_\infty\} = E\{\Delta M_T \Delta N_T\} = E\{\xi \Delta N_T\} &= E\{\frac{|\Delta N_T|1_{\{|\Delta N_T|>2c\}}}{P(|\Delta N_T| > 2c)}\} \\ &> 2c \geq c\,\|M\|_{\mathcal{H}^1}, \end{aligned}$$

which is a contradiction. On the other hand, if T is predictable, then we know that $E\{\Delta N_T|\mathcal{F}_{T-}\} = 0$ and thus we are reduced to the same calculation and the same contradiction. We conclude that $P(|\Delta N_T| > 2c) = 0$, and thus N has jumps bounded by $2c$.

Last let T be any stopping time. Let $M = N - N^T$ and $\eta = [N,N]_\infty - [N,N]_T$. Then M is in $\mathcal{H}^2$ and $[M,M]_\infty = [M,N]_\infty = \eta$. By our hypotheses it now follows that

$$\begin{aligned} E\{\eta\} = E\{[M,N]_\infty\} \le c\,\|M\|_{\mathcal{H}^1} &= cE\{\sqrt{\xi}\} \\ &= cE\{\sqrt{\eta}\,1_{\{T<\infty\}}\} \\ &\le cE\{\eta\}^{1/2}P(T<\infty)^{1/2}, \end{aligned}$$

which in turn implies

$$E\{[N,N]_\infty - [N,N]_T\} = E\{\eta\} \le c^2 P(T<\infty).$$

This then implies

$$E\{[N,N]_\infty|\mathcal{F}_T\} - [N,N]_T \le c^2 \text{ a.s.},$$

and since

$$E\{(N_\infty - N_{T-})^2|\mathcal{F}_T\} = E\{[N,N]_\infty|\mathcal{F}_T\} - [N,N]_T + N_0^2 1_{\{T=0\}} + (\Delta N_T)^2$$

almost surely, we obtain

$$E\{N_\infty^2 - N_{T-}^2\} \le 6c^2$$

which yields the result. □

During the proof of Theorem 54 we proved, *inter alia*, that the jumps of a local martingale in BMO are bounded. We formalize this result as a corollary, which in turn has its own corollary.

Corollary 1. Let N be a local martingale in BMO. Then N has bounded jumps.

Corollary 2. Let N be a local martingale in BMO. Then N is locally bounded.

Proof. Let c be a bound for the jumps of N which we know exists by the previous corollary. Next let $T_n = \inf\{t > 0 : |N_t| \ge n\}$. Then $|N^{T_n}| \le n + c$, and N is locally bounded. □

The key result concerning $\mathcal{H}^1$ and BMO is the Duality Theorem which is Theorem 55 that follows. First let us lay the foundation. For N chosen and fixed in BMO we define the operator L_N from $\mathcal{H}^1$ to $\mathbb{R}$ by

$$L_N(M) = E\{[M,N]_\infty\}$$

for all M in $\mathcal{H}^1$. Then one can easily check to see that L_N is linear, and Fefferman's inequality proves that it is bounded as well, and therefore continuous. If BMO is the Banach space dual of $\mathcal{H}^1$ then it is also complete, a fact that is apparently not easy to verify directly.

Theorem 55 (The Dual of $\mathcal{H}^1$ is BMO). *The Banach space dual of all (bounded) linear functionals on $\mathcal{H}^1$ can be identified with BMO. Moreover if L_N is such a functional then the norms $\|L_N\|$ and $\|N\|_{BMO}$ are equivalent.*

Proof. Let N be in BMO. By Fefferman's inequality we have

$$|L_N(M)| = |E\{[M,N]_\infty\}| \leq c\|N\|_{BMO}\|M\|_{\mathcal{H}^1} \tag{$*$}$$

for all M in $\mathcal{H}^1$. This shows that L_N is in the dual of $\mathcal{H}^1$ and also that $\|L_N\| \leq c\|N\|_{BMO}$. Note further that L_N cannot be trivial since $L_N(N) = E\{[N,N]_\infty\} > 0$ unless N is identically 0. Therefore the mapping $\Gamma(N) = L_N$ is an injective linear mapping from BMO into $\mathcal{H}^{1*}$, the dual of $\mathcal{H}^1$.

Let L be an arbitrary linear functional in the dual of $\mathcal{H}^1$. We have

$$|L(M)| \leq \|L\|\|M\|_{\mathcal{H}^1} \leq \|L\|\|M\|_{\mathcal{H}^2}.$$

This means that L is also a bounded linear functional on $\mathcal{H}^2$. Since $\mathcal{H}^2$ is isomorphic as a Hilbert space to the L^2 space of the terminal random variables of the martingales in $\mathcal{H}^2$, we have that there must exist a unique martingale N in $\mathcal{H}^2$ such that for any M in $\mathcal{H}^1$ we have:

$$|L_N(M)| = |E\{M_\infty N_\infty\}| = |E\{[M,N]_\infty\}| \leq \|L\|\|M\|_{\mathcal{H}^1}.$$

Clearly $L_N = L$ on $\mathcal{H}^2$, and since $\mathcal{H}^2$ is dense in $\mathcal{H}^1$ by Theorem 50, we have that L and L_N are the same functional on $\mathcal{H}^1$. This shows that BMO equipped with the norm $\|L_N\|$ is isomorphic to $\mathcal{H}^{1*}$ and thus, being the dual of a Banach space, it is itself a Banach space and in particular it is complete. Combining equation ($*$) with Theorem 54 we have that $\|L_N\|$ and $\|N\|_{BMO}$ are equivalent norms. This completes the proof. □

While Fefferman's inequality, the space of BMO martingales, and the duality of $\mathcal{H}^\infty$ and BMO are all of interest in their own right, we were motivated to present the material in order to prove the important Jacod-Yor Theorem on martingale representation, which we now present, after we recall the version of the Hahn-Banach Theorem we will use.

Theorem 56 (Hahn-Banach Theorem). *Let X be a Banach space and let Y be a closed linear subspace. Then $Y = X$ if and only if the only bounded linear functional L which has the property that $L(Y) = 0$ is the functional which is identically zero.*

Theorem 57 (Jacod-Yor Theorem on Martingale Representation). *Let $\mathcal{A}$ be a subset of $\mathcal{H}^2$ containing constant martingales. Then $\mathcal{S}(\mathcal{A})$, the stable subspace of stochastic integrals generated by $\mathcal{A}$, equals $\mathcal{H}^2$ if and only if the probability measure P is an extremal point of $\mathcal{M}^2(\mathcal{A})$, the space of probability measures making all elements of $\mathcal{A}$ square integrable martingales.*

Proof. The necessity has already been proved in Theorem 38. By the Hahn-Banach Theorem, $\mathcal{H}^1 = \mathcal{S}(\mathcal{A})$ if and only if $L(\mathcal{S}(\mathcal{A})) = 0$ implies L is identically zero, where L is a bounded linear functional. Let L be a bounded linear functional which is such that $L(\mathcal{S}(\mathcal{A})) = 0$. Then there exists a martingale N in BMO such that $L = L_N$. The local martingale N is locally bounded, so by stopping we can assume it is bounded and that $N_0 = 0$. (See the two corollaries of Theorem 54.) Let us also assume it is not identically zero, and let c be a bound for N. We can then define two new probability measures Q and R by

$$dQ = (1 - \frac{N_\infty}{2c})dP, \quad dR = (1 + \frac{N_\infty}{2c})dP.$$

Then Q and R are both in $\mathcal{M}^2(\mathcal{A})$, and $P = \frac{1}{2}Q + \frac{1}{2}R$ shows that P is not extremal in $\mathcal{M}^2(\mathcal{A})$, a contradiction. Therefore we must have that L is identically zero and we have that $\mathcal{H}^1 = \mathcal{S}(\mathcal{A})$. As far as $\mathcal{H}^2$ is concerned, it is a subspace of $\mathcal{H}^1$, hence $\mathcal{H}^2 \subset \mathcal{S}(\mathcal{A})$. But by construction of $\mathcal{S}(\mathcal{A})$, it is contained in $\mathcal{H}^2$, and we have martingale representation. □

5 Examples of Martingale Representation

In Sect. 3 we have already seen the most important example of martingale representation, that of Brownian motion. In this section we give a method to generate a family of examples which are local martingales with jumps, and which have the martingale representation property. The limitations are that the family of examples is one dimensional (so that we exclude vector-valued local martingales such as n-dimensional Brownian motion), and that the descriptions of the jumps are all of the same rather simple kind.

The idea is to construct a class of local martingales $\mathcal{H}$ such that for $X \in \mathcal{H}$ we have both that $X_0 = 0$ and the compensator of $[X, X]_t$ is $A_t = t$. If X has the martingale representation property, then there must exist a predictable process H such that

$$[X, X]_t - t = \int_0^t H_s dX_s. \qquad (*)$$

The above equation is called **Emery's structure equation**, and it is written (in a formal sense) in differential notation as

$$d[X, X]_t = dt + H_t dX_t.$$

In order to establish that solutions to Emery's structure equation actually exists we write it in a form resembling a differential equation:

$$d[X, X]_t = dt + \phi(X_{t-})dX_t \qquad (**)$$

Equation $(**)$ is unusual and different from the stochastic differential equations considered later in Chap. V, since while the unknown is of course the local martingale X (and part of the structure equation is to require that any solution X be a local martingale), no *a priori* stochastic process is given in the equation. That is, it is lacking the presence of a given stochastic driving term such as, for example, a Brownian motion, a compensated Poisson process, or more generally a Lévy process. Since therefore no probability space is specified, the only reasonable interpretation of equation (**) is that of a **weak solution**. That is, we want to show there exists a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses, and a local martingale X, such that X verifies equation (**). It would also be nice to have **weak uniqueness** which means that if X and Y are solutions of (**) for a given ϕ, possibly defined on different filtered probability spaces, then X and Y have the same distribution as processes. That means that for every Λ, a Borel set on the function space of càdlàg functions mapping $\mathbb{R}_+$ to $\mathbb{R}$, we have $P(\omega : t \mapsto X_t(\omega) \in \Lambda) = Q(\omega : t \mapsto X_t(\omega) \in \Lambda)$, where P and Q are the probability measures where X and Y are respectively defined.

Inspired by knowledge of stochastic differential equations, it is natural to conjecture that such weak solutions exist and are unique if the coefficient ϕ is Lipschitz continuous.[14] This is true for existence and was proven by P. A. Meyer [179]; see alternatively [136]. Since the proof uses weak convergence techniques which are not within the scope of this book, we omit it.

Theorem 58 (Existence of Solutions of the Structure Equation). *Let $\phi : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous. Then Emery's structure equation*

$$[X, X]_t - t = \int_0^t \phi(X_{s-})dX_s \qquad (***)$$

has a weak solution with both $(X_t)_{t\geq 0}$ and $(\int_0^t \phi(X_{s-})dX_s)_{t\geq 0}$ local martingales.

The issue of uniqueness is intriguing. Emery has shown that one has uniqueness when ϕ is linear, but uniqueness for others ϕ's, including the Lipschitz case, is open. The next theorem collects some elementary properties of a solution X.

Theorem 59. *Let X be a (weak) solution of $(***)$. Then the following hold.*

(i) $E\{X_t^2\} = E\{[X, X]_t\} = t$, and X is a square integrable martingale on compact time sets.
(ii) All jumps of X are of the form $\Delta X_t = \phi(X_{t-})$.
(iii) X has continuous paths if and only if ϕ is identically 0, in which case X is standard Brownian motion.
(iv) If a stopping time T is a jump time of X, then it is totally inaccessible.

Proof. We prove the statements in the order given. Since a solution X and the integral term are both required to be local martingales, we know

[14] Lipschitz continuity is defined and discussed in Chap. V.

there exists a sequence $(T_n)_{n\geq 1}$ of stopping times increasing to ∞ such that $\int_0^{t\wedge T_n} \phi(X_{s-})dX_s$ is in L^1. Therefore $E\{[X,X]_{t\wedge T_n}\} = E\{t\wedge T_n\}$, and applying the Monotone Convergence Theorem to each side of the equality in this equation yields $E\{[X,X]_t\} = t$ which further implies that X is a martingale on $[0,t]$ for each $t<\infty$ and that $E\{X_t^2\} = E\{[X,X]_t\} = t$.

For the second statement, recall that

$$[X,X]_t = [X,X]_t^c + \sum_{s\leq t}(\Delta X_s)^2,$$

and hence we have $\Delta[X,X]_t = (\Delta X_t)^2 = \phi(X_{t-})\Delta X_t$, and dividing both sides by ΔX_t (when it is not zero) gives the result.

For the third statement, suppose ϕ is identically zero. Then by the second statement X has no jumps and must be continuous. Since X is then a continuous local martingale with $[X,X]_t = t$, it is Brownian motion by Lévy's Theorem. For the converse, if we know that X is continuous, if it is non-trivial it has paths of infinite variation since it is a local martingale. Thus so too does the term $\int_0^t \phi(X_{s-})dX_s$. But notice that this term is the right side of equation $(***)$ which is of finite variation, and we have a contradiction, so we must have that ϕ is zero.

For the fourth statement, we implicitly stop X so that it is a uniformly integrable martingale. Next let T be a jump time of X. Then $T = T_A \wedge T_B$ where A is the accessible part of T and B is the totally inaccessible part of T. Since T_A can be covered by a countable sequence of predictable times with disjoint graphs, we can assume T_A is predictable without loss of generality. Thus it suffices to show $P(A) = 0$. However since T_A is predictable, and X is a uniformly integrable martingale, we have $E\{\Delta X_{T_A}|\mathcal{F}_{T_A-}\} = 0$. But $\Delta X_{T_A} = \phi(X_{T_A-})$ by part (ii) of this theorem, and $\phi(X_{T_A-}) \in \mathcal{F}_{T_A-}$ since T_A is predictable, which implies that the jump of X at T_A is zero, which in turn implies that $P(A) = 0$. □

Let us now consider a special class of structure equations where ϕ is assumed to be an *affine function.* That is, we assume ϕ is of the form $\phi(x) = \alpha + \beta x$. We analyze these special cases when α and β vary. Emery has named the solutions corresponding to affine structure equations the **Azéma martingales**, since J. Azéma's work on Markov processes and expansion of filtrations led him to the amazing formula of "the" Azéma martingale given later in Sect. 7. Equation $(**)$ now becomes

$$d[X,X]_t = dt + (\alpha + \beta X_{t-})dX_t,$$

and when $\beta = 0$, it reduces to

$$d[X,X]_t = dt + \alpha dX_t,$$

and *when in addition* $\alpha = 0$ *we have seen that* X *is standard Brownian motion.* Note that Lévy's Theorem gives us weak uniqueness in this case, since any solution with $\alpha = \beta = 0$ must have the same distribution, namely that of Wiener measure. We have much more, as we see in the next theorem. However it has a long and difficult proof. Rather than present it, we refer the interested reader to the excellent treatment of M. Emery, in his original paper [69] proving the result.

Theorem 60 (Emery's Uniqueness Theorem). *Let X be a local martingale solution of the structure equation*

$$d[X,X]_t = dt + (\alpha + \beta X_{t-})dX_t, \ X_0 = x_0.$$

Then X is unique in law. That is, any other solution Y must have the same distribution as does X. Moreover X is a strong Markov process.

The uniqueness is especially significant in light of the next theorem. By **martingale representation** we mean that every square integrable martingale can be represented as a stochastic integral with respect to one fundamental local martingale.

Theorem 61. *Consider the equation*

$$[X,X]_t - t = \int_0^t \phi(X_{s-})dX_s \qquad (\otimes)$$

on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ which satisfies the usual hypotheses. Then X has martingale representation for its completed natural filtration if and only if the law P is an extreme point of the convex set of all probabilities on $(\Omega, \mathcal{F}, \mathbb{F})$ for which X is a martingale and verifies the equation. Moreover if the equation with fixed initial condition X_0 has weak uniqueness of solutions, then every solution X of the equation has martingale representation with respect to the smallest filtration satisfying the usual hypotheses and to which X is adapted.

Proof. By the Jacod-Yor Theorem (Theorem 57) we need to verify that P is extremal in the set $\mathcal{M}^2$ of all probability measures such that X is a square integrable martingale. It is clearly true if P is extremal. Suppose then that P is not extremal, and let Q and R both be in $\mathcal{M}^2$, such that $P = \lambda Q + (1-\lambda)R$, with $0 < \lambda < 1$. Both Q and R are absolutely continuous with respect to P, so under Q and R the terms $[X,X]$ and $\int_0^t \phi(X_{s-})dX_s$ are the same a.s. (resp. dQ and dR). Therefore X satisfies equation $(\otimes)$ for both Q and R as well as P. Thus if P is not extremal in $\mathcal{M}^2$ then it is also not extremal in the set of probability measures such that X satisfies equation $(\otimes)$.

To prove the second statement, let Ω denote the canonical path space of càdlàg paths, with X being the projection process given by $X_t(\omega) = \omega(t)$. We have just seen that among the solutions of equation $(\otimes)$, the ones having martingale representation are those whose law constitutes an extremal probability measure. But if weak uniqueness holds, the collection of all such probabilities consists of only one, and thus extremality is trivial. □

We now examine several special cases as α and β vary. Let $\alpha = \beta = 0$, and we have seen that X is *standard Brownian motion.* Because of Theorem 61 we conclude that (one dimensional) Brownian motion has martingale representation, recovering a special case of Theorem 43.

Next suppose $\alpha = 1$ and $\beta = 0$. In this case the equation of Theorem 60 with $X_0 = 0$ becomes, in integral form,

$$[X, X]_t = t + (X_t - X_0) = t + X_t.$$

Therefore $X_t = [X, X]_t - t$ and hence X is a finite variation martingale. Moreover $\Delta X_t = 1$, so X only jumps up, with jumps always of size 1. Now let N be a standard Poisson process with arrival intensity $\lambda = 1$, and let $X_t = N_t - t$, the compensated Poisson process. Then X satisfies the equation, and by weak uniqueness all such X are compensated Poisson processes with $\lambda = 1$. We conclude that *a compensated standard Poisson process has martingale representation* with respect to its natural (completed) filtration. For general α (and not just $\alpha = 1$), but still with $\beta = 0$, it is simple to check that

$$X_t^\alpha = \alpha(N_{\frac{t}{\alpha^2}} - \frac{t}{\alpha^2})$$

is the unique solution of the equation of Theorem 60 if N is a standard Poisson process. Note that (as is well known) X^α converges (weakly) to Brownian motion as $\alpha \to 0$.

We now consider the more interesting cases where $\beta \neq 0$. We repeat the equation of Theorem 60 here (in integrated form) for ease of reference:

$$[X, X]_t = x_0 + t + \int_0^t (\alpha + \beta X_{s-}) dX_s.$$

Observe that X is a solution of the above equation if and only if $X + \frac{\alpha}{\beta}$ is a solution of

$$d[X, X]_t = dt + \beta X_{t-} dX_t \tag{$\otimes\otimes$}$$

with initial condition $X_0 = x_0 + \frac{\alpha}{\beta}$. Therefore without loss of generality *we can assume that $\alpha = 0$ and we do this from now on.*

We have two explicit examples for equation ($\otimes\otimes$). When $\beta = -1$, we take

$$X_t = \frac{2}{\sqrt{\pi}} M_t = \sqrt{2}\,\mathrm{sign}\,(B_t)\sqrt{(t - g_t)},$$

where M is **Azéma's martingale**[15], B is standard Brownian motion, and $g_t = \sup\{s \leq t : B_s = 0\}$. By Theorem 86 of Sect. 8 of this chapter, we know that $[X, X]_t^c = 0$ and $[X, X]_t = 2g_t$. Integration by parts gives $\int_0^t X_{s-} dX_s = (t - g_t) - g_t = t - [X, X]_t$, since $[X, X]_t = \sum_{s \leq t} (\Delta X_s)^2$, because of course $[X, X]_t^c = 0$. This proves the assertion that for $\beta = -1$, $X_t = \frac{2}{\sqrt{\pi}} M_t$, where

[15] Azéma's martingale is treated in detail in Sect. 8 later in this chapter.

M is Azéma's (original) martingale as presented in Sect. 8, and thus *we have martingale representation for Azéma's martingale.*

Our second (and last) explicit example is for $\beta = -2$. In this case our equation becomes

$$[X, X]_t - t = -2 \int_0^t X_{s-} dX_s$$

and using integration by parts we obtain $2\int_0^t X_{s-} dX_s = X_t^2 - [X, X]_t$. Equating terms gives $X_t^2 = t$, and since $E\{X_t\} = 0$ for all t because $x_0 = 0$ we deduce that

$$P(X_t = \sqrt{t}) = P(X_t = -\sqrt{t}) = \frac{1}{2}$$

for all $t > 0$. The jumps of X occur from a change of sign, and they arrive according to a Poisson process with intensity $\frac{1}{4t}dt$. Such a process can be seen to be a martingale (as in [204]) by constructing a process X with a distribution as above and with filtration $\mathbb{F}$. Then for $0 < s < t$,

$$\frac{X_t}{X_s} = \sqrt{\frac{t}{s}}(-1)^N$$

where N is independent of the σ-algebra $\mathcal{F}_s$ and has a Poisson distribution with parameter $\lambda = \frac{1}{4}\ln(\frac{t}{s})$. In this way it is easily seen to be a martingale. We call it the **parabolic martingale**. Once again we are able to conclude: *we have martingale representation for the parabolic martingale.*

6 Stochastic Integration Depending on a Parameter

The results of this section are of a technical nature, but they are needed for our subsequent investigation of semimartingale local times. Nevertheless they have intrinsic interest. For example, Theorems 64 and 65 are types of Fubini Theorems for stochastic integration. A more comprehensive treatment of stochastic integration depending on a parameter can be found in Stricker-Yor [219] and Jacod [103]. Throughout this section $(A, \mathcal{A})$ denotes a measurable space.

Theorem 62. *Let $Y^n(a,t,\omega)$ be a sequence of processes that are (i) $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable, and (ii) for each fixed a the process $Y^n(a,t,\omega)$ is càdlàg. Suppose $Y^n(a,t,\cdot)$ converges in ucp for each $a \in A$. Then there exists an $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable process $Y = Y(a,t,\omega)$ such that*

(a) $Y(a,t,\cdot) = \lim_{n\to\infty} Y^n(a,t,\cdot)$ with convergence in ucp;
(b) for each $a \in A$, Y is a.s. càdlàg.

Moreover there exists a subsequence $n_k(a)$ depending measurably on a such that $\lim_{n_k(a)\to\infty} Y_t^{n_k(a)} = Y_t$ uniformly in t on compacts, a.s.

Proof. Let $S^a_{u,i,j} = \sup_{t\le u} |Y^i(a,t,\cdot) - Y^j(a,t,\cdot)|$. Since Y^i is càdlàg in t the function $(a,\omega) \mapsto S^a_{u,i,j}$ is $\mathcal{A}\otimes\mathcal{F}$ measurable. By hypothesis we have $\lim_{i,j\to\infty} S^a_{u,i,j} = 0$ in probability. Let $n_0(a)=1$, and define inductively

$$n_k(a) = \inf\{m > \max(k, n_{k-1}(a)) : \sup_{i,j\ge m} P(S^a_{k,i,j} > 2^{-k}) \le 2^{-k}\}.$$

We then define

$$Z^k(a,t,\omega) = Y^{n_k(a)}(a,t,\omega).$$

Since each $a \mapsto n_k(a)$ is measurable, so also is Z^k. Define

$$T^a_{u,i,j} = \sup_{t\le u} |Z^i(a,t,\omega) - Z^j(a,t,\omega)|;$$

then also $(a,\omega) \mapsto T^a_{u,i,j}(\omega)$ is jointly measurable, since Z^i have càdlàg paths (in t). Moreover by our construction $P(T^a_{k,k,k+m} > 2^{-k}) \le 2^{-k}$ for any $m \ge 1$. The Borel-Cantelli Lemma then implies that $\lim_{i,j\to\infty} T^a_{u,i,j} = 0$ *almost surely*, which in turn implies that

$$\lim_{i\to\infty} Z^i(a,t,\cdot) \quad \text{exists a.s.},$$

with convergence uniform in t. Let Λ^a be the set where Z^i converges uniformly (note that $\Lambda^a \in \mathcal{A}\otimes\mathcal{F}$ and $P(\Lambda^a)=1$, each fixed a), and define

$$Y(a,t,\omega) = \begin{cases} \lim_{i\to\infty} Z^i(a,t,\omega), & \omega \in \Lambda^a, \\ 0, & \omega \notin \Lambda^a. \end{cases}$$

Then Y is càdlàg thanks to the uniform convergence, and it is jointly measurable. □

Theorem 63. *Let X be a semimartingale with $X_0 = 0$ a.s. and let $H(a,t,\omega) = H^a_t(\omega)$ be $\mathcal{A}\otimes\mathcal{P}$ measurable*[16] *and bounded. Then there is a function $Z(a,t,\omega)$ in $\mathcal{A}\otimes\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{F}$ such that for each $a\in A$, $Z(a,t,\omega)$ is a càdlàg, adapted version of the stochastic integral $\int_0^t H^a_s dX_s$.*

Proof. Let $\mathcal{H} = \{H \in \mathbf{b}\mathcal{A}\otimes\mathcal{P}$ such that the conclusion of the theorem holds$\}$. If $K = K(t,\omega) \in \mathbf{b}\mathcal{P}$ and $f = f(a) \in \mathbf{b}\mathcal{A}$, and if $H(a,t,\omega) = f(a)K(t,\omega)$, then

$$\int_0^t H(a,s,\cdot)dX_s = \int_0^t f(a)K(s,\cdot)dX_s = f(a)\int_0^t K(s,\cdot)dX_s,$$

and thus clearly $H = fK$ is in $\mathcal{H}$. Also note that $\mathcal{H}$ is trivially a vector space, and that H of the form $H = fK$ generate $\mathbf{b}\mathcal{A}\otimes\mathcal{P}$.

Next let $H^n \in \mathcal{H}$ and suppose that H^n converges boundedly to a process $H \in \mathbf{b}\mathcal{A}\otimes\mathcal{P}$. By Theorem 32 (for example) we have that $H^n\cdot X$ converges uniformly in t in probability on compacts, for each a. Therefore $H\in\mathcal{H}$, and an application of the Monotone Class Theorem yields the result. □

[16] Recall that $\mathcal{P}$ denotes the predictable σ-algebra.

Corollary. Let X be a semimartingale ($X_0 = 0$ a.s.), and let $H(a,t,\omega) = H_t^a(\omega) \in \mathcal{A} \otimes \mathcal{P}$ be such that for each a the process $H^a \in L(X)$. Then there exists a function $Z(a,t,\omega) = Z_t^a \in \mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ such that for each a, Z_t^a is an a.s. càdlàg version of $\int_0^t H_s^a dX_s$.

Proof. By Theorem 32 the bounded processes $Z^{a,k} = H^a 1_{\{|H^a| \le k\}} \cdot X$ converge to $H^a \cdot X$ in *ucp*, each a. But $Z^{a,k}$ can be chosen càdlàg and jointly measurable by Theorem 63. The result now follows by Theorem 43. □

Theorem 64 (Fubini's Theorem). *Let X be a semimartingale, $H_t^a = H(a,t,\omega)$ be a bounded $\mathcal{A} \otimes \mathcal{P}$ measurable function, and let μ be a finite measure on $\mathcal{A}$. Let $Z_t^a = \int_0^t H_s^a dX_s$ be $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable such that for each a, Z^a is a càdlàg version of $H^a \cdot X$. Then $Y_t = \int_A Z_t^a \mu(da)$ is a càdlàg version of $H \cdot X$, where $H_t = \int_A H_t^a \mu(da)$.*

Proof. By pre-stopping we may assume without loss of generality that $X \in \mathcal{H}^2$, and because the result holds for the finite variation part of the canonical decomposition of X by the ordinary Stieltjes Fubini Theorem, we may further assume that X is a martingale with $E\{[X,X]_\infty\} < \infty$. Next suppose H_t^a is of the form $H(a,t,\omega) = K(t,\omega)f(a)$ where $K \in \mathbf{b}\mathcal{P}$ and f is bounded, measurable. Then $K \in L(X)$ and $\int |f(a)|\mu(da) < \infty$. In this case we have $Z_t^a = f(a)K \cdot X$, and moreover

$$\begin{aligned}\int Z_t^a \mu(da) = \int f(a)K \cdot X\mu(da) &= K \cdot X \int f(a)\mu(da) \\ &= (\int f(a)\mu(da)K) \cdot X \\ &= H \cdot X.\end{aligned}$$

By linearity the same result holds for the vector space $\mathbb{V}$ generated by processes of the form $K(t,\omega)f(a)$ with $K \in \mathbf{b}\mathcal{P}$ and f bounded, measurable.

By the Monotone Class Theorem it now suffices to show that if $H_n \in \mathbb{V}$ and $\lim_{n\to\infty} H_n = H$, then the result holds for H. Let $Z_{n,t}^a = H_n^a \cdot X$, the càdlàg version. Then by Jensen's and the Cauchy-Schwarz inequalities,

$$\begin{aligned}&\frac{1}{\|\mu\|}(E\{\int_A \sup_t |Z_{n,t}^a - Z_t^a|\mu(da)\})^2 \\ &\le E\{\int_A \sup_t |Z_{n,t}^a - Z_t^a|^2\mu(da)\} = \int_A E\{\sup_t |Z_{n,t}^a - Z_t^a|^2\}\mu(da) \\ &\le 4\int_A E\{(Z_{n,\infty}^a - Z_\infty^a)^2\}\mu(da) = 4\int_A E\{[Z_n^a - Z^a, Z_n^a - Z^a]_\infty\}\mu(da)\end{aligned}$$

by Doob's quadratic inequality for the martingales Z_n^a and Z^a, and by Corollary 3 of Theorem 27 of Chap. II. Continuing, the preceding equals

$$= 4\int_A E\{\int_0^\infty (H_{n,s}^a - H_s^a)^2 d[X,X]_s\}\mu(da),$$

and the above tends to 0 by three applications of the Dominated Convergence Theorem.

We conclude from the preceding that

$$\int_A \sup_t |Z^a_{n,t} - Z^a_t|\mu(da) < \infty \ a.s.$$

and therefore $\int_A |Z^a_t|\mu(da) < \infty$ for all t, a.s. Moreover

$$E\{\sup_t |\int_A Z^a_{n,t}\mu(da) - \int_A Z^a_t\mu(da)|\} \le E\{\int_A \sup_t |Z^a_{n,t} - Z^a_t|\mu(da)\}$$

which tends to 0. Therefore taking $H_{n,t} = \int H^a_{n,t}\mu(da)$ we have $H_n \cdot X_t = \int_A Z^a_{n,t}\mu(da)$ converges in *ucp* to $\int Z^a_t\mu(da)$. Since $H_n \cdot X$ converges to $H \cdot X$ by Theorem 32, we conclude $H \cdot X = \int Z^a_t\mu(da)$. □

The version of Fubini's Theorem given in Theorem 64 suffices for the applications of it used in this book. Nevertheless it is interesting to determine under what more general conditions a Fubini-type theorem holds.

Theorem 65 (Fubini's Theorem: Second Version). *Let X be a semimartingale, let $H^a_t = H(a,t,\omega)$ be $\mathcal{A}\otimes\mathcal{P}$ measurable, let μ be a finite positive measure on A, and assume*

$$(\int_A (H^a_t)^2\mu(da))^{1/2} \in L(X).$$

Letting $Z^a_t = \int_0^t H^a_s dX_s$ be $\mathcal{A}\otimes\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{F}$ measurable and Z^a càdlàg for each a, then $Y_t = \int_A Z^a_t\mu(da)$ exists and is a càdlàg version of $H \cdot X$, where $H_t = \int_A H^a_t\mu(da)$.

Proof. By pre-stopping we may assume without loss of generality that $X \in \mathcal{H}^2$ and that $\|H^a\|_{L^2(d\mu)}$ is $(\mathcal{H}^2, X)$ integrable. Let $X = \overline{N} + \overline{A}$ be the canonical decomposition of X. Then

$$E\{\int_0^\infty \int_A (H^a_s)^2\mu(da)d[\overline{N},\overline{N}]_s\} + E\{(\int_0^\infty \|H^a_s\|_{L^2(d\mu)}|d\overline{A}_s|)^2\} < \infty.$$

Next observe that

$$\begin{aligned} E\{\int_0^\infty \|H^a_s\|_{L^2(d\mu)}|d\overline{A}_s|\} &\ge E\{\|\int_0^\infty |H^a_s||d\overline{A}_s|\|_{L^2(d\mu)}\} \\ &\ge cE\{\|\int_0^\infty |H^a_s||d\overline{A}_s|\|_{L^1(d\mu)}\} \\ &= c\int_A E\{\int_0^\infty |H^a_s||d\overline{A}_s|\}\mu(da). \end{aligned}$$

Also

$$E\{\int_0^\infty \int_A (H_s^a)^2\mu(da)d[\overline{N},\overline{N}]_s\} = \int_A E\{\int_0^\infty (H_s^a)^2 d[\overline{N},\overline{N}]_s\}\mu(da),$$

and therefore $E\{\int_0^\infty |H_s^a||d\overline{A}_s|\} < \infty$ and $E\{\int_0^\infty (H_s^a)^2 d[\overline{N},\overline{N}]_s\} < \infty$ for μ-almost all $a \in A$. Whence $H^a \in L(X)$ for μ almost all $a \in A$.

Next define $H_n = H1_{\{|H|\leq n\}}$, and the proof of Theorem 64 works as well here. □

The hypotheses of Theorem 65 are slightly unnatural, since they are not invariant under the transformation

$$H \to \frac{1}{\varphi(a)}H^a \qquad \mu \to \varphi(a)\mu(da)$$

where φ is any positive function such that $\int \varphi(a)\mu(da) < \infty$. This can be alleviated by replacing the assumption $(\int_A (H^a)^2\mu(da))^{1/2} \in L(X)$ with $(\int \frac{(H^a)^2}{\varphi(a)}\mu(da))^{1/2} \in L(X)$ for some positive $\varphi \in L^1(d\mu)$. One can also relax the assumption on μ to be σ-finite rather than finite.

Example. The hypotheses of Theorem 65 are a bit strange, however they are in some sense best possible. We give here an example of a parameterized process $(H^a)_{a\in A}$, a positive, finite measure μ on A, and a semimartingale X such that

(i) $(a,t) \to H_t^a$ is $\mathcal{A}\otimes\mathcal{P}$ measurable,
(ii) $H^a \in L(X)$, each $a \in A$, and
(iii) $\int_A |H_t^a|\mu(da) \in L(X)$

but such that if $Z_t^a = \int_0^t H_s^a dX_s$ then $\int_A Z_t^a\mu(da)$ *does not exist as a Lebesgue integral.* Thus a straightforward extension of the classical Fubini Theorem for Lebesgue integration does not hold.

Indeed, let $A = \mathbb{N} = \{1,2,3,\dots\}$ and let μ be any finite positive measure on A such that $\mu(\{a\}) > 0$ for all $a \in A$. Let X be standard Brownian motion, let $t_0 < t_1 < t_2 < \cdots$ be an increasing sequence in $[0,1]$, and define

$$H_t^a = \frac{1}{a}\mu(\{a\})^{-1}(t_a - t_{a-1})^{-1/2}1_{\{t_{a-1}<t\leq t_a\}}.$$

Then

$$H_t = \sum_{a=1}^{\infty}\frac{1}{a}(t_a - t_{a-1})^{-1/2}1_{\{t_{a-1}<t\leq t_a\}}$$

is in $L^2(dt)$, whence $H_t = \int_A |H_t^a|\mu(da) \in L(X)$, and moreover if $t \geq 1$,

$$H\cdot X_t = \sum_{a=1}^{\infty}\frac{1}{a}(t_a - t_{a-1})^{-1/2}(X_{t_a} - X_{t_{a-1}}),$$

where the sum converges in L^2. However if $t \geq 1$ then

$$Z_t^a = H^a \cdot X_t = \frac{1}{a}\mu(\{a\})^{-1}(t_a - t_{a-1})^{-1/2}(X_{t_a} - X_{t_{a-1}}),$$

and

$$\int_A |Z_t^a|\mu(da) = \sum_{a=1}^{\infty} \frac{1}{a}(t_a - t_{a-1})^{-1/2}|X_{t_a} - X_{t_{a-1}}| = \infty \quad \text{a.s.}$$

because $(t_a - t_{a-1})^{-1/2}(X_{t_a} - X_{t_{a-1}})$ is an i.i.d. sequence and $\sum_{a=1}^{\infty} a^{-1} = \infty$.

Note that this example can be modified to show that *we also cannot replace the assumption that*

$$(\int_A (H_t^a)^2\mu(da))^{1/2} \in L(X)$$

with the weaker assumption that $(\int_A (H_t^a)^p\mu(da))^{1/p} \in L(X)$ for some $p < 2$.

7 Local Times

In Chap. II we established Itô's formula (Theorem 32 of Chap. II) which showed that if $f : \mathbb{R} \to \mathbb{R}$ is C^2 and X is a semimartingale, then $f(X)$ is again a semimartingale. That is, semimartingales are preserved under C^2 transformations. This property extends slightly: semimartingales are preserved under convex transformations, as Theorem 66 below shows. (Indeed, this is the best one can do in general. If $B = (B_t)_{t\geq 0}$ is standard Brownian motion and $Y_t = f(B_t)$ is a semimartingale, then f *must* be the difference of convex functions. (See Çinlar-Jacod-Protter-Sharpe [34].) We establish a related result in Theorem 71, later in this section.) *Local times* for semimartingales appear in the extension of Itô's formula from C^2 functions to convex functions (Theorem 70).

Theorem 66. *Let $f : \mathbb{R} \to \mathbb{R}$ be convex and let X be a semimartingale. Then $f(X)$ is a semimartingale and one has*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-})dX_s + A_t$$

where f' is the left derivative of f and A is an adapted, right continuous, increasing process. Moreover $\Delta A_t = f(X_t) - f(X_{t-}) - f'(X_{t-})\Delta X_t$.

Proof. First suppose $|X|$ is bounded by n, and in $\mathcal{H}^2$, and that $X_0 = 0$. Let g be a positive C^∞ function with compact support in $(-\infty, 0]$ such that $\int_{-\infty}^{\infty} g(s)ds = 1$. Let $f_n(t) = n\int_{-\infty}^{\infty} f(t+s)g(ns)ds$. Then f_n is convex and C^2 and moreover f_n' increases to f' as n tends to ∞. By Itô's formula

$$f_n(X_t) - f_n(X_0) = \int_0^t f_n'(X_{s-})dX_s + A_t^n$$

where

$$A_t^n = \sum_{0<s\le t} \{f_n(X_s) - f_n(X_{s-}) - f_n'(X_{s-})\Delta X_s\} + \frac{1}{2}\int_0^t f_n''(X_{s-})d[X,X]_s^c.$$

The convexity of f implies that A^n is an increasing process. Letting n tend to ∞, we obtain

$$f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + A_t \tag{$*$}$$

where $\lim_{n\to\infty} A_t^n = A_t$ in L^2, and where the convergence of the stochastic integral terms is in $\mathcal{H}^2$ on $[0,t]$.

We now compare the jumps on both sides of the equation $(*)$. Since $\int_0^0 f'(X_{s-})dX_s = 0$ we have that $A_0 = 0$. When $t > 0$, the jump of the left side of $(*)$ is $f(X_t) - f(X_{t-})$, while the jump of the right side equals $f'(X_{t-})\Delta X_t + \Delta A_t$. Therefore $\Delta A_t = f(X_t) - f(X_{t-}) - f'(X_{t-})\Delta X_t$, and the theorem is established for $|X|$ bounded by n and in $\mathcal{H}^2$.

Now let X be an arbitrary semimartingale with $X_0 = 0$. By Theorem 13 we know there exists a sequence of stopping times $(T^n)_{n\ge1}$, increasing to ∞ a.s. such that $X^{T^n-} \in \mathcal{H}^2$ for each n. An examination of the proof of Theorem 13 shows that there is no loss of generality in further assuming that $|X^{T^n-}| \le n$, also. Then let $Y^n = X1_{[0,T^n)}$ and we have

$$f(Y_t^n) - f(Y_0^n) = \int_0^t f'(Y_{s-}^n)dY_s^n + A_t^n,$$

which is equivalent to saying

$$f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + A_t^n$$

on $[0,T^n)$. One easily checks that $(A^{n+1})^{T^n-} = (A^n)^{T^n-}$, and we can define $A = A^n$ on $[0,T^n)$, each n.

The above extends without difficulty to functions $g: \mathbb{R}^2 \to \mathbb{R}$ of the form $g(X_t, H)$ where H is an $\mathcal{F}_0$ measurable random variable and $x \mapsto g(x,y)$ is convex for every y. For general X we take $\hat{X}_t = X_t - X_0$, and then $f(X_t) = f(\hat{X}_t + X_0) = g(\hat{X}_t, X_0)$, where $g(x,y) = f(x+y)$. This completes the proof. □

Notation. For x a real variable let x^+, x^- be the functions $x^+ \equiv \max(x,0)$ and $x^- \equiv -\min(x,0)$. For x, y real variables, let $x \vee y \equiv \max(x,y)$ and $x \wedge y \equiv \min(x,y)$.

Corollary 1. Let X be a semimartingale. Then $|X|$, X^+, X^- are all semimartingales.

Proof. The functions $f(x) = |x|$, $g(x) = x^+$, and $h(x) = x^-$ are all convex, so the result then follows by Theorem 66. □

Corollary 2. Let X, Y be semimartingales. Then $X \vee Y$ and $X \wedge Y$ are semimartingales.

Proof. Since semimartingales form a vector space and $x \vee y = \frac{1}{2}(|x-y|+x+y)$ and $x \wedge y = \frac{1}{2}(x+y-|x-y|)$, the result is an immediate consequence of Corollary 1. □

We can summarize the surprisingly broad stability properties of semimartingales.

Theorem 67. *The space of semimartingales is a vector space, an algebra, a lattice, and is stable under C^2, and more generally under convex transformations.*

Proof. In Chap. II we saw that semimartingales form a vector space (Theorem 1), an algebra (Corollary 2 of Theorem 22: Integration by Parts), and that they are stable under C^2 transformations (Theorem 32: Itô's Formula). That they form a lattice is by Corollary 2 above, and that they are stable under convex transformations is Theorem 66. □

Definition. The **sign function** is defined to be

$$\operatorname{sign}(x) = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x \leq 0. \end{cases}$$

Note that our definition of sign is not symmetric. We further define

$$h_0(x) = |x| \quad \text{and} \quad h_a(x) = |x-a|. \tag{$*$}$$

Then $\operatorname{sign}(x)$ is the *left derivative* of $h_0(x)$, and $\operatorname{sign}(x-a)$ is the left derivative of $h_a(x)$. Since $h_a(x)$ is convex by Theorem 66 we have for a semimartingale X

$$h_a(X_t) = |X_t - a| = |X_0 - a| + \int_{0+}^{t} \operatorname{sign}(X_{s-} - a)dX_s + A_t^a, \tag{$**$}$$

where A_t^a is the increasing process of Theorem 66. Using $(*)$ and $(**)$ as defined above we can define the *local* time of an arbitrary semimartingale.

Definition. Let X be a semimartingale, and let h_a and A^a be as defined in $(*)$ and $(**)$ above. The **local time** at a of X, denoted $L_t^a = L^a(X)_t$, is defined to be the process given by

$$L_t^a = A_t^a - \sum_{0<s\leq t} \{h_a(X_s) - h_a(X_{s-}) - h_a'(X_{s-})\Delta X_s\}.$$

Notice that by the corollary of Theorem 63 the integral $\int_{0+}^{t} \operatorname{sign}(X_{s-} - a)dX_s$ in $(**)$ has a version which is jointly measurable in (a, t, ω) and càdlàg in

t. Therefore so does $(A_t^a)_{t\geq 0}$, and finally so too does the local time L_t^a. *We always choose this jointly measurable, càdlàg version of the local time,* without any special mention. We further observe that the jumps of the process A^a defined in $(**)$ are precisely $\sum_{s\leq t}\{h_a(X_s) - h_a(X_{s-}) - h_a(X_{s-})\Delta X_s\}$ (by Theorem 66), and therefore *the local time $(L_t^a)_{t\geq 0}$ is continuous in t.* Indeed, the local time L^a is the continuous part of the increasing process A^a.

The next theorem is quite simple yet crucial to proving the properties of L^a that justify its name.

Theorem 68. *Let X be a semimartingale and let L^a be its local time at a. Then*

$$(X_t - a)^+ - (X_0 - a)^+ = \int_{0+}^t 1_{\{X_{s-}>a\}} dX_s + \sum_{0<s\leq t} 1_{\{X_{s-}>a\}}(X_s - a)^- + \sum_{0<s\leq t} 1_{\{X_{s-}\leq a\}}(X_s - a)^+ + \frac{1}{2}L_t^a;$$

$$(X_t - a)^- - (X_0 - a)^- = -\int_{0+}^t 1_{\{X_{s-}\leq a\}} dX_s + \sum_{0<s\leq t} 1_{\{X_{s-}>a\}}(X_s - a)^- + \sum_{0<s\leq t} 1_{\{X_{s-}\leq a\}}(X_s - a)^+ + \frac{1}{2}L_t^a.$$

Proof. Applying Theorem 66 to the convex functions $f(x) = (x-a)^+$ and $g(x) = (x-a)^-$ we get

$$f(X_t) = f(X_0) + \int_{0+}^t f'(X_{s-})dX_s + C_t^+$$
$$g(X_t) = g(X_0) + \int_{0+}^t g'(X_{s-})dX_s + C_t^-$$

Next let

$$D_t^+ = C_t^+ - \sum_{0<s\leq t}\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\},$$
$$D_t^- = C_t^- - \sum_{0<s\leq t}\{g(X_s) - g(X_{s-}) - g'(X_{s-})\Delta X_s\}$$

and subtracting the formulas we get $C_t^+ - C_t^- = 0$ and hence $D_t^+ = D_t^-$. Also $D_t^+ + D_t^- = L_t^a$, so that $D_t^+ = D_t^- = \frac{1}{2}L_t^a$, and the proof is complete. □

The next theorem, together with the "occupation time density" formula (Corollary 1 of Theorem 70), are the traditional justifications for the terminology "local time."

Theorem 69. *Let X be a semimartingale, and let L_t^a be its local time at the level a, each $a \in \mathbb{R}$. For a.a. ω, the measure in t, $dL_t^a(\omega)$, is carried by the set $\{s : X_{s-}(\omega) = X_s(\omega) = a\}$.*

Proof. Since L_t^a has continuous paths, the measure $dL_t^a(\omega)$ is diffuse, and since $\{s : X_{s-}(\omega) = a\}$ and $\{s : X_{s-}(\omega) = X_s(\omega) = a\}$ differ by at most a countable set, it will suffice to show that $dL_t^a(\omega)$ is carried by the set $\{s : X_{s-}(\omega) = a\}$.

Suppose S, T are stopping times and that $0 < S \le T$ such that $[S,T) \subset \{(s,\omega) : X_{s-}(\omega) < a\} \equiv \{X_- < a\}$. Then $X \le a$ on $[S,T)$ as well. Hence by the first equation in Theorem 68 we have

$$(X-a)_T^+ - (X-a)_S^+ = \int_S^T 1_{\{X_{s-}>a\}} dX_s + \sum_{S<s\le T} 1_{\{X_{s-}>a\}}(X_s - a)^- + \sum_{S<s\le T} 1_{\{X_{s-}<a\}}(X_s-a)^+ + \frac{1}{2}(L_T^a - L_S^a).$$

However the left side of the above equation equals zero, and all terms on the right side except possibly $\frac{1}{2}(L_T^a - L_S^a)$ also equal zero. Therefore $\frac{1}{2}(L_T^a - L_S^a) = 0$, whence $L_T^a = L_S^a$.

Next for $r \in \mathbb{Q}$, the rationals, define the stopping times $S_r(\omega)$, $r > 0$, by

$$S_r(\omega) = \begin{cases} r, & \text{if } X_{r-}(\omega) < a, \\ \infty, & \text{if } X_{r-}(\omega) \ge a. \end{cases}$$

Then define

$$T_r(\omega) = \inf\{t > S_r(\omega) : X_{t-}(\omega) \ge a\}.$$

Then $[S_r, T_r) \subset \{X_- < a\}$, and moreover the interior of the set $\{X_- < a\}$ equals $\bigcup_{r\in\mathbb{Q},\, r>0}(S_r, T_r)$. As we have now seen, dL^a does not charge the interior of the set $\{X_- < a\}$. This set is open on the left, hence it differs from its interior by an at most countable set, and since dL^a is diffuse it doesn't charge countable sets. Thus dL^a does not charge the set $\{X_- < a\}$ itself.

Analogously one can show dL^a does not charge $\{X_- > a\}$. Hence its support is contained in the set $\{X_- = a\}$, and we are done. □

The next theorem gives a very satisfying generalization of Itô's formula (Theorem 32 of Chap. II).

Theorem 70 (Meyer-Itô Formula). *Let f be the difference of two convex functions, let f' be its left derivative, and let μ be the signed measure (when restricted to compacts) which is the second derivative of f in the generalized function sense. Then the following equation holds:*

$$f(X_t) - f(X_0) = \int_{0+}^t f'(X_{s-})dX_s + \sum_{0<s\le t}\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\} + \frac{1}{2}\int_{-\infty}^{\infty} \mu(da) L_t^a, \tag{$*$}$$

where X is a semimartingale and $L_t^a = L_t^a(X)$ is its local time at a.

Proof. Notice that equation $(*)$ is trivially true if f is an affine function (i.e., $f(x) = ax + b$). Next let us first assume that μ is a signed measure with finite total mass on a compact interval. Define a function g by

$$g(x) = \frac{1}{2} \int |x - y| \mu(dy).$$

It is well known that f and g differ by at most an affine function $h(x) = ax+b$. Therefore without loss of generality we can assume that the function f in $(*)$ is of the form $f(x) = \frac{1}{2} \int |x - y| \mu(dy)$. Then $f'(x) = \frac{1}{2} \int \operatorname{sign}(x - y) \mu(dy)$ and $f''(x) = \mu(dx)$. Moreover if

$$J_t^y = \sum_{0<s\leq t} |X_s - y| - |X_{s-} - y| - \operatorname{sign}(X_{s-} - y) \Delta X_s,$$

then

$$\frac{1}{2} \int J_t^y \mu(dy) = \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-}) \Delta X_s\}. \tag{1}$$

Also, letting $H_t^y = |X_t - y| - |X_0 - y|$, one has

$$\frac{1}{2} \int H_t^y \mu(dy) = f(X_t) - f(X_0). \tag{2}$$

Next consider $Z_t^y = \int_{0+}^t \operatorname{sign}(X_{s-} - y) dX_s$, and let Z^y be the $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ measurable version, which we know exists by Theorem 63. By Fubini's Theorem (Theorem 64) we have

$$\frac{1}{2} \int Z_t^y \mu(dy) = \int_{0+}^t f'(X_{s-}) dX_s. \tag{3}$$

Since $L_t^a = H_t^a - J_t^a - Z_t^a$, we have

$$\frac{1}{2} \int L_t^a \mu(da) = \frac{1}{2} \int H_t^a \mu(da) - \frac{1}{2} \int J_t^a \mu(da) - \frac{1}{2} \int Z_t^a \mu(da) \tag{4}$$

and combining (4) with (1), (2), and (3), yields the formula $(*)$.

It remains to reduce the general case to that of μ supported on a compact interval with finite total mass. By linearity it further suffices to treat the case where f is convex. Define

$$f_n(x) = \begin{cases} f(n) + f'(n)(x - n), & \text{if } x \geq n, \\ f(x), & \text{if } -n \leq x \leq n, \\ f(-n) + f'(-n)(x + n), & \text{if } x \leq -n. \end{cases}$$

Then f_n is also convex and its second (generalized) derivative $\mu_n(dx)$ agrees with μ on $[-n, n)$ and is of finite total mass. Let $Y^n = f_n(X)$, and $T^n =$

$\inf\{t > 0 : |X_t| \geq n\}$. Note that $Y^n = f(X)$ on $[0, T^n)$, and that $L^a_{T^n} = 0$ for all a, $|a| \geq n$, by Theorem 69. Therefore $\int L^a_t \mu_n(da) = \int L^a_t \mu(da)$ for $t \leq T^n$, and by the preceding $(*)$ holds for $Y = f_n(X) = f(X)$ on $[0, T^n)$. Therefore $(*)$ holds on $[0, T^n)$. Since the stopping times $(T^n)_{n\geq 1}$ increase to ∞ a.s., we have $(*)$ holds on all of $\mathbb{R}_+ \times \Omega$, and the theorem is proved. □

The next formula gives an interpretation of semimartingale local time as an occupation density relative to the random "clock" $d[X, X]^c_s$.[17]

Corollary 1. Let X be a semimartingale with local time $(L^a)_{a\in\mathbb{R}}$. Let g be a bounded Borel measurable function. Then a.s.

$$\int_{-\infty}^{\infty} L^a_t g(a)da = \int_0^t g(X_{s-})d[X, X]^c_s.$$

Proof. Let f be convex and C^2. Comparing $(*)$ of Theorem 70 with Itô's formula (Theorem 32 of Chap. II) shows that

$$\int_{-\infty}^{\infty} L^a_t f''(a)da = \int_0^t f''(X_{s-})d[X, X]^c_s$$

where $\mu(da)$ is of course $f''(a)da$. Since the above holds for any continuous and positive function f'', a monotone class argument shows that it must hold, up to a P-null set, for any bounded, Borel measurable function g. □

We record here an important special case of Corollary 1.

Corollary 2. Let X be a semimartingale with local time $(L^a)_{a\in\mathbb{R}}$. Then

$$[X, X]^c_t = \int_{-\infty}^{\infty} L^a_t da.$$

Corollary 3 (Meyer-Tanaka Formula). Let X be a semimartingale with continuous paths. Then

$$|X_t| = |X_0| + \int_{0+}^t \operatorname{sign}(X_s)dX_s + L^0_t.$$

Proof. This is merely Theorem 70 with $f(x) = |x|$, which implies $\mu(da) = 2\varepsilon_0(da)$, point mass at 0. The formula also follows trivially from the definition of L^0. □

If $X_t = B_t$ is a standard Brownian motion, then $M_t = \int_0^t \operatorname{sign}(B_s)dB_s$ is a continuous local martingale, and $[M, M]_t = \int_0^t \operatorname{sign}(B_s)^2 d[B, B]_s = [B, B]_t = t$. Therefore by Lévy's Theorem (Theorem 39 of Chap. II) we know that M_t

[17] Recall that $[X, X]^c$ denotes the path-by-path continuous part of $t \mapsto [X, X]_t$ with $[X, X]^c_0 = 0$.

is another Brownian motion. We therefore have, where $B_0 = x$ for standard Brownian motion,

$$|B_t| = |B_0| + \beta_t + L_t, \tag{**}$$

where $\beta_t = \int_0^t \operatorname{sign}(B_s)dB_s$ is a Brownian motion, and L_t is the local time of B at zero. Formula (**) is known as **Tanaka's Formula**.

Observe that if $f(x) = |x|$, then $f'' = 2\delta(x)$, where δ is the "delta function at 0," which of course is a generalized function, or "distribution." Thus Corollaries 1 and 3 give the intuitive interpretation of local time as $L_t^0 = \int_0^t \delta(X_s)d[X,X]_s$, and $L_t^a = \int_0^t \delta(X_s - a)d[X,X]_s$, for continuous semimartingales.

Local times also allow us to give simple examples of functions f that are only slightly more general than convex functions and are such that $f(X)$ need not be a semimartingale when X is one. Let $f(x) = |x|^\alpha$, $0 < \alpha < 1$. Then f cannot be expressed as the difference of two convex functions because the slope of f becomes vertical as x decreases to zero. The proof of this result is simpler if we restrict α to $0 < \alpha < 1/2$.

Theorem 71. *Let X be a continuous local martingale with $X_0 = 0$, and let $0 < \alpha < 1/2$. Then $Y_t = |X_t|^\alpha$ is not a semimartingale unless X is identically zero.*

Proof. Let us suppose that $Y = |X|^\alpha$ is a semimartingale. Let $\beta = 1/\alpha > 2$. We then have

$$|X_t| = Y_t^\beta = \beta \int_0^t Y_s^{\beta-1} dY_s + \frac{\beta(\beta-1)}{2} \int_0^t Y_s^{\beta-2} d[Y,Y]_s,$$

since Y has continuous paths. Using Theorem 69 and the Meyer-Tanaka formula (Corollary 3 of Theorem 70) we have, where $L_t = L_t^0$ is the local time of X at 0,

$$\begin{aligned} L_t &= \int_0^t 1_{\{X_s=0\}} dL_s \\ &= -\int_0^t 1_{\{X_s=0\}} \operatorname{sign}(X_s)dX_s + \int_0^t 1_{\{X_s=0\}} d|X|_s \end{aligned}$$

The integral $\int_0^t 1_{\{X_s=0\}} \operatorname{sign}(X_s)dX_s$ is identically zero. For if T^n is a sequence of stopping times increasing to ∞ a.s. such that X^{T^n} is a bounded martingale, then

$$\begin{aligned} E\{(\int_0^{t\wedge T^n} 1_{\{X_s=0\}} \operatorname{sign}(X_s)dX_s)^2\} &= E\{\int_0^{t\wedge T^n} 1_{\{X_s=0\}} d[X,X]_s\} \\ &= E\{\int_{-\infty}^{\infty} L_{t\wedge T^n}^a 1_{\{0\}}(a)da\} = 0. \end{aligned}$$

Therefore $L_t = L_t^0 = \int_0^t 1_{\{X_s=0\}} d|X|_s$. Since $\{X_s = 0\}$ equals $\{Y_s = 0\}$, this becomes

$$\begin{aligned} L_t &= \int_0^t 1_{\{Y_s=0\}} d|X|_s \\ &= \beta \int_0^t Y_s^{\beta-1} 1_{\{Y_s=0\}} dY_s + \frac{\beta(\beta-1)}{2} \int_0^t Y_s^{\beta-2} 1_{\{Y_s=0\}} d[Y,Y]_s \\ &= \beta \int_0^t 0\, dY_s + \frac{\beta(\beta-1)}{2} \int_0^t 0\, d[Y,Y]_s \\ &= 0. \end{aligned}$$

Using the Meyer-Tanaka formula again we conclude that

$$|X_t| = \int_0^t \operatorname{sign}(X_s) dX_s.$$

Since X is a continuous local martingale, so also is the stochastic integral on the right side above (Theorem 30); it is also non-negative, and equal to zero at 0. Such a local martingale must be identically zero. This completes the proof. □

It is worth noting that if X is a bounded, continuous martingale with $X_0 = 0$, then $Y_t = |X_t|^\alpha$, $0 < \alpha < 1/2$ is an example of an *asymptotic martingale*, or AMART, which is not a semimartingale.[18]

We next wish to determine when there exists a version of the local time L_t^a which is jointly continuous in $(a,t) \mapsto L_t^a$ a.s., or jointly right continuous in $(a,t) \mapsto L_t^a$. We begin with the classical result of Kolmogorov which gives a sufficient condition for joint continuity. There are several versions of Kolmogorov's Lemma. We give here a quite general one because we will use it often in Chap. V. In this section we use only its corollary, which can also be proved directly in a fairly simple way.

Before the statement of the theorem we establish some notation. Let Δ denote the dyadic rational points of the unit cube $[0,1]^n$ in $\mathbb{R}^n$, and let Δ_m denote all $x \in \Delta$ whose coordinates are of the form $k2^{-m}$, $0 \le k \le 2^m$.

Theorem 72 (Kolmogorov's Lemma). *Let (E,d) be a complete metric space, and let U^x be an E-valued random variable for all x dyadic rationals in $\mathbb{R}^n$. Suppose that for all x, y, we have $d(U^x,U^y)$ is a random variable and that there exist strictly positive constants ε, C, β such that*

$$E\{d(U^x,U^y)^\varepsilon\} \le C\|x-y\|^{n+\beta}.$$

Then for almost all ω the function $x \mapsto U^x$ can be extended uniquely to a continuous function from $\mathbb{R}^n$ to E.

---

[18] For a more elementary example of an asymptotic martingale that is not a semimartingale, see Gut [86, page 7].

Proof. We prove the theorem for the unit cube $[0,1]^n$ and leave the extension to $\mathbb{R}^n$ to the reader. Two points x and y in Δ_m are *neighbors* if $\sup_i |x^i y^i| = 2^{-m}$. We use Chebyshev's inequality on the inequality hypothesized to get

$$P\{d(U^x, U^y) \geq 2^{-\alpha m}\} \leq C 2^{\alpha \varepsilon m} 2^{-m(n+\beta)}.$$

Let

$$\Lambda_m = \{\omega : \exists \text{ neighbors } x, y \in \Delta_m \text{ with } d(U^x(\omega), U^y(\omega)) \geq 2^{-\alpha m}\}.$$

Since each $x \in \Delta_m$ has at most 3^n neighbors, and the cardinality of Δ_m is 2^{mn}, we have

$$P(\Lambda_m) \leq c 2^{m(\alpha\varepsilon - \beta)}$$

where the constant $c = 3^n C$. Take α sufficiently small so that $\alpha\varepsilon < \beta$. Then

$$P(\Lambda_m) \leq c 2^{-m\delta}$$

where $\delta = \beta - \alpha\varepsilon > 0$. The Borel-Cantelli Lemma then implies $P(\Lambda_m$ infinitely often$) = 0$. That is, there exists an m_0 such that for $m \geq m_0$ and every pair (u, v) of points of Δ_m that are neighbors,

$$d(U^u, U^v) \leq 2^{-\alpha m}.$$

We now use the preceding to show that $x \mapsto U^x$ is uniformly continuous on Δ and hence extendable uniquely to a continuous function on $[0,1]^n$. To this end, let $x, y \in \Delta$ be such that $\|x - y\| \leq 2^{-k-1}$. We will show that $d(U^x, U^y) \leq c 2^{-\alpha k}$ for a constant c, and this will complete the proof.

Without loss of generality assume $k \geq m_0$. Then $x = (x^1, \ldots, x^n)$ and $y = (y^1, \ldots, y^n)$ in Δ with $\|x - y\| \leq 2^{-k-1}$ have dyadic expansions of the form

$$x^i = u^i + \sum_{j>k} a^i_j 2^{-j}$$
$$y^i = v^i + \sum_{j>k} b^i_j 2^{-j}$$

where a^i_j, b^i_j are each 0 or 1, and u, v are points of Δ_k which are either equal or neighbors.

Next set $u_0 = u$, $u_1 = u_0 + a_{k+1} 2^{-k-1}$, $u_2 = u_1 + a_{k+2} 2^{-k-2}, \ldots$. We also make analogous definitions for $v_0, v_1, v_2, \ldots$. Then u_{i-1} and u_i are equal or neighbors in Δ_{k+i}, each i, and analogously for v_{i-1} and v_i. Hence

$$d(U^x(\omega), U^u(\omega)) \leq \sum_{j=k}^{\infty} 2^{-\alpha j}$$
$$d(U^y(\omega), U^v(\omega)) \leq \sum_{j=k}^{\infty} 2^{-\alpha j}$$

and moreover

$$d(U^u(\omega), U^v(\omega)) \leq 2^{-\alpha k}.$$

The result now follows by the triangle inequality. □

Comment. If the complete metric space (E, d) in Theorem 72 is *separable*, then the hypothesis that $d(U^x, U^y)$ be measurable is satisfied. Often the metric spaces chosen are one of $\mathbb{R}$, $\mathbb{R}^d$, or the function space $\mathcal{C}$ with the sup norm and these are separable.

A complete metric space that arises often in Chap. V is the space $E = \mathcal{D}^n$ of càdlàg functions mapping $[0, \infty)$ into $\mathbb{R}^n$, topologized by uniform convergence on compacts. While this is a complete metric space, *it is not separable.* Indeed, a compatible metric is

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n}(1 \wedge \sup_{0 \leq s \leq n} |f(s) - g(s)|).$$

However if $f_\alpha(t) = 1_{[\alpha,\infty)}(t)$, then $d(f_\alpha, f_\beta) = 1/2$ for all α, β with $0 \leq \alpha < \beta \leq 1$, and since there are uncountably many such α, β, the space is not separable. Fortunately, however, the condition that $d(U^x, U^y)$ be measurable is nevertheless satisfied in this case, due to the path regularity of the functions in the function space $\mathcal{D}^n$. (Note that in many other contexts the space $\mathcal{D}^n$ is endowed with the Skorohod topology, and with this topology $\mathcal{D}^n$ is a complete metric space which is also separable; see for example Ethier-Kurtz [71] or Jacod-Shiryaev [110].)

We state as a corollary the form of Kolmogorov's Lemma (also known as Kolmogorov's continuity criterion) that we will use in our study of local times.

Corollary 1 (Kolmogorov's Continuity Criterion). Let $(X_t^a)_{t \geq 0, a \in \mathbb{R}^n}$ be a parameterized family of stochastic processes such that $t \mapsto X_t^a$ is càdlàg a.s., each $a \in \mathbb{R}^n$. Suppose that

$$E\{\sup_{s \leq t} |X_s^a - X_s^b|^\varepsilon\} \leq C(t)\|a - b\|^{n+\beta}$$

for some $\varepsilon, \beta > 0$, $C(t) > 0$. Then there exists a version $\hat{X}_t^a$ of X_t^a which is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{F}$ measurable and which is càdlàg in t and uniformly continuous in a on compacts and is such that for all $a \in \mathbb{R}^n$, $t \geq 0$,

$$\hat{X}_t^a = X_t^a \quad \text{a.s.}$$

(The null set $\{\hat{X}_t^a \neq X_t^a\}$ can be chosen independently of t.) In this section we will use the above corollary for parameterized processes X^a which are *continuous* in t. In this case the process obtained from the corollary of Kolmogorov's Lemma, $\hat{X}^a$, *will be jointly continuous in* (a, t) almost surely.

In particular, Kolomogorov's Lemma can be used to prove that the paths of standard Brownian motion are continuous.

Corollary 2. Let B be standard Brownian motion. Then there is a version of B with continuous paths, a.s.

Proof. Since $B_t - B_s$ is Gaussian with mean zero and variance $t-s$, we know that $E\{|B_t - B_s|^4\} \leq c(t-s)^2$. (One can give a cute proof of this moment estimate using the scaling property of Brownian motion.) If we think of time as the parameter and the process as being constant in time, we see that the exponent 4 is strictly positive, and that the exponent on the right, 2, is strictly bigger than the dimension, which is of course 1. Corollary 2 now follows from Corollary 1. □

Hypothesis A. For the remainder of this section we let X denote a semimartingale with the restriction that $\sum_{0<s\leq t} |\Delta X_s| < \infty$ a.s., each $t > 0$.

Observe that if $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ is a probability space where $(\mathcal{F}_t)_{t\geq 0}$ is the completed minimal filtration of a Brownian motion $B = (B_t)_{t\geq 0}$, *then all semimartingales on this Brownian space verify Hypothesis A.* Indeed, by Corollary 1 of Theorem 43 all the local martingales are continuous. Thus if X is a semimartingale, let $X = M + A$ be a decomposition with M a local martingale and A an FV process. Then the jump processes ΔX and ΔA are equal, hence

$$\sum_{0<s\leq t} |\Delta X_s| = \sum_{0<s\leq t} |\Delta A_s| < \infty,$$

since A is of finite variation on compacts.

Let X be a semimartingale satisfying Hypothesis A, and let

$$J_t = \sum_{0\leq s\leq t} \Delta X_s;$$

that is, J is the process which is the sum of the jumps. Because of our hypothesis J is an FV process and thus also a semimartingale. Moreover $Y_t = X_t - J_t$ is a continuous semimartingale with $Y_0 = 0$, and we let

$$Y = M + A$$

be its (unique) decomposition, where M is a continuous local martingale and A is a continuous FV process with $M_0 = A_0 = 0$. Such a process M is then uniquely determined and we can write

$$M = X^c,$$

the *continuous local martingale part of X*.

Notation. We assume given a semimartingale X satisfying Hypothesis A. If Z is any other semimartingale we write

$$\hat{Z}_t^a = \int_{0+}^t 1_{\{X_{s-}>a\}} dZ_s \qquad (a \in \mathbb{R}).$$

This notation should not be confused with that of Kolmogorov's Lemma (Theorem 72). It is always assumed that we take the $\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{F}$ measurable, càdlàg version of Z^a (cf., Theorem 63).

Before we can prove our primary regularity property (Theorem 75), we need two preliminary results. The first is a very special case of a family of martingale inequalities known as the Burkholder-Davis-Gundy inequalities.

Theorem 73 (Burkholder's Inequality). *Let X be a continuous local martingale with $X_0 = 0$, $2 \le p < \infty$, and T a finite stopping time. Then*

$$E\{(X_T^*)^p\} \le C_p E\{[X,X]_T^{p/2}\}$$

where $C_p = \{q^p(\frac{p(p-1)}{2})\}^{p/2}$, with $\frac{1}{p}+\frac{1}{q}=1$.

Proof. By stopping, it suffices to consider the case where X and $[X,X]$ are bounded. By Itô's formula we have

$$|X_T|^p = p\int_0^T \operatorname{sign}(X_s)|X_s|^{p-1}dX_s + \frac{p(p-1)}{2}\int_0^T |X_s|^{p-2}d[X,X]_s.$$

By Doob's inequalities (Theorem 20 of Chap. I) we have (with $\frac{1}{p}+\frac{1}{q}=1$)

$$\begin{aligned} E\{(X_T^*)^p\} &\le q^p E\{|X_T|^p\} \\ &= q^p E\{\frac{p(p-1)}{2}\int_0^T |X_s|^{p-2}d[X,X]_s\} \\ &\le q^p\left(\frac{p(p-1)}{2}\right)E\{(X_T^*)^{p-2}[X,X]_T\} \\ &\le q^p\left(\frac{p(p-1)}{2}\right)E\{(X_T^*)^p\}^{1-\frac{2}{p}}E\{[X,X]_T^{p/2}\}^{2/p}, \end{aligned}$$

with the last inequality by Hölder's inequality. Since $E\{(X_T^*)^p\}^{1-\frac{2}{p}} < \infty$, we divide both sides by it to obtain

$$E\{(X_T^*)^p\}^{2/p} \le q^p\left(\frac{p(p-1)}{2}\right)E\{[X,X]_T^{p/2}\}^{2/p},$$

and raising both sides to the power $p/2$ gives the result. □

Actually much more is true. Indeed for any local martingale (continuous or not) it is known that there exist constants c_p, C_p such that for a finite stopping time T

$$E\{(X_T^*)^p\}^{1/p} \le c_p E\{[X,X]_T^{p/2}\}^{1/p} \le C_p E\{(X_T^*)^p\}^{1/p}$$

for $1 \le p < \infty$. See Sect. 3 of Chap. VII of Dellacherie-Meyer [46] for these and related results. When the local martingales are continuous some results even hold for $0 < p < 1$ (see, e.g., Barlow-Jacka-Yor [9, Table 4.1 page 162]).

Theorem 74. *Let X be a semimartingale satisfying Hypothesis A. There exists a version of $(\hat{X}^c)_t^a$ such that $(a,t,\omega) \mapsto (\hat{X}^c)_t^a(\omega)$ is $\mathcal{B}(\mathbb{R}) \otimes \mathcal{P}$ measurable, and everywhere jointly continuous in (a,t).*

Proof. Without loss of generality we may assume $X - X_0 \in \mathcal{H}^2$. If it is not, we can stop $X - X_0$ at T^n-. The continuous local martingale part of X^{T^n-} is then just $(X^c)^{T^n}$. Suppose $-\infty < a < b < \infty$, and let

$$\alpha_t(a,b) = E\{(\int_0^t 1_{\{b \geq X_{s-} > a\}} d[X,X]_s^c)^2\}.$$

By Corollary 1 of Theorem 70 we have

$$\begin{aligned}\alpha_t(a,b) = E\{(\int_a^b L_t^u du)^2\} &= (b-a)^2 E\{(\frac{1}{b-a}\int_a^b L_t^u du)^2\} \\ &\leq (b-a)^2 E\{\frac{1}{b-a}\int_a^b (L_t^u)^2 du\},\end{aligned}$$

by the Cauchy-Schwarz inequality. The above implies

$$\alpha_t(a,b) \leq (b-a)^2 \sup_{u \in (a,b)} E\{(L_t^u)^2\}.$$

By the definition,

$$L_t^u \leq A_t^u \leq |X_t - X_0| - \int_{0+}^t \operatorname{sign}(X_{s-} - u) dX_s$$

and therefore

$$\begin{aligned}E\{(L_t^u)^2\} &\leq 2E\{|X_t - X_0|^2\} + 2E\{(\int_{0+}^t \operatorname{sign}(X_{s-} - u) dX_s)^2\} \\ &\leq 4\|X - X_0\|_{\mathcal{H}^2}^2 + 4\|\operatorname{sign}(X_{s-} - u) \cdot (X - X_0)\|_{\mathcal{H}^2}^2 \\ &\leq 8\|X - X_0\|_{\mathcal{H}^2}^2 < \infty,\end{aligned}$$

and the bound is independent of u. Therefore

$$\alpha_t(a,b) \leq (b-a)^2 \Gamma,$$

for a constant $\Gamma < \infty$, and *independent of* t. Next using Burkholder's inequality (Theorem 73) we have

$$\begin{aligned}E\{\sup_s |(\hat{X}^c)_s^b - (\hat{X}^c)_s^a|^4\} &\leq C_4 E\{(\int_0^\infty 1_{\{b \geq X_{s-} > a\}} d[X,X]_s^c)^2\} \\ &\leq C_4 \sup_t \alpha_t(a,b) \\ &\leq C_4 \Gamma (b-a)^2.\end{aligned}$$

The result now follows by applying Kolmogorov's Lemma (the corollary of Theorem 72). □

We can now establish our primary result.

Theorem 75. *Let X be a semimartingale satisfying Hypothesis A. Then there exists a $\mathcal{B}(\mathbb{R})\otimes\mathcal{P}$ measurable version of $(a,t,\omega)\mapsto L_t^a(\omega)$ which is everywhere jointly right continuous in a and continuous in t. Moreover a.s. the limits $L_t^{a-}=\lim_{b\to a,\,b<a}L_t^b$ exist.*

Proof. Since X satisfies Hypothesis A, the process $J_t=\sum_{0\le s\le t}\Delta X_s$ is an FV semimartingale, and $Y=X-J$ is a continuous semimartingale. We let $Y=M+A$ be the (unique) decomposition of Y, with $M_0=A_0=0$.

Then $X=M+A+J$. Further define

$$S_t^a=\sum_{0<s\le t}1_{\{X_{s-}>a\}}(X_s-a)^-+\sum_{0<s\le t}1_{\{X_{s-}\le a\}}(X_s-a)^+.$$

Observe that $|S_t^a|\le\sum_{0<s\le t}|\Delta X_s|<\infty$. By Theorem 68 we have

$$(X_t-a)^+-(X_0-a)^+=(\hat{M})_t^a+(\hat{A})_t^a+(\hat{J})_t^a+S_t^a+\frac{1}{2}L_t^a.$$

Consider

$$(\hat{A})_t^a=\int_{0+}^t 1_{\{X_{s-}>a\}}dA_s.$$

We have

$$\lim_{\substack{a\to b\\a>b}}(\hat{A})_t^a=\int_{0+}^t 1_{\{X_{s-}>b\}}dA_s$$

and

$$\lim_{\substack{a\to b\\a<b}}(\hat{A})_t^a=\int_{0+}^t 1_{\{X_{s-}\ge b\}}dA_s$$

where the convergence is uniform in t on $[0,\tau]$, and $\tau>0$. We have analogous results for $(\hat{J})_t^a$ and also for S_t^a, because it is dominated by $\sum_{0<s\le t}|\Delta X_s|<\infty$. Since we already know that $(\hat{M})_t^a$ is continuous, the proof is complete. □

The next three corollaries are simple consequences of Theorem 75. Recall that for a semimartingale X satisfying Hypothesis A, we let $J_t=\sum_{0\le s\le t}\Delta X_s$, $Y=X-J$, and $Y=M+A$ be the (unique) decomposition of Y, with $A_0=0$.

Corollary 1. Let X be a semimartingale satisfying Hypothesis A. Let $X=M+A+J$ be a decomposition with M and A continuous and J the jump process of X, and let $(L_t^a)_{t\ge0}$ be its local time at the level a. Then

$$\begin{aligned}L_t^a-L_t^{a-}&=2\int_0^t 1_{\{X_{s-}=a\}}dA_s\\&=2\int_0^t 1_{\{X_s=a\}}dA_s.\end{aligned}$$

An example of a semimartingale satisfying Hypothesis A but having a discontinuous local time is $X_t = |B_t|$ where B is standard Brownian motion with $B_0 = 0$. Here $L^a(X) = L^a(B) + L^{-a}(B)$ for $a > 0$, $L^0(X) = L^0(B)$, and $L^a(X) = 0$ for $a < 0$. J. Walsh [227] has given a more interesting (but more complicated) example.

Corollary 2. Let X be a semimartingale satisfying Hypothesis A. Let A be as in Corollary 1. The local time (L_t^a) is continuous in t and is continuous at $a = a_0$ if and only if

$$\int_0^\infty 1_{\{X_s=a_0\}}|dA_s| = 0.$$

Observe that if $X = B$, a Brownian motion (or any continuous local martingale), then $A = 0$ and the local time of X can be taken everywhere jointly continuous.

Corollary 3. Let X be a semimartingale satisfying Hypothesis A. Then for every (a,t) we have

$$L_t^a = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_0^t 1_{\{a\leq X_s\leq a+\varepsilon\}} d[X,X]_s^c, \quad \text{a.s.}$$

and

$$L_t^{a-} = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\int_0^t 1_{\{a-\varepsilon\leq X_s\leq a\}} d[X,X]_s^c, \quad \text{a.s.}$$

The lack of symmetry in the above formula stems from the definition of local time, where we defined $\operatorname{sign}(x)$ in an asymmetric way:

$$\operatorname{sign}(x) = \begin{cases} 1, & x > 0, \\ -1, & x \leq 0. \end{cases}$$

A symmetrized result follows trivially, namely

$$\frac{L_t^a + L_t^{a-}}{2} = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_0^t 1_{\{|X_s-a|\leq\varepsilon\}} d[X,X]_s^c, \quad \text{a.s.}$$

Corollary 3 is intuitively very appealing, and justifies thinking of local time as an *occupation time density*. Also if $X = B$, a Brownian motion, then Corollary 3 becomes the classical result

$$L_t^a = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon}\int_0^t 1_{(a-\varepsilon,a+\varepsilon)}(B_s)ds, \quad \text{a.s.}$$

Local times have interesting properties when viewed as processes with "time" fixed, and the space variable "a" as the parameter. The Ray-Knight Theorem giving the distribution of $(L_T^{1-a})_{0\leq a<1}$, the Brownian local time sampled at the hitting time of 1, as the square of a two dimensional Bessel diffusion, is one example.

The Meyer-Itô formula (Theorem 70) can be extended in a different direction, which gives rise to the Bouleau-Yor formula, allowing non-convex functions of semimartingales. The key idea is that the function $a \mapsto L_U^a$ induces a measure on $\mathbb{R}$ if L is the local time of a semimartingale X satisfying Hypothesis A and U is a positive random variable.

Theorem 76. *Let X be a semimartingale satisfying Hypothesis A, U a positive random variable, L^a the local times of X. Then the operation*

$$f \mapsto \sum_{i=1}^{n} f_i(L_U^{a_{i+1}} - L_U^{a_i}),$$

where $f(x) = \sum f_i 1_{(a_i, a_{i+1}]}(x)$, can be extended uniquely to a vector measure on $\mathcal{B}(\mathbb{R})$ with values in L^0.

Proof. Recall that L^0 denotes finite-valued random variables. By Theorem 68 we know that

$$\begin{aligned}\frac{1}{2}L_t^a = (X_t - a)^- - (X_0 - a)^- + \int_{0+}^t 1_{\{X_{s-} \leq a\}} dX_s \\ - \sum_{0<s\leq t} 1_{\{X_{s-}>a\}}(X_s - a)^- - \sum_{0<s\leq t} 1_{\{X_{s-}\leq a\}}(X_s - a)^+.\end{aligned}$$

We write S_t^a for the last two sums on the right side of the equation above. Note first that $(X_U - a)^- - (X_0 - a)^-$ is Lipschitz continuous in a and hence absolutely continuous in a. Therefore it induces a measure in a. Also the function $a \mapsto S_U^a$ is càdlàg and of finite variation in a, and moreover $\int |d_a S_U^a| \leq 2\sum_{0<s\leq U} |\Delta X_s|$. Finally consider the stochastic integral. Let $X_t^a = \int_{0+}^t 1_{\{X_{s-}\leq a\}} dX_s$. Then

$$\begin{aligned}\sum_i f_i(X_U^{a_{i+1}} - X_U^{a_i}) &= \int_{0+}^U \sum_i f_i 1_{\{a_i < X_{s-} \leq a_{i+1}\}} dX_s \\ &= \int_{0+}^U f(X_{s-}) dX_s.\end{aligned}$$

Therefore by the Dominated Convergence Theorem for stochastic integrals (Theorem 32) we have that $\sum_i f_i(X_U^{a_{i+1}} - X_U^{a_i})$ extends to a measure with values in L^0 and moreover $\int f(a) d_a X_U^a = \int_{0+}^U f(X_{s-}) dX_s$, for f bounded, Borel measurable on $\mathbb{R}$. Thus $d_a L_U^a$ can be defined as the sum of three L^0-valued measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and the theorem is proved. $\square$

In the proof of Theorem 76 we proved the following result, which is important enough to state as a corollary.

Corollary. Let X be a semimartingale satisfying Hypothesis A, U a positive random variable, and $X_t^a = \int_0^t 1_{\{X_{s-} \le a\}} dX_s$. Then $d_a X_U^a$ can be defined as an L^0-valued measure, and for $f \in b\mathcal{B}(\mathbb{R})$ we have

$$\int_{\mathbb{R}} f(a) d_a X_U^a = \int_{0+}^{U} f(X_{s-}) dX_s.$$

Combining Theorem 76 and its corollary yields the Bouleau-Yor formula. We leave its proof to the reader.

Theorem 77 (Bouleau-Yor Formula). *Let X be a semimartingale satisfying Hypothesis A, U a positive random variable, f a bounded, Borel function, and $F(x) = \int_0^x f(u)du$. Then*

$$\begin{aligned} F(X_U) - F(X_0) = &\int_{0+}^{U} f(X_{s-}) dX_s - \frac{1}{2} \int f(a) d_a L_U^a \\ &+ \sum_{0<s\le U} \{F(X_s) - F(X_{s-}) - f(X_{s-})\Delta X_s\}. \end{aligned}$$

Combining Theorem 77 with the Meyer-Itô formula (Theorem 70) we obtain the following relationship.

Corollary. Let X be a semimartingale satisfying Hypothesis A, and let T be a finite-valued random variable. Let L^a be the local times of X and let f be a $\mathcal{C}^1$ function. Then

$$\int_{-\infty}^{\infty} f'(a) L_T^a da = - \int_{-\infty}^{\infty} f(a) d_a L_T^a.$$

Remark. The Meyer-Itô formula (Theorem 70) and the Bouleau-Yor formula can be extended in interesting ways. This was first done in [77], which inspired subsequent work by N. Eisenbaum [62] and others (e.g., [6]). See [80] for a compilation of selected results in these areas and also new ones.

8 Azéma's Martingale

In order to prove the regularity properties of local times in Sect. 7, we needed *Hypothesis A*: a semimartingale X satisfies Hypothesis A if X is such that $\sum_{0<s\le t} |\Delta X_s| < \infty$ a.s., each $t > 0$. This hypothesis was needed to prove Theorems 74 through 77 and their corollaries. We present here a counterexample, known as Azéma's martingale, that shows that these results do not hold for general semimartingales. In particular, Theorems 81 and 82 show that Azéma's martingale has a local time that is zero at every level a except $a = 0$, and therefore there is no regular version in the space variable.

Let $(\Omega, \mathcal{F}, \mathbb{F}, (B_t)_{t\ge 0}, P)$ be a standard Brownian motion, $B_0 = 0$, with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$ completed (and hence right continuous). We define

$$\operatorname{sign}(x) = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x \le 0. \end{cases}$$

Set $\mathcal{G}_t^0 = \sigma\{\operatorname{sign}(B_s); s \le t\}$, and let $\mathbb{G} = (\mathcal{G}_t)_{t\ge0}$ denote the filtration $(\mathcal{G}_t^0)_{t\ge0}$ completed. (It is then right continuous as a consequence of the strong Markov property of Brownian motion.)

Let $M_t = E\{B_t|\mathcal{G}_t\}, t \ge 0$. Then for $s < t$,

$$E\{M_t|\mathcal{G}_s\} = E\{B_t|\mathcal{G}_s\} = E\{E\{B_t|\mathcal{F}_s\}|\mathcal{G}_s\} = E\{B_s|\mathcal{G}_s\} = M_s,$$

and M is a $\mathbb{G}$ martingale. We always take the càdlàg version of M.

Definition. The process $M_t = E\{B_t|\mathcal{G}_t\}$ is called **Azéma's martingale.**

Fundamentally related to Azéma's martingale is the process which gives *the last exit from zero* before time t. We define

$$g_t = \sup\{s \le t : B_s = 0\}.$$

To study the process g we need the *reflection principle for Brownian motion*, which we proved in Chap. I (Theorem 33 and its corollary) and which we recall here for convenience.

Theorem 78 (Reflection Principle for Brownian Motion). *Let B be standard Brownian motion, $B_0 = 0$, and $c > 0$. Then*

$$P(\sup_{s\le t} B_s > c) = 2P(B_t > c).$$

An elementary conditioning argument provides a useful corollary.

Corollary. Let B be standard Brownian motion, $B_0 = 0$, and $0 < s < t$. Then

$$P(\sup_{s<u\le t} B_u > 0, B_s < 0) = 2P(B_t > 0, B_s < 0).$$

Theorem 79 is known as *Lévy's arcsine law* for the last exit from zero.

Theorem 79. *The process g_t is $\mathbb{G}$ adapted, and $P(g_t \le s) = \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}$, $0 \le s \le t$.*

Proof. For $s \le t$ we have almost surely

$$\{g_t \le s\} = \Big(\bigcap_{\substack{u\in\mathbb{Q}\\ s<u<t}} \{B_u > 0\}\Big) \bigcup \Big(\bigcap_{\substack{u\in\mathbb{Q}\\ s<u<t}} \{B_u < 0\}\Big)$$

from which it follows that g_t is $\mathcal{G}_t$ adapted.

To calculate the distribution of g_t we observe for $s < t$ (using the symmetry of B),

$$
\begin{aligned}
P(g_t > s) &= 2P(g_t > s, B_s < 0) \\
&= 2P(\sup_{s<u\le t} B_u > 0, B_s < 0) \\
&= 4P(B_t > 0, B_s < 0),
\end{aligned}
$$

by the corollary of Theorem 78. Since (B_s, B_t) are jointly Gaussian, there exist X, Y that are independent $N(0,1)$ random variables with

$$B_s = \sqrt{s}\,X \qquad \text{and} \qquad B_t = \sqrt{s}\,X + \sqrt{t-s}\,Y.$$

Then $\{B_s < 0\} = \{X < 0\}$, and $\{B_t > 0\} = \{Y > -\frac{\sqrt{s}}{\sqrt{t-s}}X\}$. To calculate $P(\{Y > -\frac{\sqrt{s}}{\sqrt{t-s}}X, X < 0\})$, use polar coordinates:

$$
\begin{aligned}
P(B_t > 0, B_s < 0) &= \frac{1}{2\pi}\int_{\pi/2}^{\pi-\arcsin\sqrt{\frac{s}{t}}}\int_0^\infty e^{-r^2/2} r\,dr\,d\theta \\
&= \frac{1}{2\pi}\Big(\frac{\pi}{2} - \arcsin\sqrt{\frac{s}{t}}\Big) \\
&= \frac{1}{4} - \frac{1}{2\pi}\arcsin\sqrt{\frac{s}{t}}.
\end{aligned}
$$

Therefore $4P(B_t > 0, B_s < 0) = 1 - \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}$ and $P(g_t \le s) = \frac{2}{\pi}\arcsin\sqrt{\frac{s}{t}}$. □

Theorem 80. *Azéma's martingale M is given by $M_t = \mathrm{sign}(B_t)\sqrt{\frac{\pi}{2}}\sqrt{t-g_t}$, $t \ge 0$.*

Proof. By definition,

$$M_t = E\{B_t|\mathcal{G}_t\} = E\{\mathrm{sign}(B_t)|B_t||\mathcal{G}_t\} = \mathrm{sign}(B_t)E\{|B_t||\mathcal{G}_t\}.$$

However $E\{|B_t||\mathcal{G}_t\} = E\{|B_t||g_t; \mathrm{sign}(B_s), s \le t\}$. But the process $\mathrm{sign}(B_s)$ is independent of $|B_t|$ for $s < g_t$, and $\mathrm{sign}(B_s) = \mathrm{sign}(B_t)$ for $g_t < s \le t$. Hence $\mathrm{sign}(B_s)$ is constant after g_t. It follows that

$$E\{|B_t||\mathcal{G}_t\} = E\{|B_t||g_t\}.$$

Therefore if we can show $E\{|B_t||g_t = s\} = \sqrt{\frac{\pi}{2}}\sqrt{t-s}$, the proof will be complete.

Given $0 < s < t$ we define

$$T = t \wedge \inf\{u > s : B_u = 0\}.$$

Note that T is a bounded stopping time. Then

$$\begin{aligned}E\{|B_t|; g_t \le s\} &= 2E\{B_t; g_t \le s \text{ and } B_s > 0\}\\ &= 2E\{B_t; T = t \text{ and } B_s > 0\}\\ &= 2E\{B_T; B_s > 0\}\\ &= 2E\{B_s; B_s > 0\},\end{aligned}$$

since $B_T = 0$ on $\{T \neq t\}$, and the last step uses that B is a martingale, and $s \le T$. The last term above equals $E\{|B_s|\} = \sqrt{\frac{2}{\pi}}\sqrt{s}$. Therefore

$$\begin{aligned}E\{|B_t||g_t = s\} &= \frac{\frac{d}{ds}E\{|B_t|; g_t \le s\}}{\frac{d}{ds}P(g_t \le s)}\\ &= \sqrt{\frac{\pi}{2}}\sqrt{t-s},\end{aligned}$$

since $\frac{d}{ds}P(g_t \le s) = \frac{1}{\pi\sqrt{s(t-s)}}$, by Theorem 79. □

Let $\mathcal{H}_t^0 = \sigma\{M_s; s \le t\}$, the minimal filtration of Azéma's martingale M, and let $\mathbb{H} = (\mathcal{H}_t)_{t \ge 0}$ be the completed filtration.

Corollary. The two filtrations $\mathbb{H}$ and $\mathbb{G}$ are equal.

Proof. By definition $M_t \in \mathcal{G}_t$, each $t \ge 0$. Hence $\mathcal{H}_t \subset \mathcal{G}_t$. However, by Theorem 80, $\operatorname{sign}(M_t) = \operatorname{sign}(B_t)$, so that $\mathcal{G}_t \subset \mathcal{H}_t$. □

Next we turn our attention to the local times of Azéma's martingale. For a semimartingale $(X_t)_{t\ge 0}$, let $(L_t^a(X))_{t \ge 0} = L^a(X)$ denote its local time at the level a.

Theorem 81. *The local time at zero of Azéma's martingale is not identically zero.*

Proof. By the Meyer-Itô formula (Theorem 70) for $f(x) = |x|$ we have

$$|M_t| = \int_0^t \operatorname{sign}(M_{s-})dM_s + L_t^0(M) + \sum_{0<s\le t}\{|M_s| - |M_{s-}| - \operatorname{sign}(M_{s-})\Delta M_s\}.$$

Set $Y_t = |M_t|$. Since $\operatorname{sign}(M_{s-})\Delta M_s = \Delta Y_s$, we have that $|M_t| - L_t^0(M)$ is a martingale. If $L^0(M)$ were the zero process, then $Y = |M|$ would be a non-negative martingale with $Y_0 = 0$, hence identically zero. Since Y is not identically zero, the theorem is proved. □

Theorem 82. *The local times at all levels except zero of Azéma's martingale are* 0. *That is,* $L_t^a(M) = 0$, $t \ge 0$, *if* $a \neq 0$.

Proof. By Theorem 69 we know that the measure $dL^a_s(\omega)$ on $\mathbb{R}_+$ is carried by the set $\{s : M_{s-}(\omega) = M_s(\omega) = a\}$, a.s., for each a. However if $a \neq 0$, this set is countable. Since $s \mapsto L^a_s$ is a.s. continuous, it cannot charge a countable set. Therefore dL^a_s is the zero measure for $a \neq 0$. Since $L^a_0 = 0$, we have $L^a_t \equiv 0$ for $a \neq 0$. □

Theorems 81 and 82 together show that Azéma's martingale provides a counterexample to a general version of Theorem 75, for example. Therefore a hypothesis such as Hypothesis A is necessary, and also we see that Azéma's martingale does not satisfy Hypothesis A. Since all semimartingales for the Brownian motion filtration $\mathbb{F}$ do satisfy Hypothesis A, as noted in Sect. 7, we conclude the following.

Theorem 83. *Azéma's martingale M is not a semimartingale for the Brownian filtration $\mathbb{F}$.*

While the local time at zero, $L^0(M)$, is non-trivial by Theorem 81, more information is available. It is actually the same as the Brownian local time, as Theorem 85 below shows. First we need a preliminary result.

Theorem 84. *The local time $L^0(B)$ of Brownian motion is $\mathbb{G}$ adapted.*

Theorem 84 will be proved if one can express $L^0(B)$ as the a.s. limit of measurable functionals of the zero set of Brownian motion. Such a result is classical: see, e.g., Kingman [126, page 730].

Theorem 85. *The local times at zero of Brownian motion and Azéma's martingale are the same. That is, $L^0(B) = L^0(M)$.*

Proof. As we saw in the proof of Theorem 81, $|M_t| - L^0_t(M)$ is a martingale. Since the process $L^0(M)$ is non-decreasing, continuous and adapted, it is natural. If A is another natural process such that $|M| - A$ is a martingale, then $L^0(M) = A$ by (for example) Theorem 30 of Chap. III. Thus, it suffices to show that $|M| - L^0(B)$ is a martingale for the filtration $\mathbb{G}$.

By Tanaka's formula, $|B_t| - L^0_t(B) = N_t$, where N is an $\mathbb{F}$ martingale. Then

$$E\{|B_t| - L^0_t(B)|\mathcal{G}_t\} = E\{N_t|\mathcal{G}_t\}$$

and by Theorem 84

$$E\{|B_t||\mathcal{G}_t\} - L^0_t(B) \text{ is a martingale.}$$

Since

$$\begin{aligned} E\{|B_t||\mathcal{G}_t\} = E\{\operatorname{sign}(B_t)B_t|\mathcal{G}_t\} &= \operatorname{sign}(B_t)E\{B_t|\mathcal{G}_t\} \\ &= \operatorname{sign}(B_t)M_t \\ &= |M_t|, \end{aligned}$$

we have $|M_t| - L^0_t(B)$ is a martingale, and the theorem is proved. □

Corollary 1. The process $|M| - L^0(B)$ is a $\mathbb{G}$ martingale.

Proof. This corollary is proved at the end of the proof of Theorem 85. $\square$

Corollary 2. The process $1_{\{B_t>0\}}\sqrt{\frac{\pi}{2}}\sqrt{t-g_t} - \frac{1}{2}L_t^0(B)$ is a $\mathbb{G}$ martingale.

Proof. Since $1_{\{M_{s-}>0\}}M_s^- = 1_{\{M_{s-}\leq 0\}}M_s^+ = 0$, it follows from Theorem 68 that $M_t^+ - \frac{1}{2}L_t^0(M)$ is a martingale. However $M_t^+ = 1_{\{B_t>0\}}\sqrt{\frac{\pi}{2}}\sqrt{t-g_t}$ by Theorem 80, and $L_t^0(M) = L_t^0(B)$ by Theorem 85, and the theorem follows. $\square$

The next theorem shows that Azéma's martingale is quadratic pure jump.

Theorem 86. *For Azéma's martingale M, $[M,M]^c \equiv 0$, and $[M,M]_t = \frac{\pi}{2}g_t$.*

Proof. By Corollary 2 of Theorem 70,

$$[M,M]_t^c = \int_{-\infty}^{\infty} L_t^a da.$$

By Theorem 82, $[M,M]^c \equiv 0$. We conclude that $[M,M]_t = \sum_{0<s\leq t}(\Delta M_s)^2$, and it follows from Theorem 80 that $[M,M]_t = \frac{\pi}{2}g_t$. $\square$

Theorem 86 allows us to compute the compensation of $[M,M]$; that is, the unique predictable, increasing process A such that $[M,M]-A$ is a martingale. This process A is denoted $\langle M,M\rangle$ in the literature, and it was defined in Sect. 4 of Chap. III.

Theorem 87. *Let M be Azéma's martingale. Then $[M,M]_t - \frac{\pi}{4}t$ is a martingale. That is, $\langle M,M\rangle_t = \frac{\pi}{4}t$.*

Proof. Recall that $M_t^2 - [M,M]_t$ is a martingale (cf., Theorem 27 of Chap. II). By Theorems 80 and 86,

$$\begin{aligned} M_t^2 - [M,M]_t &= \frac{\pi}{2}\{(t-g_t) - g_t\} \\ &= \frac{\pi}{2}t - \pi g_t. \end{aligned}$$

Since the above is a martingale, it remains only to observe that $\frac{\pi}{2}g_t - \frac{\pi}{4}t$ is a martingale. $\square$

Note that if $X_t = \operatorname{sign}(B_t)\sqrt{2}\sqrt{t-g_t}$, then X is a quadratic pure jump martingale with $\langle X,X\rangle_t = t$. Recall that if N is a Poisson process with intensity one, then $N_t - t$ is a martingale. Since $[N,N]_t = N_t$, we have $[N,N]_t - t$ is a martingale; thus $\langle N,N\rangle_t = t$. Another example of a martingale Y with $\langle Y,Y\rangle_t = t$ is Brownian motion. Since $[B,B]_t = t$ is continuous, $[B,B]_t = \langle B,B\rangle_t = t$.

9 Sigma Martingales

We saw with Emery's example (which precedes Theorem 34) a stochastic integral with respect to an $\mathcal{H}^2$ martingale that was not even a local martingale. Yet "morally" we feel that it should be one. It is therefore interesting to consider the space of stochastic integrals with respect to local martingales. We know these are of course semimartingales, but we want to characterize the class of processes that arise in a more specific yet reasonable way.

By analogy, recall that in measure theory a σ-finite measure is a measure μ on a space $(S, \mathcal{B})$ such that there exists a sequence of measurable sets $\Lambda_i \in \mathcal{B}$ such that $\bigcup_i \Lambda_i = S$ and $\mu(\Lambda_i) < \infty$, each i. Such a measure can be thought of as a special case of a countable sum of probability measures. It is essentially a similar phenomenon that brings us out of the class of local martingales.

Definition. An $\mathbb{R}^d$-valued semimartingale X is called a **sigma martingale** if there exists an $\mathbb{R}^d$-valued martingale M and a predictable $\mathbb{R}_+$-valued process $H \in L(M)$ such that $X = H \cdot M$.

In the next theorem, H will always denote a strictly positive predictable process.

Theorem 88. *Let X be a semimartingale. The following are equivalent:*

(i) X is a sigma martingale;
(ii) $X = H \cdot M$ where M is a local martingale;
(iii) $X = H \cdot M$ where M is a martingale;
(iv) $X = H \cdot M$ where M is a martingale in $\mathcal{H}^1$.

Proof. It is clear that (iv) implies (iii), that (iii) implies (ii), and that (iii) obviously implies (i), so we need to prove only that (i) implies (iv). Without loss assume that $M_0 = 0$. Since X is a sigma martingale it has a representation of the form $X = H \cdot M$ where M is a martingale. We know that we can localize M in $\mathcal{H}^1$ by Theorem 51, so let T_n be a sequence of stopping times tending to ∞ a.s. such that $M^{T_n} \in \mathcal{H}^1$ for each n. Set $T_0 = 0$ and let $N^n = 1_{(T_{n-1}, T_n]} \cdot M^{T_n}$. Let α_n be a strictly positive sequence of real numbers such that $\sum_n \alpha_n \|N^n\|_{\mathcal{H}^1} < \infty$. Define $N = \sum_n \alpha_n N^n$ and one can check that N is an $\mathcal{H}^1$ martingale. Set $J = 1_{\{H=0\}} + H \sum_n \alpha_n^{-1} 1_{(T_{n-1}, T_n]}$. Then it is simple to check that $X = J \cdot N$ and that J is strictly positive. □

Corollary 1. A local sigma martingale is a sigma martingale.

Proof. Let X be a local sigma martingale, so that there exists a sequence of stopping times tending to ∞ a.s. such that X^{T_n} is a sigma martingale for each n. Since X^{T_n} is a sigma martingale, there exists a martingale M^n in $\mathcal{H}^1$ such that $X^{T_n} = H^n \cdot M^n$, for each n, where H^n is positive predictable. Choose ϕ^n positive, predictable so that $\|\phi^n H^n \cdot M^{T_n}\|_{\mathcal{H}^1} < 2^{-n}$. Set $T_0 = 0$ and $\phi^0 = \phi^1 1_{\{0\}}$. Then $\phi = \phi^0 + \sum_{n \geq 1} \phi^n 1_{(T_{n-1}, T_n]}$ is strictly positive and predictable. Moreover $\phi \cdot X$ is an $\mathcal{H}^1$ martingale, whence $X = \frac{1}{\phi} \cdot \phi \cdot X$, and the corollary is established. □

Corollary 2. A local martingale is a sigma martingale.

Proof. This is simply a consequence of the fact that a local martingale is locally a martingale, and trivially a martingale is a sigma martingale. □

The next theorem gives a satisfying stability result, showing that sigma martingales are stable with respect to stochastic integration.

Theorem 89. *If X is a local martingale and $H \in L(X)$, then the stochastic integral $H \cdot X$ is a sigma martingale. Moreover if X is a sigma martingale (and a fortiori a semimartingale) and $H \in L(X)$, then $H \cdot X$ is a sigma martingale.*

Proof. Clearly it suffices to prove the second statement, since a local martingale is already a sigma martingale. But the second statement is simple. Since X is a sigma martingale we know an equivalent condition is that there exists a strictly positive predictable process ϕ and a local martingale M such that $X = \phi \cdot M$. Since H is predictable, the processes $H^1 = H1_{\{H>0\}} + 1$ and $H^2 = -H1_{\{H<0\}} + 1$ are both strictly positive and predictable. Moreover $H = H^1 - H^2$. The processes H^1 and H^2 are both in $L(X)$ because H is. By associativity and linearity of the stochastic integral,

$$H \cdot X = H(\phi \cdot M) = (H\phi) \cdot M = H^1\phi \cdot M - H^2\phi \cdot M,$$

and since $H^i\phi$ are both strictly positive for $i = 1, 2$ we have by Theorem 88 that $H^i\phi \cdot M$ is a sigma martingale for each $i = 1, 2$. Sigma martingales form a vector space, and we are done. □

We also would like to have sufficient conditions for a sigma martingale to be a local martingale.

Theorem 90. *A sigma martingale which is also a special semimartingale is a local martingale.*

Proof. Any sigma martingale is a semimartingale by definition; here we also assume it is special. Thus it has a canonical decomposition $X = M + A$ where M is a local martingale and A is a process of finite variation on compacts which is also predictable. We want to show $A = 0$. We assume $A_0 = 0$. Choose H predictable, $H > 0$ everywhere, such that $H \cdot X$ is a martingale, which we can easily do using Theorem 88. Then $(H \wedge 1) \cdot X = (\frac{H \wedge 1}{H}) \cdot (H \cdot X)$ is a local martingale, and hence without loss of generality we can assume H is bounded. Then $H \cdot A$ is a finite variation predictable process, and we have $H \cdot A = H \cdot X - H \cdot M$ which is a local martingale. Thus we conclude that $H \cdot A = 0$. If we replace H with JH where J is predictable, $|J| \leq 1$, and is such that $\int_0^t J_s dA_s = \int_0^t |dA_s|$, then we conclude that $\int_0^t H_s |dA_s| = 0$, and since $H > 0$ everywhere, we conclude that $A = 0$ establishing the result. □

We now have that any criterion that ensures that the semimartingale X is special, will also imply that the sigma martingale X is a local martingale.

Corollary 1. If a sigma martingale X has continuous paths, then it is a local martingale.

Corollary 2. If X is a sigma martingale and if either $X_t^* = \sup_{s \le t} |X_s|$ or $Y_t = \sup_{s \le t} |\Delta X_s|$ is locally integrable, then X is a local martingale.

Note that as a corollary to Corollary 2 we have the following.

Corollary 3. If X is a sigma martingale with bounded jumps, then X is a local martingale.

Bibliographic Notes

The extension by continuity of stochastic integration from processes in $\mathbb{L}$ to predictable processes is presented here for essentially the first time. However the procedure was indicated earlier in Protter [201, 202]. Other approaches which are closely related are given by Jacod [103] and Chou-Meyer-Stricker [31] (see an exposition in Dellacherie-Meyer [46, page 381]).

The "$\mathcal{H}$" in the space $\mathcal{H}^2$ of semimartingales comes from Hardy spaces, and this bears explaining. When the semimartingale $X \in \mathcal{H}^2$ is actually a martingale, the space of $\mathcal{H}^p$ martingales has a theory analogous to that of Hardy spaces. While this is not the case for semimartingales, the $\mathcal{H}^p$ norms for semimartingales are, in a certain sense, a natural generalization of the $\mathcal{H}^p$ norms for martingales, and so the name has been preserved. The $\mathcal{H}^p$ norm was first introduced by Emery [64], though it was implicit in Protter [197].

Most of the treatment of stochastic integration in Sect. 2 of this chapter is new, though essentially all of the results have been proved elsewhere with different methods. The author benefited greatly throughout by discussions with S. Janson. Theorem 1 (that $\mathcal{H}^2$ is a Banach space) is due to Emery [66]. The equivalence of the two pseudonorms for $\mathcal{H}^2$ (the corollary of Theorem 24) is originally due to Yor [241], while the fact that $\mathcal{H}^2(Q) \subset \mathcal{H}^2(P)$ if $\frac{dQ}{dP}$ is bounded is originally due to Lenglart [144]. Theorem 25 is originally due to Jacod [103, page 228]. The concept of a predictable σ-algebra and its importance to stochastic integration is due to Meyer [166]. The local behavior of the stochastic integral was first investigated by McShane for his integral [154], and then independently by Meyer [171], who established Theorem 26 and its corollary for the semimartingale integral. The first Dominated Convergence Theorem for the semimartingale integral appearing in print seems to be in Jacod [103].

The theory of martingale representation dates back to Itô's work on multiple stochastic integrals [96], and the theory for $\mathbf{M}^2$ presented here is largely due to Kunita-Watanabe [134]. A more powerful theory (for $\mathbf{M}^1$) is presented in Dellacherie-Meyer [46]. Absolutely continuous spaces are defined here for the first time. The author benefited from discussions with Yan Zeng on this subject. That the dual of $\mathcal{H}^1$ is BMO for martingales is due to the combined efforts of many researchers, culminating in Meyer's paper [170]; we follow the treatment in [87]. The Jacod-Yor Theorem is from [111], and Sect. 9 largely follows Emery [69].

The theory of stochastic integration depending on a parameter is due to the fundamental work of Stricker-Yor [219], who used a key idea of Doléans-Dade [48]. The Fubini Theorems for stochastic integration have their origins in the book of Doob [55] and Kallianpur-Striebel [120] for the Itô integral, Kailath-Segall-Zakai [119] for martingale integrals, and Jacod [103] for semimartingales. The counterexample to a general Fubini Theorem presented here is due to Janson.

The theory of semimartingale local time is of course abstracted from Brownian local time, which is due to Lévy [149]. It was related to stochastic integration by Tanaka (see McKean [153]) for Brownian motion and the Itô integral. The theory presented here for semimartingale local time is due largely to Meyer [171]. See also Millar [181]. The measure theory needed for the rigorous development of semimartingale local time was developed by Stricker-Yor [219], and the Meyer-Itô formula (Theorem 70) was formally presented in Yor [242], as well as in Jacod [103]. Theorem 71 is due to Yor [244], and has been extended by Ouknine [188]. A more general version (but more complicated) is in Çinlar-Jacod-Protter-Sharpe [34]. The proof of Kolmogorov's Lemma given here is from Meyer [177]. The results proved under Hypothesis A are all due to Yor [243] except the Bouleau-Yor formula [22]. See Bouleau [21] for more on this formula.

Azéma's martingale is of course due to Azéma [4], though our presentation of it is new and it is due largely to Janson. For many more interesting results concerning Azéma's martingale, see Azéma-Yor [5], Emery [69], and Meyer [179].

Sigma martingales date back to the work of Emery [67], and to Chou [29]. The importance of sigma martingales was clarified through the work in mathematical finance of Delbaen and Schachermayer [38]. See [110] for a more comprehensive treatment.

Exercises for Chapter IV

Exercise 1. Let X be a semimartingale. Show that X is special if and only if the increasing process $C_t = [X,X]_t$ is locally integrable; that is, there exists a sequence of stopping times T_n increasing to infinity a.s. such that $E\{[X,X]_{T_n}\} < \infty$ for each time T_n.

***Exercise 2.** Let X be a special semimartingale with canonical decomposition $X_t = X_0 + M_t + A_t$. Show that the following two inequalities hold:

$$E\{[A,A]_\infty\} \leq E\{[X,X]_\infty\} \text{ and } E\{[M,M]_\infty\} \leq 4E\{[X,X]_\infty\}.$$

Exercise 3. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let X_n be a sequence of random variables such that $P(|X_n| < \infty) = 1$, all n. Show that there exists another probability Q which is equivalent to P in the sense that the null sets are the same for both measures, with $\frac{dQ}{dP}$ bounded, and every random variable X_n is in $L^1(dQ)$.

Exercise 4. Let X be a semimartingale. Show that there is a probability measure Q, equivalent to P, such that under Q, X has a decomposition $X = M + A$, where M is a martingale with $E_Q([M,M]_t) < \infty$ for each t, $0 \leq t < \infty$, and A is a predictable process of finite variation on compacts with $E_Q(\int_0^t |dA_s|) < \infty$ for each t, $0 \leq t < \infty$. (*Hint:* Use Exercises 1, 2, and 3.)

***Exercise 5.** Let X^n be a sequence of semimartingales. Show that there exists one probability measure Q, equivalent to P, such that under Q, each X^n has a decomposition $X^n = M^n + A^n$, where M^n is a martingale with $E_Q([M^n, M^n]_t) < \infty$ for each t, $0 \leq t < \infty$ and each n, and A^n is a predictable process of finite variation on compacts with $E_Q(\int_0^t |dA_s^n|) < \infty$ for each t, $0 \leq t < \infty$ and each n.

Exercise 6. Let $\mathcal{S}$ denote the space of all square integrable martingales with continuous paths a.s., on a given filtered complete probability space satisfying the usual conditions. Show that $\mathcal{S}$ is a stable subspace. (This exercise was referred to in Definition 3.)

Exercise 7. Let Z be a Lévy process with $E\{|Z_t|\} < \infty$ and $E\{Z_t\} = 0$ for all $t \geq 0$. Show that Z is a martingale, and that Z^c is either 0 or Brownian motion, where Z^c is defined in Exercise 6. Conclude that any martingale Lévy process Z that has no Brownian component is purely discontinuous.

Exercise 8. Let N be a standard Poisson process with arrival intensity λ. Show that the martingale $N_t - \lambda t$ is purely discontinuous. (Note that this gives an example of a purely discontinuous martingale whose sample paths are not purely discontinuous in the sense that they do not change only by jumps.)

Exercise 9. Let M be a square integrable martingale. Show that $[M,M]_t^c = [M^c, M^c]_t = [M^c, M]_t$, all $t \geq 0$, almost surely.

Exercise 10 (example of orthogonal projection). Let T be a stopping time. Show that the collection $\mathcal{S}$ of all square integrable martingales stopped at T is a stable subspace. Also show that if M is an arbitrary square integrable martingale, then its orthogonal decomposition with respect to this subspace is $M = M^T + (M - M^T)$. Give a description of the orthocomplement of $\mathcal{S}$.

Exercise 11 (example of orthogonal projection). Let T be a totally inaccessible stopping time and let M be a square integrable martingale. Let $U_t = \Delta M_T 1_{\{t \geq T\}}$ and let L be the martingale of U minus its compensator. We denote L by $\varPi M$. Show that $\varPi M$ and $M - \varPi M$ are orthogonal. Let $\mathcal{S}$ denote the space of all square integrable martingales which are continuous at the instant T. Show that $\mathcal{S}$ is a stable subspace, and that $\varPi M$ represents the orthogonal projection of M onto the orthocomplement of $\mathcal{S}$.

***Exercise 12.** Let T be an arbitrary stopping time, $T > 0$, and let X be a random variable in L^1. Show that $M = X1_{\{t \geq T\}}$ is a uniformly integrable martingale if and only if $E\{X|\mathcal{F}_{T-}\} = 0$.

***Exercise 13.** Let T be a stopping time, $T > 0$, and let M be a square integrable martingale. Show that the decomposition

$$M_t = [(M_T - E\{M_T|\mathcal{F}_{T-}\})1_{\{t \geq T\}}] + [M_t - (M_T - E\{M_T|\mathcal{F}_{T-}\})1_{\{t \geq T\}})]$$

is an orthogonal decomposition. Show that one of the stable subspaces in question is the space of all square integrable martingales M such that M_T is $\mathcal{F}_{T-}$ measurable. (*Hint:* Use Exercise 12.)

***Exercise 14.** Let Z be a Lévy process on a complete probability space $(\Omega, \mathcal{F}, P)$ and let $\mathbb{F}$ be its minimal filtration completed (and satisfying the usual hypotheses). Show that $(\Omega, \mathcal{F}, \mathbb{F}, P)$ is an absolutely continuous space.

Exercise 15. Consider the **structure equation** $d[X, X]_t = dt + \beta X_{t-} dX_t$. Show that this equation has a **scaling property**: if X is a martingale solution, then so too is $\frac{1}{\lambda} X_{\lambda^2 t}$ for every $\lambda \neq 0$. Note that by the uniqueness in distribution of the solution, the processes $(X_t)_{t \geq 0}$ and $(\frac{1}{\lambda} X_{\lambda^2 t})_{t \geq 0}$ have the same distribution.

***Exercise 16.** Let X be a martingale solution of the structure equation of Exercise 15. Show that $X^c = 0$, where X^c is defined in Exercise 6.

Exercise 17. Let M be a local martingale in $\mathcal{H}^p$ for some $p \geq 1$. Show that M is a uniformly integrable martingale. Show further that if $p = 1$ then $E\{M_\infty\} \leq c\|M\|_{\mathcal{H}^1}$ where the constant c is universal. That is, c does not depend on M.

Exercise 18. Suppose that M is a bounded martingale. Show that

$$\|M\|_{BMO} \leq \sqrt{5}\|M_\infty\|_{L^\infty}.$$

This can improved to replace $\sqrt{5}$ with 2.

Exercise 19. Show that there are martingales with finite BMO norm which are not bounded. (*Hint:* Let M be a bounded martingale and H be a predictable process such that $|H| \leq 1$, and let $N = H \cdot M$. Then it is easy to see that N is in BMO, and to construct such an N that is not bounded.)

Exercise 20. Prove this version of the Fundamental Theorem of Local Martingales. Let M be a uniformly integrable martingale. Show that there exist N and U such that N is in BMO, U is of integrable variation, and $M = U + V$. If M is a local martingale this holds locally.

Exercise 21. Show that an arbitrary local martingale is a semimartingale by using the Burkholder-Davis-Gundy inequalities instead of the Fundamental Theorem of Local Martingales. (See [61].)

Exercise 22. Let M be a local martingale. If T is a predictable stopping time, then $1_{[0,T)}(s) \in \mathbf{b}\mathcal{P}$ so that $M_t^{T-} = \int_0^t 1_{[0,T)}(s)dM_s$ is a local martingale. Give an example of a local martingale M and a stopping time T such that M^{T-} is not a local martingale.

Exercise 23. Let X be a continuous semimartingale. Show that $[|X|, X] = [X, X]$. (*Hint:* Use local times.)

***Exercise 24.** Let B be standard Brownian motion and let $\alpha > 0$. Show that $\int_0^1 |B_t|^{-\alpha} dt < \infty$ a.s. if and only if $\alpha < 1$. (*Hint:* Use local times.)

Exercise 25. Let X be a semimartingale. Show that $L_t^0(X) = \frac{1}{2}L_t^0(X^+)$ where $X^+ = \max(X, 0)$ and $L_t^0(X)$ denotes the local time of X at time t and level 0.

Exercise 26. Let A be adapted, continuous, and of finite variation on compacts, and let B be standard Brownian motion. Let λ denote Lebesgue measure on the line. Show that $\lambda(s : B_s = A_s) = 0$ a.s.

Exercise 27. Let X be a continuous semimartingale. Show that $L_t^0 \in \mathcal{F}_t^{|X|}$, each $t \geq 0$, where $\mathcal{F}_t^{|X|}$ denotes the smallest right continuous completed filtration generated by the process $(|X_t|)_{t\geq 0}$.

Exercise 28. Let X be a semimartingale. Show that for each fixed (ω, t) the section $x \mapsto L_t^x(X)$ has compact support.

Exercise 29 (Skorohod's Lemma). This result will be useful for the next two exercises. Let x be a real valued continuous function on $[0, \infty)$ such that $x(0) \geq 0$. Show there exists a pair of functions (y, a) on $[0, \infty)$ such that $x = y + a$, y is positive, and a is increasing, continuous, $a(0) = 0$, and the measure da_s induced by a is carried by the set $\{s : y(s) = 0\}$. Show further that the function a is given by $a(t) = \sup_{s\leq t}(-x(s) \vee 0)$.

Exercise 30. Let M be a continuous local martingale with $M_0 = 0$. Let $|M_t| = N_t + L_t^0(M)$. Show that N is a continuous local martingale with $N_0 = 0$, and that $L_t^0(M) = \sup_{s\leq t}(-N_s)$.

***Exercise 31.** Let X be a continuous semimartingale, and suppose Y satisfies the equation $dY_t = -\operatorname{sign}(Y_t)dX_t$. Show that $|Y_t| = X_t^* - X_t$, where $X_t^* = \sup_{u\leq t} X_u$. (*Hint:* Use Tanaka's formula and Exercise 29.)

Exercise 32. Suppose that X and Y are continuous semimartingales such that $L_t^0(Y - X) = 0$. Show that

$$L_t^0(X \vee Y) = \int_0^t 1_{\{Y_s\leq 0\}} dL_s^0(X) + \int_0^t 1_{\{X_s<0\}} dL_s^0(Y).$$

Exercise 33. Let X and Y be continuous semimartingales. Show that

$$L_t^0(X \wedge Y)$$
$$= \int_0^t 1_{\{Y_s>0\}} dL_s^0(X) + \int_0^t 1_{\{X_s\geq 0\}} dL_s^0(Y) + \int_0^t 1_{\{X_s=Y_s=0\}} dL_s^0(Y^+ - X^+).$$

Exercise 34 (Ouknine's Formula). Let X and Y be continuous semimartingales. Show that

$$L_t^0(X \vee Y) + L_t^0(X \wedge Y) = L_t^0(X) + L_t^0(Y).$$

*__Exercise 35 (Emery-Perkins Theorem).__ Let B be standard Brownian motion and set $L_t = L_t^0(B)$. Let $X = B + cL$, $c \in \mathbb{R}$ fixed. Assume that there exists a set $D \subset \mathbb{R}_+ \times \Omega$ such that (i) D is predictable for $\mathcal{F}^X$; (ii) $P(\omega : (t,\omega) \in D) = 1$ each $t > 0$; and (iii) $P(\omega : (T_t(\omega), \omega) \in D) = 0$ each $t > 0$, where $T_t = \inf(s > 0 : L_s > t)$.

(a) Show that $B_t = \int_0^t 1_D(s) dX_s$. Also show that $B_t \in \mathcal{F}^X$, each $t > 0$, and conclude that $\mathcal{F}^B = \mathcal{F}^X = \mathcal{F}^{B+cL}$, all $c \in \mathbb{R}$.
(b) Let

$$D = \left\{ (t,\omega) : \lim_{n\to\infty} \sum_{k=1}^{n} 1_{\{X_{t-2^{-k}}(\omega) - X_{t-2^{-(k-1)}}(\omega) > 0\}} = \frac{1}{2} \right\}.$$

Use the fact that $s \mapsto (B_{s\wedge T_t}, L_{s\wedge T_t})$ and $s \mapsto (-B_{T_{t-s}}, L_{T_t} - L_{T_{t-s}})$ have the same joint distribution to show that D satisfies (i), (ii), and (iii).

*__Exercise 36.__ With the notation of Exercise 35, let $Y = |B| + cL$. Show that $\mathcal{F}^Y \subset \mathcal{F}^{|B|}$, for all c.

*__Exercise 37.__ With the notation of Exercise 36, **Pitman's Theorem** states that Y is a Bessel process of order 3, which in turn implies that for $\mathcal{F}^Y$, with $c = 1$,

$$Y_t = \beta_t + \int_0^t \frac{1}{Y_s} ds$$

for an $\mathcal{F}^Y$-Brownian motion β. Use this to show that $\mathcal{F}^Y \neq \mathcal{F}^{|B|}$ for $c = 1$. (*Note:* One can show that $\mathcal{F}^Y = \mathcal{F}^{|B|}$ for all $c \neq 1$.)

*__Exercise 38 (Barlow's Theorem).__ Let X solve the equation

$$X_t = X_0 + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds$$

where B is a standard Brownian motion. The above equation is said to have a **unique weak solution** if whenever X and Y are two solutions, and if μ_X is defined by $\mu_X(\Lambda) = P\{\omega : t \mapsto X_t(\omega) \in \Lambda\}$, then $\mu_X = \mu_Y$. The above equation is said to have a **unique strong solution** if whenever both X and Y are two solutions on the probability space on which the Brownian motion B is defined, then $P\{\omega : t \mapsto X_t(\omega) = t \mapsto Y_t(\omega)\} = 1$. Assume that the equation has a unique weak solution with the property that if X, Y are any two solutions defined on the same probability space, then for all $t \geq 0$, $L_t^0(X - Y) = 0$. Show that the equation has a unique strong solution.

Exercise 39. Consider the situation of Exercise 38, and suppose that σ and b are bounded Borel. Further, suppose that $(\sigma(x)-\sigma(y))^2 \leq \rho(|x-y|)$ for all x, y, where $\rho : [0,\infty) \to [0,\infty)$ is increasing with $\int_{0+}^{\varepsilon} \frac{1}{\rho(u)} du = +\infty$ for every $\varepsilon > 0$. Show that if X, Y are two solutions, then $L_t^0(X-Y) = 0$.

***Exercise 40.** Let B be a standard Brownian motion and consider the equation

$$X_t = \int_0^t \operatorname{sign}(X_s) dB_s.$$

Show that the equation has a unique weak solution, but that if X is a solution then so too is $-X$. (See Exercise 38 for a definition of weak and strong solutions.) Hence, the equation does not have a unique strong solution. Finally, show that if X is a solution, then X is *not* adapted to the filtration $\mathcal{F}^B$. (*Hint:* Show first that $B_t = \int_0^t 1_{\{|X_s| \neq 0\}} d|X_s|$.)

Exercise 41. Let X be a sigma martingale and suppose X is bounded below. Show that X is in fact a local martingale.

***Exercise 42.** Show that any Lévy process which is a sigma martingale is actually a martingale (and not just a local martingale).

Exercise 43. Let $(U_n)_{n\geq 1}$ be i.i.d. random variables with $P(U_1 = 1) = P(U_1 = -1) = 1/2$, and let $X = \sum_{n\geq 1} 2^{-n} U_n 1_{\{t \geq q_n\}}$ where $(q_n)_{n\geq 1}$ is an enumeration of the rationals in $(0,1)$. Let $H_t = \sum_{n\geq 1} \frac{2^n}{n} 1_{\{t\geq q_n\}}$ and show that $X \in \mathcal{H}^2$ and $H \in L(X)$, but also that X is of finite variation and $Y = H \cdot X$ has infinite variation. *Thus the space of finite variation processes is not closed under stochastic integration!*

***Exercise 44.** Let X^n be a sequence of semimartingales on a filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfying the usual hypotheses. Show there exists one probability Q, which is equivalent to P, with a bounded density such that under Q each X^n is a semimartingale in $\mathcal{H}^2$.

Exercise 45. Let B be a standard Brownian motion and let L^x be its local time at the level x. Let

$$\Lambda = \{\omega : dL_s^x(\omega) \text{ is singular as a measure with respect to } ds\}$$

where of course ds denotes Lebesgue measure on $\mathbb{R}_+$. Show that $P(\Lambda) = 1$. Conclude that *almost surely the paths $s \mapsto L_s^x$ are not absolutely continuous.* (*Hint:* Use Theorem 69.)

Exercise 46. Let B be a standard Brownian motion and let N be a Poisson process with parameter $\lambda = 1$, with B and N independent as processes. Let L be the local time process for B at level 0, and let X be the vector process $X_t = (B_t, N_{L_t})$. Let $\mathbb{F}$ be the minimal filtration of X, completed in the usual way.

(a) Show that X is a strong Markov process.

(b) Let T be the first jump time of X. Show that the compensator of the process $C_t = 1_{\{t \geq T\}}$ is the process $A_t = L_{t \wedge T}$.

Note that this gives an example of a compensator of the first jump time of a strong Markov process which has paths that are almost surely not absolutely continuous.

****Exercise 47.** Let Z be a Lévy process. Show that if Z is a sigma martingale, then Z is a martingale. (*Note:* This exercise is closely related to Exercise 29 of Chap. I.)

V

Stochastic Differential Equations

1 Introduction

A *diffusion* can be thought of as a strong Markov process (in $\mathbb{R}^n$) with continuous paths. Before the development of Itô's theory of stochastic integration for Brownian motion, the primary method of studying diffusions was to study their transition semigroups. This was equivalent to studying the infinitesimal generators of their semigroups, which are partial differential operators. Thus Feller's investigations of diffusions (for example) were actually investigations of partial differential equations, inspired by diffusions.

The primary tool to study diffusions was Kolmogorov's differential equations and Feller's extensions of them. Such approaches did not permit an analysis of the paths of diffusions and their properties. Inspired by Lévy's investigations of sample paths, Itô studied diffusions that could be represented as solutions of stochastic differential equations[1] of the form

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds \tag{$*$}$$

where B is a Brownian motion in $\mathbb{R}^n$, σ is an $n \times n$ matrix, and b is an n-vector of appropriately smooth functions to ensure the existence and uniqueness of solutions. This gives immediate intuitive meaning. If $(\mathcal{F}_t)_{t\geq 0}$ is the underlying filtration for the Brownian motion B, then for small $\varepsilon > 0$

$$\begin{aligned}
&E\{X^i_{t+\varepsilon} - X^i_t|\mathcal{F}_t\} = b^i(X_t)\varepsilon + o(\varepsilon)\\
&E\{(X^i_{t+\varepsilon} - X^i_t - \varepsilon b^i(X_t))(X^j_{t+\varepsilon} - X^j_t - \varepsilon b^j(X_t))|\mathcal{F}_t\} = (\sigma\sigma')^{ij}(X_t)\varepsilon + o(\varepsilon),
\end{aligned}$$

where σ' denotes the transpose of the matrix σ.

Itô's differential "dB" was found to have other interpretations as well. In particular, "dB" can be thought of as "white noise" in statistical communication theory. Thus if ξ_t is white noise at time t, $B_t = \int_0^t \xi_s ds$, an equation

[1] More properly (but less often) called "stochastic integral equations."

which can be given a rigorous meaning using the theory of generalized functions (cf., e.g., Arnold [2]). Here the Markov nature of the solutions is not as important, and coefficients that are functionals of the paths of the solutions can be considered.

Finally it is now possible to consider semimartingale driving terms (or "semimartingale noise"), and to study stochastic differential equations in full generality. Since "dB" and "dt" are semimartingale differentials, they are always included in our results as special cases. While our treatment is very general, it is not always the most general available. We have at times preferred to keep proofs simple and non-technical rather than to achieve maximum generality.

The study of stochastic differential equations (SDEs) driven by general semimartingales (rather than just by dB, dt, dN, and combinations thereof, where N is a Poisson process) allows one to see which properties of the solutions are due to certain special properties of Brownian motion, and which are true in general. For example, in Sect. 6 we see that the Markov nature of the solutions is due to the independence of the increments of the differentials. In Sects. 8 and 10 we see precisely how the homeomorphic and diffeomorphic nature of the flow of the solution is a consequence of path continuity. In Sect. 5 we study Fisk-Stratonovich equations which reveal that the "correction" term is due to the continuous part of the quadratic variation of the differentials. In Sect. 11 we illustrate when standard moment estimates on solutions of SDEs driven by Brownian motion and dt can be extended to solutions of SDEs driven by Lévy processes.

2 The $\underline{\underline{H}}^p$ Norms for Semimartingales

We defined an $\mathcal{H}^2$ norm for semimartingales in Chap. IV as follows. If X is a special semimartingale with $X_0 = 0$ and canonical decomposition $X = \overline{N} + \overline{A}$, then

$$\|X\|_{\mathcal{H}^2} = \|[\overline{N}, \overline{N}]_\infty^{1/2}\|_{L^2} + \|\int_0^\infty |d\overline{A}_s|\|_{L^2}.$$

We now use an *equivalent* norm. To avoid confusion, we write $\underline{\underline{H}}^2$ instead of $\mathcal{H}^2$. Moreover we will define $\underline{\underline{H}}^p$, $1 \le p \le \infty$. We begin, however, with a different norm on the space $\mathbb{D}$ (i.e., the space of adapted càdlàg processes). For a process $H \in \mathbb{D}$ we define

$$H^* = \sup_t |H_t|,$$

$$\|H\|_{\underline{\underline{S}}^p} = \|H^*\|_{L^p}.$$

Occasionally if H is in $\mathbb{L}$ (adapted and càglàd) we write $\|H\|_{\underline{\underline{S}}^p}$ as well, where the meaning is clear.

If A is a semimartingale with paths of finite variation, a natural definition of a norm would be $\|A\|_p = \| \int_0^\infty |dA_s| \|_{L^p}$, where $|dA_s(\omega)|$ denotes the total variation measure on $\mathbb{R}_+$ induced by $s \mapsto A_s(\omega)$. Since semimartingales do not in general have such nice paths, however, such a norm is not appropriate. *Throughout this chapter, we will let Z denote a semimartingale with $Z_0 = 0$,* a.s. Let Z be an arbitrary semimartingale (with $Z_0 = 0$). By the Bichteler-Dellacherie Theorem (Theorem 43 of Chap. III) we know there exists at least one decomposition $Z = N + A$, with N a local martingale and A an adapted, càdlàg process, with paths of finite variation (also $N_0 = A_0 = 0$ a.s.). For $1 \leq p \leq \infty$ we set

$$j_p(N, A) = \| [N, N]_\infty^{1/2} + \int_0^\infty |dA_s| \|_{L^p}.$$

Definition. Let Z be a semimartingale. For $1 \leq p \leq \infty$ define

$$\|Z\|_{\underline{\underline{H}}^p} = \inf_{Z=N+A} j_p(N, A)$$

where the infimum is taken over all possible decompositions $Z = N + A$ where N is a local martingale, $A \in \mathbb{D}$ with paths of finite variation on compacts, and $A_0 = N_0 = 0$.

The corollary of Theorem 1 below shows that this norm generalizes the $\underline{\underline{H}}^p$ norm for local martingales, which has given rise to a martingale theory analogous to the theory of Hardy spaces in complex analysis. We do not pursue this topic (cf., e.g., Dellacherie-Meyer [46]).

Theorem 1. *Let Z be a semimartingale ($Z_0 = 0$). Then $\| [Z, Z]_\infty^{1/2} \|_{L^p} \leq \|Z\|_{\underline{\underline{H}}^p}$, $(1 \leq p \leq \infty)$.*

Proof. Let $Z = M + A$, $M_0 = A_0 = 0$, be a decomposition of Z. Then

$$\begin{aligned}
[Z, Z]_\infty^{1/2} &\leq [M, M]_\infty^{1/2} + [A, A]_\infty^{1/2} \\
&= [M, M]_\infty^{1/2} + \Big(\sum_s (\Delta A_s)^2\Big)^{1/2} \\
&\leq [M, M]_\infty^{1/2} + \sum_s |\Delta A_s| \\
&\leq [M, M]_\infty^{1/2} + \int_0^\infty |dA_s|,
\end{aligned}$$

where the equality above holds because A is a quadratic pure jump semimartingale. Taking L^p norms yields $\| [Z, Z]_\infty^{1/2} \|_{L^p} \leq j_p(M, A)$ and the result follows. □

Corollary. If Z is a local martingale ($Z_0 = 0$), then $\|Z\|_{\underline{\underline{H}}^p} = \| [Z, Z]_\infty^{1/2} \|_{L^p}$.

Proof. Since Z is a local martingale, we have that $Z = Z+0$ is a decomposition of Z. Therefore

$$\|Z\|_{\underline{\underline{H}}^p} \leq j_p(Z,0) = \|[Z,Z]_\infty^{1/2}\|_{L^p}.$$

By Theorem 1 we have $\|[Z,Z]_\infty^{1/2}\|_{L^p} \leq \|Z\|_{\underline{\underline{H}}^p}$, hence we have equality. □

Theorem 2 is analogous to Theorem 5 of Chap. IV. For most of the proofs which follow we need only the case $p = 2$. Since this case does not need Burkholder's inequalities, we distinguish it from the other cases in the proof.

Theorem 2. *For $1 \leq p < \infty$ there exists a constant c_p such that for any semimartingale Z, $Z_0 = 0$, $\|Z\|_{\underline{\underline{S}}^p} \leq c_p\|Z\|_{\underline{\underline{H}}^p}$.*

Proof. A semimartingale Z is in $\mathbb{D}$, so $\|Z\|_{\underline{\underline{S}}^p}$ makes sense. Let $Z = M + A$ be a decomposition with $M_0 = A_0 = 0$. Then

$$\|Z\|_{\underline{\underline{S}}^p}^p = E\{(Z_\infty^*)^p\} \leq E\{(M_\infty^* + \int_0^\infty |dA_s|)^p\} \leq c_p E\{(M_\infty^*)^p + (\int_0^\infty |dA_s|)^p\},$$

using $(a+b)^p \leq 2^{p-1}(a^p + b^p)$.

In the case $p = 2$ we have by Doob's maximal quadratic inequality that

$$E\{(M_\infty^*)^2\} \leq 4E\{M_\infty^2\} = 4E\{[M,M]_\infty\}.$$

For general p, $1 \leq p < \infty$, we need Burkholder's inequalities, which state

$$E\{(M_\infty^*)^p\} \leq c_p E\{[M,M]_\infty^{p/2}\}$$

for a universal constant c_p which depends only on p and not on the local martingale M. For *continuous* local martingales and $p \geq 2$ we proved this using Itô's formula in Chap. IV (Theorem 73). For general local martingales and for all finite $p \geq 1$ see, for example, Dellacherie-Meyer [46, page 287].

Continuing, letting the constant c_p vary from line to line we have

$$\begin{aligned}
\|Z\|_{\underline{\underline{S}}^p}^p &\leq c_p E\{(M_\infty^*)^p + (\int_0^\infty |dA_s|)^p\} \\
&\leq c_p E\{[M,M]_\infty^{p/2} + (\int_0^\infty |dA_s|)^p\} \\
&\leq c_p [j_p(M,A)]^p,
\end{aligned}$$

and taking p-th roots yields the result. □

Corollary. On the space of semimartingales, the $\underline{\underline{H}}^p$ norm is stronger than the $\underline{\underline{S}}^p$ norm, $1 \leq p < \infty$.

Theorem 3 (Emery's Inequality). *Let Z be a semimartingale, $H \in \mathbb{L}$, and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ $(1 \leq p \leq \infty, 1 \leq q \leq \infty)$. Then*

$$\|\int_0^\infty H_s dZ_s\|_{\underline{\underline{H}}^r} \leq \|H\|_{\underline{\underline{S}}^p}\|Z\|_{\underline{\underline{H}}^q}.$$

Proof. Let $H \cdot Z$ denote $(\int_0^t H_s dZ_s)_{t\geq 0}$. Recall that we always assume $Z_0 = 0$ a.s., and let $Z = M + A$ be a decomposition of Z with $M_0 = A_0 = 0$ a.s. Then $H \cdot M + H \cdot A$ is a decomposition of $H \cdot Z$. Hence

$$\|H \cdot Z\|_{\underline{\underline{H}}^r} \leq j_r(H \cdot M, H \cdot A).$$

Next recall that $[H \cdot M, H \cdot M] = \int H_s^2 d[M, M]_s$, by Theorem 29 of Chap. II. Therefore

$$\begin{aligned} j_r(H \cdot M, H \cdot A) &= \|(\int_0^\infty H_s^2 d[M,M]_s)^{1/2} + \int_0^\infty |H_s||dA_s|\|_{L^r} \\ &\leq \|H_\infty^*([M,M]_\infty^{1/2} + \int_0^\infty |dA_s|)\|_{L^r} \\ &\leq \|H_\infty^*\|_{L^p}\|([M,M]_\infty^{1/2} + \int_0^\infty |dA_s|)\|_{L^q} \\ &= \|H\|_{\underline{\underline{S}}^p} j_q(M, A), \end{aligned}$$

where the last inequality above follows from Hölder's inequality. The foregoing implies that

$$\|H \cdot Z\|_{\underline{\underline{H}}^r} \leq \|H\|_{\underline{\underline{S}}^p} j_q(M, A)$$

for any such decomposition $Z = M + A$. Taking infimums over all such decompositions yields the result. □

For a process $X \in \mathbb{D}$ and a stopping time T, recall that

$$\begin{aligned} X^T &= X_t 1_{[0,T)} + X_T 1_{[T,\infty)}, \\ X^{T-} &= X_t 1_{[0,T)} + X_{T-} 1_{[T,\infty)}. \end{aligned}$$

A property holding locally was defined in Chap. I, and a property holding prelocally was defined in Chap. IV. Recall that a property π is said to hold **locally** for a process X if $X^{T^n} 1_{\{T^n>0\}}$ has property π for each n, where T^n is a sequence of stopping times tending to ∞ a.s. If the process X is zero at zero (i.e., $X_0 = 0$ a.s.) then the property π is said to hold **prelocally** if $X^{T^n -}$ has property π for each n.

Definition. A process X is **locally** in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) if there exist stopping times $(T^n)_{n\geq 1}$ increasing to ∞ a.s. such that $X^{T_n} 1_{\{T^n>0\}}$ is in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) for each n, $1 \leq p \leq \infty$. If $X_0 = 0$ then X is said to be **prelocally** in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) if $X^{T^n -}$ is in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) for each n.

While there are many semimartingales which are not locally in $\underline{\underline{H}}^p$, *all* semimartingales are prelocally in $\underline{\underline{H}}^p$. The proof of Theorem 4 below closely parallels the proof of Theorem 13 of Chap. IV.

Theorem 4. *Let Z be a semimartingale ($Z_0 = 0$). Then Z is prelocally in $\underline{\underline{H}}^p$, $1 \leq p \leq \infty$.*

Proof.. By the Fundamental Theorem of Local Martingales (Theorem 25 of Chap. III) and the Bichteler-Dellacherie Theorem (Theorem 43 of Chap. III) we know that for given $\varepsilon > 0$, Z has a decomposition $Z = M+A$, $M_0 = A_0 = 0$ a.s., such that the jumps of the local martingale M are bounded by ε. Define inductively

$$T_0 = 0,$$
$$T_{k+1} = \inf\{t \geq T_k : [M,M]_t^{1/2} + \int_0^t |dA_s| \geq k+1\}.$$

The sequence $(T_k)_{k\geq 1}$ are stopping times increasing to ∞ a.s. Moreover

$$Z^{T_k-} = (M^{T_k}) + (A^{T_k-} - \Delta M_{T_k} 1_{[T_k,\infty)}) = N + C$$

is a decomposition of Z^{T_k-}. Also, since $[M,M]_{T_k} = [M,M]_{T_k-} + (\Delta M_{T_k})^2$, we conclude

$$\begin{aligned} j_\infty(N,C) &= \| [N,N]_\infty^{1/2} + \int_0^\infty |dC_s| \|_{L^\infty} \\ &= \| ([M,M]_{T_k-} + (\Delta M_{T_k})^2)^{1/2} + \int_0^{T_k} |dC_s| \|_{L^\infty} \\ &\leq \| (k^2+\varepsilon^2)^{1/2} + (k+\varepsilon) \|_{L^\infty} < \infty. \end{aligned}$$

Therefore $Z^{T_k-} \in \underline{\underline{H}}^\infty$ and hence it is in $\underline{\underline{H}}^p$ as well, $1 \leq p \leq \infty$. □

Definition. Let Z be a semimartingale in $\underline{\underline{H}}^\infty$ and let $\alpha > 0$. A finite sequence of stopping times $0 = T_0 \leq T_1 \leq \cdots \leq T_k$ is said to **α-slice** Z if $Z = Z^{T_k-}$ and $\|(Z - Z^{T_i})^{T_{i+1}-}\|_{\underline{\underline{H}}^\infty} \leq \alpha$, $0 \leq i \leq k-1$. If such a sequence of stopping times exists, we say Z is **α-sliceable**, and we write $Z \in \mathcal{S}(\alpha)$.

Theorem 5. *Let Z be a semimartingale with $Z_0 = 0$ a.s.*

(i) For $\alpha > 0$, if $Z \in \mathcal{S}(\alpha)$ then for every stopping time T, $Z^T \in \mathcal{S}(\alpha)$ and $Z^{T-} \in \mathcal{S}(2\alpha)$.
(ii) For every $\alpha > 0$, there exists an arbitrarily large stopping time T such that $Z^{T-} \in \mathcal{S}(\alpha)$.

Proof. Since $Z^{T-} = M^T + (A^{T-} - \Delta M_T 1_{[T,\infty)})$, and since $\|Z^T\|_{\underline{\underline{H}}^\infty} \leq \|Z\|_{\underline{\underline{H}}^\infty}$ always, one concludes $\|Z^{T-}\|_{\underline{\underline{H}}^\infty} \leq 2\|Z\|_{\underline{\underline{H}}^\infty}$, so that (i) follows.

Next consider (ii). If semimartingales Z and Y are α-sliceable, let T_i^z and T_j^y be two sequences of stopping times respectively α-slicing Z and Y. By reordering the points T_i^z and T_j^y and using (i), we easily conclude that $Z+Y$ is 8α-sliceable. Next let $Z = M + A$, $M_0 = A_0 = 0$ a.s., with the local martingale M having jumps bounded by the constant $\beta = \alpha/24$. By the preceding observation it suffices to consider M and A separately.

For A, let $T_0 = 0$, $T_{k+1} = \inf\{t \geq T_k : \int_{T_k}^t |dA_s| \geq \alpha/8$ or $\int_0^t |dA_s| \geq k\}$. Then $A^{T_k-} \in \mathcal{S}(\alpha/8)$ for each k, and the stopping times (T_k) increase to ∞ a.s.

For M, let $R_0 = 0$, $R_{k+1} = \inf\{t \geq R_k : [M,M]_t - [M,M]_{R_k} \geq \beta^2$ or $[M,M]_t \geq k\}$. Then $M^{R_k-} \in \underline{\underline{H}}^\infty$, each k, and moreover

$$(M - M^{R_k})^{R_{k+1}-} = M^{R_{k+1}} - M^{R_k} - (\Delta M_{R_{k+1}} 1_{\{R_{k+1} > R_k\}}) 1_{[R_{k+1},\infty)}.$$

Hence

$$\begin{aligned}
&\|(M - M^{R_k})^{R_{k+1}-}\|_{\underline{\underline{H}}^\infty} \\
&\quad \leq \|([M,M]_{R_{k+1}} - [M,M]_{R_k})^{1/2} + |\Delta M_{R_{k+1}}|\|_{L^\infty} \\
&\quad = \|((\Delta M_{R_{k+1}})^2 + [M,M]_{R_{k+1}-} - [M,M]_{R_k})^{1/2} + |\Delta M_{R_{k+1}}|\|_{L^\infty} \\
&\quad \leq \|(\beta^2 + \beta^2)^{1/2} + \beta\|_{L^\infty} = (1 + \sqrt{2})\beta.
\end{aligned}$$

Thus for each k, $M^{R_k-} \in \mathcal{S}((1+\sqrt{2})\beta)$, and since $\beta = \alpha/24$, the result follows. □

3 Existence and Uniqueness of Solutions

In presenting theorems on the existence and uniqueness of solutions of stochastic differential equations, there are many choices to be made. First, we do not present the most general conditions known to be allowed; in exchange we are able to give simpler proofs. Moreover the conditions we do give are extremely general and are adequate for the vast majority of applications. For more general results the interested reader can consult Jacod [103, page 451]. Second, we consider only Lipschitz-type hypotheses and thus obtain strong solutions. There is a vast literature on weak solutions (cf., e.g., Stroock-Varadhan [220]). However, weak solutions are more natural (and simpler) when the differentials are the Wiener process and Lebesgue measure, rather than general semimartingales.

A happy consequence of our approach to stochastic differential equations is that it is just as easy to prove theorems for coefficients that depend not only on the state X_t of the solution at time t (the traditional framework), but on the past history of the process X before t as well.

We begin by stating a theorem whose main virtue is its simplicity. It is a trivial corollary of Theorem 7 which follows it. Recall that a process H is in $\mathbb{L}$ if it has càglàd paths and is adapted.

Theorem 6. *Let Z be a semimartingale with $Z_0 = 0$ and let $f : \mathbb{R}_+ \times \Omega \times \mathbb{R} \to \mathbb{R}$ be such that*

(i) for fixed x, $(t,\omega) \mapsto f(t,\omega,x)$ is in $\mathbb{L}$; and

(ii) for each (t,ω), $|f(t,\omega,x)-f(t,\omega,y)| \leq K(\omega)|x-y|$ for some finite random variable K.

Let X_0 be finite and $\mathcal{F}_0$ measurable. Then the equation

$$X_t = X_0 + \int_0^t f(s,\cdot,X_{s-})dZ_s$$

admits a solution. The solution is unique and it is a semimartingale.

Of course one could state such a theorem for a finite number of differentials dZ^j, $1 \leq j \leq d$, and for a finite system of equations.

In the theory of (non-random) ordinary differential equations, coefficients are typically Lipschitz continuous, which ensures the existence and the uniqueness of a solution. In stochastic differential equations we are led to consider more general coefficients that arise, for example, in control theory. There are enough different definitions to cause some confusion, so we present all the definitions here in ascending order of generality. Note that we add, for technical reasons, the non-customary condition (ii) below to the definition of Lipschitz which follows.

Definition. A function $f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is **Lipschitz** if there exists a (finite) constant k such that

(i) $|f(t,x) - f(t,y)| \leq k|x-y|$, each $t \in \mathbb{R}_+$, and
(ii) $t \mapsto f(t,x)$ is right continuous with left limits, each $x \in \mathbb{R}^n$.

f is said to be **autonomous** if $f(t,x) = f(x)$, all $t \geq 0$.

Definition. A function $f : \mathbb{R}_+ \times \Omega \times \mathbb{R}^n \to \mathbb{R}$ is **random Lipschitz** if f satisfies conditions (i) and (ii) of Theorem 6.

Let $\mathbb{D}^n$ denote the space of processes $\mathbf{X} = (X^1,\dots,X^n)$ where each $X^i \in \mathbb{D}$ $(1 \leq i \leq n)$.

Definition. An operator F from $\mathbb{D}^n$ into $\mathbb{D}^1 = \mathbb{D}$ is said to be **process Lipschitz** if for any $\mathbf{X}$, $\mathbf{Y}$ in $\mathbb{D}^n$ the following two conditions are satisfied:

(i) for any stopping time T, $\mathbf{X}^{T-} = \mathbf{Y}^{T-}$ implies $F(\mathbf{X})^{T-} = F(\mathbf{Y})^{T-}$, and
(ii) there exists an adapted process $K \in \mathbb{L}$ such that

$$\|F(\mathbf{X})_t - F(\mathbf{Y})_t\| \leq K_t \|\mathbf{X}_t - \mathbf{Y}_t\|.$$

Definition. An operator F mapping $\mathbb{D}^n$ to $\mathbb{D}^1 = \mathbb{D}$ is **functional Lipschitz** if for any $\mathbf{X}$, $\mathbf{Y}$ in $\mathbb{D}^n$ the following two conditions are satisfied:

(i) for any stopping time T, $\mathbf{X}^{T-} = \mathbf{Y}^{T-}$ implies $F(\mathbf{X})^{T-} = F(\mathbf{Y})^{T-}$, and
(ii) there exists an increasing (finite) process $K = (K_t)_{t \geq 0}$ such that $|F(\mathbf{X})_t - F(\mathbf{Y})_t| \leq K_t \|\mathbf{X} - \mathbf{Y}\|_t^*$ a.s., each $t \geq 0$.

Note that if $g(t,x)$ is a Lipschitz function, then $f(t,x) = g(t-,x)$ is random Lipschitz. A Lipschitz, or a random Lipschitz, function induces a process Lipschitz operator, and if an operator is process Lipschitz, then it is also functional Lipschitz.

An autonomous function with a bounded derivative is Lipschitz by the Mean Value Theorem. If a function f has a continuous but not bounded derivative, f will be **locally Lipschitz**; such functions are defined and considered in Sect. 7 of this chapter.

Let $A = (A_t)_{t\geq 0}$ be continuous and adapted. Then a linear coefficient such as $f(t,\omega,x) = A_t(\omega)x$ is an example of a process Lipschitz coefficient. A functional Lipschitz operator F will typically be of the form $F(X) = f(t,\omega; X_s, s \leq t)$, where f is defined on $[0,t] \times \Omega \times D[0,t]$ for each $t \geq 0$; here $D[0,t]$ denotes the space of càdlàg functions defined on $[0,t]$. Another example is a generalization of the coefficients introduced by Itô and Nisio [101], namely

$$F(X)_t = \int_0^t g(u,\omega,X_u)\mu(\omega,du)$$

for a random signed measure μ and a bounded Lipschitz function g with constant $C(\omega)$. In this case, the Lipschitz process for F is given by $K_t(\omega) = C(\omega)\|\mu(\omega)_t\|$, where $\|\mu(\omega)_t\|$ denotes the total mass of the measure $\mu(\omega,du)$ on $[0,t]$.

Lemmas 1 and 2 which follow are used to prove Theorem 7. We state and prove them in the one dimensional case, their generalizations to n dimensions being simple.

Lemma 1. Let $1 \leq p < \infty$, let $J \in \underline{\underline{S}}^p$, let F be functional Lipschitz with $F(0) = 0$, and suppose $\sup_t |K_t(\omega)| \leq k$ a.s. Let Z be a semimartingale in $\underline{\underline{H}}^\infty$ such that $\|Z\|_{\underline{\underline{H}}^\infty} \leq \frac{1}{2c_p k}$. Then the equation

$$X_t = J_t + \int_0^t F(X)_{s-}dZ_s$$

has a solution in $\underline{\underline{S}}^p$. It is unique, and moreover

$$\|X\|_{\underline{\underline{S}}^p} \leq 2\|J\|_{\underline{\underline{S}}^p}.$$

Proof. Define $\Lambda : \underline{\underline{S}}^p \to \underline{\underline{S}}^p$ by $\Lambda(X)_t = J_t + \int_0^t F(X)_s dZ_s$. Then by Theorems 2 and 3 the operator is 1/2 Lipschitz, and the fixed point theorem gives existence and uniqueness. Indeed

$$\begin{aligned}
\|X\|_{\underline{\underline{S}}^p} &\leq \|J\|_{\underline{\underline{S}}^p} + \|\int F(X)_{s-}dZ_s\|_{\underline{\underline{S}}^p} \\
&\leq \|J\|_{\underline{\underline{S}}^p} + c_p\|F(X)\|_{\underline{\underline{S}}^p}\|Z\|_{\underline{\underline{H}}^\infty} \\
&\leq \|J\|_{\underline{\underline{S}}^p} + \frac{1}{2k}\|F(X)\|_{\underline{\underline{S}}^p}.
\end{aligned}$$

Since $\|F(X)\|_{\underline{\underline{S}}^p} = \|F(X) - F(0)\|_{\underline{\underline{S}}^p}$, we have $\|X\|_{\underline{\underline{S}}^p} \leq \|J\|_{\underline{\underline{S}}^p} + \frac{1}{2}\|X\|_{\underline{\underline{S}}^p}$, which yields the estimate. $\square$

Lemma 2. Let $1 \leq p < \infty$, let $J \in \underline{\underline{S}}^p$, let F be functional Lipschitz with $F(0) = 0$, and suppose $\sup_t |K_t(\omega)| \leq k < \infty$ a.s. Let Z be a semimartingale such that $Z \in \mathcal{S}(\frac{1}{2c_p k})$. Then the equation

$$X_t = J_t + \int_0^t F(X)_{s-} dZ_s$$

has a solution in $\underline{\underline{S}}^p$. It is unique, and moreover $\|X\|_{\underline{\underline{S}}^p} \leq C(k, Z)\|J\|_{\underline{\underline{S}}^p}$, where $C(k, Z)$ is a constant depending only on k and Z.

Proof. Let $z = \|Z\|_{\underline{\underline{H}}^\infty}$ and $j = \|J\|_{\underline{\underline{S}}^p}$. Let $0 = T_0, T_1, \ldots, T_\ell$ be the slicing times for Z, and consider the equations, indexed by $i = 0, 1, 2, \ldots$,

$$X = J^{T_i-} + \int F(X)_{s-} dZ_s^{T_i-}, \qquad i = 0, 1, 2, \ldots. \tag{i}$$

Equation (i) has the trivial solution $X \equiv 0$ since $J^{0-} = Z^{0-} = 0$ for all t, and its $\underline{\underline{S}}^p$ norm is 0. Assume that equation (i) has a unique solution X^i, and let $x^i = \|X^i\|_{\underline{\underline{S}}^p}$. Stopping next at T_i instead of T_i-, let Y^i denote the unique solution of $Y^i = J^{T_i} + \int F(Y^i)_{s-} dZ_s^{T_i}$, and set $y^i = \|Y^i\|_{\underline{\underline{S}}^p}$. Since $Y^i = X^i + \{\Delta J_{T_i} + F(X^i)_{T_i-} \Delta Z_{T_i}\} 1_{[T_i, \infty)}$, we conclude that

$$\begin{aligned}
\|Y^i\|_{\underline{\underline{S}}^p} &\leq \|X^i\|_{\underline{\underline{S}}^p} + 2\|J\|_{\underline{\underline{S}}^p} + \|F(X^i)\|_{\underline{\underline{S}}^p} \|Z\|_{\underline{\underline{H}}^\infty} \\
&\leq x^i + 2j + kx^i z \\
&= x^i(1 + kz) + 2j;
\end{aligned}$$

hence

$$y^i \leq 2j + x^i(1 + kz). \tag{$*$}$$

We set for $U \in \mathbb{D}$, $D_i U = (U - U^{T_i})^{T_{i+1}-}$. Since each solution X of equation $(i+1)$ satisfies $X^{T_i} = Y^i$ on $[0, T_{i+1})$, we can change the unknown by $U = X - (Y^i)^{T_{i+1}-}$, to get the equations $U = D_i J + \int F(Y^i + U)_{s-} dD_i Z_s$. However since $F(Y^i + 0)$ need not be 0 we define $G_i(\cdot) = F(Y^i + \cdot) - F(Y^i)$, and thus the above equation can be equivalently expressed as

$$U = (D_i J + \int F(Y^i)_{s-} dD_i Z_s) + \int G_i(U)_{s-} dD_i Z_s.$$

We can now apply Lemma 1 to this equation to find that it has a unique solution in $\underline{\underline{S}}^p$, and its norm u^i is majorized by

$$\begin{aligned}
u^i &\leq 2\|D_i J + \int F(Y^i)_{s-} dD_i Z_s\|_{\underline{\underline{S}}^p} \\
&\leq 2(2j + c_p k y^i \frac{1}{2c_p k}) \leq 4j + y^i.
\end{aligned}$$

We conclude equation $(i+1)$ has a unique solution in $\underline{\underline{S}}^p$ with norm x^{i+1} dominated by (using $(*)$)

$$x^{i+1} \leq u^i + y^i \leq 4j + 2y^i \leq 8j + 2(1+kz)x^i.$$

Next we iterate from $i = 0$ to $\ell - 1$ to conclude that

$$x^\ell \leq 8\{\frac{(2+2kz)^\ell - 1}{1+2kz}\}j.$$

Finally, since $Z = Z^{T_\ell -}$, we have seen that the equation $X = J + \int F(X)_{s-}dZ_s$ has a unique solution in $\underline{\underline{S}}^p$, and moreover $X = X^\ell + J - J^{T_\ell -}$. Therefore $\|X\|_{\underline{\underline{S}}^p} \leq x^\ell + 2j$, and hence $C(k, Z) \leq 2 + 8\{\frac{(2+2kz)^\ell - 1}{1+2kz}\}$. □

Theorem 7. *Given a vector of semimartingales* $\mathbf{Z} = (Z^1, \ldots, Z^d)$, $\mathbf{Z}_0 = 0$ *processes* $J^i \in \mathbb{D}$, $1 \leq i \leq n$, *and operators* F^i_j *which are functional Lipschitz* $(1 \leq i \leq n, 1 \leq j \leq d)$, *the system of equations*

$$X^i_t = J^i_t + \sum_{j=1}^d \int_0^t F^i_j(\mathbf{X})_{s-}dZ^j_s$$

$(1 \leq i \leq n)$ *has a solution in* $\mathbb{D}^n$, *and it is unique. Moreover if* $(J^i)_{i \leq n}$ *is a vector of semimartingales, then so is* $(X^i)_{i \leq n}$.

Proof. The proof for systems is the same as the proof for one equation provided we take F to be matrix-valued and X, J and Z to be vector-valued. Hence we give here the proof for $n = d = 1$. Thus we will consider the equation

$$X_t = J_t + \int_0^t F(X)_{s-}dZ_s. \qquad (*)$$

Assume that $\max_{i,j} \sup_t K^{i,j}_t(\omega) \leq k < \infty$ a.s. Also, by considering the equation

$$X_t = (J_t + \int_0^t F(0)_{s-}dZ_s) + \int_0^t G(X)_{s-}dZ_s,$$

where $G(X) = F(X) - F(0)$, it suffices to consider the case where $F(0) = 0$. We also need Lemmas 1 and 2 only for $p = 2$. In this case $c_2 = \sqrt{8}$.

Let T be an arbitrarily large stopping time such that $J^{T-} \in \underline{\underline{S}}^2$ and such that $Z^{T-} \in \mathcal{S}(\frac{1}{4\sqrt{8}k})$. Then by Lemma 2 there exists a unique solution in $\underline{\underline{S}}^2$ of

$$X(T)_t = J^{T-}_t + \int_0^t F(X(T))_{s-}dZ^{T-}_s.$$

By the uniqueness in $\underline{\underline{S}}^2$ one has, for $R > T$, that $X(R)^{T-} = X(T)^{T-}$, and therefore we can define a process X on $\Omega \times [0, \infty)$ by $X = X(T)$ on $[0, T)$. Thus we have existence.

Suppose next Y is another solution. Let S be arbitrarily large such that $(X - Y)^{S-}$ is bounded, and let $R = \min(S,T)$, which can also be taken arbitrarily large. Then X^{R-} and Y^{R-} are both solutions of

$$U = J^{R-} + \int_0^t F(U)_{s-} dZ_s^{R-},$$

and since $Z^{R-} \in \mathcal{S}(\frac{1}{2\sqrt{8}k})$, we know that $X^{R-} = Y^{R-}$ by the uniqueness established in Lemma 2. Thus $X = Y$, and we have uniqueness.

We have assumed that $\max_{i,j} \sup_t K_t^{i,j}(\omega) \leq k < \infty$ a.s. By proving existence and uniqueness on $[0, t_0]$, for t_0 fixed, we can reduce the Lipschitz processes $K_t^{i,j}$ to the random constants $K_{t_0}^{i,j}(\omega)$, which we replace with $K(\omega) = \max_{i,j} K_{t_0}^{i,j}(\omega)$. Thus without loss of generality we can assume we have a Lipschitz constant $K(\omega) < \infty$ a.s. Then we can choose a constant c such that $P(K \leq c) > 0$. Let $\Omega_n = \{K \leq c + n\}$, each $n = 1, 2, 3, \ldots$. Define a new probability P_n by $P_n(A) = P(A \cap \Omega_n)/P(\Omega_n)$, and note that $P_n \ll P$. Moreover for $n > m$ we have $P_m \ll P_n$. From now on assume $n > m$ in the rest of this proof. Therefore we know that all P semimartingales and all P_n semimartingales are P_m semimartingales, and that on Ω_m a stochastic integral calculated under P_m agrees with the the same one calculated under P_n, by Theorem 14 of Chap. II. Let Y^n be the unique solution with respect to P_n which we know exists by the foregoing. We conclude $Y^n = Y^m$ on Ω_m, a.s. (dP_m). Define

$$Y_t = \sum_{n=1}^{\infty} Y_t^n 1_{\{\Omega_n \setminus \Omega_{n-1}\}}$$

and we have $Y = Y^n$ a.s. (dP_n) on Ω_n, and hence also a.s. (dP) on Ω_n, each n. Since $\Omega = \cup_{n=1}^{\infty}(\Omega_n \setminus \Omega_{n-1})$ a.s. (dP), we have that on Ω_n:

$$\begin{aligned} Y_t &= J_t + \int_0^t F(Y^n)_{s-} dZ_s \\ &= J_t + \int_0^t F(Y)_{s-} dZ_s \end{aligned}$$

a.s. (dP), for each n. This completes the proof. □

Theorem 7 can be generalized by weakening the Lipschitz assumption on the coefficients. If the coefficients are Lipschitz on compact sets, for example, in general one has unique solutions existing only up to a stopping time T; at this time one has $\limsup_{t \to T} |\mathbf{X}_t| = \infty$. Such times are called **explosion times**, and they can be finite or infinite. Coefficients that are Lipschitz on compact sets are called *locally Lipschitz*. Simple cases are treated in Sect. 7 of this chapter (cf., Theorems 38, 39, and 40), where they arise naturally in the study of flows.

We end this section with the remark that we have already met a fundamental stochastic differential equation in Chap. II, that of the **stochastic exponential equation**

$$X_t = X_0 + \int_0^t X_{s-} dZ_s, \quad Z_0 = 0.$$

There we obtained a formula for its solution (thus *a fortiori* establishing the existence of a solution), namely

$$X_t = X_0 \exp\{Z_t - \frac{1}{2}[Z, Z]_t\} \prod_{0<s\le t} (1 + \Delta Z_s) \exp\{-\Delta Z_s + \frac{1}{2}(\Delta Z_s)^2\}.$$

The uniqueness of this solution is a consequence of Theorem 7, or of Theorem 6.

A traditional way to show the existence and uniqueness of solutions of ordinary differential equations is the Picard iteration method. One might well wonder if Picard-type iterations converge in the case of stochastic differential equations. As it turns out, the following theorem is quite useful.

Theorem 8. *Let the hypotheses of Theorem 7 be satisfied, and in addition let* $(X^0)^i = H^i$ *be processes in* $\mathbb{D}$ $(1 \le i \le n)$. *Define inductively*

$$(X_t^{m+1})^i = J_t^i + \sum_{j=1}^d \int_0^t F_j^i(\mathbf{X}^m)_{s-} dZ_s^j$$

and let $\mathbf{X}$ *be the solution of*

$$X_t^i = J_t^i + \sum_{j=1}^d \int_0^t F_j^i(\mathbf{X})_{s-} dZ_s^j \ (1 \le i \le n).$$

Then $\mathbf{X}^m$ *converges to* $\mathbf{X}$ *in ucp.*

Proof. We give the proof for $d = n = 1$. It is easy to see that if $\lim_{m\to\infty} X^m = X$ prelocally in $\underline{\underline{S}}^2$, then $\lim_{m\to\infty} X^m = X$ in *ucp*; in any event this is proved in Theorem 12 in Sect. 4. Thus we will show that $\lim_{m\to\infty} X^m = X$ prelocally in $\underline{\underline{S}}^2$. We first assume $\sup_t K_t \le a < \infty$ a.s. Without loss of generality we can assume $Z \in \mathcal{S}(\alpha)$, with $\alpha = \frac{1}{2\sqrt{8a}}$, and that $J \in \underline{\underline{S}}^2$. Let $0 = T_0 \le T_1 \le \cdots \le T_k$ be the stopping times that α-slice Z. Then $(X^m)^{T_1-}$ and X^{T_1-} are in $\underline{\underline{S}}^2$ by Lemma 1. Then

$$\|(X^{m+1} - X)^{T_1-}\|_{\underline{\underline{S}}^2} \le \sqrt{8}\|(F(X^m) - F(X))^{T_1-}\|_{\underline{\underline{S}}^2} \|Z^{T_1-}\|_{\underline{\underline{H}}^\infty}$$

by Theorems 2 and 3. Therefore

$$\begin{aligned}\|(X^{m+1}-X)^{T_1-}\|_{\underline{\underline{S}}^2} &\leq \frac{1}{2}\|(X^m-X)^{T_1-}\|_{\underline{\underline{S}}^2} \\ &\leq \frac{1}{2^m}\|(X^1-X)^{T_1-}\|_{\underline{\underline{S}}^2}\end{aligned}$$

so that $\lim_{m\to\infty}\|(X^{m+1}-X)^{T_1-}\|_{\underline{\underline{S}}^2}=0$. We next analyze the jump at T_1. Since

$$X_{T_1}^{m+1}=X_{T_1-}^{m+1}+F(X^m)_{T_1-}\Delta Z_{T_1},$$

we have

$$\|(X^{m+1}-X)^{T_1}\|_{\underline{\underline{S}}^2}\leq\|(X^{m+1}-X)^{T_1-}\|_{\underline{\underline{S}}^2}+az\|(X^m-X)^{T_1-}\|_{\underline{\underline{S}}^2},$$

where $z=\|Z\|_{\underline{\underline{H}}^\infty}<\infty$. Therefore

$$\|(X^{m+1}-X)^{T_1}\|_{\underline{\underline{S}}^2}\leq\frac{1}{2^{m-1}}(1+az)\|(X^1-X)^{T_1-}\|_{\underline{\underline{S}}^2}.$$

Next suppose we know that

$$\|(X^{m+1}-X)^{T_\ell}\|_{\underline{\underline{S}}^2}\leq\frac{(m+1)^{\ell-1}(1+az)^\ell\gamma}{2^{m-1}}$$

where $\gamma=\|(X^1-X)^{T_k-}\|_{\underline{\underline{S}}^2}$. Then

$$(X_t^{m+1}-X_t)^{T_{\ell+1}-}=(X_t^{m+1}-X_t)^{T_\ell}+\int_{T_\ell}^t(F(X^m)_{s-}-F(X)_{s-})d\hat{Z}_s^{\ell+1}$$

where $\hat{Z}^{\ell+1}=(Z-Z^{T_\ell})^{T_{\ell+1}-}$. Therefore by iterating on ℓ, $0\leq\ell\leq k$,

$$\begin{aligned}\|(X^{m+1}-X)^{T_{\ell+1}-}\|_{\underline{\underline{S}}^2}&\leq\|(X^{m+1}-X)^{T_\ell}\|_{\underline{\underline{S}}^2}+a\sqrt{8}\alpha\|(X^m-X)^{T_{\ell+1}-}\|_{\underline{\underline{S}}^2}\\ &\leq\frac{(m+1)^{\ell-1}(1+az)^\ell}{2^{m-1}}+a\sqrt{8}\alpha\|(X^m-X)^{T_{\ell+1}-}\|_{\underline{\underline{S}}^2}\\ &\leq\frac{(m+1)^{\ell-1}(1+az)^\ell\gamma}{2^{m-1}}.\end{aligned}$$

Note that the above expression tends to 0 as m tends to ∞. Therefore X^m tends to X prelocally in $\underline{\underline{S}}^2$ by a (finite) induction, and hence $\lim_{m\to\infty}X^m=X$ in *ucp*.

It remains to remove the assumption that $\sup_t K_t\leq a<\infty$ a.s. Fix a $t<\infty$; we will show *ucp* on $[0,t]$. Since t is arbitrary, this will imply the result. As we did at the end of the proof of Theorem 7, let $c>0$ be such that $P(K_t\leq c)>0$. Define $\Omega_n=\{\omega: K_t(\omega)\leq c+n\}$, and $P_n(\Lambda)\equiv P(\Lambda|\Omega_n)$. Then $P_n\ll P$, and under P_n, $\lim_{m\to\infty}X^m=X$ in *ucp* on $[0,t]$. For $\varepsilon>0$, choose N such that $n\geq N$ implies $P(\Omega_n^c)<\varepsilon$. Then

$$P((X^m-X)_t^*>\delta)\leq P_n((X^m-X)_t^*>\delta)+\varepsilon,$$

hence $\lim_{m\to\infty}P((X^m-X)_t^*>\delta)\leq\varepsilon$. Since $\varepsilon>0$ was arbitrary, the limit must be zero. □

4 Stability of Stochastic Differential Equations

Since one is never exactly sure of the accuracy of a proposed model, it is important to know how robust the model is. That is, if one perturbs the model a bit, how large are the resulting changes? Stochastic differential equations are stable with respect to perturbations of the coefficients, or of the initial conditions. Perturbations of the differentials, however, are a more delicate matter. One must perturb the differentials in the right way to have stability. Not surprisingly, an $\underline{\underline{H}}^p$ perturbation is the right kind of perturbation. In this section we will be concerned with equations of the form

$$X_t^n = J_t^n + \int_0^t F^n(X^n)_{s-} dZ_s^n, \tag{$*n$}$$

$$X_t = J_t + \int_0^t F(X)_{s-} dZ_s, \tag{$*$}$$

where J^n, J are in $\mathbb{D}$, Z^n, Z are semimartingales, and F^n, F are functional Lipschitz, with Lipschitz processes K_n, K, respectively. *We will assume that the Lipschitz processes K_n, K are each uniformly bounded by the same constant, and that the semimartingale differentials Z^n, Z are always zero at* 0 (that is, $Z_0^n = 0$ a.s., $n \geq 1$, and $Z_0 = 0$ a.s.).

For simplicity we state and prove the theorems in this section for one equation (rather than for finite systems), with one semimartingale driving term (rather than a finite number), and for $p = 2$. The generalizations are obvious, and the proofs are exactly the same except for notation. We say a functional Lipschitz operator F is **bounded** if for all $H \in \mathbb{D}$, there exists a non-random constant $c < \infty$ such that $F(H)^* < c$.

Theorem 9. *Let J, $J^n \in \mathbb{D}$; Z, Z^n be semimartingales; and F, F^n be functional Lipschitz with constants K, K_n, respectively. Assume that*

(i) J, J^n are in $\underline{\underline{S}}^2$ (resp. $\underline{\underline{H}}^2$) and $\lim_{n\to\infty} J^n = J$ in $\underline{\underline{S}}^2$ (resp. $\underline{\underline{H}}^2$);
(ii) F^n are all bounded by the same constant c, and $\lim_{n\to\infty} F^n(X) = F(X)$ in $\underline{\underline{S}}^2$, where X is the solution of $(*)$; and
(iii) $\max(\sup_n K_n, K) \leq a < \infty$ a.s. (a not random), $Z \in \mathcal{S}(\frac{1}{2\sqrt{8a}})$, $(Z^n)_{n\geq 1}$ are in $\underline{\underline{H}}^2$, and $\lim_{n\to\infty} Z^n = Z$ in $\underline{\underline{H}}^2$.[2]

Then $\lim_{n\to\infty} X^n = X$ in $\underline{\underline{S}}^2$ (resp. in $\underline{\underline{H}}^2$), where X^n is the solution of $(*n)$ and X is the solution of $(*)$.

Proof. We use the notation $H \cdot Z_t$ to denote $\int_0^t H_s dZ_s$, and $H \cdot Z$ to denote the process $(H \cdot Z_t)_{t\geq 0}$. We begin by supposing that J, $(J^n)_{n\geq 1}$ are in $\underline{\underline{S}}^2$ and J^n converges to J in $\underline{\underline{S}}^2$. Then

[2] $\mathcal{S}(\alpha)$ is defined in Sect. 2 on page 248.

$$X - X^n = J - J^n + (F(X) - F^n(X))_- \cdot Z \\ + (F^n(X) - F^n(X^n))_- \cdot Z + F^n(X^n)_- \cdot (Z - Z^n).$$

Let $Y^n = (F(X) - F^n(X))_- \cdot Z + F^n(X^n) \cdot (Z - Z^n)$. Then

$$X - X^n = J - J^n + Y^n + (F^n(X) - F^n(X^n))_- \cdot Z. \tag{$**$}$$

For $U \in \mathbb{D}$ define G^n by

$$G^n(U) = F^n(X) - F^n(X - U).$$

Then $G^n(U)$ is functional Lipschitz with constant a and $G^n(0) = 0$. Take $U = X - X^n$, and $(**)$ becomes

$$U = (J - J^n + Y^n) + G^n(U)_- \cdot Z.$$

By Lemma 2 preceding Theorem 7 we have

$$\|X - X^n\|_{\underline{\underline{S}}^2} \leq C(a, Z)\|J - J^n + Y^n\|_{\underline{\underline{S}}^2} \leq C(a, Z)\{\|J - J^n\|_{\underline{\underline{S}}^2} + \|Y^n\|_{\underline{\underline{S}}^2}\}.$$

Since $C(a, Z)$ is independent of n and $\lim_{n\to\infty} \|J - J^n\|_{\underline{\underline{S}}^2} = 0$ by hypothesis, it suffices to show $\lim_{n\to\infty} \|Y^n\|_{\underline{\underline{S}}^2} = 0$. But

$$\begin{aligned} \|Y^n\|_{\underline{\underline{S}}^2} &\leq \|(F(X) - F^n(X))_- \cdot Z\|_{\underline{\underline{S}}^2} + \|F^n(X^n)_- \cdot (Z - Z^n)\|_{\underline{\underline{S}}^2} \\ &\leq \sqrt{8}\|F(X) - F^n(X)\|_{\underline{\underline{S}}^2}\|Z\|_{\underline{\underline{H}}^\infty} \\ &\quad + \sqrt{8}\|F^n(X^n)_-\|_{\underline{\underline{S}}^\infty}\|Z - Z^n\|_{\underline{\underline{H}}^2} \end{aligned} \tag{$***$}$$

by Theorem 2 and Emery's inequality (Theorem 3). Since $\|Z\|_{\underline{\underline{H}}^\infty} < \infty$ by hypothesis and since

$$\lim_{n\to\infty} \|F(X) - F^n(X)\|_{\underline{\underline{S}}^2} = \lim_{n\to\infty} \|Z - Z^n\|_{\underline{\underline{H}}^2} = 0,$$

again by hypothesis, we are done.

Note that if we knew J^n, $J \in \underline{\underline{H}}^2$ and that J^n converged to J in $\underline{\underline{H}}^2$, then

$$\|X - X^n\|_{\underline{\underline{H}}^2} \leq \|J - J^n\|_{\underline{\underline{H}}^2} + \|Y^n\|_{\underline{\underline{H}}^2} + C\|X - X^n\|_{\underline{\underline{S}}^2}\|Z\|_{\underline{\underline{H}}^\infty}.$$

We have seen already that $\lim_{n\to\infty} \|X - X^n\|_{\underline{\underline{S}}^2} = 0$, hence it suffices to show $\lim_{n\to\infty} \|Y^n\|_{\underline{\underline{H}}^2} = 0$. Proceeding as in $(***)$ we obtain the result. □

Comment. Condition (ii) in Theorem 9 seems very strong; it is the perturbation of the semimartingale differentials that make it necessary. Indeed, the hypothesis cannot be relaxed in general, as the following example shows. We take $\Omega = [0, 1]$, P to be Lebesgue measure on $[0, 1]$, and $(\mathcal{F}_t)_{t\geq 0}$ equal to $\mathcal{F}$, the Lebesgue sets on $[0, 1]$. Let $\phi(t) = \min(t, 1)$, $t \geq 0$. Let $f_n(\omega) \geq 0$ and set $Z_t^n(\omega) = \phi(t)f_n(\omega)$, $\omega \in [0, 1]$, and $Z_t(\omega) = \phi(t)f(\omega)$.

Let $F^n(X) = F(X) \equiv X$, and finally let $J^n_t = J_t = 1$, all $t \geq 0$. Thus the equations $(*n)$ and $(*)$ become respectively

$$X^n_t = 1 + \int_0^t X^n_{s-} dZ^n_s,$$
$$X_t = 1 + \int_0^t X_{s-} dZ_s,$$

which are elementary continuous exponential equations and have solutions

$$X^n_t = \exp\{Z^n_t\} = \exp\{f_n(\omega)\phi(t)\},$$
$$X_t = \exp\{Z_t\} = \exp\{f(\omega)\phi(t)\}.$$

We can choose f_n such that $\lim_{n\to\infty} E\{f_n^2\} = 0$ but $\lim_{n\to\infty} E\{f_n^p\} \neq 0$ for $p > 2$. Then the Z^n converge to 0 in $\underline{\underline{H}}^2$ but X^n does not converge to $X = 1$ (since $f = 0$) in $\underline{\underline{S}}^p$, for any $p \geq 1$. Indeed, $\lim_{n\to\infty} E\{f_n^p\} \neq 0$ for $p > 2$ implies $\lim_{n\to\infty} E\{e^{tf_n}\} \neq 1$ for any $t > 0$.

The next result *does not require that the coefficients be bounded*, because there is only one, fixed, semimartingale differential. Theorem 10, 11, and 13 all have $\underline{\underline{H}}^2$ as well as $\underline{\underline{S}}^2$ versions as in Theorem 9, but we state and prove only the $\underline{\underline{S}}^2$ versions.

Theorem 10. *Let J, $J^n \in \mathbb{D}$; Z be a semimartingale; F, F^n be functional Lipschitz with constants K, K_n, respectively; and let X^n, X be the unique solutions of equations $(*n)$ and $(*)$, respectively. Assume that*

(i) J^n, J are in $\underline{\underline{S}}^2$ and $\lim_{n\to\infty} J^n = J$ in $\underline{\underline{S}}^2$;
(ii) $\lim\limits_{n\to\infty} F^n(X) = F(X)$ in $\underline{\underline{S}}^2$, where X is the solution of $(*)$; and
(iii) $\max(\sup_n K_n, K) \leq a < \infty$ a.s. for a non-random constant a, and $Z \in \mathcal{S}(\frac{1}{2\sqrt{8a}})$.

Then $\lim_{n\to\infty} X^n = X$ in $\underline{\underline{S}}^2$ where X^n is the solution of $(*n)$ and X is the solution of $(*)$.

Proof. Let X^n and X be the solutions of equations $(*n)$ and $(*)$, respectively. Then

$$X - X^n = J - J^n + (F(X) - F^n(X))_- \cdot Z + (F^n(X) - F^n(X^n))_- \cdot Z.$$

We let $Y^n = (F(X) - F^n(X))_- \cdot Z$, and we define a new functional Lipschitz operator G^n by

$$G^n(U) = F^n(X) - F^n(X - U).$$

Then $G^n(0) = 0$. If we set $U = X - X^n$, we obtain the equation

$$U = J - J^n + Y^n + G^n(U)_- \cdot Z.$$

Since $Z \in \underline{\underline{H}}^{\infty}$, by Emery's inequality (Theorem 3) we have $Y^n \to 0$ in $\underline{\underline{H}}^2$, and hence also in $\underline{\underline{S}}^2$ (Theorem 2). In particular $\|Y^n\|_{\underline{\underline{S}}^2} < \infty$, and therefore by Lemma 2 in Sect. 3 we have

$$\|U\|_{\underline{\underline{S}}^2} \leq C(a,Z)\|J - J^n + Y^n\|_{\underline{\underline{S}}^2},$$

where $C(a,Z)$ is independent of n, and where the right side tends to zero as $n \to \infty$. Since $U = X - X^n$, we are done. □

We now wish to localize the results of Theorems 9 and 10 so that they hold for general semimartingales and exogenous processes J^n, J. We first need a definition, which is consistent with our previous definitions of properties holding locally and prelocally (defined in Chap. IV, Sect. 2).

Definition. Processes M^n are said to **converge locally (resp. prelocally) in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$)** to M if M^n, M are in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) and if there exists a sequence of stopping times T_k increasing to ∞ a.s. such that $\lim_{n\to\infty}\|(M^n - M)^{T_k}1_{\{T_k>0\}}\|_{\underline{\underline{S}}^p} = 0$ (resp. $\lim_{n\to\infty}\|(M^n - M)^{T_k-}\|_{\underline{\underline{S}}^p} = 0$) for each $k \geq 1$ (resp. $\underline{\underline{S}}^p$ replaced by $\underline{\underline{H}}^p$).

Theorem 11. *Let J, $J^n \in \mathbb{D}$; Z be a semimartingale ($Z_0 = 0$); and F, F^n be functional Lipschitz with Lipschitz processes K, K_n, respectively. Let X^n, X be solutions respectively of*

$$X_t^n = J_t^n + \int_0^t F^n(X^n)_{s-}dZ_s, \tag{$*n$}$$

$$X_t = J_t + \int_0^t F(X)_{s-}dZ_s. \tag{$*$}$$

Assume that

(i) J^n converge to J prelocally in $\underline{\underline{S}}^2$;
(ii) $F^n(X)$ converges to $F(X)$ prelocally in $\underline{\underline{S}}^2$ where X is the solution of ($*$); and
(iii) $\max(\sup_n K_n, K) \leq a < \infty$ *a.s. (a not random).*

Then $\lim_{n\to\infty} X^n = X$ *prelocally in $\underline{\underline{S}}^2$ where X^n is the solution of ($*n$) and X is the solution of ($*$).*

Proof. By stopping at $T-$ for an arbitrarily large stopping time T we can assume without loss of generality that $Z \in \mathcal{S}(\frac{1}{2\sqrt{8}a})$ by Theorem 5, and that J^n converges to J in $\underline{\underline{S}}^2$ and $F(X^n)$ converges to $F(X)$ in $\underline{\underline{S}}^2$, by hypothesis. Next we need only to apply Theorem 10. □

We can recast Theorem 11 in terms of convergence in *ucp* (uniform convergence on compacts, in probability), which we introduced in Sect. 4 of Chap. II in order to develop the stochastic integral.

Corollary. Let J^n, $J \in \mathbb{D}$; Z be a semimartingale ($Z_0 = 0$); and F, F^n be functional Lipschitz with Lipschitz processes K, K_n, respectively. Let X, X^n be as in Theorem 11. Assume that

(i) J^n converges to J in *ucp*,
(ii) $F^n(X)$ converges to $F(X)$ in *ucp*, and
(iii) $\max(\sup_n K_n, K) \leq a < \infty$ a.s. (a not random).

Then $\lim_{n\to\infty} X^n = X$ in *ucp*.

Proof. Recall that convergence in *ucp* is metrizable; let d denote a distance compatible with it. If X^n does not converge to 0 in *ucp*, we can find a subsequence n' such that $\inf_{n'} d(X^{n'}, 0) > 0$. Therefore no sub-subsequence $(X^{n''})$ can converge to 0 in *ucp*, and hence $X^{n''}$ cannot converge to 0 prelocally in $\underline{\underline{S}}^2$ as well. Therefore to establish the result we need to show only that for any subsequence n', there exists a further subsequence n'' such that $X^{n''}$ converges prelocally to 0 in $\underline{\underline{S}}^2$. This is the content of Theorem 12 which follows, so the proof is complete. □

Theorem 12. *Let H^n, $H \in \mathbb{D}$. For H^n to converge to H in ucp it is necessary and sufficient that there exist a subsequence n' such that $\lim_{n'\to\infty} H^{n'} = H$, prelocally in $\underline{\underline{S}}^2$.*

Proof. We first show the necessity. Without loss of generality, we assume that $H = 0$. We construct by iteration a decreasing sequence of subsets $(\mathbb{N}_k)$ of $\mathbb{N} = \{1, 2, 3, \dots\}$, such that

$$\lim_{\substack{n\to\infty \\ n\in\mathbb{N}_k}} \sup_{0\leq s\leq k} |H_s^n| = 0 \quad \text{a.s.}$$

By Cantor's diagonalization procedure we can find an infinite subset $\mathbb{N}'$ of $\mathbb{N}$ such that

$$\lim_{\substack{n\to\infty \\ n\in\mathbb{N}'}} \sup_{0\leq s\leq k} |H_s^n| = 0 \quad \text{a.s.},$$

each integer $k > 0$. By replacing $\mathbb{N}$ with $\mathbb{N}'$ we can assume without loss of generality that H^n tends to 0 uniformly on compacts, *almost surely*. We next define

$$T_n = \inf\{t \geq 0 : |H_t^n| \geq 1\},$$
$$S_n = \inf_{m\geq n} T_m.$$

Then T_n and S_n are stopping times and the S_n increase to ∞ a.s. Indeed, for each k there exists $N(\omega)$ such that for $n \geq N(\omega)$, $\sup_{0\leq s\leq k} |H_s^n(\omega)| < 1$. Hence $T_n(\omega) \geq k$ and $S_{N(\omega)}(\omega) \geq k$. Next, define

$$L^m = (H^m)^{(S_n\wedge n)-}.$$

Then

$$(L^m)^* \leq \sup_{0\leq s\leq n} |H_s^m|,$$

which tends to 0 a.s. Moreover, when $m \geq n$,

$$(L^m)^* \leq \sup_{0\leq s<S_n} |H_s^m| \leq \sup_{0\leq s<T_m} |H_s^m| \leq 1.$$

Hence by the Dominated Convergence Theorem we have that $(L^m)^*$ tends to 0 in L^2, which implies the result.

For the sufficiency, suppose that H^n does not converge to H in *ucp*. Then there exist $t_0 > 0$, $\delta > 0$, $\varepsilon > 0$, and a subsequence n' such that $P\{(H^{n'} - H)_{t_0}^* > \delta\} > \varepsilon$ for each n'. Then $\|(H^{n'} - H)_{t_0}^*\|_{L^2} \geq \delta\sqrt{\varepsilon}$ for all n'. Let T^k tend to ∞ a.s. such that $(H^{n'} - H)^{T^k-}$ tends to 0 in $\underline{\underline{S}}^2$, each k. Then there exists $K > 0$ such that $P(T^k < t_0) < \frac{\delta\sqrt{\varepsilon}}{2}$ for all $k > K$. Hence

$$\lim_{n'\to\infty} \|(H^n - H)_{t_0}^*\|_{L^2} \leq \lim_{n'\to\infty} \|(H^{n'} - H)^{T^k-}\|_{\underline{\underline{S}}^2} + \frac{\delta\sqrt{\varepsilon}}{2}.$$

We conclude $\delta\sqrt{\varepsilon} \leq \frac{\delta\sqrt{\varepsilon}}{2}$, a contradiction. Whence H^n converges to H in *ucp*. □

Recall that we have stated and proven our theorems for the simplest case of one equation (rather than finite systems) and one semimartingale driving term (rather than a finite number). The extensions to systems and several driving terms is simple and essentially only an exercise in notation. We leave this to the reader.

An interesting consequence of the preceding results is that prelocal $\underline{\underline{S}}^p$ and prelocal $\underline{\underline{H}}^p$ convergence *are not topological* in the usual sense. If they were, then one would have that a sequence converged to zero if and only if every subsequence had a sub-subsequence that converged to zero. To see that this is not the case for $\underline{\underline{S}}^2$ for instance, consider the example given in the comment following Theorem 9. In this situation, the solutions X^n converge to X in *ucp*. By Theorem 12, this implies the existence of a subsequence n' such that $\lim_{n'\to\infty} X^{n'} = X$, prelocally in $\underline{\underline{S}}^2$. However we saw in the comment that X^n does not converge to X in $\underline{\underline{S}}^2$. It is still *a priori* possible that X^n converges to X *prelocally* in $\underline{\underline{S}}^2$, however. In the framework of the example a stopping time is simply a non-negative random variable. Thus our counterexample is complete with the following real analysis result (see Protter [199, page 344] for a proof). There exist non-negative functions f_n on $[0,1]$ such that $\lim_{n\to\infty} \int_0^1 f_n(x)^2 dx = 0$ and $\limsup_{n\to\infty} \int_\Lambda (f_n(x))^p dx = +\infty$ for all $p > 2$ and all Lebesgue sets Λ with strictly positive Lebesgue measure. In conclusion, this counterexample gives a sequence of semimartingales X^n such that every subsequence has a sub-subsequence converging prelocally in $\underline{\underline{S}}^p$, but the sequence itself does not converge prelocally in $\underline{\underline{S}}^p$, $(1 \leq p < \infty)$.

Finally, we observe that such non-topological convergence is not as unusual as one might think at first. Indeed, let X_n be random variables which converge to zero in probability but not a.s. Then every subsequence has a subsubsequence which converges to zero a.s., and thus almost sure convergence is also not topological in the usual sense.

As an example of Theorem 10, let us consider the equations

$$X_t^n = J_t^n + \int_0^t F^n(X^n)_s dW_s + \int_0^t G^n(X^n)_s ds,$$
$$X_t = J_t + \int_0^t F(X)_s dW_s + \int_0^t G(X)_s ds$$

where W is a standard Wiener process. If $(J^n - J)_t^*$ converges to 0 in L^2, each $t > 0$, and if F^n, G^n, F, G are all functional Lipschitz with constant $K < \infty$ and are such that $(F^n(X) - F(X))_t^*$ and $(G^n(X) - G(X))_t^*$ converge to 0 in L^2, each $t > 0$, then $(X^n - X)_t^*$ converges to 0 in L^2 as well, each $t > 0$. Note that we require only that $F^n(X)$ and $G^n(X)$ converge respectively to $F(X)$ and $G(X)$ for the one X that is the solution, and not for all processes in $\mathbb{D}$.

One can weaken the hypothesis of Theorem 9 and still let the differentials vary, provided the coefficients stay bounded, as the next theorem shows.

Theorem 13. *Let J^n, $J \in \mathbb{D}$; Z, Z^n be semimartingales ($Z_0^n = Z_0 = 0$ a.s.); and F, F^n be functional Lipschitz with Lipschitz processes K, K_n, respectively. Let X^n, X be solutions of $(*n)$ and $(*)$, respectively. Assume that*

(i) J^n converges to J prelocally in $\underline{\underline{S}}^2$;
(ii) $F^n(X)$ converges to $F(X)$ prelocally in $\underline{\underline{S}}^2$, and the coefficients F^n, F are all bounded by $c < \infty$;
(iii) Z^n converges to Z prelocally in $\underline{\underline{H}}^2$; and
(iv) $\max(\sup_n K_n, K) \le a < \infty$ a.s. (a not random).

Then $\lim_{n\to\infty} X^n = X$ prelocally in $\underline{\underline{S}}^2$.

Proof. By stopping at $T-$ for an arbitrarily large stopping time T we can assume without loss that $Z \in \mathcal{S}(\frac{1}{2\sqrt{8}a})$ by Theorem 5, and that J^n converges to J in $\underline{\underline{S}}^2$, $F^n(X)$ converges in $\underline{\underline{S}}^2$ to $F(X)$, and Z^n converges to Z in $\underline{\underline{H}}^2$, all by hypothesis. We then invoke Theorem 9, and the proof is complete. □

The assumptions of prelocal convergence are a bit awkward. This type of convergence, however, leads to a topology on the space of semimartingales which is the natural topology for convergence of semimartingale differentials, just as *ucp* is the natural topology for processes related to stochastic integration. This is exhibited in Theorem 15.

Before defining a topology on the space of semimartingales, let us recall that we can define a "distance" on $\mathbb{D}$ by setting, for Y, $Z \in \mathbb{D}$

$$r(Y) = \sum_{n>0} 2^{-n} E\{1 \wedge \sup_{0 \le t \le n} |Y_t|\},$$

and $d(Y,Z) = r(Y-Z)$. This distance is compatible with uniform convergence on compacts in probability, and it was previously defined in Sect. 4 of Chap. II.

Using stochastic integration we can define, for a semimartingale X,

$$\hat{r}(X) = \sup_{|H| \le 1} r(H \cdot X)$$

where the supremum is taken over all predictable processes bounded by one. The **semimartingale topology** is then defined by the distance $\hat{d}(X,Y) = \hat{r}(X-Y)$.

The semimartingale topology can be shown to make the space of semimartingales a topological vector space which is complete. Furthermore, the following theorem relates the semimartingale topology to convergence in $\underline{\underline{H}}^p$. For its proof and a general treatment of the semimartingle topology, see Emery [66] or Protter [201].

Theorem 14. *Let $1 \le p < \infty$, let X^n be a sequence of semimartingales, and let X be a semimartingale.*

(i) If X^n converges to X in the semimartingale topology, then there exists a subsequence which converges prelocally in $\underline{\underline{H}}^p$.
(ii) If X^n converges to X prelocally in $\underline{\underline{H}}^p$, then it converges to X in the semimartingale topology.

In Chap. IV we established the equivalence of the norms $\|X\|_{\underline{\underline{H}}^2}$ and $\sup_{|H|\le 1} \|H \cdot X\|_{\underline{\underline{S}}^2}$ in the corollary to Theorem 24 in Sect. 2 of that chapter. Given this result, Theorem 14 can be seen as a uniform version of Theorem 12.

We are now able once again to recast a result in terms of *ucp* convergence. Theorem 13 has the following corollary.

Corollary. Let J^n, $J \in \mathbb{D}$; Z^n, Z be semimartingales ($Z_0^n = Z_0 = 0$); and F^n, F be functional Lipschitz with Lipschitz processes K, K_n, respectively. Let X^n, X be solutions of $(*n)$ and $(*)$, respectively. Assume that

(i) J^n converges to J in *ucp*,
(ii) $F^n(X)$ converges to $F(X)$ in *ucp* where X is the solution of $(*)$, and moreover all the coefficients F^n are bounded by a random $c < \infty$,
(iii) Z^n converges to Z in the semimartingale topology, and
(iv) $\max(\sup_n K_n, K) \le a < \infty$ a.s.

Then $\lim_{n\to\infty} X^n = X$ in *ucp*.

Proof. Since Z^n converges to Z in the semimartingale topology, by Theorem 14 there exists a subsequence n' such that $Z^{n'}$ converges to Z prelocally in $\underline{\underline{H}}^2$. Then by passing to further subsequences if necessary, by Theorem 12 we may assume without loss that J^n converges to J and $F^n(X)$ converges to $F(X)$ both prelocally in $\underline{\underline{S}}^2$, where X is the solution of $(*)$. Therefore, by Theorem 13, X^n converges to X prelocally in $\underline{\underline{S}}^2$ for this subsequence. We have shown that for the sequence (X^n) there is always a subsequence that converges prelocally in $\underline{\underline{S}}^2$. We conclude by Theorem 12 that X^n converges to X in *ucp*. □

The next theorem extends Theorem 9 and the preceding corollary by relaxing the hypotheses on convergence and especially the hypothesis that all the coefficients be bounded.

Theorem 15. *Let J^n, $J \in \mathbb{D}$; Z^n, Z be semimartingales ($Z_0^n = Z_0 = 0$); and F^n, F be functional Lipschitz with Lipschitz process K, the same for all n. Let X^n, X be solutions respectively of*

$$X_t^n = J_t^n + \int_0^t F^n(X^n)_{s-} dZ_s^n, \tag{$*n$}$$

$$X_t = J_t + \int_0^t F(X)_{s-} dZ_s. \tag{$*$}$$

Assume that

(i) J^n converges to J in ucp;
(ii) $F^n(X)$ converges to $F(X)$ in ucp, where X is the solution of $(*)$; and
(iii) Z^n converges to Z in the semimartingale topology.

Then X^n converges to X in ucp.

Proof. First we assume that $\sup_t K_t(\omega) \le a < \infty$. We remove this hypothesis at the end of the proof. By Theorem 12, it suffices to show that there exists a subsequence n' such that $X^{n'}$ converges to X prelocally in $\underline{\underline{S}}^2$. Then by Theorem 12 we can assume with loss of generality, by passing to a subsequence if necessary, that J^n converges to J and $F^n(X)$ converges to $F(X)$ both prelocally in $\underline{\underline{S}}^2$. Moreover by Theorem 14 we can assume without loss, again by passing to a subsequence if necessary, that Z^n converges to Z prelocally in $\underline{\underline{H}}^2$, and that $Z \in \mathcal{S}(\frac{1}{4\sqrt{8}a})$. Thus all the hypotheses of Theorem 13 are satisfied except one. We do not assume that the coefficients F^n are bounded. However by pre-stopping we can assume without loss that $|F(X)|$ is uniformly bounded by a constant $c < \infty$.

Let us introduce **truncation operators** T^x defined (for $x \ge 0$) by

$$T^x(Y) = \min(x, \sup(-x, Y)).$$

Then T^x is functional Lipschitz with Lipschitz constant 1, for each $x \ge 0$. Consider the equations

$$Y_t^n = J_t^n + \int_0^t (T^{a+c+1} F^n)(Y^n)_{s-} dZ_s^n.$$

Then, by Theorem 13, Y^n converges to X prelocally in $\underline{\underline{S}}^2$. By passing to yet another subsequence, if necessary, we may assume that $F^n(X)$ tends to $F(X)$ and Y^n tends to X uniformly on compacts almost surely. Next we define

$$S^k = \inf\{t \geq 0 : \exists\, m \geq k : |Y_t^m - X_t| + |F^m(X)_t - F(X)_t| \geq 1\}.$$

The stopping times S^k increase a.s. to ∞. By stopping at S^k-, we have for $n \geq k$ that $(Y^n - X)^*$ and $(F^n(X) - F(X))^*$ are a.s. bounded by 1. (Note that stopping at S^k- changes Z to being in $\mathcal{S}(\frac{1}{2\sqrt{8}a})$ instead of $\mathcal{S}(\frac{1}{4\sqrt{8}a})$ by Theorem 5.) Observe that

$$\begin{aligned} |F^n(Y^n)| &\leq |F^n(Y^n) - F^n(X)| + |F^n(X) - F(X)| + |F(X)| \\ &\leq a(Y^n - X)^* + (F^n(X) - F(X))^* + F(X)^* \\ &\leq a + 1 + c, \end{aligned}$$

whence $(T^{a+c+1}F^n)(Y^n) = F^n(Y^n)$. We conclude that, for an arbitrarily large stopping time R, with J^n and Z^n stopped at R, Y^n is a solution of

$$Y_t^n = J_t^n + \int_0^t F^n(Y^n)_{s-} dZ_s^n,$$

which is equation $(*n)$. By the uniqueness of solutions we deduce $Y^n = X^n$ on $[0, R)$. Since Y^n converges to X prelocally in $\underline{\underline{S}}^2$, we thus conclude X^n converges to X prelocally in $\underline{\underline{S}}^2$.

It remains only to remove the hypothesis that $\sup_t K_t(\omega) \leq a < \infty$. Since we are dealing with local convergence, it suffices to consider $\sup_{s \leq t} K_s \leq K_t$, for a fixed t. Since $K_t < \infty$ a.s., let $\alpha > 0$ be such that $P(K_t \leq \alpha) > 0$, and define

$$\Omega_m = \{\omega : K_t(\omega) \leq \alpha + m\}.$$

Then Ω_m increase to Ω a.s. and as in the proof of Theorem 7 we define P_m by $P_m(A) = P(A|\Omega_m)$, and define $\mathcal{F}_t^m = \mathcal{F}_t|_{\Omega_m}$, the trace of $\mathcal{F}_t$ on Ω_m. Then $P_m \ll P$, so that by Lemma 2 preceding Theorem 25 in Chap. IV, if Z^n converges to Z prelocally in $\underline{\underline{H}}^2(P)$, then Z^n converges to Z prelocally in $\underline{\underline{H}}^2(P_m)$ as well. Therefore by the first part of this proof, X^n converges to X in *ucp* under P_m, each $m \geq 1$.

Choose $\varepsilon > 0$ and m so large that $P(\Omega_m^c) < \varepsilon$. Then

$$\begin{aligned} \lim_{n\to\infty} P((X^n - X)_t^* > \delta) &\leq \lim_{n\to\infty} P_m((X^n - X)_t^* > \delta) + P(\Omega_m^c) \\ &\leq \lim_{n\to\infty} P_m((X^n - X)_t^* > \delta) + \varepsilon \\ &\leq \varepsilon, \end{aligned}$$

and since $\varepsilon > 0$ was arbitrary, we conclude that X^n converges to X in *ucp* on $[0, t]$. Finally since t was arbitrary, we conclude X^n converges to X in *ucp*. □

Another important topic is how to approximate solutions by difference solutions. Our preceding convergence results yield two consequences (Theorem 16 and its corollary).

The next lemma is a type of Dominated Convergence Theorem for stochastic integrals and it is used in the proof of Theorem 16.

Lemma (Dominated Convergence Theorem). Let p, q, r be given such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, where $1 < r < \infty$. Let Z be a semimartingale in $\underline{\underline{H}}^q$, and let $H^n \in \underline{\underline{S}}^p$ such that $|H^n| \leq Y \in \underline{\underline{S}}^p$, all $n \geq 1$. Suppose $\lim_{n\to\infty} H^n_{t-}(\omega) = 0$, all (t, ω). Then

$$\lim_{n\to\infty} \| \int H^n_{s-} dZ_s \|_{\underline{\underline{H}}^r} = 0.$$

Proof. Since $Z \in \underline{\underline{H}}^q$, there exists a decomposition of Z, $Z = N + A$, such that

$$j_q(N, A) = \| [N, N]^{1/2}_\infty + \int_0^\infty |dA_s| \|_{L^q} < \infty.$$

Let C^n be the random variable given by

$$C^n = \left(\int_0^\infty (H^n_{s-})^2 d[N, N]_s\right)^{1/2} + \int_0^\infty |H^n_{s-}||dA_s|.$$

The hypothesis that $|H^n| \leq Y$ implies

$$C^n \leq Y^*\left([N, N]^{1/2}_\infty + \int_0^\infty |dA_s|\right) \quad \text{a.s.}$$

However

$$\begin{aligned} \|Y^*([N, N]^{1/2}_\infty + \int_0^\infty |dA_s|)\|_{L^r} &\leq \|Y^*\|_{L^p} \|[N, N]^{1/2}_\infty + \int_0^\infty |dA_s|\|_{L^q} \\ &= \|Y\|_{\underline{\underline{S}}^p} j_q(N, A) < \infty. \end{aligned}$$

Thus C^n is dominated by a random variable in L^r and hence by the Dominated Convergence Theorem C^n tends to 0 in L^r. □

We let σ_n denote a sequence of random partitions tending to the identity.[3] Recall that for a process Y and a random partition $\sigma = \{0 = T_0 \leq T_1 \leq \cdots \leq T_{k_n}\}$, we define

$$Y^\sigma \equiv Y_0 1_{\{0\}} + \sum_k Y_{T_k} 1_{(T_k, T_{k+1}]}.$$

Note that if Y is adapted, càdlàg (i.e, $Y \in \mathbb{D}$), then $(Y^\sigma_s)_{s \geq 0}$ is left continuous with right limits (and adapted, of course). It is convenient to have a version of $Y^\sigma \in \mathbb{D}$, occasionally, so we define

$$Y^{\sigma+} = \sum_k Y_{T_k} 1_{[T_k, T_{k+1})}.$$

[3] Random partitions tending to the identity are defined in Chap. II, preceding Theorem 21.

Theorem 16. *Let $J \in \underline{\underline{S}}^2$, let F be process Lipschitz with Lipschitz process $K \leq a < \infty$ a.s. and $F(0) \in \underline{\underline{S}}^2$. Let Z be a semimartingale in $\mathcal{S}(\frac{1}{2\sqrt{8}a})$, and let $X(\sigma)$ be the solution of*

$$X_t = J_t + \int_0^t F(X^{\sigma+})_s^\sigma dZ_s, \tag{$*\sigma$}$$

for a random partition σ. If σ_n is sequence of random partitions tending to the identity, then $X(\sigma_n)$ tends to X in $\underline{\underline{S}}^2$, where X is the solution of $(*)$ of Theorem 15.

Proof. For the random partition σ_n (n fixed), define an operator G^n on $\mathbb{D}$ by

$$G^n(H) = F(H^{\sigma_n+})^{\sigma_n+}.$$

Note that $G^n(H) \in \mathbb{D}$ for each $H \in \mathbb{D}$ and that $G^n(H)_- = F(H^{\sigma_n+})^{\sigma_n}$. Then G^n is functional Lipschitz with constant K and sends $\underline{\underline{S}}^2$ into itself, as the reader can easily verify. Since $F(0) \in \underline{\underline{S}}^2$, so also are $G^n(0) \in \underline{\underline{S}}^2$, $n \geq 1$, and an argument analogous to the proof of Theorem 10 (though a bit simpler) shows that it suffices to show that $\int_0^t G^n(X)_{s-} dZ_s$ converges to $\int_0^t F(X)_{s-} dZ_s$ in $\underline{\underline{S}}^2$, where X is the solution of

$$X_t = J_t + \int_0^t F(X)_{s-} dZ_s.$$

Towards this end, fix (t, ω) with $t > 0$, and choose $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|X_u(\omega) - X_{t-}(\omega)| < \varepsilon$ for all $u \in [t - 2\delta, t)$. If $\operatorname{mesh}(\sigma) < \delta$, then also $|X_u^{\sigma+}(\omega) - X_u(\omega)| < 2\varepsilon$ for all $u \in [t - \delta, t)$. This then implies that $|F(X^{\sigma+})(\omega) - F(X)(\omega)| < 2a\varepsilon$. Therefore $(F(X^{\sigma_n+}) - F(X))_t^{\sigma_n}(\omega)$ tends to 0 as $\operatorname{mesh}(\sigma_n)$ tends to 0. Since, on $(0, \infty)$,

$$\lim_{\operatorname{mesh}(\sigma_n) \to 0} F(X)^{\sigma_n} = F(X)_-,$$

we conclude that

$$\lim_{\operatorname{mesh}(\sigma_n) \to 0} F(X^{\sigma_n+})^{\sigma_n} = F(X)_-,$$

where convergence is pointwise in (t, ω). Thus

$$\lim_{\operatorname{mesh}(\sigma_n) \to 0} G^n(X)_- = F(X)_-.$$

However we also know that

$$\begin{aligned} |G^n(X)_-| = |F(X^{\sigma_n+})^{\sigma_n}| \leq F(X^{\sigma_n+})^* &\leq F(0)^* + K(X^{\sigma_n+})^* \\ &\leq F(0)^* + aX^* \end{aligned}$$

which is in $\underline{\underline{S}}^2$ and is independent of σ_n. Therefore using the preceding lemma, the Dominated Convergence Theorem for stochastic integrals, we obtain the convergence in $\underline{\underline{S}}^2$ of $\int_0^t G^n(X)_{s-} dZ_s$ to $\int_0^t F(X)_{s-} dZ_s$, and the proof is complete. $\square$

Remark. In Theorem 16 and its corollary which follows, we have assumed that F is *process Lipschitz*, and not functional Lipschitz. Indeed, Theorem 16 is not true in general for functional Lipschitz coefficients. Let $J_t = 1_{\{t\geq 1\}}$, $Z_t = t \wedge 2$, and $F(Y) = Y_1 1_{\{t\geq 1\}}$. Then X, the solution of $(*)$ of Theorem 15 is given by $X_t = (t\wedge 2)1_{\{t\geq 1\}}$, but if σ is any random partition such that $T_k \neq 1$ a.s., then $(X(\sigma)^{\sigma+})_t = 0$ for $t \leq 1$, and therefore $F(X(\sigma)^{\sigma+}) = 0$, and $X(\sigma)_t = J_t = 1_{\{t\geq 1\}}$. (Here $X(\sigma)$ denotes the solution to equation $(*\sigma)$ of Theorem 16.)

Corollary. Let $J \in \mathbb{D}$; F be process Lipschitz; Z be a semimartingale; and let σ_n be a sequence of random partitions tending to the identity. Then

$$\lim_{n\to\infty} X(\sigma_n) = X \qquad \text{in } ucp$$

where $X(\sigma_n)$ is the solution of $(*\sigma)$ and X is the solution of $(*)$, as in Theorem 16.

Proof. First assume $K \leq a < \infty$, a.s. Fix $t > 0$ and $\varepsilon > 0$. By Theorem 5 we can find a stopping time T such that $Z^{T-} \in \mathcal{S}(\frac{1}{8\sqrt{8a}})$, and $P(T < t) < \varepsilon$. Thus without loss of generality we can assume that $Z \in \mathcal{S}(\frac{1}{8\sqrt{8a}})$. By letting $S^k = \inf\{t \geq 0 : |J_t| > k\}$, we have that S^k is a stopping time, and $\lim_{k\to\infty} S^k = \infty$ a.s. By now stopping at S^k- we have that J is bounded, hence also in $\underline{\underline{S}}^2$, and $Z \in \mathcal{S}(\frac{1}{4\sqrt{8a}})$. An analogous argument gives us that $F(0)$ can be assumed bounded (and hence in $\underline{\underline{S}}^2$) as well; hence $Z \in \mathcal{S}(\frac{1}{2\sqrt{8a}})$. We now can apply Theorem 16 to obtain the result.

To remove the assumption that $K \leq a < \infty$ a.s., we need only apply an argument like the one used at the end of the proofs of Theorems 7, 8 and 15. $\square$

Theorem 16 and its corollary give us a way to approximate the solution of a general stochastic differential equation with finite differences. Indeed, let X be the solution of

$$X_t = J_t + \int_0^t F(X)_{s-} dZ_s \qquad (*)$$

where Z is a semimartingale and F is process Lipschitz. For each random partition $\sigma_n = \{0 = T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n\}$, we see that the random variables $X(\sigma_n)_{T_k^n}$ verify the relations (writing σ for σ_n, X for $X(\sigma_n)$, T_k for T_k^n)

$$\begin{aligned}
&X_{T_0} = J_0,\\
&X_{T_{k+1}} = X_{T_k} + J_{T_{k+1}} - J_{T_k} + F(X^{\sigma+})_{T_k}(Z_{T_{k+1}} - Z_{T_k}).
\end{aligned}$$

Then the solution of the finite difference equation above converges to the solution of $(*)$, under the appropriate hypotheses.

As an example we give a means to approximate the stochastic exponential.

Theorem 17. *Let Z be a semimartingale and let $X = \mathcal{E}(Z)$, the stochastic exponential of Z. That is, X is the solution of*

$$X_t = 1 + \int_0^t X_{s-} dZ_s.$$

Let σ_n be a sequence of random partitions tending to the identity. Let

$$X^n = \prod_{i=1}^{k_n-1} (1 + (Z^{T^n_{i+1}} - Z^{T^n_i})).$$

Then $\lim_{n\to\infty} X^n = X$ in ucp.

Proof. Let Y^n be the solution of

$$Y_t = 1 + \int_0^t Y_s^{\sigma_n} dZ_s,$$

equation $(*\sigma)$ of Theorem 16. By the corollary of Theorem 16 we know that Y^n converges to $X = \mathcal{E}(Z)$ in *ucp*. Thus it suffices to show $Y^n = X^n$.

Let $\sigma_n = \{0 = T^n_0 \leq T^n_1 \leq \cdots \leq T^n_{k_n}\}$. On $(T^n_i, T^n_{i+1}]$ we have

$$\begin{aligned} Y^n_t &= Y^n_{T^n_i} + Y^n_{T^n_i}(Z^{T^n_{i+1}} - Z^{T^n_i}) \\ &= Y^n_{T^n_i}(1 + (Z^{T^n_{i+1}} - Z^{T^n_i})). \end{aligned}$$

Inducting on i down to 0 we have

$$Y^n_t = \prod_{j\leq i}(1 + (Z^{T^n_{j+1}} - Z^{T^n_j})),$$

for $T^n_i < t \leq T^n_{i+1}$. Since $Z^{T^n_{j+1}} - Z^{T^n_j} = 0$ for all $j > i$ when $T^n_i < t \leq T^n_{i+1}$, we have that $Y^n = X^n$, and the theorem is proved. □

5 Fisk-Stratonovich Integrals and Differential Equations

In this section we extend the notion of the Fisk-Stratonovich integral given in Chap. II, Sect. 7, and we develop a theory of stochastic differential equations with Fisk-Stratonovich differentials. We begin with some results on the quadratic variation of stochastic processes.

Definition. Let H, J be adapted, càdlàg processes. The **quadratic covariation process** of H, J denoted $[H, J] = ([H, J]_t)_{t\geq 0}$, if it exists, is defined to be the adapted, càdlàg process of finite variation on compacts, such that for any sequence σ_n of random partitions tending to the identity,

$$\begin{aligned}\lim_{n\to\infty} S_{\sigma_n}(H,J) &= \lim_{n\to\infty} H_0J_0 + \sum_i (H^{T^n_{i+1}} - H^{T^n_i})(J^{T^n_{i+1}} - J^{T^n_i}) \\ &= [H,J]\end{aligned}$$

with convergence in *ucp*, where σ_n is the sequence $0 = T^n_0 \le T^n_1 \le \cdots \le T^n_{k_n}$. A process H in $\mathbb{D}$ is said to have **finite quadratic variation** if $[H,H]_t$ exists and is finite a.s., each $t \ge 0$.

If H, J, and $H+J$ in $\mathbb{D}$ have finite quadratic variation, then the *polarization identity* holds:

$$[H,J] = \frac{1}{2}([H+J, H+J] - [H,H] - [J,J]).$$

For X a semimartingale, in Chap. II we defined the quadratic variation of X using the stochastic integral. However Theorem 22 of Chap. II shows every semimartingale is of finite quadratic variation and that the two definitions are consistent.

Notation. For H of finite quadratic variation we let $[H,H]^c$ denote the continuous part of the (non-decreasing) paths of $[H,H]$. Thus,

$$[H,H]_t = [H,H]^c_t + \sum_{0\le s\le t} \Delta[H,H]_s,$$

where $\Delta[H,H]_t = [H,H]_t - [H,H]_{t-}$, the jump at t.

The next definition extends the definition of the Fisk-Stratonovich integral given in Chap. II, Sect. 7.

Definition. Let $H \in \mathbb{D}$, X be a semimartingale, and assume $[H,X]$ exists. The **Fisk-Stratonovich integral** of H with respect to X, denoted $\int_0^t H_{s-} \circ dX_s$, is defined to be

$$\int_0^t H_{s-} \circ dX_s \equiv \int_0^t H_{s-}dX_s + \frac{1}{2}[H,X]^c_s.$$

To consider properly general Fisk-Stratonovich differential equations, we need a generalization of Itô's formulas (Theorems 32 and 33 of Chap. II). Since Itô's formula is proved in detail there, we only sketch the proof of this generalization.

Theorem 18 (Generalized Itô's Formula). *Let $\mathbf{X} = (X^1, \dots, X^n)$ be an n-tuple of semimartingales, and let $f : \mathbb{R}_+ \times \Omega \times \mathbb{R}^n \to \mathbb{R}$ be such that*

(i) there exists an adapted FV process A and a function g such that

$$f(t,\omega,\mathbf{x}) = \int_0^t g(s,\omega,\mathbf{x})dA_s,$$

$(s,\omega) \mapsto g(s,\omega,\mathbf{x})$ is an adapted, jointly measurable process for each $\mathbf{x}$, and $\int_0^t \sup_{\mathbf{x}\in K} |g(s,\omega,\mathbf{x})||dA_s| < \infty$ a.s. for compact sets K.

(ii) *the function g of (i) is C^2 in $\mathbf{x}$ uniformly in s on compacts. That is,*

$$\sup_{s\le t}\{|g(s,\omega,\mathbf{y})-\sum_{i=1}^{n} g_{xi}(s,\omega,\mathbf{x})(y_i-x_i) - \sum_{1\le i,j\le n} g_{x_ix_j}(s,\omega,\mathbf{x})(y_i-x_i)(y_j-x_j)|\} \le r_t(\omega,\|\mathbf{x}-\mathbf{y}\|)\|\mathbf{x}-\mathbf{y}\|^2$$

a.s., where $r_t:\Omega\times\mathbb{R}_+\to\mathbb{R}_+$ is an increasing function with $\lim_{u\downarrow 0} r_t(u)=0$ a.s., provided $\mathbf{x}$ ranges through a compact set (r_t depends on the compact set chosen).

(iii) *the partial derivatives f_{x_i}, $f_{x_ix_j}$, $1\le i,\ j\le n$ all exist and are continuous, and moreover*

$$f_{x_i}(t,\omega,\mathbf{x})=\int_0^t g_{x_i}(s,\omega,\mathbf{x})dA_s,$$
$$f_{x_ix_j}(t,\omega,\mathbf{x})=\int_0^t g_{x_ix_j}(s,\omega,\mathbf{x})dA_s.$$

Then

$$\begin{aligned}&f(t,\omega,\mathbf{X}_t)\\ &=f(0,\omega,\mathbf{X}_0)+\int_0^t g(s,\omega,\mathbf{X}_s)dA_s+\sum_{i=1}^n\int_0^t f_{x_i}(s-,\omega,\mathbf{X}_{s-})dX_s^i\\ &+\frac{1}{2}\sum_{i,j=1}^n\int_{0+}^t f_{x_ix_j}(s-,\omega,\mathbf{X}_{s-})d[X^i,X^j]_s^c+\sum_{0<s\le t}\{f(s,\omega,\mathbf{X}_s)-f(s-,\omega,\mathbf{X}_{s-})\\ &-g(s,\omega,\mathbf{X}_s)\Delta A_s-\sum_{i=1}^n f_{x_i}(s-,\omega,\mathbf{X}_{s-})\Delta X_s^i\}.\end{aligned}$$

Proof. We sketch the proof for $n=1$. We have, letting $0=t_0\le t_1\le\cdots\le t_m=t$ be a partition of $[0,t]$, and assuming temporarily $|X|\le k$ for all $s\le t$, k a constant,

$$\begin{aligned}&f(t,\omega,X_t)-f(0,\omega,X_0)\\ &=\sum_{k=0}^{m-1} f(t_{k+1},\omega,X_{t_{k+1}})-f(t_k,\omega,X_{t_{k+1}})+\sum_{k=0}^{m-1} f(t_k,\omega,X_{t_{k+1}})-f(t_k,\omega,X_{t_k})\\ &=\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} g(u,\omega,X_{t_{k+1}})dA_u+\sum_{k=0}^{m-1} f(t_k,\omega,X_{t_{k+1}})-f(t_k,\omega,X_{t_k}) \qquad (*)\end{aligned}$$

Consider first the term $\sum_{k=0}^{m-1}\int_{t_k}^{t_{k+1}} g(u,\omega,X_{t_{k+1}})dA_u$. The integrand is not adapted, however one can interpret this integral as a path-by-path Stieltjes integral since A is an FV process. Expanding the integrand for fixed (u, ω) by the Mean Value Theorem yields

$$g(u, \omega, X_{t_{k+1}}) = g(u, \omega, X_u) + g_x(u, \omega, \hat{X}_u)(X_{t_{k+1}} - X_u)$$

where $\hat{X}_u$ is in between X_u and $X_{t_{k+1}}$. Therefore

$$\begin{aligned}&\sum_k \int_{t_k}^{t_{k+1}} g(u, \omega, X_{t_{k+1}}) dA_u \\ &\quad = \sum_k \int_{t_k}^{t_{k+1}} g(u, \omega, X_u) dA_u + \sum_k \int_{t_k}^{t_{k+1}} g_x(u, \omega, \hat{X}_u)(X_{t_{k+1}} - X_u) dA_u,\end{aligned}$$

and since A is of finite variation and X is right continuous, the second sum tends to zero as the mesh of the partitions tends to zero. Therefore letting π_n denote a sequence of partitions of $[0, t]$ with $\lim_{n\to\infty} \operatorname{mesh}(\pi_n) = 0$,

$$\lim_{n\to\infty} \sum_{t_k \in \pi_n} \int_{t_k}^{t_{k+1}} g(u, \omega, X_{t_{k+1}}) dA_u = \int_0^t g(u, \omega, X_u) dA_u(\omega) \quad \text{a.s.}$$

Next consider the second term on the right side of $(*)$, namely

$$\sum_{k=0}^{m-1} f(t_k, \omega, X_{t_{k+1}}) - f(t_k, \omega, X_{t_k}).$$

Here we proceed analogously to the proof of Theorem 32 of Chap. II. Given $\varepsilon > 0$, $t > 0$, let $A(\varepsilon, t)$ be a set of jumps of X that has a.s. a finite number of times s, and let $B = B(\varepsilon, t)$ be such that $\sum_{s\in B} (\Delta X_s)^2 \le \varepsilon^2$, where $A \cup B$ exhaust the jumps of X on $(0, t]$. Then

$$\begin{aligned}&\sum_{k=0}^{m-1} f(t_k, \omega, X_{t_{k+1}}) - f(t_k, \omega, X_{t_k}) \\ &\quad = \sum_{k,A} f(t_k, \omega, X_{t_{k+1}}) - f(t_k, \omega, X_{t_k}) + \sum_{k,B} f(t_k, \omega, X_{t_{k+1}}) - f(t_k, \omega, X_{t_k})\end{aligned}$$

where $\sum_{k,A}$ denotes $\sum_{t_k\in\pi_n} 1_{\{A\cap(t_k, t_{k+1}]\neq\emptyset\}}$. Thus

$$\lim_{n\to\infty} \sum_{k,A} f(t_k, \omega, X_{t_{k+1}}) - f(t_k, \omega, X_{t_k}) = \sum_{s\in A} f(s-, \omega, X_s) - f(s-, \omega, X_{s-}).$$

By Taylor's formula, and letting $\Delta_k X$ denote $X_{t_{k+1}} - X_{t_k}$,

$$\sum_{k,B} f(t_k,\omega,X_{t_{k+1}}) - f(t_k,\omega,X_{t_k}) \tag{$**$}$$
$$= \sum_k f_x(t_k,\omega,X_{t_k})\Delta_k X + \frac{1}{2}\sum_k f_{xx}(t_k,\omega,X_{t_k})(\Delta_k X)^2$$
$$- \sum_{k,A} f_x(t_k,\omega,X_{t_k})\Delta_k X + \frac{1}{2} f_{xx}(t_k,\omega,X_{t_k})(\Delta_k X)^2$$
$$+ \sum_{k,B} R(t_k,\omega,X_{t_k},X_{t_{k+1}}).$$

By Theorems 21 and 30 of Chap. II, the first sums on the right above converge in *ucp* to $\int_0^t f_x(s-,\omega,X_{s-})dX_s$ and $\frac{1}{2}\int_0^t f_{xx}(s-,\omega,X_{s-})d[X,X]_s$, respectively. The third sum converges a.s. to

$$-\sum_{s\in A}\{f_x(s-,\omega,X_{s-})\Delta X_s + \frac{1}{2} f_{xx}(s-,\omega,X_{s-})(\Delta X_s)^2\}.$$

By condition (ii) on the function g, we have

$$\limsup_n \sum_{t_k\in\pi_n,B} R(t_k,\omega,X_{t_k},X_{t_{k+1}}) \le r_t(\omega,\varepsilon+)[X,X]_t$$

where

$$r_t(\omega,\varepsilon+) = \limsup_{\delta\downarrow\varepsilon} r_t(\omega,\delta).$$

Next we let ε tend to 0; then $r_t(\omega,\varepsilon+)[X,X]_t$ tends to 0 a.s., and finally combining the two series indexed by A we see that

$$\sum_{s\in A(\varepsilon,t)} \{f(s-,\omega,X_s) - f(s-,\omega,X_{s-}) - f_x(s-,\omega,X_{s-})\Delta X_s$$
$$- \frac{1}{2} f_{xx}(s-,\omega,X_{s-})(\Delta X_s)^2\}$$

tends to the series

$$\sum_{0<s\le t} \{f(s-,\omega,X_s) - f(s-,\omega,X_{s-}) - f_x(s-,\omega,X_{s-})\Delta X_s$$
$$- \frac{1}{2} f_{xx}(s-,\omega,X_{s-})(\Delta X_s)^2\}$$

which is easily seen to be an absolutely convergent series (cf., the proof of Theorem 32 of Chap. II). Incorporating $-\sum_{0<s\le t}\frac{1}{2}f_{xx}(s-,\omega,X_{s-})(\Delta X_s)^2$ into $\frac{1}{2}\int_0^t f_{xx}(s-,\omega,X_{s-})d[X,X]_s$ yields the term $\frac{1}{2}\int_0^t f_{xx}(s-,\omega,X_{s-})d[X,X]_s^c$.

We further note that $f(s,\omega,X_s) - f(s-,\omega,X_s) = g(s,\omega,X_s)\Delta A_s$, and since $\sum_{0<s\le t} g(s,\omega,X_s)\Delta A_s$ is absolutely convergent, the theorem is proved. The assumption that $|X_s|\le k$ for $s\le t$ is removed as it was in Theorem 32 of Chap. II. □

Note that a consequence of Theorem 18 is that for f satisfying the hypotheses, we have $f(t,\cdot,\mathbf{X}_t)$ is a semimartingale when $\mathbf{X} = (X^1,\dots,X^n)$ is an n-tuple of semimartingales. Also, an important special case is when the process $A_s \equiv s$; in this case the hypotheses partly reduce to assuming f is absolutely continuous in t.

The next theorem allows an improvement of the Fisk-Stratonovich change of variables formula given in Chap. II (Theorem 34).

Theorem 19. *Let $\mathbf{X} = (X^1,\dots,X^d)$ be a vector of semimartingales and let $f : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ be such that*

(i) there exists an adapted FV process A and a function g such that

$$f(t,\omega,\mathbf{x}) = \int_0^t g(s,\omega,\mathbf{x})dA_s$$

where $(s,\omega) \mapsto g(s,\omega,\mathbf{x})$ is an adapted and jointly measurable process for each $\mathbf{x}$.

(ii) for each compact set K, $\int_0^t \sup_{\mathbf{x}\in K} |g(s,\omega,\mathbf{x})||dA_s| < \infty$ a.s.

(iii) f_{x_i} exists and is continuous in $\mathbf{x}$ and $f_{x_i}(t,\omega,\mathbf{x}) = \int_0^t g_{x_i}(s,\omega,\mathbf{x})dA_s$.

Let

$$Y_t = f(t,\omega,X^1,\dots,X^d).$$

Then Y is an adapted process of finite quadratic variation, and moreover

$$[Y,Y]_t = \sum_{1\le i,j\le d} \int_0^t \frac{\partial f}{\partial x_i}(s-,\omega,\mathbf{X}_{s-})\frac{\partial f}{\partial x_j}(s-,\omega,\mathbf{X}_{s-})d[X^i,X^j]_s^c + \sum_{0\le s\le t} \Delta Y_s^2.$$

Proof. First assume that for fixed (t,ω) the function f and all its first partials are bounded functions of $\mathbf{x}$. Then by optional stopping at times $T-$ we can assume without loss of generality that f and all its first partials (in $\mathbf{x}$) are bounded functions. Let f_k be a sequence of functions on $\mathbb{R}_+ \times \Omega \times \mathbb{R}^d$ which are C^2 on $\mathbb{R}^d$ such that f_k and its first partials converge uniformly on $\mathbb{R}^d$ respectively to f and its corresponding first partials. Moreover we can take all the f_k bounded and Lipschitz continuous with Lipschitz constant c, independent of k. For simplicity take $d = 1$. Let σ_n be a sequence of random partitions tending to the identity. We write, for a process Z,

$$S_{\sigma_n}(Z) = \sum_i (Z^{T_{i+1}^n} - Z^{T_i^n})^2$$

where

$$\sigma_n = \{0 = T_0^n \le T_1^n \le \cdots \le T_{k_n}^n\}.$$

Since f_k and their first partials converge to f and its first partials uniformly, we have $f_k - f_\ell$ is Lipschitz with constant $\varepsilon_{k\ell}$, which tends to 0 as k,ℓ tend to ∞. Therefore, letting τ denote σ_n for a given n and writing $f_k(X)$ for $f_k(t,\omega,X_t)$, we have

$$\begin{aligned}
&|S_\tau(f_k(X)) - S_\tau(f_\ell(X))| \\
&\qquad \le S_\tau((f_k - f_\ell)(X)) + 2\{S_\tau(f_k(X))S_\tau((f_k - f_\ell)(X))\}^{1/2} \qquad (*) \\
&\qquad \le (\varepsilon_{k\ell}^2 + 2c\varepsilon_{k\ell})S_\tau(X)
\end{aligned}$$

by the Lipschitz properties. Since $f_k(X)$ and $f_\ell(X)$ are semimartingales, we know (by Theorem 22 of Chap. II) that $S_{\sigma_n}(f_k(X))$ and $S_{\sigma_n}(f_\ell(X))$ converge in *ucp* respectively to $[f_k(X), f_k(X)]$ and $[f_\ell(X), f_\ell(X)]$. By restricting our attention to an interval $(s,t]$, we then have from $(*)$ that

$$\begin{aligned}
&|\{[f_k(X), f_k(X)]_t - [f_k(X), f_k(X)]_s\} - \{[f_\ell(X), f_\ell(X)]_t - [f_\ell(X), f_\ell(X)]_s\}| \\
&\qquad\qquad \le (\varepsilon_{k\ell}^2 + 2c\varepsilon_{k\ell})\{[X,X]_t - [X,X]_s\}.
\end{aligned}$$

Since this is true for all $0 < s < t < \infty$, we deduce that

$$\int_0^t |d[f_k(X), f_k(X)]_s - d[f_\ell(X), f_\ell(X)]_s| \le (\varepsilon_{k\ell}^2 + 2c\varepsilon_{k\ell})[X,X]_t.$$

Therefore $d[f_k(X), f_k(X)]$ is a Cauchy sequence of random measures, converging in total variation norm on $(0,\infty)$, a.s. In addition, by Itô's formula (Theorem 18) and Theorem 29 of Chap. II we have that

$$[f_k(X), f_k(X)]_t^c = \int_0^t \frac{\partial f_k}{\partial x}(s-, \omega, X_{s-})^2 d[X,X]_s^c;$$

from Theorem 23 of Chap. II we also know that

$$(\Delta f_k(X))^2 = \Delta[f_k(X), f_k(X)].$$

Therefore

$$[f_k(X), f_k(X)]_t = \int_0^t \frac{\partial f_k}{\partial x}(s-, \omega, X_{s-})^2 d[X,X]_s^c + \sum_{0 \le s \le t} (\Delta f_k(X_s))^2,$$

and since we have a.s. convergence in total variation norm, we can pass to the limit to obtain

$$V_t = \int_0^t f'(X_{s-})^2 d[X,X]_s^c + \sum_{0 \le s \le t} (\Delta f(X_s))^2.$$

Our identification of the limit removes the dependence on the representation of Y by f, because

$$\begin{aligned}
|S_{\sigma_n}(f(X))_t - V_t| &\le |S_{\sigma_n}(f(X))_t - S_{\sigma_n}(f_k(X))_t| + |V_t - [f_k(X), f_k(X)]_t| \\
&\quad + |S_{\sigma_n}(f_k(X))_t - [f_k(X), f_k(X)]_t|,
\end{aligned}$$

and by taking k large enough the first two terms on the right side can be taken small in probability independently of n; one then takes n large enough in the third term to make it small in probability, and we have $\lim_{n\to\infty} S_{\sigma_n}(f(X))_t = V_t$, in probability. Therefore $V_t = [Y,Y]_t$ and this completes the proof for f and its first partials bounded functions of $\mathbf{x}$, and $d=1$. The proof for general d is exactly analogous. For general f satisfying the hypotheses, let $g_m : \mathbb{R}^d \to \mathbb{R}^d$ be $\mathcal{C}^\infty$ such that $g_m(\mathbf{x}) = \mathbf{x}$ if $|\mathbf{x}| \le m$, and g_m has compact support. Defining $Y_t^m = f(t,\omega,g_m(X_t))$, let $T^m(\omega) = \inf\{t \ge 0 : Y_t(\omega) \ne Y_t^m(\omega)\}$. Then T^m increase to ∞ a.s. Also, since the quadratic variation is a path property, $[Y^m,Y^m]_t(\omega) = [Y,Y]_t(\omega)$, for all $t < T^m(\omega)$. Therefore Y is of finite quadratic variation, and since by the preceding (for $d=1$)

$$[Y^m,Y^m]_t$$
$$= \int_0^t \frac{\partial f}{\partial x}(s-,\omega,g_m(X_{s-}))g_m'(X_{s-})d[X,X]_s^c + \sum_{0<s\le t} \Delta f(s,\omega,g_m(X_s))^2$$

which, on $[0,T_m)$, is equal to

$$= \int_0^t \frac{\partial f}{\partial x}(s-,\omega,X_{s-})d[X,X]_s^c + \sum_{0<s\le t} \Delta f(s,X_s)^2;$$

whence the result. □

Note that the subspace of processes Y that have a representation $Y = f(X^1,\dots,X^d)$, where $\mathbf{X} = (X^1,\dots,X^d)$ is a finite-dimensional semimartingale (d is not fixed) and $f \in \mathcal{C}^1$, is a vector subspace of the space of processes of finite quadratic variation. As an immediate consequence of Theorem 19 we have the following extension of Theorem 34 of Chap. II.

Theorem 20. *Let X be a semimartingale and let f be $\mathcal{C}^2$. Then*

$$f(X_t)-f(X_0) = \int_{0+}^t f'(X_{s-})\circ dX_s + \sum_{0<s\le t}\{f(X_s)-f(X_{s-})-f'(X_{s-})\Delta X_s\}.$$

Proof. Note that f' is $\mathcal{C}^1$, so that $f'(X)$ is in the domain of definition of the F-S integral by Theorem 19. Also by definition we have

$$\int_{0+}^t f'(X_{s-})\circ dX_s = \int_{0+}^t f'(X_{s-})dX_s + \frac{1}{2}[f'(X),X]_t^c.$$

By Theorem 19 and polarization we know that

$$[f'(X),X]_t^c = \int_0^t f''(X_{s-})d[X,X]_s^c,$$

and thus the result follows by Itô's formula (Theorem 32, Chap. II). □

Theorem 20 has an obvious generalization to the multi-dimensional case. We omit the proof which is an analogous consequence of Theorem 33 of Chap. II.

Theorem 21. *Let $\mathbf{X} = (X^1, \dots, X^n)$ be an n-tuple of semimartingales, and let $f : \mathbb{R}^n \to \mathbb{R}$ have second order continuous partial derivatives. Then $f(\mathbf{X})$ is a semimartingale and the following formula holds:*

$$f(\mathbf{X}_t) - f(\mathbf{X}_0)$$
$$= \sum_{i=1}^{n} \int_{0+}^{t} \frac{\partial f}{\partial x_i}(\mathbf{X}_{s-}) \circ dX_s^i + \sum_{0<s\leq t} \{f(\mathbf{X}_s) - f(\mathbf{X}_{s-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(\mathbf{X}_{s-})\Delta X_s^i\}.$$

As a corollary of Theorem 21, we have the *Stratonovich integration by parts formula*.

Corollary 1 (Stratonovich Integration by Parts Theorem). Let X and Y be semimartingales. Then

$$X_t Y_t - X_0 Y_0 = \int_{0+}^{t} X_{s-} \circ dY_s + \int_{0+}^{t} Y_{s-} \circ dX_s + \sum_{0<s\leq t} \Delta X_s \Delta Y_s.$$

Proof. Let $f(x, y) = xy$ and apply Theorem 21. □

Corollary 2. Let X and Y be semimartingales, with at least one of X or Y continuous. Then

$$X_t Y_t - X_0 Y_0 = \int_{0+}^{t} X_{s-} \circ dY_s + \int_{0+}^{t} Y_{s-} \circ dX_s.$$

Recall that since $X_{0-} = 0$ by convention for a càdlàg process X, we did not really need to write $\int_{0+}^{t}$; the formula also holds for $\int_0^t$.

For stochastic differential equations with Fisk-Stratonovich differentials we are limited as to the coefficients we can consider, because the integrands must be of finite quadratic variation. Nevertheless we can still obtain reasonably general results. We first describe the coefficients.

Definition. A function $f : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R}$ is said to be **Fisk-Stratonovich acceptable** if

(i) there exists an adapted FV process A and a function g such that

$$f(t, \omega, \mathbf{x}) = \int_0^t g(s, \omega, \mathbf{x}) dA_s$$

where $(s, \omega) \mapsto g(s, \omega, \mathbf{x})$ is an adapted, jointly measurable process for each $\mathbf{x}$;

(ii) for each compact set K, $\int_0^t \sup_{x \in K} |g(s, \omega, \mathbf{x})||dA_s| < \infty$ a.s.;

(iii) f_{x_i} is C^1 and

$$f_{x_i}(t, \omega, \mathbf{x}) = \int_0^t g_{x_i}(s, \omega, \mathbf{x}) dA_s;$$

(iv) for each fixed (t,ω), the functions $\mathbf{x} \mapsto f(t,\omega,\mathbf{x})$ and $\mathbf{x} \mapsto (\frac{\partial f}{\partial x^i} \cdot f)(t,\omega,\mathbf{x})$ are all Lipschitz continuous with Lipschitz constant $K(\omega)$, $K < \infty$ a.s. $(1 \leq i \leq d)$.[4]

We will often write "F-S acceptable" in place of "Fisk-Stratonovich acceptable."

Theorem 22. *Given a vector of semimartingales $Z = (Z^1, \ldots, Z^k)$, semimartingales J^i $(1 \leq i \leq d)$, and F-S acceptable functions f_j^i $(1 \leq i \leq d, 1 \leq j \leq k)$, then the system of equations*

$$X_t^i = J_t^i + \sum_{j=1}^{k} \int_0^t f_j^i(s-,\omega,\mathbf{X}_{s-}) \circ dZ_s^j \tag{$*$}$$

has a unique semimartingale solution. Moreover the solution $\mathbf{X}$ of $(*)$ is also the (unique) solution of

$$\begin{aligned} X_t^i = J_t^i &+ \sum_{j=1}^{k} \int_0^t f_j^i(s-,\omega,\mathbf{X}_{s-}) dZ_s^j \\ &+ \frac{1}{2} \sum_{j,n=1}^{k} \sum_{m=1}^{d} \int_0^t \Big(\frac{\partial f_j^i}{\partial x^m} \cdot f_n^m\Big)(s-,\omega,\mathbf{X}_{s-}) d[Z^n, Z^j]_s^c \\ &+ \frac{1}{2} \sum_{m=1}^{d} \int_0^t \frac{\partial f_j^i}{\partial x^m}(s-,\omega,X_{s-}) d[J^m, Z^j]_s^c. \end{aligned} \tag{$**$}$$

Proof. We note that equation $(**)$ has a unique solution as a trivial consequence of Theorem 7. Since $\mathbf{X}$ is a d-dimensional semimartingale, we know that $f_j^i(s-,\cdot,\mathbf{X}_{s-})$ is in the domain of definition of the F-S integral by Theorem 19. Further, as a consequence of Theorem 19, we have that[5]

$$\begin{aligned} [f_j^i(\cdot,\omega,\mathbf{X}_\cdot), Z_\cdot^j]_t^c &= \sum_{m=1}^{d} \int_0^t \frac{\partial f_j^i}{\partial x^m}(s,\omega,\mathbf{X}_s)(d[X^m, Z^j]_s^c + d[J^m, Z^j]_s^c) \\ &= \sum_{n=1}^{k} \sum_{m=1}^{d} \int_0^t \frac{\partial f_j^i}{\partial x^m}(s,\omega,\mathbf{X}_s) f_n^m(s,\omega,\mathbf{X}_s) d[Z^n, Z^j]_s^c, \end{aligned}$$

and the equivalence of $(*)$ and $(**)$ follows. Therefore the existence of a unique semimartingale solution of $(**)$ is equivalent to the existence of a unique semimartingale solution of $(*)$, and we are done. Note that if J^m is of finite variation, the terms involving $[J^m, Z^j]^c$ disappear. □

[4] By $(\frac{\partial f}{\partial x^i} \cdot f)(t,\omega,\mathbf{x})$ we mean the usual product of functions $\frac{\partial f}{\partial x^i}(t,\omega,\mathbf{x}) \cdot f(t,\omega,\mathbf{x})$.

[5] Since we are taking the continuous parts of the quadratic variations, we need not write $f(s-,\omega,X_{s-})$, etc.

In Chap. II we studied the stochastic exponential of a semimartingale. The F-S integral allows us to give a version that has a more natural appearance.

Theorem 23. *Let Z be a semimartingale, $Z_0 = 0$. The unique solution of the equation*

$$X_t = X_0 + \int_0^t X_{s-} \circ dZ_s$$

is given by

$$X_t = X_0 \exp\{Z_t\} \prod_{0<s\leq t} (1+\Delta Z_s)e^{-\Delta Z_s},$$

and it is called the Fisk-Stratonovich exponential, $X_0 \mathcal{E}_{F\text{-}S}(Z)$.

Proof. By Theorem 22 the equation above is equivalent to

$$\begin{aligned} X_t &= X_0 + \int_0^t X_{s-} dZ_s + \frac{1}{2}\int_0^t X_{s-} d[Z,Z]_s^c \\ &= X_0 + \int_0^t X_{s-} d(Z_s + \frac{1}{2}[Z,Z]_s^c). \end{aligned}$$

Therefore

$$\begin{aligned} \mathcal{E}_{\text{F-S}}(Z) &= \mathcal{E}(Z + \frac{1}{2}[Z,Z]^c) \\ &= \mathcal{E}(Z)\mathcal{E}(\frac{1}{2}[Z,Z]^c) \\ &= \mathcal{E}(Z)\exp\{\frac{1}{2}[Z,Z]^c\}, \end{aligned}$$

where the second equality is by Theorem 38 of Chap. II. Using the formula for the stochastic exponential established in Theorem 37 of Chap. II, the result follows. □

The most interesting case is when the semimartingale is continuous.

Corollary 1. Let Z be a continuous semimartingale, $Z_0 = 0$. Then the unique solution, $\mathcal{E}_{\text{F-S}}(Z)$, of the exponential equation

$$X_t = 1 + \int_0^t X_s \circ dZ_s$$

is given by $\mathcal{E}_{\text{F-S}}(Z)_t = \exp\{Z_t\}$.

Corollary 2. Let B be a Brownian motion with $B_0 = 0$. Then the unique solution of

$$X_t = X_0 + \int_0^t X_s \circ dB_s$$

is given by $X_t = X_0 \exp\{B_t\}$.

The simplicity gained by using the F-S integral can be surprisingly helpful. As an example let $\mathbf{Z} = (Z^1, \ldots, Z^n)$ be an n-dimensional *continuous* semimartingale and let $\mathbf{x} = (x^1, \ldots, x^n)'$ be a column vector in $\mathbb{R}^n$. (Here $'$ denotes transpose.) Let I be the identity $n \times n$ matrix and define

$$a(\mathbf{x}) = I - \frac{\mathbf{x}\mathbf{x}'}{|\mathbf{x}|^2}.$$

(The matrix $a(\mathbf{x})$ represents projection onto the hyperplane normal to $\mathbf{x}$.) Note that $\mathbf{x}'a(\mathbf{x}) = \mathbf{0}$. We want to study the system of F-S differential equations

$$\mathbf{X}_t = \mathbf{x}_0 + \int_0^t a(\mathbf{X}_s) \circ d\mathbf{Z}_s, \qquad 1 \le i \le n. \qquad (***)$$

where $\mathbf{X} = (X^1, \ldots, X^n)'$ and $\mathbf{Z} = (Z^1, \ldots, Z^n)'$.

Theorem 24. *The solution* $\mathbf{X}$ *of equation* $(***)$ *above exists, is unique, and it always stays on the sphere of center* $\mathbf{0}$ *and radius* $\|\mathbf{x}_0\|$.

Proof. Since a is singular at the origin, we need to modify it slightly. Let $g(\mathbf{x})$ be a $\mathcal{C}^\infty$ function equal to $a(\mathbf{x})$ outside of a ball centered at the origin, N_0, such that $\mathbf{x}_0 \notin N_0$. Let $\mathbf{Y}$ be the solution of

$$\mathbf{Y}_t = \mathbf{x}_0 + \int_0^t g(\mathbf{Y}_s) \circ dZ_s.$$

Then $\mathbf{Y}$ exists and is unique by Theorem 22. Let $T = \inf\{t > 0 : Y_t \in N_0\}$. Since $\mathbf{Y}_0 = \mathbf{x}_0$ and $\mathbf{Y}$ is continuous, $P(T > 0) = 1$. However for $s < T$, $g(\mathbf{Y}_s) = a(\mathbf{Y}_s)$. Therefore it suffices to show that the function $t \mapsto \|\mathbf{Y}_t\|$ is constant, since $\|\mathbf{Y}_0\| = \|\mathbf{x}_0\|$. This would imply that $\mathbf{Y}$ always stays on the sphere of center 0 and radius $\|\mathbf{x}_0\|$, and hence $P(T = \infty) = 1$ and $g(\mathbf{Y}) = a(\mathbf{Y})$ always. To this end, let $f(\mathbf{x}) = \sum_{i=1}^n (x^i)^2$, where $\mathbf{x} = (x^1, \ldots, x^n)'$. Then it suffices to show that $df(\mathbf{Y}_t) = 0$, $(t \ge 0)$. Note that f is C^2. We have that for $t < T$,

$$df(\mathbf{Y}_t) = 2\mathbf{Y}_t' \circ d\mathbf{Y} = 2\mathbf{Y}_t' \cdot a(\mathbf{Y}_t) \circ d\mathbf{Z}_t = \mathbf{0}.$$

Therefore $\|Y_t\| = \|x_0\|$ for $t < T$ and thus by continuity $T = \infty$ a.s. □

Corollary. Let $\mathbf{B} = (B^1, \ldots, B^n)'$ be n-dimensional Brownian motion, let $a(\mathbf{x}) = I - \frac{\mathbf{x}\mathbf{x}'}{|\mathbf{x}|^2}$, and let $\mathbf{X}$ be the solution of

$$\mathbf{X}_t = \mathbf{x}_0 + \int_0^t a(\mathbf{X}_s) \circ d\mathbf{B}_s.$$

Then $\mathbf{X}$ is a Brownian motion on the sphere of radius $\|\mathbf{x}_0\|$.

Proof. In Theorem 36 of Sect. 6 we show that the solution $\mathbf{X}$ is a diffusion. By Theorem 24 we know that $\mathbf{X}$ always stays on the sphere of center $\mathbf{0}$ and radius $\|\mathbf{x}_0\|$. It thus remains to show only that $\mathbf{X}$ is a *rotation invariant* diffusion, since this characterizes Brownian motion on a sphere. Let U be an orthogonal matrix. Then $U\mathbf{B}$ is again a Brownian motion, and thus it suffices to show

$$d(U\mathbf{X}) = a(U\mathbf{X}) \circ d(U\mathbf{B}).$$

The above equation shows that $U\mathbf{X}$ is statistically the same diffusion as is $\mathbf{X}$, and hence $\mathbf{X}$ is rotation invariant. The coefficient a satisfies

$$a(U\mathbf{x}) = Ua(\mathbf{x})U',$$

and therefore

$$\begin{aligned} d(U\mathbf{X}) = U \circ d\mathbf{X} = Ua(\mathbf{X}) \circ d\mathbf{B} &= Ua(\mathbf{X})U'U \circ d\mathbf{B} \\ &= a(U\mathbf{X}) \circ d(U\mathbf{B}), \end{aligned}$$

and we are done. □

The F-S integral can also be used to derive explicit formulas for solutions of stochastic differential equations in terms of solutions of (non-random) ordinary differential equations. As an example, consider the equation

$$X_t = x_0 + \int_0^t f(X_s) \circ dZ_s + \int_0^t g(X_s)ds, \qquad (*4)$$

where Z is a continuous semimartingale, $Z_0 = 0$. (Note that $\int_0^t g(X_s) \circ ds = \int_0^t g(X_s)ds$, so we have not included *Itô's circle* in this term.) Assume that f is C^2 and that f, g, and ff' are all Lipschitz continuous. Let $u = u(x,z)$ be the unique solution of

$$\begin{aligned} &\frac{\partial u}{\partial z}(x,z) = f(u(x,z)) \\ &u(x,0) = x. \end{aligned}$$

Then $\frac{\partial}{\partial z}\frac{\partial u}{\partial x} = f'(u(x,z))\frac{\partial u}{\partial x}$, and $\frac{\partial u}{\partial x}(x,0) = 1$, from which we conclude that

$$\frac{\partial u}{\partial x}(x,z) = \exp\{\int_0^z f'(u(x,v))dv\}.$$

Let $Y = (Y_t)_{t\geq 0}$ be the solution of

$$Y_t = x_0 + \int_0^t \exp\{-\int_0^{Z_s} f'(u(Y_s,v))dv\} g(u(Y_s,Z_s))ds$$

which we assume exists. For example if $\dfrac{g(u(x,Z_s))}{\frac{\partial u}{\partial x}(x,Z_s)}$ is Lipschitz, this would suffice.

Theorem 25. *With the notation and hypotheses given above, the solution X of $(*4)$ is given by*

$$X_t = u(Y_t, Z_t).$$

Proof. Using the F-S calculus we have

$$\begin{aligned} u(Y_t, Z_t) &= u(x_0, 0) + \int_0^t \frac{\partial u}{\partial x}(Y_s, Z_s) \circ dY_s + \int_0^t \frac{\partial u}{\partial z}(Y_s, Z_s) \circ dZ_s \\ &= x_0 + \int_0^t \exp\{\int_0^{Z_s} f'(u(Y_s, v))dv\} \circ dY_s + \int_0^t f(u(Y_s, Z_s)) \circ dZ_s. \end{aligned}$$

Since

$$\frac{dY_s}{ds} = \exp\{-\int_0^{Z_s} f'(u(Y_s, v))dv\} g(u(Y_s, Z_s)),$$

we deduce

$$u(Y_t, Z_t) = x_0 + \int_0^t g(u(Y_s, Z_s))ds + \int_0^t f(u(Y_s, Z_s)) \circ dZ_s.$$

By the uniqueness of the solution, we conclude that $X_t = u(Y_t, Z_t)$. □

We consider the special case of a simple Stratonovich equation driven by a (one dimensional) Brownian motion B. As a corollary of Theorem 25 we obtain that the simple Stratonovich equation

$$X_t = x_0 + \int_0^t f(X_s) \circ dB_s$$

has a solution $X_t = h^{-1}(B_t + h(X_0))$, where

$$h(x) = \int_0^x \frac{1}{f(s)} ds + C.$$

The corresponding Itô stochastic differential equation is

$$dX_t = \frac{1}{2} f(X_t) f'(X_t) dt + f(X_t) dB_t.$$

By explicitly solving the analogous ordinary differential equation (without using any probability theory) and composing it, we can obtain examples of stochastic differential equations with explicit solutions. This can be useful when testing simulations and numerical solution procedures. We give a few examples.

Example. The equation

$$dX_t = -\frac{1}{2} a^2 X_t dt + a\sqrt{1 - X_t^2} dB_t$$

has solution $X_t = \sin(aB_t + \arcsin(X_0))$.

Example. The following equation has only locally Lipschitz coefficients and thus can have explosions: for $m \neq 1$,

$$dX_t = \frac{1}{2}a^2 m X_t^{2m-1} dt + aX^m dB_t$$

has solution $X_t = (X_0^{1-m} - a(m-1)B_t)^{\frac{1}{1-m}}$.

Example. Our last example can also have explosions:

$$dX_t = X_t(1+X_t^2)dt + (1+X_t^2)dB_t$$

has the solution $X_t = \tan(t + B_t + \arctan(X_0))$.

The Fisk-Stratonovich integrals also have an interpretation as limits of sums, as Theorems 26 through 29 illustrate. These theorems are then useful in turn for approximating solutions of stochastic differential equations.

Theorem 26. *Let H be càdlàg, adapted, and let X be a semimartingale. Assume $[H,X]$ exists. Let $\sigma_n = \{0 = T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n\}$ be a sequence of random partitions tending to the identity. If H and X have no jumps in common (i.e., $\sum_{0<s\leq t} \Delta H_s \Delta X_s = 0$, all $t \geq 0$), then*

$$\lim_{n\to\infty} \sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n})(X^{T_{i+1}^n} - X^{T_i^n})$$

equals the F-S integral $\int_0^\cdot H_{s-} \circ dX_s$ in ucp.

Proof. It follows easily from the definition of $[H,X]$ at the beginning of this section that $\Delta[H,X]_t = \Delta H_t \Delta X_t$ and $\lim_{n\to\infty} \sum (H_{T_{i+1}^n} - H_{T_i^n})(X^{T_{i+1}^n} - X^{T_i^n}) = [H,X] - H_0 X_0$. Thus if H and X have no jumps in common we conclude $[H,X] = [H,X]^c + H_0 X_0$. Observing that

$$\sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n})(X^{T_{i+1}^n} - X^{T_i^n})$$
$$= \sum_i H_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n}) + \frac{1}{2}\sum_i (H_{T_{i+1}^n} - H_{T_i^n})(X^{T_{i+1}^n} - X^{T_i^n}),$$

the result follows from Theorems 21 and 22 of Chap. II. □

Corollary 1. If either H or X in Theorem 21 is continuous, then

$$\lim_{n\to\infty} \sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n})(X^{T_{i+1}^n} - X^{T_i^n}) = \int_0^\cdot H_{s-} \circ dX_s,$$

where $H_{0-} = 0$, in *ucp*.

Corollary 2. Let X and Y be continuous semimartingales, and let $\sigma_n = \{0 = T_0^n \leq T_1^n \leq \cdots \leq T_{k_n}^n\}$ be a sequence of random partitions tending to the identity. Then

$$\lim_{n\to\infty} \sum_i \frac{1}{2}(Y_{T_i^n} + Y_{T_{i+1}^n})(X^{T_{i+1}^n} - X^{T_i^n}) = \int_0^{\cdot} Y_{s-} \circ dX_s,$$

with convergence in *ucp*.

Proof. By Theorem 22 of Chap. II, any semimartingale Y has finite quadratic variation. Thus Corollary 2 is a special case of Theorem 26 (and of Corollary 1). □

Theorem 27. *Let H be càdlàg, adapted, of finite quadratic variation, and suppose $\sum_{0<s\leq t} |\Delta H_s| < \infty$ a.s., each $t > 0$. Let X be a semimartingale, and let $\sigma_n = \{T_i^n\}_{0\leq i\leq k_n}$ be a sequence of random partitions tending to the identity. Assume $[H, X]$ exists. Then*

$$\lim_{n\to\infty} \{\sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n} - \sum_{T_i^n < s \leq T_{i+1}^n} \Delta H_s)(X^{T_{i+1}^n} - X^{T_i^n})\} = \int_0^{\cdot} H_{s-} \circ dX_s,$$

with convergence in ucp.

Proof. First observe that $\hat{H}_t = H_t - \sum_{0<s\leq t} \Delta H_s$ defines a continuous process of finite quadratic variation and that $[\hat{H}, X] = [H, X]^c + H_0 X_0$. Next we note that

$$\begin{aligned}\frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n} - \sum_{T_i^n < s \leq T_{i+1}^n} \Delta H_s) &= H_{T_i^n} + \frac{1}{2}(H_{T_{i+1}^n} - H_{T_i^n} - \sum_{T_i^n < s \leq T_{i+1}^n} \Delta H_s) \\ &= H_{T_i^n} + \frac{1}{2}(\hat{H}_{T_{i+1}^n} - \hat{H}_{T_i^n}).\end{aligned}$$

Therefore by Theorems 21 and 22 of Chap. II, or directly from the definition as in the proof of Theorem 26, we have

$$\begin{aligned}&\lim_{n\to\infty} \{\sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n} - \sum_{T_i^n < s \leq T_{i+1}^n} \Delta H_s)(X^{T_{i+1}^n} - X^{T_i^n})\} \\ &= \lim_{n\to\infty} \sum_i H_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n}) + \lim_{n\to\infty} \frac{1}{2} \sum_i (\hat{H}_{T_{i+1}^n} - \hat{H}_{T_i^n})(X^{T_{i+1}^n} - X^{T_i^n}) \\ &= \int_{0+}^{\cdot} H_{s-} dX_s + \frac{1}{2}([\hat{H}, X]_\cdot + H_0 X_0) \\ &= \int_0^{\cdot} H_{s-} \circ dX_s + \frac{1}{2}[H, X]^c_\cdot = \int_0^{\cdot} H_{s-} \circ dX_s, \quad \text{where } H_{0-} = H_0,\end{aligned}$$

and the result follows. □

For the general case we have a more complicated result. Let H be a càdlàg, adapted process of finite quadratic variation. For each $\varepsilon > 0$ we define

$$J_t^\varepsilon = \sum_{0<s\leq t} \Delta H_s 1_{\{|\Delta H_s|>\varepsilon\}}.$$

Then J^ε is also a càdlàg, adapted process, and it has paths of finite variation on compacts.

Theorem 28. *Let H be càdlàg, adapted, of finite quadratic variation, and let X be a semimartingale. Assume $[H,X]$ exists. Let $\sigma_n = \{T_i^n\}$ be a sequence of random partitions tending to the identity. Let J^ε be as defined above. Then*

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\{\sum_i \frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n} + (J^\varepsilon_{T_{i+1}^n} - J^\varepsilon_{T_i^n}))(X^{T_{i+1}^n} - X^{T_i^n})\}$$
$$= \int_0^\cdot H_{s-} \circ dX_s,$$

with convergence in ucp.

Proof. Let $H^\varepsilon = H - J^\varepsilon$. Then the jumps of H^ε are bounded by ε. Moreover,

$$\frac{1}{2}(H_{T_i^n} + H_{T_{i+1}^n} + (J^\varepsilon_{T_{i+1}^n} - J^\varepsilon_{T_i^n})) = H_{T_i^n} + \frac{1}{2}(H^\varepsilon_{T_{i+1}^n} - H^\varepsilon_{T_i^n}).$$

Therefore the first limit at $t > 0$ is

$$\int_{0+}^t H_{s-}dX_s + \frac{1}{2}[H,X]_t^c + \frac{1}{2}\sum_{0<s\leq t}\Delta H_s^\varepsilon \Delta X_s = \int_0^t H_{s-}\circ dX_s + \frac{1}{2}\sum_{0<s\leq t}\Delta H_s^\varepsilon \Delta X_s,$$

and letting ε tend to 0 gives the result. That is, since

$$\sum_{0<s\leq t} |\Delta H_s^\varepsilon \Delta X_s| \leq [H^\varepsilon, H^\varepsilon]_t^{1/2}[X,X]_t^{1/2} < \infty,$$

and since $[H^\varepsilon, H^\varepsilon]_t = \sum_{0<s\leq t}(\Delta H_s^\varepsilon)^2$, we have $\lim_{\varepsilon\to 0}\sum_{0<s\leq t}|\Delta H_s^\varepsilon \Delta X_s| = 0$. □

The next result is useful, for example, for approximating integrals of the form $\int_0^t f(B_s)\circ dB_s$, where B is a Brownian motion. It can be generalized to Y a continuous, adapted process of finite quadratic variation, but we do not do so here.

Theorem 29. *Let X be a semimartingale and Y a continuous semimartingale, and let f be C^1. Let μ be a probability measure on $[0,1]$ let $\alpha = \int \lambda\mu(d\lambda)$, and let $\sigma_n = \{T_i^n\}_{i\geq 0}$ be a sequence of random partitions tending to the identity. Then*

$$\lim_{n\to\infty}\sum_i \int_0^1 f(Y_{T_i^n}+\lambda(Y_{T_{i+1}^n}-Y_{T_i^n}))\mu(d\lambda)(X^{T_{i+1}^n}-X^{T_i^n})$$
$$=\int_{0+}^{\cdot} f(Y_s)dX_s+\alpha\int_{0+}^{\cdot} f'(Y_s)d[Y,X]_s,$$

with convergence in ucp. In particular if $\alpha=1/2$ then the limit is the F-S integral $\int_{0+}^{\cdot} f(Y_s)\circ dX_s$.

Proof. We begin by observing that

$$\sum_i \int_0^1 f(Y_{T_i^n}+\lambda(Y_{T_{i+1}^n}-Y_{T_i^n}))\mu(d\lambda)(X^{T_{i+1}^n}-X^{T_i^n})$$
$$=\sum_i f(Y_{T_i^n})(X^{T_{i+1}^n}-X^{T_i^n})$$
$$+\sum_i \int_0^1 \mu(d\lambda)\{f(Y_{T_i^n}+\lambda(Y_{T_{i+1}^n}-Y_{T_i^n}))-f(Y_{T_i^n})\}(X^{T_{i+1}^n}-X^{T_i^n}).$$

The first sum on the right side of the above equation tends to $\int_{0+}^{\cdot} f(Y_{s-})dX_s$ in *ucp*. Using the Fundamental Theorem of Calculus, the second sum on the right above equals

$$\sum_i \int_0^1 \mu(d\lambda)\int_0^1 ds\lambda f'(Y_{T_i^n}+\lambda s(Y_{T_{i+1}^n}-Y_{T_i^n}))(Y_{T_{i+1}^n}-Y_{T_i^n})(X^{T_{i+1}^n}-X^{T_i^n}),$$

which in turn equals

$$\alpha\sum_i f'(Y_{T_i^n})(Y_{T_{i+1}^n}-Y_{T_i^n})(X^{T_{i+1}^n}-X^{T_i^n})$$
$$+F_{\sigma_n}(\omega)\sum_i |(Y_{T_{i+1}^n}-Y_{T_i^n})(X^{T_{i+1}^n}-X^{T_i^n})|,$$

where

$$F_{\sigma_n}(\omega)\le\sup_i\{|f'(Y_{T_i^n}+\lambda s(Y_{T_{i+1}^n}-Y_{T_i^n}))-f'(Y_{T_i^n})|\}.$$

Since f' and Y are continuous, $F_{\sigma_n}(\omega)$ tends to 0 on compact time intervals. Also, on $[0,t]$,

$$\limsup_{n\to\infty}\sum_i |(Y_{T_{i+1}^n}-Y_{T_i^n})(X^{T_{i+1}^n}-X^{T_i^n})|\le[Y,Y]_t^{1/2}[X,X]_t^{1/2},$$

and the result follows by Theorem 30 of Chap. II. □

Corollary. Let X be a semimartingale, Y be a continuous semimartingale, and $f\in\mathcal{C}^1$. Let $\sigma_n=\{T_i^n\}_{0\le i\le k_n}$ be a sequence of random partitions tending to the identity. Then

$$\lim_{n\to\infty}\sum_i f(\frac{1}{2}(Y_{T_i^n}+Y_{T_{i+1}^n}))(X^{T_{i+1}^n}-X^{T_i^n})=\int_0^{\cdot} f(Y_{s-})\circ dX_s$$

with convergence in *ucp*.

Proof. Let $\mu(d\lambda)=\varepsilon_{\{1/2\}}(d\lambda)$, point mass at $1/2$. Then $\alpha=\int_0^1\lambda\mu(d\lambda)=1/2$, and we need only apply Theorem 29. □

For Brownian motion, the Fisk-Stratonovich integral is sometimes defined as a limit of the form

$$\lim\sum_i f(B_{\frac{t_i+t_{i+1}}{2}})(B_{t_{i+1}}-B_{t_i})=\int_0^t f(B_s)\circ dB_s;$$

that is, the sampling times are averaged. Such an approximation does not hold in general even for continuous semimartingales (see Yor [238, page 524]), but it does hold with a supplementary hypothesis on the quadratic covariation, as Theorem 30 reveals.

Theorem 30. *Let X be a semimartingale and Y be a continuous semimartingale. Let μ be a probability measure on $[0,1]$ and let $\alpha=\int\lambda\mu(d\lambda)$. Further suppose that $[X,Y]_t=\int_0^t J_s ds$; that is, the paths of $[X,Y]$ are absolutely continuous. Let $\sigma_n=\{t_i^n\}$ be a sequence of non-random partitions tending to the identity. Let f be C^1. Then*

$$\lim_{n\to\infty}\sum_i\int_0^1 f(Y_{t_i+\lambda(t_{i+1}-t_i)})\mu(d\lambda)(X^{t_{i+1}}-X^{t_i})$$
$$=\int_{0+}^{\cdot} f(Y_s)dX_s+\alpha\int_{0+}^{\cdot} f'(Y_s)d[Y,X]_s,$$

with convergence in ucp. In particular if $\alpha=1/2$ then the limit is the F-S integral $\int_{0+}^{\cdot} f(Y_s)\circ dX_s$.

Proof. We begin with a real analysis result. We let t_i^λ denote $t_i+\lambda(t_{i+1}-t_i)$, where the t_i are understood to be in σ_n. Suppose a is continuous on $[0,t]$. Then

$$\left|\sum_i\int_{t_i}^{t_i^\lambda} a(s)ds-\sum_i a(t_i)\lambda(t_{i+1}-t_i)\right|\le\sum_i\int_{t_i}^{t_i^\lambda}|a(s)-a(t_i)|ds$$

which tends to 0. Therefore

$$\lim_{n\to\infty}\sum_i\int_{t_i}^{t_i^\lambda} a(s)ds=\lambda\int_0^t a(s)ds. \tag{$*$}$$

Moreover since continuous functions are dense in $L^1([0,t],ds)$, the limiting result $(*)$ holds for all a in $L^1([0,t],ds)$.

Next suppose H is a continuous, adapted process, and set

$$H^{\lambda,n} = \sum_i H_{t_i} 1_{(t_i, t_i^\lambda]}.$$

Then taking limits in *ucp* we have

$$\lim_n \sum_i H_{t_i}\{[X,Y]^{t_i^\lambda} - [X,Y]^{t_i}\} = \lim_n \int_0^t H_s^{\lambda,n} d[X,Y]_s,$$

and using integration by parts, this equals

$$\lim_n \{H^{\lambda,n} \cdot (XY) - (H^{\lambda,n} X_-) \cdot Y - (H^{\lambda,n} Y) \cdot X\}.$$

By Theorem 21 of Chap. II, this equals

$$\lim_n \{\sum_i H_{t_i}(X^{t_i^\lambda} Y^{t_i^\lambda} - X^{t_i} Y^{t_i}) - \sum_i H_{t_i} X_{t_i}(Y^{t_i^\lambda} - Y^{t_i}) - \sum_i H_{t_i} Y_{t_i}(X^{t_i^\lambda} - X^{t_i})\}$$
$$= \lim_n \sum_i H_{t_i}(X^{t_i^\lambda} - X^{t_i})(Y^{t_i^\lambda} - Y^{t_i}).$$

However since $[X,Y]_t = \int_0^t J_s ds$, by the result $(*)$ we conclude

$$\begin{aligned} \lim_n \sum_i H_{t_i}(X^{t_i^\lambda} - X^{t_i})(Y^{t_i^\lambda} - Y^{t_i}) &= \lim_n \sum_i H_{t_i}\{[X,Y]^{t_i^\lambda} - [X,Y]^{t_i}\} \qquad (**) \\ &= \lim_n \sum_i H_{t_i} \int_{t_i}^{t_i^\lambda} J_s ds \\ &= \lambda \int_0^t H_s d[X,Y]_s. \end{aligned}$$

We now turn our attention to the statement of the theorem. Using the Mean Value Theorem we have

$$\sum_i \int_0^1 f(Y_{t_i^\lambda})\mu(d\lambda)(X^{t_{i+1}} - X^{t_i}) \qquad (***)$$
$$= \sum_i f(Y_{t_i})(X^{t_{i+1}} - X^{t_i}) + \sum_i \int_0^1 \mu(d\lambda) f'(Y_{t_i})(Y_{t_i^\lambda} - Y_{t_i})(X^{t_{i+1}} - X^{t_i})$$
$$+ \sum_i \int_0^1 \mu(d\lambda) \int_0^1 ds\{f'(Y_{t_i} + s(Y_{t_i^\lambda} - Y_{t_i})) - f'(Y_{t_i})\}(Y_{t_i^\lambda} - Y_{t_i})(X^{t_{i+1}} - X^{t_i}).$$

The first sum tends in *ucp* to $\int f(Y_s) dX_s$ by Theorem 21 of Chap. II. The second sum on the right side of $(***)$ can be written as

$$\sum_i \int_0^1 \mu(d\lambda) f'(Y_{t_i})(Y_{t_i^\lambda} - Y_{t_i})(X^{t_i^\lambda} - X^{t_i})$$
$$+ \sum_i \int_0^1 \mu(d\lambda) f'(Y_{t_i})(Y_{t_i^\lambda} - Y_{t_i})(X^{t_{i+1}} - X^{t_i^\lambda}).$$

The first sum above converges to

$$\int_0^1 \mu(d\lambda)\lambda \int_0^t f'(Y_s) d[Y,X]_s$$

by (**), and the second sum can be written as

$$K^{\lambda,n} \cdot X,$$

where

$$K_s^{\lambda,n} = \sum_i f'(Y_{t_i})(Y_{t_i^\lambda} - Y_{t_i}) 1_{(t_i^\lambda, t_{i+1}]}(s).$$

Then $\lim_n K^{\lambda,n} \cdot X$ converges to 0 locally in *ucp* by the Dominated Convergence Theorem (Theorem 32 of Chap. IV).

Finally consider the third sum on the right side of $(***)$. Let

$$F_n(\omega) = \sup_{t_i \in \sigma_n} \sup_{s \in [0,1]} |f'(Y_{t_i} + s(Y_{t_i^\lambda} - Y_{t_i})) - f'(Y_{t_i})|.$$

Then

$$\lim_n \sum_i \int_0^1 \mu(d\lambda) F_n(\omega) |Y_{t_i^\lambda} - Y_{t_i}||X^{t_{i+1}} - X^{t_i}|$$
$$\leq \lim_n \int_0^1 \mu(d\lambda) F_n \{\sum_i (Y_{t_i^\lambda} - Y_{t_i})^2\}^{1/2} \{\sum_i (X_{t_{i+1}} - X_{t_i})^2\}^{1/2}$$
$$= 0$$

since $\lim F_n = 0$ a.s. and the summations stay bounded in probability. This completes the proof. Since Y is continuous, $[Y,X] = [Y,X]^c$, whence if $\alpha = 1/2$ we obtain the Fisk-Stratonovich integral. □

Corollary. Let X be a semimartingale, Y a continuous semimartingale, and f be $\mathcal{C}^1$. Let $[X,Y]$ be absolutely continuous and let $\sigma_n = \{t_i^n\}$ be a sequence of non-random partitions tending to the identity. Then

$$\lim_{n\to\infty} \sum_i f(Y_{\frac{t_i + t_{i+1}}{2}})(X^{t_{i+1}} - X^{t_i}) = \int_{0+}^{\cdot} f(Y_s) \circ dX_s$$

with convergence in *ucp*.

Proof. Let $\mu(d\lambda) = \varepsilon_{\{1/2\}}(d\lambda)$, point mass at 1/2. Then $\int \lambda\mu(d\lambda) = 1/2$, and apply Theorem 30. □

Note that if Y is a continuous semimartingale and B is standard Brownian motion, then $[Y, B]$ is absolutely continuous as a consequence of the Kunita-Watanabe inequality. Therefore, if f is C^1 and σ_n are partitions of $[0, t]$, then

$$\lim_{n\to\infty} \sum f(Y_{\frac{t_i+t_{i+1}}{2}})(B_{t_{i+1}} - B_{t_i}) = \int_0^t f(Y_s) \circ dB_s,$$

with convergence in probability.

6 The Markov Nature of Solutions

One of the original motivations for the development of the stochastic integral was to study continuous strong Markov processes (that is, **diffusions**), as solutions of stochastic differential equations. Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}^n$. K. Itô studied systems of differential equations of the form

$$X_t = X_0 + \int_0^t f(s, X_s)dB_s + \int_0^t g(s, X_s)ds,$$

and under appropriate hypotheses on the coefficients f, g he showed that a unique continuous solution exists and that it is strong Markov.

Today we have semimartingale differentials, and it is therefore natural to replace dB and ds with general semimartingales and to study any resulting Markovian nature of the solution. If we insist that the solution itself be Markov then the semimartingale differentials should have independent increments (see Theorem 32); but if we need only to relate the solution to a Markov process, then more general results are available.

For convenience we recall here the Markov property of a stochastic process which we have already treated in Chap. I. Assume as given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the *usual hypotheses*.[6]

Definition. A process Z with values in $\mathbb{R}^d$ and adapted to $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is a **simple Markov process** with respect to $\mathbb{F}$ if for each $t \geq 0$ the σ-fields $\mathcal{F}_t$ and $\sigma\{Z_u; u \geq t\}$ are conditionally independent given Z_t.

Thus one can think of the Markov property as a weakening of the property of independent increments. It is easy to see that the simple Markov property is equivalent to the following. For $u \geq t$ and for every f bounded, Borel measurable,

$$E\{f(Z_u)|\mathcal{F}_t\} = E\{f(Z_u)|\sigma\{Z_t\}\}. \qquad (*)$$

One thinks of this as "the best prediction of the future given the past and the present is the same as the best prediction of the future given the present."

[6] See Chap. I, Sect. 1 for a definition of the "usual hypotheses" (page 3).

Using the equivalent relation $(*)$, one can define a **transition function** for a Markov process as follows, for $s < t$ and f bounded, Borel measurable, let

$$P_{s,t}(Z_s, f) = E\{f(Z_t)|\mathcal{F}_s\}.$$

Note that if $f(x) = 1_A(x)$, the indicator function of a set A, then the preceding equality reduces to

$$P(Z_t \in A|\mathcal{F}_s) = P_{s,t}(Z_s, 1_A).$$

Identifying 1_A with A, we often write $P_{s,t}(Z_s, A)$ on the right side above. When we speak of a Markov process without specifying the filtration of σ-algebras $(\mathcal{F}_t)_{t\geq 0}$, we mean implicitly that $\mathcal{F}_t^0 = \sigma\{Z_s; s \leq t\}$, the natural filtration generated by the process.

It often happens that the transition function satisfies the relationship

$$P_{s,t} = P_{t-s}$$

for $t \geq s$. In this case we say the Markov process is **time homogeneous**, and the transition functions are a semigroup of operators, known as the **transition semigroup** $(P_t)_{t\geq 0}$. In the time homogeneous case, the Markov property becomes

$$P(Z_{t+s} \in A|\mathcal{F}_t) = P_s(Z_t, A).$$

A stronger requirement that is often satisfied is that the Markov property hold for stopping times.

Definition. A time homogeneous simple Markov process is **strong Markov** if for any stopping time T with $P(T < \infty) = 1$, $s \geq 0$,

$$P(Z_{T+s} \in A|\mathcal{F}_T) = P_s(Z_T, A)$$

or equivalently

$$E\{f(Z_{T+s})|\mathcal{F}_T\} = P_s(Z_T, f),$$

for any bounded, Borel measurable function f.

The fact that we defined the strong Markov property only for time homogeneous processes is not much of a restriction, since if X is an $\mathbb{R}^d$-valued simple Markov process, then it is easy to see that the process $Z_t = (X_t, t)$ is an $\mathbb{R}^{d+1}$-valued time homogeneous simple Markov process.

Examples of strong Markov processes (with respect to their natural filtrations of σ-algebras) are Brownian motion, the Poisson process, and indeed any Lévy process by Theorem 32 of Chap. I. The results of this section will give many more examples as the solutions of stochastic differential equations.

Since we have defined strong Markov processes for time homogeneous processes only, it is convenient to take the coefficients of our equations to be autonomous. We could let them be non-autonomous, however, and then with an extra argument we can conclude that if X is the solution then the process $Y_t = (X_t, t)$ is strong Markov.

We recall a definition from Sect. 3 of this chapter.

Definition. A function $f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is said to be **Lipschitz** if there exists a finite constant k such that

(i) $|f(t,x) - f(t,y)| \leq k|x-y|$, each $t \in \mathbb{R}_+$, and
(ii) $t \mapsto f(t,x)$ is right continuous with left limits, each $x \in \mathbb{R}^n$.

f is said to be **autonomous** if $f(t,x) = f(x)$.

In order to allow arbitrary initial conditions, we need (in general) a larger probability space than the one on which Z is defined. We therefore define

$$\begin{aligned}\overline{\Omega} &= \mathbb{R}^n \times \Omega \\ \overline{\mathcal{F}}^0_t &= \mathcal{B} \otimes \mathcal{F}_t \\ \overline{P}^y &= \varepsilon_y \times P\end{aligned}$$

where $\mathcal{B}$ denotes the Borel sets of $\mathbb{R}^n$, and ε_y denotes the Dirac point mass measure at y. For a point $\overline{\omega} = (y, w) \in \overline{\Omega}$, we further define

$$X_0(\overline{\omega}) = y, \qquad \text{when } \overline{\omega} = (y, \omega). \tag{$*$}$$

Finally let $\overline{\mathcal{F}}_t = \bigcap_{u>t} \overline{\mathcal{F}}^0_u$. A random variable Z defined on Ω is considered to be extended automatically to $\overline{\Omega}$ by the rule $Z(\overline{\omega}) = Z(\omega)$, when $\overline{\omega} = (y, \omega)$.

We begin with a measurability result which is an easy consequence of Sect. 3 of Chap. IV.

Theorem 31. *Let Z^j be semimartingales $(1 \leq j \leq d)$, H^x a vector of adapted processes in $\mathbb{D}$ for each $x \in \mathbb{R}^n$, and suppose $(x,t,\omega) \mapsto H^x_t(\omega)$ is $\mathcal{B} \otimes \mathcal{B}_+ \otimes \mathcal{F}$ measurable.*[7] *Let F^i_j be functional Lipschitz and for each $x \in \mathbb{R}^n$, X^x is the unique solution of*

$$(X^x_t)^i = (H^x_t)^i + \sum_{j=1}^d \int_0^t F^i_j(X^x)_{s-} dZ^j_s.$$

There exists a version of X^x such that $(x,t,\omega) \mapsto X^x_t(\omega)$ is $\mathcal{B} \otimes \mathcal{B}_+ \otimes \mathcal{F}$ measurable, and for each x, X^x_t is a càdlàg solution of the equation.

Proof. Let $X^0(x,t,\omega) = H^x_t(\omega)$ and define inductively

$$X^{n+1}(x,t,\omega)^i = H^x_t + \sum_{j=1}^d \int_0^t F^i_j(X^n(x,\cdot,\cdot))_{s-} dZ^j_s.$$

The integrands above are in $\mathbb{L}$, hence by Theorem 63 in Chap. IV there exists measurable, càdlàg versions of the stochastic integrals. By Theorem 8 the processes X^n converge *ucp* to the solution X for each x. Then an application of Theorem 62 of Chap. IV yields the result. □

[7] $\mathcal{B}$ denotes the Borel sets on $\mathbb{R}^n$; $\mathcal{B}_+$ the Borel sets on $\mathbb{R}_+$.

We state and prove the next theorem for one equation. An analogous result (with a perfectly analogous proof) holds for finite systems of equations.

Theorem 32. *Let $\mathbf{Z} = (Z^1, \dots, Z^d)$ be a vector of independent Lévy processes, $\mathbf{Z}_0 = \mathbf{0}$, and let (f^i_j) $1 \le j \le d$, $1 \le i \le n$, be Lipschitz functions. Let X_0 be as in $(*)$ and let X be the solution of*

$$X^i_t = X_0 + \sum_{j=1}^{d} \int_0^t f^i_j(s-, X_{s-}) dZ^j_s. \qquad (**)$$

Then X is a Markov process, under each $\overline{P}^y$ and X is strong Markov if the f^i_j are autonomous.

Proof. We treat only the case $n = 1$. Let T be an $\mathbb{F}$ stopping time, $T < \infty$ a.s. Define $\mathcal{G}^T = \sigma\{Z^j_{T+u} - Z^j_T; u \ge 0, 1 \le j \le d\}$. Then $\mathcal{G}^T$ is independent of $\mathcal{F}_T$ under $\overline{P}^y$, since the Z^j are Lévy processes, as a consequence of Theorem 32 of Chap. I. Choose a stopping time $T < \infty$ a.s. and let it be fixed. For $u \ge 0$ define inductively

$$Y^0(x, T, u) = x,$$

$$Y^{n+1}(x, T, u) = x + \sum_{j=1}^{d} \int_T^{T+u} f_j(v-, Y^n(x, T, v-)) dZ^j_v.$$

Also, let $X(x, T, u)$ denote the unique solution of

$$X(x, T, u) = x + \sum_{j=1}^{d} \int_T^{T+u} f_j(v-, X(x, T, v-)) dZ^j_v,$$

taking the jointly measurable version (cf., Theorem 31). By Theorem 8 we know that $X(x, T, u)$ is $\mathcal{G}^T$ measurable. By approximating the stochastic integral as a limit of sums, we see by induction that $Y^n(x, T, u)$ is $\mathcal{G}^T$ measurable as well. Under $\overline{P}^x$ we have $X(X_0, T, u) = X(x, T, u)$ a.s., and $Y^n(X_0, T, u) = Y^n(x, T, u)$ $\overline{P}^x$-a.s., also. By uniqueness of solutions and using Theorem 31, for all $u \ge 0$ a.s.

$$X(X_0, 0, T + u) = X(X(X_0, 0, T), T, u).$$

There is no problem with sets of probability zero, due to (for example) the continuity of the flows. (See Theorem 37.) Writing E^x to denote expectation on $\overline{\Omega}$ with respect to $\overline{P}^x$, and using the independence of $\mathcal{F}_T$ and $\mathcal{G}^T$ (as well as of $\overline{\mathcal{F}}_T$ and $\overline{\mathcal{G}}^T$), we have for any bounded, Borel function h

$$\begin{aligned} E^x\{h(X(X_0, 0, T + u)) | \overline{\mathcal{F}}_T\} &= E\{h(X(x, 0, T + u)) | \mathcal{F}_T\} 1_{\mathbb{R}} \\ &= E\{h(X(X(x, 0, T), T, u))\} 1_{\mathbb{R}} \\ &= j(X(x, 0, T)) 1_{\mathbb{R}}, \end{aligned}$$

where $j(y) = E\{h(X(y,T,u)\}$. The last equality follows from the elementary fact that $E\{F(H,\cdot)|\mathcal{H}\} = f(H)$, where $f(h) = E\{F(h,\cdot)\}$, if F is independent of $\mathcal{H}$ and H is $\mathcal{H}$ measurable. This completes the proof, since the fact that $E^x\{h(X(X_0,0,T+u))|\overline{\mathcal{F}}_T\}$ is a function only of $X(x,0,T)$ implies that

$$E^x\{h(X(X_0,0,T+u))|\overline{\mathcal{F}}_T\} = E^x\{h(X(X_0,0,T+u))|X(X_0,0,T)\}. \quad \square$$

It is interesting to note that Theorem 32 remains true with Fisk-Stratonovich differentials. To see this we need a preliminary result.

Theorem 33. *Let* $\mathbf{Z} = (Z^1,\dots,Z^d)$ *be a vector of independent Lévy processes,* $\mathbf{Z}_0 = \mathbf{0}$. *Then* $[Z^i,Z^j]^c = 0$ *if* $i \neq j$, *and* $[Z^i,Z^i]_t^c = \alpha t$, *where* $\alpha = E\{[Z^i,Z^i]_1^c\}$.

Proof. First assume that the jumps of each Z^i are bounded. Then the moments of Z^i of all orders exist (Theorem 34 of Chap. I), and in particular $M_t^i \equiv Z_t^i - E\{Z_t^i\}$ is an L^2 martingale for each i, with $E\{Z_t^i\} = tE\{Z_1^i\}$. By independence M^iM^j is also a martingale and hence $[M^i,M^j] = 0$ by Corollary 2 of Theorem 27 of Chap. II. Therefore $[Z^i,Z^j]_t^c = [M^i,M^j]_t^c = 0$ as well.

Next consider $A_t^i \equiv [Z^i,Z^i]_t = [M^i,M^i]_t$. It is an immediate consequence of approximation by sums (Theorem 22 of Chap. II) that A^i also has independent increments. Since

$$A_t^i = [M^i,M^i]_t^c + \sum_{0<s\leq t}(\Delta M_s^i)^2,$$

and the process $J_t^i = \sum_{0<s\leq t}(\Delta M_s^i)^2$ also clearly has independent increments, we deduce that $[M^i,M^i]_t^c$ has independent increments as well. Therefore $[M^i,M^i]_t^c - E\{[M^i,M^i]_t^c\}$ is a finite variation continuous martingale, hence constant by Theorem 27 of Chap. II. Since $[M^i,M^i]_0^c = 0$, we deduce that $[Z^i,Z^i]_t^c = [M^i,M^i]_t^c = E\{[M^i,M^i]_t^c\}$. Since $[M^i,M^i]_t$ is also a Lévy process we have $E\{[M^i,M^i]_t\} = tE\{[M^i,M^i]_1\}$ (by the stationarity of the increments), and also

$$E\{\sum_{0<s\leq t}(\Delta M_s^i)^2\} = E\{(\sum_{0<s\leq t}\Delta M_s^i)^2\} = tE\{\sum_{0<s\leq 1}(\Delta M_s^i)^2\}$$

by the same reasoning. Therefore $[Z^i,Z^i]_t^c = tE\{[Z^i,Z^i]_1^c\}$ by subtraction.

If the jumps of $\mathbf{Z}$ are not bounded, let $J_t^i = \sum_{0<s\leq t}\Delta Z_s^i 1_{\{|\Delta Z_s^i|\geq 1\}}$. Then $\hat{Z}^i \equiv Z^i - J^i$ and J^i are independent Lévy processes, and J^i is a quadratic pure jump semimartingale. It therefore follows that $[Z^i,Z^j] = [\hat{Z}^i,\hat{Z}^j] = 0$ and that $[Z^i,Z^i]^c = [\hat{Z}^i,\hat{Z}^i]^c$, and the theorem is proved. $\square$

Theorem 34. *Let* $\mathbf{Z} = (Z^1,\dots,Z^d)$ *be a vector of independent Lévy processes,* $\mathbf{Z}_0 = \mathbf{0}$, *and let* (f_j^i) $1 \leq j \leq d$, $1 \leq i \leq n$, *be non-random F-S acceptable functions.*[8] *Let* X_0 *be as in Theorem 32 and* X *be the solution of*

[8] F-S acceptable functions are defined on page 278.

$$X_t^i = X_0^i + \sum_{j=1}^{d} \int_0^t f_j^i(s-, X_{s-}) \circ dZ_s^j.$$

Then X is a Markov process and X is strong Markov if the coefficients f_j^i are autonomous.

Proof. We treat only the case $n = 1$. By Theorem 33 $[Z^i, Z^j]_t^c = 0$ if $i \neq j$ and $[Z^i, Z^i]_t^c = \alpha_i t$, $\alpha_i \geq 0$. Therefore, by Theorem 22 the equation is equivalent to

$$X_t = X_0 + \sum_{j=1}^{d} \int_0^t f_j(s-, X_{s-}) dZ_s^j + \frac{1}{2} \sum_{j=1}^{d} \alpha_j \int_0^t (\frac{\partial f_j}{\partial x} \cdot f_j)(s-, X_{s-}) ds,$$

and since the process $Y_t = t$ is a Lévy process we need only to invoke Theorem 32 to complete the proof. □

If the differentials are not Lévy processes but only strong Markov processes which are semimartingales, then the solutions of equations such as (**) in Theorem 32 need not be Markov processes. One does, however, have the following result.

Theorem 35. *Let Z be a strong Markov processes with values in $\mathbb{R}$, $Z_0 = 0$, such that Z is a semimartingale. Let f and g be Lipschitz functions. Let X_0 be as in Theorem 32 and X be the solution of*

$$X_t = X_0 + \int_0^t f(s-, X_{s-}) dZ_s + \int_0^t g(s, X_{s-}) ds.$$

Then the vector process (X, Z) is Markov under $\overline{P}^y$, each $y \in \mathbb{R}$, and strong Markov if f and g are autonomous.

Proof. First recall that X is defined on $\overline{\Omega} = \mathbb{R} \times \Omega$ and Z is automatically extended to $\overline{\Omega}$ as explained at the beginning of this section. The probability $\overline{P}^y = \varepsilon_y \times P$ is such that $\overline{P}^y(X_0 = y) = 1$, each $y \in \mathbb{R}$. As in the proof of Theorem 32 for a fixed stopping time T ($T < \infty$ a.s.) and $x \in \mathbb{R}$ define inductively for $u \geq 0$

$$\begin{aligned} Y^0(x, T, u) &= x \\ Y^{n+1}(x, T, u) &= x + \int_T^{T+u} f(v-, Y^n(x, T, v-)) dZ_v \\ &\quad + \int_T^{T+u} g(v-, Y^n(x, T, v-)) dv, \end{aligned}$$

and let $X(x, T, u)$ denote the unique solution of

$$X(x, T, u) = x + \int_T^{T+u} f(v-, X(x, T, v-)) dZ_v + \int_T^{T+u} g(v-, X(x, T, v-)) dv.$$

Next define $\mathcal{G}^T = \sigma\{Z_{T+u} - Z_T; u \geq 0\}$ for the same stopping time T. As in the proof of Theorem 32 we see that $Y^n(x, T, u)$ is $\mathcal{G}^T$ measurable for each $n \geq 1$ and that $X(x, T, u)$ is $\mathcal{G}^T$ measurable as well. Next let h be Borel measurable and bounded, and let $G \in \mathcal{G}^T$ and bounded.

Observe that

$$\begin{aligned} E\{h(X_T)G|\mathcal{F}_T\} &= h(X_T)E\{G|\mathcal{F}_T\} \\ &= h(X_T)E\{G|Z_T\} \\ &= j(X_T, Z_T), \end{aligned}$$

where $j : \mathbb{R}^2 \to \mathbb{R}$ is Borel. Since $X_{T+u} = X(X_T, T, u)$ by the uniqueness of the solution, the theorem now follows from the Monotone Class Theorem (Theorem 8 of Chap. I). □

Theorem 35 can also be shown to be true with Stratonovich differentials, but the proof is more complicated since the quadratic variation process is an additive functional of Z, rather than a deterministic process (as is the case when Z is a Lévy process). For this type of result we refer the reader to Çinlar-Jacod-Protter-Sharpe [34].

Each of Theorems 32, 34 and 35 can be interpreted with X having an arbitrary initial distribution μ. Indeed for μ a probability measure on $(\mathbb{R}, \mathcal{B})$, define $\overline{P}^\mu$ on $\overline{\Omega}$ by $\overline{P}^\mu(\Lambda) = \int_{\mathbb{R}} \overline{P}^x(\Lambda)\mu(dx)$, for $\Lambda \in \mathcal{B} \otimes \mathcal{F}$. Then the conclusions of the three theorems are trivially still valid for $\overline{P}^\mu$, and the distribution of X_0 is μ.

Traditionally the most important Markovian solutions of stochastic differential equations are diffusions. Suppose as given a space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ satisfying the usual hypotheses.

Definition. An adapted process X with values in $\mathbb{R}^n$ is a **diffusion**[9] if it has continuous sample paths and if it satisfies the strong Markov property.

A restatement of Theorems 32 and 34 yields the following.

Theorem 36. *Let* $\mathbf{B} = (B^1, \ldots, B^d)$ *be a standard Brownian motion on* $\mathbb{R}^d$, $\mathbf{B}_0 = \mathbf{0}$, *and let* (f^i_j) $1 \leq j \leq d$, $1 \leq i \leq n$, g^i *be autonomous Lipschitz functions. Let* X^i_0 *be as in Theorem 32 and let* X *be the solution of*

$$X^i_t = X^i_0 + \sum_{j=1}^{d} \int_0^t f^i_j(X_s)dB^j_s + \int_0^t g^i(X_s)ds. \qquad (***)$$

Then X *is a diffusion. If the coefficients* $(f^i_j)_{1\leq j\leq d}$, $1 \leq i \leq n$, *are non-random F-S acceptable functions and if* Y *is the solution of*

$$Y^i_t = Y^i_0 + \sum_{j=1}^{d} \int_0^t f^i_j(Y_s) \circ dB^j_s + \int_0^t g^i(Y_s)ds,$$

[9] The definition of a diffusion is not yet standardized. We give a general definition.

then Y is a diffusion.

In equation $(***)$ of Theorem 36 the coefficients f^i_j are called the **diffusion coefficients** and the coefficients g^i are called the **drift coefficients**.

Note that a diffusion need not be semimartingale, even though of course the solutions of equations such as $(***)$ are semimartingales. Indeed any deterministic continuous function with paths of unbounded variation is a diffusion which is not a semimartingale. Another interesting example is provided by Tanaka's formula. The process

$$|B_t| = \int_0^t \operatorname{sign}(B_s)dB_s + L^0_t$$

is a diffusion and it is a semimartingale, where B is standard Brownian motion on $\mathbb{R}$. Since the paths of L^0_t are singular with respect to Lebesgue measure, and since a semimartingale decomposition is unique for continuous processes (the corollary of Theorem 31 of Chap. III), we see that $|B_t|$ cannot be represented as a solution of $(***)$. Another example is that of $|B_t|^{1/3}$ which is a time homogeneous diffusion but not a semimartingale (Theorem 71 of Chap. IV).

An intuitive notion of a diffusion is to imagine a pollen grain floating downstream in a river. The grain is subject to two forces: the current of the river (drift), and the aggregate bombardment of the grain by the surrounding water molecules (diffusion). The coefficient $f(t,x)$ then represents the sensitivity of the particle at time t and place x to the diffusion forces. For example, if part of the river water is warmer at certain times and places (due to sunlight or industrial effluents, for example), then f might be larger. Analogously g would be larger when the river was flowing faster due to a steeper incline.

We give three simple examples of diffusions.

Example. The stochastic exponential $\exp\{B_t - \frac{1}{2}t\}$ is a diffusion where $f(t,x) = x$ and $g(t,x) = 0$.

Example. Consider the simple system

$$V_t = V_0 + \int_0^t \sigma dB_s + \int_0^t \alpha V_s ds,$$
$$X_t = X_0 + \int_0^t V_s ds.$$

The process X can be used as a model of Brownian motion alternative to Einstein's. It is called the Ornstein-Uhlenbeck Brownian motion, or simply the **Ornstein-Uhlenbeck process**. Note that here the process X has paths of finite variation and hence the process $(V_t)_{t\geq 0}$ is a true velocity process for X. Using integration by parts we can verify that

$$V_t = e^{\alpha t}(V_0 + \int_0^t e^{-\alpha s}\sigma dB_s)$$

is an explicit solution for V. Indeed

$$\begin{aligned}e^{-\alpha t}V_t &= V_0 + \int_0^t V_s(-\alpha e^{-\alpha s})ds + \int_0^t e^{-\alpha s}dV_s \\ &= V_0 - \alpha \int_0^t V_s e^{-\alpha s}ds + \int_0^t e^{-\alpha s}\alpha V_s ds + \int_0^t e^{-\alpha s}\sigma dB_s \\ &= V_0 + \int_0^t e^{-\alpha s}\sigma dB_s,\end{aligned}$$

and we are done. Since V_0 and the Brownian motion B are independent (by construction), we see that when V_0 has a Gaussian distribution then V is a Gaussian process. If α is negative and we take $\sigma^2 = -1/(2\alpha)$, then V is a *stationary* Gaussian process.

Example. Consider next the equation

$$X_t = \int_0^t \frac{-X_s}{1-s}ds + B_t, \qquad (0 \leq t < 1)$$

for B a standard Brownian motion.

For each t_0, $0 < t_0 < 1$, there is a solution which is a diffusion on $[0, t_0]$. By the uniqueness of solutions if $t_1 < t_0$ then the solution for $[0, t_0]$ agrees with the solution for $[0, t_1]$ on the interval $[0, t_1]$. Thus we have a solution on $[0, 1)$. If we can show that $\lim_{t\to 1} X_t = 0$, then the solution extends by continuity to $[0, 1]$ and we will have constructed a diffusion X on $[0, 1]$ with $X_0 = X_1 = 0$, known as the **Brownian bridge.**[10]

An application of integration by parts shows that the solution of the Brownian bridge equation is given by

$$X_t = (1-t)\int_0^t \frac{1}{1-s}dB_s \qquad (0 \leq t < 1).$$

Write $X_t = f(t)M_t$, where $f(t) = (1-t)$ and $M_t = \int_0^t (1-s)^{-1}dB_s$. To see that $\lim_{t\to 1} X_t = 0$, we study M_t, making the change of variables $t = u(1+u)^{-1}$. Then $0 \leq t < 1$ corresponds to $0 \leq u < \infty$, and define

$$N_u = M_{\frac{u}{1+u}}, \quad \mathcal{G}_u = \mathcal{F}_{\frac{u}{1+u}}, \; 0 \leq u < \infty.$$

N is clearly a continuous $\mathcal{G}_u$ martingale, and moreover $[N, N]_u = u$; hence N is a standard Brownian motion by Lévy's Theorem (Theorem 39 of Chap. II). It is then easy to show, as a consequence of the Strong Law of Large Numbers, that[11]

$$\lim_{t\to\infty} \frac{N_t}{t} = 0 \quad \text{a.s.}$$

[10] The Brownian bridge is also known as *tied down Brownian motion*, and alternatively as *pinned Brownian motion.*

[11] See, for example, Breiman [23, page 265].

Let $g(t,\omega) = N_t(\omega)/t$ so that $\lim_{t\to 0} g(t^{-1},\omega) = 0$ a.s. Thus, $\lim_{t\to 0} tN_{1/t} = 0$ a.s. We then have, replacing t with $(1-t)$

$$\lim_{t\to 1}(1-t)N_{1/(1-t)} = \lim_{t\to 1}(1-t)M_{1/(2-t)} = 0 \quad \text{a.s.},$$

and therefore $\lim_{u\to 1}(1-u)M_u = 0$ a.s., and hence $\lim_{t\to 1} X_t = 0$ a.s.

We now know that X is a continuous diffusion on $[0,1]$, and that $X_0 = X_1 = 0$ a.s. Also X is clearly a semimartingale on $[0,1)$, but it is not obvious that X is a semimartingale on $[0,1]$. One needs to show that the integral $\int_0^1 \frac{|X_s|}{1-s}ds < \infty$ a.s. To see this calculate $E\{X_t^2\}$, $0 \le t < 1$,

$$\begin{aligned} E\{X_t^2\} = f(t)^2E\{M_t^2\} &= f(t)^2E\{[M,M]_t\} \\ &= (1-t)^2\int_0^t \frac{1}{(1-s)^2}ds \\ &= (1-t)^2\left(\frac{1}{1-t} - 1\right). \end{aligned}$$

By the Cauchy-Schwarz inequality

$$E\{|X_t|\} \le E\{X_t^2\}^{1/2} = \sqrt{t(1-t)}.$$

Therefore

$$\begin{aligned} E\{\int_0^1 \frac{|X_s|}{1-s}ds\} &= \int_0^1 \frac{E\{|X_s|\}}{1-s}ds \\ &\le \int_0^1 \frac{\sqrt{s(1-s)}}{1-s}ds < \infty, \end{aligned}$$

whence $\int_0^1 \frac{|X_s|}{1-s}ds < \infty$ a.s. Therefore the solution X is a semimartingale on $[0,1]$.

Finally we remark that one can construct a similar Brownian bridge from any value a to any value b, on an interval of arbitrary finite length τ. Using the interval $[0,\tau]$ one has the equation

$$X_t = a + \int_0^t \frac{b-X_s}{\tau - s}ds + B_t, \qquad 0 \le t < \tau,$$

with solution

$$X_t = \begin{cases} a(1-\frac{t}{\tau}) + \frac{bt}{\tau} + (\tau - t)\int_0^t \frac{1}{\tau - s}dB_s, & 0 \le t < \tau, \\ X_\tau = b. \end{cases}$$

7 Flows of Stochastic Differential Equations: Continuity and Differentiability

Consider a stochastic differential equation of the form

$$X_t = x + \int_0^t F(X)_{s-} dZ_s.$$

Obviously there is a dependence on the initial condition, and we can write the solution in the form $X(t, \omega, x)$, or $X_t^x(\omega)$. The study of the *flow of a stochastic differential equation* is the study of the functions $\phi : x \to X(t, \omega, x)$ which can be considered as mapping $\mathbb{R}^n \to \mathbb{R}^n$ for (t, ω) fixed, or as mapping $\mathbb{R}^n \to \mathcal{D}^n$, where $\mathcal{D}^n$ denotes the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}^n$, equipped with the topology of uniform convergence on compacts. It is important to distinguish between $\mathcal{D}^n$ and $\mathbb{D}^n$. The former is a *function space*, and it is associated in the literature with weak convergence results (see, e.g., Billingsley [17], or Ethier-Kurtz [71], or Kurtz-Protter [136] for recent results in a semimartingale context); the latter is the space of *stochastic processes* with càdlàg paths, and which are adapted to the underlying filtration.

Note that we have already encountered flows in Sect. 6 (cf., Theorem 31), where we proved measurability of the solution with respect to a parameter (which can be taken to be, of course, the initial condition). We will be interested in several properties of the flows: continuity, differentiability, injectivity, and when the flows are diffeomorphisms of $\mathbb{R}^n$.

We begin with continuity. We consider a general system of equations of the form

$$X_t^x = H_t^x + \int_0^t F(X^x)_{s-} dZ_s, \tag{$*$}$$

where X_t^x and H_t^x are column vectors in $\mathbb{R}^n$, Z is a column vector of m semimartingales with $\mathbf{Z}_0 = \mathbf{0}$, and F is an $n \times m$ matrix with elements (F_α^i). For x fixed, for each y we have that $\overline{X}_t = X_t^y - X_t^x$ is a solution of the equation

$$\overline{X}_t = H_t^y - H_t^x + \int_0^t \overline{F}(\overline{X})_{s-} dZ_s, \tag{$**$}$$

where $\overline{F}(Y) = F(X^x + Y) - F(X^x)$.

Theorem 37. *Let H^x be processes in $\mathbb{D}^n$, and let $x \mapsto H^x : \mathbb{R}^n \to \mathbb{D}^n$ be prelocally Lipschitz continuous in $\underline{\underline{S}}^p$, some $p > n$. Let F be an $n \times m$ matrix of functional Lipschitz operators (F_α^i), $1 \le i \le n$, $1 \le \alpha \le m$.*[12] *Then there exists a function $X(t, \omega, x)$ on $\mathbb{R}_+ \times \Omega \times \mathbb{R}^n$ such that*

(i) for each x the process $X_t^x(\omega) = X(t, \omega, x)$ is a solution of $(*)$, and
(ii) for almost all ω, the flow $x \mapsto X(\cdot, \omega, x)$ from $\mathbb{R}^n$ into $\mathcal{D}^n$ is continuous in the topology of uniform convergence on compacts.

[12] See page 250 for the definition of functional Lipschitz.

Proof. We recall the method of proof used to show the existence and uniqueness of a solution (Theorem 7). By stopping at a fixed time t_0, we can assume the Lipschitz process is just a random variable K which is finite a.s. Then by conditioning[13] we can assume without loss of generality that this Lipschitz constant is non-random, and we call it $c < \infty$. By replacing H_t^x with $X_t^x + \int_0^t F(0)_{s-} dZ_s$, and then by replacing F with G given by $G(Y)_t = F(Y)_t - F(0)_t$, we can further assume without loss of generality that $F(0) = 0$. Then for $\beta = C_p(c)$, by Theorem 5 we can find an arbitrarily large stopping time T such that $Z^{T-} \in \mathcal{S}(\beta)$, and H^x is Lipschitz continuous in $\underline{\underline{S}}^p$ on $[0, T)$. Then by Lemma 2 (preceding Theorem 7) we have that for the solution $\overline{X}$ of $(**)$

$$\|\overline{X}^{T-}\|_{\underline{\underline{S}}^p} \leq C_p(c, Z)\|(H^x - H^y)^{T-}\|_{\underline{\underline{S}}^p}, \qquad \text{for any } p \geq 2,$$

and some (finite) constant $C_p(c, Z)$. Choose $p > n$, and we have

$$E\{\sup_{s<T} |X_s^x - X_s^y|^p\} \leq C_p(c, Z)K\|x - y\|^p, \tag{$***$}$$

due to the Lipschitz hypothesis on $x \mapsto H^x$. By Kolmogorov's Lemma (Theorem 72 of Chap. IV) we have the result on $\mathbb{R}^n \times [0, T)$. However since T was arbitrarily large, the result holds as well on $\mathbb{R}^n \times \Omega \times \mathbb{R}_+$. $\square$

For the remainder of this section we will consider less general equations. Indeed, the following will be our basic hypotheses, which we will supplement as needed.

Hypothesis (H1). Z^α are given semimartingales with $Z_0^\alpha = 0$, $1 \leq \alpha \leq m$.

Hypothesis (H2). $f_\alpha^i : \mathbb{R}^n \to \mathbb{R}$ are given functions, $1 \leq i \leq n$, $1 \leq \alpha \leq m$, and $f(x)$ denotes the $n \times m$ matrix $(f_\alpha^i(x))$.

We will study the system of equations

$$X_t^i = x^i + \sum_{\alpha=1}^m \int_0^t f_\alpha^i(X_{s-})dZ_s^\alpha, \quad 1 \leq i \leq n \tag{$\otimes$}$$

which we also write

$$X_t = x + \int_0^t f(X_{s-})dZ_s \tag{$\otimes\otimes$}$$

where it is understood that X_t and x are column vectors in $\mathbb{R}^n$, $f(X_{s-})$ is an $n \times m$ matrix, and Z is a column vector of m semimartingales.

To study the differentiability of the flow we will need a more general theorem on existence and uniqueness of solutions. Indeed, we wish to replace our customary (global) Lipschitz condition with a local Lipschitz condition.

[13] See the proofs of Theorems 7, 8, or 15 for this argument.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **locally Lipschitz** if there exists an increasing sequence of open sets Λ_k such that $\bigcup_k \Lambda_k = \mathbb{R}^n$ and f is Lipschitz with a constant K_k on each Λ_k.

For example, if f has continuous first partial derivatives, then it is locally Lipschitz, while if its continuous first partials are *bounded*, then it is Lipschitz. If f, g, are both Lipschitz, then their product fg is locally Lipschitz. These coefficients arise naturally in the study of Fisk-Stratonovich equations (see Sect. 5).

Theorem 38. *Let Z be as in (H1) and let the functions (f^i_α) in (H2) be locally Lipschitz. Then there exists a function $\zeta(x,\omega) : \mathbb{R}^n \times \Omega \to [0,\infty]$ such that for each x $\zeta(x,\cdot)$ is a stopping time, and there exists a unique solution of*

$$X_t = x + \int_0^t f(X_{s-})dZ_s \qquad (\otimes\otimes)$$

up to $\zeta(x,\cdot)$ with $\limsup_{t\to\zeta(x,\cdot)} \|X_t\| = \infty$ a.s. on $\{\zeta < \infty\}$. Moreover $x \mapsto \zeta(x,\omega)$ is lower semi-continuous, strictly positive, and the flow of X is continuous on $[0,\zeta(x,\cdot))$.

Remark. Before proving the theorem we comment that for each x fixed the stopping time $T(\omega) = \zeta(x,\omega)$ is called an **explosion time**. Thus Theorem 38 assures the existence and uniqueness of a solution *up to an explosion time*; and at that time, the solution does indeed explode in the sense that $\limsup_{t\to T} \|X_t\| = +\infty$ on $\{T < \infty\}$. Note however that if the coefficients (f^i_α) in (H2) are (globally) Lipschitz, then a.s. $\zeta = \infty$ for all x (Theorem 7).

Proof. Let Λ_ℓ be open sets increasing to $\mathbb{R}^n$ such that there exist (h_ℓ), a sequence of C^∞ functions with compact support mapping $\mathbb{R}^n$ to $[0,1]$ such that $h_\ell = 1$ on Λ_ℓ. For each ℓ we let $X_\ell(t,\omega,x)$ denote the solution (continuous in x) of the equation

$$X_t = x + \int_0^t g_\ell(X_{s-})dZ_s$$

where the matrix $g_\ell(x)$ is defined by $h_\ell(x)f(x)$.

Define stopping times, for x fixed,

$$S_\ell(\omega,x) = \inf\{t > 0 : X_\ell(t,\omega,x) \notin \Lambda_\ell \text{ or } X_\ell(t-,\omega,x) \notin \Lambda_\ell\}.$$

By Theorem 37 the flow is uniformly continuous on compact sets, hence the functions S_ℓ are lower semi-continuous. Then on $[0,S_\ell)$ we have that $X_\ell = X_{\ell+1}$ by the uniqueness of the solutions, since they both satisfy equation $(\otimes\otimes)$. We wish to show that the relation

$$X_\ell(\cdot,\omega,x) = X_{\ell+1}(\cdot,\omega,x) \quad \text{on } [0,S_\ell(\omega,x)) \qquad (L)$$

holds for all $x \in \mathbb{R}^n$ simultaneously. Choose an x and let

$$A^x = \{\omega : X_\ell(\cdot,\omega,x) = X_{\ell+1}(\cdot,\omega,x) \text{ on } [0,S_\ell(\omega,x))\}.$$

Then $P(A^x) = 1$. We set $A = \bigcup_{x\in\mathbb{Q}^n} A^x$, where $\mathbb{Q}^n$ denotes the rationals in $\mathbb{R}^n$. Then $P(A) = 1$ as well, and without loss of generality we take $A = \Omega$. Next for arbitrary $y \in \mathbb{R}^n$, let $y_n \to y$, with $y_n \in \mathbb{Q}^n$. Then

$$S_\ell(\omega,y) \le \liminf_{n\to\infty} S_\ell(\omega,y_n),$$

and therefore the relation (L) holds for y as well, by the previously established continuity.

Observe that $S_\ell(\omega,x) \le S_{\ell+1}(\omega,x)$, and let $\zeta(x,\omega) = \sup_\ell S_\ell(\omega,x)$, for each x. Then ζ is lower semi-continuous because the S_ℓ are, and it is strictly positive. Further, we have shown that there exists a unique function $X(\cdot,\omega,x)$ on $[0,\zeta(x,\omega))$ which is a solution of $(\otimes\otimes)$, and it is equal to $X_\ell(\cdot,\omega,x)$ on $[0,S_\ell(\omega,x))$, each $\ell \ge 1$. Indeed, on $[0,S_\ell(\omega,x))$ we have $X(\cdot,\omega,x) = X_\ell(\cdot,\omega,x)$, and since $X_\ell(\cdot,\omega,x) \in \Lambda_\ell$, we have $h_\ell(X_\ell(\cdot,\omega,x)) = 1$, wherefore X_ℓ is a solution of $(\otimes\otimes)$ on $[0,S_\ell)$.

By our construction we have that $X(S_\ell(\omega,x),\omega,x)$ belongs to Λ_ℓ^c. Indeed, letting $S_\ell(\omega,x) = s$, then $X(s-,\omega,x) = X_\ell(s-,\omega,x) = u \in \overline{\Lambda_\ell}$, for some value u. Since $u \in \overline{\Lambda_\ell}$ we have $h_\ell(u) = 1$, and thus X and X_ℓ have the same jump at s and we can conclude $X(s,\omega,x) = X_\ell(s,\omega,x)$. But $X_\ell(s,\omega,x)$ or $X_\ell(s-,\omega,x)$ must be in Λ_ℓ^c by right continuity and the definition of S_ℓ; therefore $X(s,\omega,x)$ or $X(s-,\omega,x)$ is in Λ_ℓ^c. This shows that $\limsup_{t\to\zeta}\|X_t\| = \infty$ on $\{\zeta < \infty\}$. □

It is tempting to conclude that if there are no explosions at a finite time for each initial condition x a.s. (null set depending on x), then also $\zeta(x,\cdot) = \infty$ a.s. (with the same null set for all initial values x). Unfortunately this is not true as the next example shows.

Example. Let B be a complex Brownian motion. That is, let B^1 and B^2 be two independent Brownian motions on $\mathbb{R}$ and let $B_t \equiv B_t^1 + iB_t^2$, where $i^2 = -1$. Consider the stochastic differential equation

$$Z_t = z - \int_0^t Z_s^2 dB_s,$$

where the initial value z is complex. This equation has a closed form solution given by

$$Z(t,\omega,z) = \frac{z}{1 + zB_t(\omega)}.$$

Indeed, if we set $f(x) = z(1+zx)^{-1}$ and $Z_t = f(B_t)$, then $f'(B_t) = -Z_t^2$, and since $B_t = B_t^1 + iB_t^2$, we have by Itô's formula

$$Z_t = z - \int_0^t Z_s^2 dB_s + \frac{1}{2}\int_0^t f''(B_s)(d[B^1,B^1]_s + i^2 d[B^2,B^2]_s);$$

since $d[B^1,B^1]_s = d[B^2,B^2]_s = ds$, we see that $Z_t = z - \int_0^t Z_s^2 dB_s$.

For z fixed we know that $P(\exists\, t : B_t = -1/z) = 0$, since Brownian motion in $\mathbb{R}^2$ a.s. does not hit a specified point. Therefore Z does not have explosions in a finite time for each fixed initial condition $Z_0 = z$. On the other hand for *any* ω_0 and t_0 fixed we have $B_{t_0}(\omega_0) = z_0$ for some $z_0 \in \mathbb{C}$, and thus for the initial value $z = -1/z_0$ we have an explosion at the chosen finite time t_0. Thus each trajectory has initial conditions such that it will explode at a finite time, and $P\{\omega : \exists\, z \text{ with } \zeta(\omega, z) < \infty\} = 1$.

We next turn our attention to the differentiability of the flows. To this end we consider the system of $n+n^2$ equations (assuming that the coefficients (f^i_α) are at least C^1)

$$\begin{aligned} X^i_t &= x_i + \sum_{\alpha=1}^m \int_0^t f^i_\alpha(X_{s-})dZ^\alpha_s \\ D^i_{kt} &= \delta^i_k + \sum_{\alpha=1}^m \sum_{j=1}^n \int_0^t \frac{\partial f^i_\alpha}{\partial x_j}(X_{s-})D^j_{ks-}dZ^\alpha_s, \end{aligned} \tag{D}$$

$(1 \le i \le n)$ where D denotes an $n \times n$ matrix-valued process and $\delta^i_k = 1$ if $i = k$ and 0 otherwise (Kronecker's delta). A convenient convention, sometimes called the *Einstein convention*, is to leave the summations implicit. Thus the system of equations (D) can be alternatively written as

$$\begin{aligned} X^i_t &= x_i + \int_0^t f^i_\alpha(X_{s-})dZ^\alpha_s \\ D^i_{kt} &= \delta^i_k + \int_0^t \frac{\partial f^i_\alpha}{\partial x_j}(X_{s-})D^j_{ks-}dZ^\alpha_s. \end{aligned} \tag{D}$$

We will use the Einstein convention when there is no ambiguity. Note that in equations (D) if X is already known, then the second system is *linear* in D. Also note that the coefficients for the system (D) are not *globally* Lipschitz, but if the first partials of the (f^i_α) are locally Lipschitz, then so also are the coefficients of (D).

Theorem 39. *Let Z be as in (H1) and let the functions (f^i_α) in (H2) have locally Lipschitz first partial derivatives. Then for almost all ω there exists a function $X(t,\omega,x)$ which is continuously differentiable in the open set $\{x : \zeta(x,\omega) > t\}$, where ζ is the explosion time (cf., Theorem 38). If (f^i_α) are globally Lipschitz then $\zeta = \infty$. Let $D_k(t,\omega,x) \equiv \frac{\partial}{\partial x_k}X(t,\omega,x)$. Then for each x the process $(X(\cdot,\omega,x), D(\cdot,\omega,x))$ is identically càdlàg, and it is the solution of equations (D) on $[0,\zeta(x,\cdot))$.*

Proof. We will give the proof in several steps. In *Step 1* we will reduce the problem to one where the coefficients are globally Lipschitz. We then resolve the first system (for X) of (D), and in *Step 2* we will show that, given X, there exists a "nice" solution D of the second system of equations, which depends continuously on x. In *Step 3* we will show that D^i_k is the partial derivative in x_k of X^i in the *distributional sense.*[14] Then since it is continuous (in x), we can conclude that it is the true partial derivative.

Step 1. Choose a constant N. Then the open set $\{x : \zeta(x,\omega) > N\}$ is a countable union of closed balls. Therefore it suffices to show that if B is one of these balls, then on the set $\Gamma = \{\omega : \forall x \in B, \zeta(x,\omega) > N\}$ the function $x \mapsto X(t,\omega,x)$ is continuously differentiable on B. However by Theorem 38 we know that for each $\omega \in \Gamma$, the image of X as x runs through B is compact in $\mathcal{D}^n$ with $0 \le t \le N$, hence it is contained in a ball of radius R in $\mathbb{R}^n$, for R sufficiently large.

We fix the radius R and we denote by K the ball of radius R of $\mathbb{R}^n$ centered at 0. Let

$$\Lambda = \{\omega : \text{ for } x \in B \text{ and } 0 \le t \le N, X(t,\omega,x) \in K\}.$$

We then condition on Λ. That is, we replace P by P_Λ, where $P_\Lambda(A) \equiv P(A|\Lambda) = \frac{P(A\cap\Lambda)}{P(\Lambda)}$. Then $P_\Lambda \ll P$, so Z is still a semimartingale with respect to P_Λ (Theorem 2 of Chap. II). This allows us to make, without loss of generality, the following simplifying assumption: *if $x \in B$ then $\zeta(x,\omega) > N$, and $X(t,\omega,x) \in K$, $0 \le t \le N$.*

Next let $h : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^\infty$ with compact support and such that $h(x) = 1$ if $x \in K$ and replace f with fh. Let Z be implicitly stopped at the (constant stopping) time N. (That is, Z^N replaces Z.) With these assumptions and letting P_Λ replace P, *we can therefore assume—without loss of generality—that the coefficients in the first equation in (D) are globally Lipschitz and bounded.*

Step 2. In this step we assume that the simplifying assumptions of *Step 1* hold. We may also assume by Theorem 5 that $Z \in \mathcal{S}(\beta)$ for a β which will be specified later. If we were to proceed to calculate formally the derivative with respect to x_k of X^i, we would get

$$\begin{aligned}\frac{\partial X^i_t}{\partial x_k} &= 1 + \sum_{\alpha=1}^m \frac{\partial}{\partial x_k}\int_0^t f^i_\alpha(X_{s-})dZ^\alpha_s \\ &= 1 + \sum_{\alpha=1}^m \int_0^t \sum_{j=1}^n \frac{\partial f^i_\alpha}{\partial x_j}(X_{s-})\frac{\partial}{\partial x_k}(X^j_{s-})dZ^\alpha_s.\end{aligned}$$

Therefore our best candidate for the partial derivative with respect to x_k is the solution of the system

$$D^i_{kt} = \delta^i_k + \sum_{\alpha=1}^m \int_0^t \sum_{j=1}^n \frac{\partial f^i_\alpha}{\partial x_j}(X_{s-})D^j_{ks-}dZ^\alpha_s, \tag{$*4$}$$

[14] These derivatives are also known as *derivatives in the generalized function sense.*

and let D be the matrix (D^i_k). Naturally $X_s = X(s,\omega,x)$, and we can make explicit this dependence on x by rewriting the above equation as

$$D^{ix}_{kt} = \delta^i_k + \int_0^t \frac{\partial f^i_\alpha}{\partial x_j}(X^x_{s-})D^{jx}_{ks-}dZ^\alpha_s,$$

where the summations over α and j are implicit (Einstein convention). We now show that $D_k = (D^1_k,\dots,D^n_k)$ is continuous in x. Fix $x,y \in \mathbb{R}^n$ and let $V_s(\omega) = D_k(s,\omega,x) - D_k(s,\omega,y)$. Then

$$\begin{aligned}
V^i_t &= \int_0^t \left\{\frac{\partial f^i_\alpha}{\partial x_j}(X^x_{s-})D^{jx}_{ks-} - \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-})D^{jy}_{ks-}\right\}dZ^\alpha_s \\
&= \int_0^t \left\{\frac{\partial f^i_\alpha}{\partial x_j}(X^x_{s-}) - \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-})\right\}D^{jx}_{ks-}dZ^\alpha_s \\
&\quad + \int_0^t \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-})\left\{D^{jx}_{ks-} - D^{jy}_{ks-}\right\}dZ^\alpha_s \\
&= H^i_t(x,y,k) + \int_0^t \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-})V^j_{s-}dZ^\alpha_s \\
&= H^i_t + \int_0^t V^j_{s-}dY^{iy}_{js},
\end{aligned}$$

where $Y^{iy}_{js} = \sum_{\alpha=1}^m \int_0^s \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{u-})dZ^\alpha_u$. Note that by *Step 1* we know that $\frac{\partial f^i_\alpha}{\partial x_j}(X^y_{u-})$ is bounded; therefore since $Z^\alpha \in \mathcal{S}(\beta)$, the Y^i_j are in $\mathcal{S}(c\beta)$ for a constant c. If β is small enough, by Lemma 2 (preceding Theorem 7) we have that for each $p \geq 2$ there exists a constant $C_p(Z)$ such that

$$\|V\|_{\underline{\underline{S}}^p} \leq C_p(Z)\|H\|_{\underline{\underline{S}}^p}.$$

However we can also estimate $\|H\|_{\underline{\underline{S}}^p}$. If we let

$$J^i_{\alpha s} = \sum_{j=1}^n \left\{\frac{\partial f^i_\alpha}{\partial x_j}(X^x_{s-}) - \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-})\right\}D^{jx}_{ks-},$$

then $H^i_t = \sum_{\alpha=1}^m \int_0^t J^i_{\alpha s}dZ^\alpha_s$, and therefore by Emery's inequality (Theorem 3) we have that

$$\|H\|_{\underline{\underline{S}}^p} \leq c_p\|J\|_{\underline{\underline{S}}^p}\|Z\|_{\underline{\underline{H}}^\infty},$$

which in turn implies

$$\|V\|_{\underline{\underline{S}}^p} \leq \hat{C}_p(Z)\|J\|_{\underline{\underline{S}}^p}.$$

We turn our attention to estimating $\|J\|_{\underline{\underline{S}}^p}$. Consider first the terms

$$\frac{\partial f^i_\alpha}{\partial x_j}(X^x_{s-}) - \frac{\partial f^i_\alpha}{\partial x_j}(X^y_{s-}).$$

By the simplifying assumptions of *Step 1*, the functions $\frac{\partial f_\alpha^i}{\partial x_j}$ are Lipschitz in K, and X^x takes its values in K. Therefore Theorem 37 applies, and as we saw in its proof (inequality $(***)$) we have that

$$\left\|\frac{\partial f_\alpha^i}{\partial x_j}(X_{s-}^x) - \frac{\partial f_\alpha^i}{\partial x_j}(X_{s-}^y)\right\|_{\underline{\underline{S}}^{2p}} \leq K\|X^x - X^y\|_{\underline{\underline{S}}^{2p}} \leq \widetilde{K}(p,Z)\|x-y\|.$$

Next consider the terms D_{ks}^{jx}. We have seen that these terms are solutions of the system of equations

$$D_{kt}^{jx} = \delta_k^j + \int_0^t D_{ks-}^{\ell x} \frac{\partial f_\alpha^j}{\partial x_\ell}(X_{s-}^x)dZ_s^\alpha$$

and therefore they can be written as solutions of the exponential system

$$D_{kt}^{jx} = \delta_k^j + \int_0^t D_{ks-}^{\ell x} dY_{\ell s}^{jx},$$

with $Y_{\ell s}^{jx} = \int_0^s \frac{\partial f_\alpha^j}{\partial x_\ell}(X_{u-}^x)dZ_u^\alpha$. As before, by Lemma 2,

$$\|D_k^{jx}\|_{\underline{\underline{S}}^{2p}} \leq C_{2p}(Z).$$

Recalling the definition of J and using the Cauchy-Schwarz inequality gives

$$\begin{aligned}\|J\|_{\underline{\underline{S}}^p} &\leq \|\frac{\partial f_\alpha^i}{\partial x_j}(X_{s-}^x) - \frac{\partial f_\alpha^i}{\partial x_j}(X_{s-}^y)\|_{\underline{\underline{S}}^{2p}}\|D_k^{jx}\|_{\underline{\underline{S}}^{2p}} \\ &\leq \hat{C}_{2p}(Z)\|x-y\|,\end{aligned}$$

which in turn combined with previous estimates yields

$$\|V\|_{\underline{\underline{S}}^p} \leq C(p,Z)\|x-y\|.$$

Since V was defined to be $V_s(\omega) = D_k(s,\omega,x) - D_k(s,\omega,y)$, we have shown that (with $p > n$)

$$E\left\{\sup_{s\leq N}\|D_k(s,\omega,x) - D_k(s,\omega,y)\|^p\right\} \leq C(p,Z)^p\|x-y\|^p,$$

and therefore by Kolmogorov's Lemma (Theorem 72 of Chap. IV) we have the continuity in x of $D_k(t,\omega,x)$.

Step 3. In this step we first show that $D_k(t,\omega,x)$, the solution of equations $(*4)$ (and hence also the solution of the n^2 equations of the second line of (D)) is the partial derivative of X in the variable x_k in the sense of distributions (i.e., generalized functions). Since in *Step 2* we established the continuity of D_k in x, we can conclude that D_k is the true partial derivative.

Let us first note that with the continuity established in *Step 2*, by increasing the compact ball K, we can assume further that $D_k(s,\omega,x) \in K$ also, for $s \leq N$ and all $x \in B$.

We now make use of Cauchy's method of approximating solutions of differential equations, established for stochastic differential equations in Theorem 16 and its corollary. Note that by our simplifying assumptions, $Y = (X, D)$ takes its values in a compact set, and therefore the coefficients are (globally) Lipschitz. The process Y is the solution of a stochastic differential equation, which we write in vector and matrix form as

$$Y_t = y + \int_0^t f(Y_{s-})dZ_s.$$

Let σ_r be a sequence of partitions tending to the identity, and with the notation of Theorem 16 let $Y(\sigma) = (X(\sigma), D(\sigma))$ denote the solution of the equation of the form

$$Y_t = y + \int_0^t f(Y^{\sigma+})_s^\sigma dZ_s.$$

For each (σ_r) the equations (D) become difference equations, and thus trivially

$$\frac{\partial X(\sigma_r)}{\partial x_k} = D_k(\sigma_r).$$

The proof of the theorem will be finished if we can find a subsequence r_q such that $\lim_{q\to\infty} X(\sigma_{r_q}) = X$ and $\lim_{q\to\infty} D_k(\sigma_{r_q}) = D_k$, in the sense of distributions, considered as functions of x.

We now enlarge our space Ω exactly as in Sect. 6 (immediately preceding Theorem 31):

$$\overline{\Omega} = \mathbb{R}^{2n} \times \Omega$$
$$\overline{\mathcal{F}}_t^0 = \mathcal{B}^{2n} \otimes \mathcal{F}_t, \ \overline{\mathcal{F}}_t = \bigcap_{u>t} \overline{\mathcal{F}}_t^0,$$

where $\mathcal{B}^{2n}$ denotes the Borel sets of $\mathbb{R}^{2n}$. Let λ be normalized Lebesgue measure of K. Finally define

$$\overline{P} = \lambda \times P.$$

We can as in the proof of *Step 2* assume that $Z \in \mathcal{S}(\beta)$ for β small enough and then,

$$\lim_{r\to\infty} (X(\sigma_r), D(\sigma_r) = (X, D) \ \text{ in } \underline{\underline{S}}^2,$$

by Theorem 16. Therefore there exists a subsequence r_q such that

$$M = \sum_{q=1}^{\infty} \sup_t \|(X(r_q), D(r_q)) - (X, D)\| \in L^1(d\overline{P}).$$

The function $M = M(\omega, x)$ is in $L^1(\lambda \times P)$, and therefore for P-almost all ω the function $x \mapsto M(\omega, x) \in L^1(d\lambda)$. For ω not in the exceptional set, and t fixed it follows that

$$\lim_{q\to\infty} (X(r_q), D(r_q)) = (X, D)$$

λ a.e. Further, it is bounded by the function $M(\omega,\cdot) + \|(X(t,\omega,\cdot), D(t,\omega,\cdot))\|$ which is integrable by hypothesis. This gives convergence in the distributional sense, and the proof is complete. □

We state the following corollary to Theorem 39 as a theorem.

Theorem 40. *Let Z be as in (H1) and let the functions (f^i_α) in (H2) have locally Lipschitz derivatives up to order N, for some N, $0 \le N \le \infty$. Then there exists a solution $X(t, w, x)$ to*

$$X^i_t = x_i + \sum_{\alpha=1}^{m} \int_0^t f^i_\alpha(X_{s-}) dZ^\alpha_s, \; 1 \le i \le n,$$

which is N times continuously differentiable in the open set $\{x : \zeta(x, w) > t\}$, where ζ is the explosion time of the solution. If the coefficients (f^i_α) are globally Lipschitz, then $\zeta = \infty$.

Proof. If $N = 0$, then Theorem 40 is exactly Theorem 38. If $N = 1$, then Theorem 40 is Theorem 39. If $N > 1$, then the coefficients of equations (D) have locally Lipschitz derivatives of order $N - 1$ at least. Induction yields $(X, D) \in \mathcal{C}^{N-1}$, whence $X \in \mathcal{C}^N$. □

Note that the coefficients (f^i_α) in Theorem 40 are locally Lipschitz of order N if, for example, they have $N + 1$ continuous partial derivatives; that is, if $f^i_\alpha \in \mathcal{C}^{N+1}(\mathbb{R}^n)$, for each i and α, then (f^i_α) are locally Lipschitz of order N.

8 Flows as Diffeomorphisms: The Continuous Case

In this section we will study a system of differential equations of the form

$$X^i_t = x_i + \sum_{\alpha=1}^{m} \int_0^t F^i_\alpha(X)_{s-} dZ^\alpha_s, \quad 1 \le i \le n, \tag{$*$}$$

where the *semimartingales Z^α are assumed to have continuous paths* with $Z_0 = 0$. The continuity assumption leads to pleasing results. In Sect. 10 we consider the general case where the semimartingale differentials can have jumps. The flow of an equation such as $(*)$ is considered to be an $\mathbb{R}^n$-valued function $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ given by $\varphi(x) = X(t, \omega, x)$, for each (t, w). We first consider the possible injectivity of φ, of which there are two forms.

Definition. The flow φ of equation $(*)$ is said to be **weakly injective** if for each fixed $x, y \in \mathbb{R}^n$, $x \neq y$,

$$P\{w : \exists t : X(t, \omega, x) = X(t, \omega, y)\} = 0.$$

Definition. The flow φ of equation $(*)$ is said to be **strongly injective** (or, simply, **injective**) if for almost all ω the function $\varphi : x \to X(t, \omega, x)$ is injective for all t.

For convenience we recall here a definition from Sect. 3 of this chapter.

Definition. An operator F from $\mathbb{D}^n$ into $\mathbb{D}$ is said to be **process Lipschitz** if for any $\mathbf{X}, \mathbf{Y} \in \mathbb{D}^n$ the following two conditions are satisfied.

(i) For any stopping time T, $\mathbf{X}^{T-} = \mathbf{Y}^{T-}$ implies $F(\mathbf{X})^{T-} = F(\mathbf{Y})^{T-}$.
(ii) There exists an adapted process $K \in \mathbb{L}$ such that

$$\|F(\mathbf{X})_t - F(\mathbf{Y})_t\| \leq K_t \|\mathbf{X}_t - \mathbf{Y}_t\|.$$

Actually, process Lipschitz is only slightly more general than random Lipschitz. The norm symbols in the above definition denote Euclidean norm, and not sup norm. Note that if F is process Lipschitz then F is also functional Lipschitz and all the theorems we have proven for functional Lipschitz coefficients hold as well for process Lipschitz coefficients. If f is a function which is Lipschitz (as defined at the beginning of Sect. 3) then f induces a process Lipschitz operator. Finally, observe that by Theorem 37 we know that the flow of equation $(*)$ is continuous from $\mathbb{R}^n$ into $\mathbb{R}^n$ or from $\mathbb{R}^n$ into $\mathcal{D}^n$ a.s., where $\mathcal{D}^n$ has the topology of uniform convergence on compacts.

Theorem 41. *Let Z^α be continuous semimartingales, $1 \leq \alpha \leq m$, H a vector of adapted càdlàg processes, and F an $n \times m$ matrix of process Lipschitz operators. Then the flow of the solution of*

$$X_t = x + H_t + \int_0^t F(X)_{s-} dZ_s$$

is weakly injective.[15]

Proof. Let x, $y \in \mathbb{R}^n$, $x \neq y$. Let X^x, X^y denote the solutions of the above equation with initial conditions x, y respectively. We let $u = x - y$ and $U = X^x - X^y$. We must show $P\{\omega : \exists t : U_t(\omega) = 0\} = 0$. Set $V = F(X^x)_- - F(X^y)_-$. Then $V \in \mathbb{L}$ and $|V| \leq K|U|$. Further, the processes U and V are related by

$$U_t = u + \int_0^t V_s dZ_s.$$

Let $T = \inf\{t > 0 : U_t = 0\}$; the aim is to show $P(T = \infty) = 1$. Since U is continuous the stopping time T is the limit of a sequence of increasing stopping

[15] We are using the Einstein convention on sums.

times S^k strictly less than T. Therefore the process $1_{[0,T)} = \lim_{k\to\infty} 1_{[0,S_k]}$ is predictable.

We use Itô's formula (Theorem 32 of Chap. II) on $[0,T)$ for the function $f(x) = \log\|x\|$. Note that

$$\frac{\partial f}{\partial x_i} = \frac{1}{\|x\|}\frac{\partial\|x\|}{\partial x_i} = \frac{1}{\|x\|}\frac{x_i}{\|x\|} = \frac{x_i}{\|x\|^2};$$
$$\frac{\partial^2 f}{\partial x_i^2} = \frac{1}{\|x\|^2} - \frac{2x_i}{\|x\|^3}\frac{\partial\|x\|}{\partial x_i} = \frac{1}{\|x\|^2} - \frac{2x_i^2}{\|x\|^4};$$
$$\frac{\partial^2 f}{\partial x_i \partial x_j} = -\frac{2x_i x_j}{\|x\|^4}.$$

Therefore on $[0,T)$,

$$\begin{aligned}&\log\|U_t\| - \log\|u\| \\ &= \sum_i \int_0^t \frac{1}{\|U_s\|^2} U_s^i dU_s^i + \frac{1}{2}\sum_i \int_0^t \frac{1}{\|U_s\|^2} d[U^i,U^i]_s - \sum_{i,j}\int_0^t \frac{U_s^i U_s^j}{\|U_s\|^4} d[U^i,U^j]_s.\end{aligned}$$

Since $dU^i = \sum_\alpha V^{i,\alpha} dZ^\alpha$, the foregoing equals

$$\begin{aligned}&= \sum_{i,\alpha}\int_0^t \frac{U_s^i V_s^{i,\alpha}}{\|U_s\|^2} dZ_s^\alpha + \frac{1}{2}\sum_{i,\alpha,\beta}\int_0^t \frac{V_s^{i,\alpha}V_s^{i,\beta}}{\|U_s\|^2} d[Z^\alpha,Z^\beta]_s \\ &\qquad - \sum_{i,j,\alpha,\beta}\int_0^t \frac{U_s^i U_s^j V_s^{i,\alpha} V_s^{j,\beta}}{\|U_s\|^4} d[Z^\alpha,Z^\beta]_s.\end{aligned}$$

All the integrands on the right side are predictable and since $\|V\| \le K\|U\|$ they are moreover bounded by K and K^2 in absolute value. However on $\{T<\infty\}$ the left side of the equation, $\log\|U_t\| - \log\|u\|$, tends to $-\infty$ as t increases to T; the right side is a well-defined non-exploding semimartingale on all of $[0,\infty)$. Therefore $P(T<\infty) = 0$, and the proof is complete. □

In the study of strong injectivity the stochastic exponential of a semimartingale (introduced in Theorem 37 of Chap. II) plays an important role. Recall that if Z is a *continuous* semimartingale, then $X_0\mathcal{E}(Z)$ denotes the (unique) solution of the equation

$$X_t = X_0 + \int_0^t X_s dZ_s,$$

and $\mathcal{E}(Z)_t = \exp\{Z_t - \frac{1}{2}[Z,Z]_t\}$. In particular, $P(\inf_{s\le t}\mathcal{E}(Z)_s > 0) = 1$.

Theorem 42. *For $x \in \mathbb{R}^n$, let H^x be in $\mathbb{D}^k$ such that they are locally bounded uniformly in x. Assume further that there exists a sequence of stopping times $(T_\ell)_{\ell\ge 1}$ increasing to ∞ a.s. such that $\|(H_-^x - H_-^y)^{T_\ell}\|_{\underline{\underline{S}}^r} \le K\|x-y\|$, each*

$\ell \geq 1$, for a constant K and for some $r > n$. Let $Z = (Z^1, \dots, Z^k)$ be k semimartingales. Then the functions

$$x \mapsto \int_0^t H_s^x dZ_s$$
$$x \mapsto [H^x \cdot Z, H^x \cdot Z]_t$$

have versions which are continuous as functions from $\mathbb{R}^n$ into $\mathcal{D}$, with $\mathcal{D}$ having the topology of uniform convergence on compacts.

Proof. By Theorem 5 there exists an arbitrarily large stopping time T such that $Z^{T-} \in \underline{\underline{H}}^\infty$. Thus without loss of generality we can assume that $Z \in \underline{\underline{H}}^\infty$, and that H^x is bounded by some constant K, uniformly in x. Further we assume $\|H_-^x - H_-^y\|_{\underline{\underline{S}}^r} \leq K\|x-y\|$. Then

$$\begin{aligned} E\{\sup_t \| \int_0^t H_{s-}^x dZ_s - \int_0^t H_{s-}^y dZ_s\|^r\} &\leq CE\{\sup_t \|H_{t-}^x - H_{t-}^y\|^r\}\|Z\|_{\underline{\underline{H}}^\infty}^r \\ &\leq \widetilde{K}\|x-y\|^r\|Z\|_{\underline{\underline{H}}^\infty}^r, \end{aligned}$$

where we have used Emery's inequality (Theorem 3). The result for $\int H_{s-}^x dZ_s$ now follows from Kolmogorov's Lemma (Theorem 72 of Chap. IV). For the second result we have

$$\begin{aligned} &\|[H_-^x \cdot Z, H_-^x \cdot Z] - [H_-^y \cdot Z, H_-^y \cdot Z]\|_{\underline{\underline{S}}^r} \\ &\quad = \| \int \{(H_{s-}^x)^2 - (H_{s-}^y)^2\} d[Z,Z]_s\|_{\underline{\underline{S}}^r} \\ &\quad = \| \sum_{i,j=1}^k \int (H_{s-}^{xi}) + H_{s-}^{yi}(H_{s-}^{xj} - H_{s-}^{yj}) d[Z^i, Z^j]_s\|_{\underline{\underline{S}}^r}, \\ &\quad \leq 2K\|Z\|_{\underline{\underline{H}}^\infty}^2 \|H^x - H^y\|_{\underline{\underline{S}}^r}, \end{aligned}$$

and the result follows. □

Theorem 43. *Let F be a matrix of process Lipschitz operators and X^x the solution of $(*)$ with initial condition x, for continuous semimartingales Z^α, $1 \leq \alpha \leq m$. Fix x, $y \in \mathbb{R}^n$. For $r \in \mathbb{R}$ there exist for every x, $y \in \mathbb{R}^n$ with $x \neq y$ (uniformly) locally bounded predictable processes $H^\alpha(x,y)$, $J^{\alpha,\beta}(x,y)$, which depend on r, such that*

$$\|X_t^x - X_t^y\|^r = \|x-y\|^r \mathcal{E}(\Lambda_r(x,y))_t$$

where

$$\Lambda_r(x,y)_t = \int_0^t H_s^\alpha(x,y) dZ_s^\alpha + \int_0^t J_s^{\alpha,\beta}(x,y) d[Z^\alpha, Z^\beta]_s.$$

Proof. Fix x, $y \in \mathbb{R}^n$ and let $U = X^x - X^y$, $V = F(X^x)_- - F(X^y)_-$. Itô's formula applies since U is never zero by weak injectivity (Theorem 36). Using the Einstein convention,

$$\|U\|^r = \|x-y\|^r + \int r\|U_s\|^{r-2}U_s^i dU_s^i$$
$$+ \frac{1}{2}\int r\{(r-2)\|U_s\|^{r-4}U_s^i U_s^j + \delta_j^i \|U_s\|^{r-2}\}d[U^i, U^j]_s.$$

Let $(\cdot,\cdot)$ denote Euclidean inner product on $\mathbb{R}^n$. It suffices to take

$$H_s^\alpha(x,y) = r\|U_s\|^{-2}(U_s, V_s^\alpha),$$

(where V^α is the α-th column of V); and

$$J_s^{\alpha,\beta}(x,y) = \frac{1}{2}r\{(r-2)\|U_s\|^{-4}(U_s, V_s^\alpha)(U_s, V_s^\beta) + \|U_s\|^{-2}(V_s^\alpha, V_s^\beta)\}.$$

One checks that these choices work by observing that $dU_t^i = \sum_{\alpha=1}^m V_t^{i,\alpha} dZ_t^\alpha$. Finally the above allows us to conclude that

$$\|U_t\|^r = \|x-y\|^r + \int_0^t \|U_s\|^r dA_r(x,y)_s,$$

and the result follows. □

Before giving a key corollary to Theorem 43, we need a lemma. Let $\widetilde{H}^\infty$ be the space of *continuous* semimartingales X with $X_0 = 0$ such that X has a (unique) decomposition

$$X = N + A$$

where N is a continuous local martingale, A is a continuous process of finite variation, $N_0 = A_0 = 0$, *and such that* $[N,N]_\infty$ *and* $\int_0^\infty |dA_s|$ *are in* L^∞. Further, let us define

$$\|X\|_{\widetilde{H}^\infty} = \| [N,N]_\infty^{1/2} + \int_0^\infty |dA_s| \|_{L^\infty}.$$

Lemma. For every $p, a < \infty$, there exists a constant $C(p,a) < \infty$ such that if $\|X\|_{\widetilde{H}^\infty} \le a$, then $\|\mathcal{E}(X)\|_{\underline{\underline{S}}^p} \le C(p,a)$.

Proof. Let $X = N + A$ be the (unique) decomposition of X. Then

$$\begin{aligned}\|\mathcal{E}(X)\|_{\underline{\underline{S}}^p}^p &= E\{\sup_t \exp\{p(X_t - \frac{1}{2}[X,X]_t)\}\} \\ &\le E\{e^{pX^*}\} \qquad \text{(recall that } X^* = \sup_t |X_t|) \\ &\le E\{\exp\{pN^* + pa\}\} \\ &= e^{pa}E\{e^{pN^*}\},\end{aligned}$$

since $|A_t| \le a$, a.s. We therefore need to prove only an exponential maximal inequality for continuous martingales. By Theorem 42 of Chap. II, since N is a continuous martingale, it is a time change of a Brownian motion. That is, $N_t = B_{[N,N]_t}$, where B is a Brownian motion defined with a different filtration. Therefore since $[N,N]_\infty \le a^2$, we have

$$N^* \le B^*_{[N,N]_\infty} \le B^*_{a^2},$$

and hence $E\{e^{pN^*}\} \le E\{\exp\{pB^*_{a^2}\}\}$. Using the reflection principle (Theorem 33 of Chap. I) we have

$$E\{\exp\{pB^*_{a^2}\}\} \le 2E\{\exp\{pB_{a^2}\}\} = 2\exp\{\frac{p^2a^2}{2}\}.$$

□

Note that in the course of the proof of the lemma we obtained $C(p,a) = 2^{1/p}\exp\{a + pa^2/2\}$.

Corollary. Let $-\infty < r < \infty$ and $p < \infty$, and let $(\Lambda_r(x,y)_t)_{t\ge 0}$ be as given in Theorem 43. Then $\mathcal{E}(\Lambda_r(x,y))$ is locally in $\underline{\underline{S}}^p$, uniformly in x, y.

Proof. We need to show that there exists a sequence of stopping times T_ℓ increasing to ∞ a.s., and constants $C_\ell < \infty$, such that $\|\mathcal{E}(\Lambda_r(x,y))^{T_\ell}\|_{\underline{\underline{S}}^p} \le C_\ell$ for all $x, y \in \mathbb{R}^n$, $x \ne y$.

By stopping, we may assume that Z^α and $[Z^\alpha, Z^\beta]$ are in $\underline{\underline{\widetilde{H}}}^\infty$ and that $|H^\alpha|$ and $|J^{\alpha,\beta}| \le b$ for all (x,y), $1 \le \alpha, \beta \le m$. Therefore $\|\Lambda_r(x,y)\|_{\widetilde{H}^\infty} \le C$ for a constant C, since if $X \in \underline{\underline{\widetilde{H}}}^\infty$ and if K is bounded, predictable, then $K \cdot X \in \widetilde{H}^\infty$ and $\|K \cdot X\|_{\widetilde{H}^\infty} \le \|K\|_{\underline{\underline{S}}^\infty}\|X\|_{\widetilde{H}^\infty}$, as can be proved exactly analogously to Emery's inequalities (Theorem 3). The result now follows by the preceding lemma. □

Comment. A similar result in the right continuous case is proved by a different method in the proof of Theorem 62 in Sect. 10.

Theorem 44. *Let Z^α be continuous semimartingales, $1 \le \alpha \le m$, and F an $n \times m$ matrix of process Lipschitz operators. Then the flow of the solution of*

$$X_t = x + \int_0^t F(X)_{s-}dZ_s \qquad (*)$$

is strongly injective on $\mathbb{R}^n$.

Proof. It suffices to show that for any compact set $C \subset \mathbb{R}^n$, for each N, there exists an event of probability zero outside of which for every $x, y \in C$ with $x \ne y$,

$$\inf_{s\le N} \|X(s,\omega,x) - X(s,\omega,y)\| > 0.$$

Let x, y, s have rational coordinates. By Theorem 43 a.s.

$$\|X(s,\omega,x)-X(s,\omega,y)\|^r = \|x-y\|^r \mathcal{E}(\Lambda_r(x,y))_s.$$

The left side of the equation is continuous (Theorem 37). As for the right side, $\mathcal{E}(\Lambda_r(x,y))$ will be continuous if we can show that the processes $H_s^\alpha(x,y)$ and $J_s^{\alpha,\beta}(x,y)$, given in Theorem 43, verify the hypotheses of Theorem 42. To this end, let B be any relatively compact subset of $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x)\}$ (e.g., $B = B_1 \times B_2$ where B_1, B_2 are open balls in $\mathbb{R}^n$ with disjoint closures). Then $\|x-y\|^r$ is bounded on B for any real number r. Without loss we take $r = 1$ here. Let $U(x,y) = X^x - X^y$, $V(x,y) = F(X^x)_- - F(X^y)_-$, and let $V^\alpha(x,y)$ be the α-th column of V. Then for (x,y) and (x',y') in B we have

$$\begin{aligned} &H^\alpha(x,y) - H^\alpha(x',y') \qquad (**)\\ &\quad = (\|U(x,y)\|^{-2} - \|U(x',y')\|^{-2})(U(x,y), V^\alpha(x,y))\\ &\qquad + \|U(x',y')\|^{-2}(U(x,y)-U(x',y'), V^\alpha(x,y))\\ &\qquad\quad + \|U(x',y')\|^{-2}(U(x',y'), V^\alpha(x,y) - V^\alpha(x',y')). \end{aligned}$$

The first term on the right side of $(**)$ above is dominated in absolute value by

$$\begin{aligned} &\frac{|\|U(x,y)\| - \|U(x',y')\||(\|U(x,y)\| + \|U(x',y')\|)}{\|U(x,y)\|^2\|U(x',y')\|^2}\|U(x,y)\|\|V^\alpha(x,y)\|\\ &\quad \le K\|U(x,y)-U(x',y')\|(\|U(x,y)\| + \|U(x',y')\|)\|U(x',y')\|^{-2}, \end{aligned}$$

where we are assuming (by stopping), that F has a Lipschitz constant K. Since $U(x,y) - U(x',y') = U(x,x') - U(y,y')$, the above is less than

$$\begin{aligned} &K(\|U(x,x')\| + \|U(y,y')\|)(\|U(x,y)\| + \|U(x',y')\|)\|U(x',y')\|^{-2}\\ &\quad = K(\|x-x'\|\mathcal{E}(\Lambda_1(x,x')) + \|y-y'\|\mathcal{E}(\Lambda_1(y,y')))\cdot\\ &\qquad (\|x-y\|\mathcal{E}(\Lambda_1(x,y)) + \|x'-y'\|\mathcal{E}(\Lambda_1(x',y')))\|x'-y'\|^{-2}\mathcal{E}(\Lambda_{-2}(x',y'))\\ &\quad \le K_1\|(x,y)-(x',y')\|(\mathcal{E}(\Lambda_1(x,x')) + \mathcal{E}(\Lambda_1(y,y')))(\mathcal{E}(\Lambda_1(x,y))\\ &\qquad + \mathcal{E}(\Lambda_1(x',y')))\mathcal{E}(\Lambda_{-2}(x',y')). \end{aligned}$$

By the lemma following Theorem 43, and by Hölder's and Minkowski's inequalities we may, for any $p < \infty$, find stopping times T_ℓ increasing to ∞ a.s. such that the last term above is dominated in $\underline{\underline{S}}^p$ norm by $K_\ell\|(x,y)-(x',y')\|$ for a constant K_ℓ corresponding to T_ℓ. We get analogous estimates for the second and third terms on the right side of $(**)$ by similar (indeed, slightly simpler) arguments. Therefore H^α satisfies the hypotheses of Theorem 42, for $(x,y) \in B$. The same is true for $J^{\alpha,\beta}$, and therefore Theorem 42 shows that Λ_r and $[\Lambda_r, \Lambda_r]$ are continuous in (x,y) on B. (Actually we are using a local version of Theorem 42 with $(x,y) \in B \subset \mathbb{R}^{2n}$ instead of all of $\mathbb{R}^{2n}$; this is not a problem since Theorem 42 extends to the case $x \in W$ open in $\mathbb{R}^n$, because

Kolmogorov's Lemma does—recall that continuity is a local property.) Finally since Λ_r and $[\Lambda_r, \Lambda_r]$ are continuous in $(x, y) \in B$ we deduce that $\mathcal{E}(\Lambda_r(x, y))$ is continuous in $\{(x, y) \in \mathbb{R}^{2n} : x \neq y\}$.

We have shown that both sides of

$$\|X(s, w, x) - X(s, w, y)\|^r = \|x - y\|^r \mathcal{E}(\Lambda_r(x, y))_s$$

can be taken jointly continuous. Therefore except for a set of probability zero the equality holds for all $(x, y, s) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+$. The result follows because $\mathcal{E}(\Lambda_r(x, y))_t$ is defined for all t finite and it is never zero. □

Theorem 45. *Let Z^α be continuous semimartingales, $1 \leq \alpha \leq m$, and let F be an $n \times m$ matrix of process Lipschitz operators. Let X be the solution of $(*)$. Then for each $N < \infty$ and almost all ω*

$$\lim_{\|x\| \to \infty} \inf_{s \leq N} \|X(s, \omega, x)\| = \infty.$$

Proof. By Theorem 43 the equality

$$\|X_t^x - X_t^y\|^r = \|x - y\|^r \mathcal{E}(\Lambda_r(x, y))_t$$

is valid for all $r \in \mathbb{R}$.

For $x \neq 0$ let $Y^x = \|X^x - X^0\|^{-1}$. (Note that Y^x is well-defined by Theorem 41.) Then

$$\begin{aligned} |Y^x| &= \|x\|^{-1} \mathcal{E}(\Lambda_{-1}(x, 0)) \qquad (***) \\ |Y^x - Y^y| &\leq \|X^x - X^y\| \|X^x - X^0\|^{-1} \|X^y - X^0\|^{-1} \\ &= \|x - y\| \|x\|^{-1} \|y\|^{-1} \mathcal{E}(\Lambda_1(x, y)) \mathcal{E}(\Lambda_{-1}(x, 0)) \mathcal{E}(\Lambda_{-1}(y, 0)). \end{aligned}$$

Define $Y^\infty = 0$. The mapping $x \mapsto x\|x\|^{-2}$ inspires a distance d on $\mathbb{R}^n \setminus \{0\}$ by $d(x, y) = \frac{\|x-y\|}{\|x\|\|y\|}$. Indeed,

$$\left\| \frac{x}{\|x\|^2} - \frac{y}{\|y\|^2} \right\|^2 = \left(\frac{\|x - y\|}{\|x\| \|y\|} \right)^2.$$

By Hölder's inequality we have that

$$\begin{aligned} &\|\mathcal{E}(\Lambda_1(x, y)) \mathcal{E}(\Lambda_{-1}(x, 0) \mathcal{E}(\Lambda_{-1}(y, 0))\|_{\underline{\underline{S}}^r} \\ &\qquad \leq \|\mathcal{E}(\Lambda_1(x, y))\|_{\underline{\underline{S}}^{3r}} \|\mathcal{E}(\Lambda_{-1}(x, 0))\|_{\underline{\underline{S}}^{3r}} \|\mathcal{E}(\Lambda_{-1}(y, 0))\|_{\underline{\underline{S}}^{3r}} \end{aligned}$$

and therefore by the corollary to Theorem 43 we can find a sequence of stopping times $(T_\ell)_{\ell \geq 1}$ increasing to ∞ a.s. such that there exist constants C_ℓ with (using $(***)$)

$$\|(Y^x - Y^y)^{T_\ell}\|_{\underline{\underline{S}}^r} \leq d(x, y) C_\ell.$$

Next set

$$\hat{Y}^x = \begin{cases} Y^{x\|x\|^{-2}}, & 0 < \|x\| < \infty, \\ Y^\infty = 0, & \|x\| = 0. \end{cases}$$

Then $\|(\hat{Y}^x - \hat{Y}^y)^{T_\ell}\|^r_{\underline{\underline{S}}^r} \leq C^r_\ell \|x - y\|^r$ on $\mathbb{R}^n$, and by Kolmogorov's Lemma (Theorem 72 of Chap. IV), there exists a jointly continuous version of $(t, x) \mapsto \hat{Y}^x_t$, on $\mathbb{R}^n$. Therefore $\lim_{\|x\|\to 0} \hat{Y}^x$ exists and equals 0. Since $(\hat{Y}^x)^{-1} = \|X^{x\|x\|^{-2}} - X^0\|$, we have the result. □

Theorem 46. *Let Z^α be continuous semimartingales, $1 \leq \alpha \leq m$, and F be an $n \times m$ matrix of process Lipschitz operators. Let X be the solution of $(*)$. Let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be the flow $\varphi(x) = X(t, \omega, x)$. Then for almost all ω one has that for all t the function φ is surjective and moreover it is a homeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$.*

Proof. As noted preceding Theorem 41, the flow φ is continuous from $\mathbb{R}^n$ to $\mathcal{D}^n$, topologized by uniform convergence on compacts; hence for a.a. ω it is continuous from $\mathbb{R}^n$ to $\mathbb{R}^n$ for all t.

The flow φ is injective a.s. for all t by Theorem 44.

Next observe that the image of $\mathbb{R}^n$ under φ, denoted $\varphi(\mathbb{R}^n)$, is closed. Indeed, let $\overline{\varphi(\mathbb{R}^n)}$ denote its closure and let $y \in \overline{\varphi(\mathbb{R}^n)}$. Let (x_k) denote a sequence such that $\lim_{k\to\infty} \varphi(x_k) = y$. By Theorem 45, $\limsup_{k\to\infty} \|x_k\| < \infty$, and hence the sequence (x_k) has a limit point $x \in \mathbb{R}^n$. Continuity implies $\varphi(x) = y$, and we conclude that $\varphi(\mathbb{R}^n) = \overline{\varphi(\mathbb{R}^n)}$; that is, $\varphi(\mathbb{R}^n)$ is closed. Then, as we have seen, the set $\{x_k\}$ is bounded. If x_k does not converge to x, there must exist a limit point $z \neq x$. But then $\varphi(z) = y = \varphi(x)$, and this violates the injectivity, already established. Therefore φ^{-1} is continuous.

Since φ is a homeomorphism from $\mathbb{R}^n$ to $\varphi(\mathbb{R}^n)$, the subspace $\varphi(\mathbb{R}^n)$ of $\mathbb{R}^n$ is homeomorphic to a manifold of dimension n in $\mathbb{R}^n$; therefore by the theorem of the invariance of the domain (see, e.g., Greenberg [84, page 82]), the space $\varphi(\mathbb{R}^n)$ is open in $\mathbb{R}^n$. But $\varphi(\mathbb{R}^n)$ is also closed and non-empty. There is only one such set in $\mathbb{R}^n$ that is open and closed and non-empty and it is the entire space $\mathbb{R}^n$. We conclude that $\varphi(\mathbb{R}^n) = \mathbb{R}^n$. □

Comment. The proof of Theorem 46 can be simplified as follows: extend φ to the Alexandrov compactification $\mathbb{R}^n_* = \mathbb{R}^n \cup \{\infty\}$ of $\mathbb{R}^n$ to $\overline{\varphi}$ by

$$\overline{\varphi}(x) = \begin{cases} \varphi(x), & x \in \mathbb{R}^n, \\ \infty, & x = \infty. \end{cases}$$

Then $\overline{\varphi}$ is continuous on $\mathbb{R}^n_*$ by Theorem 45, and obviously it is still injective. Since $\mathbb{R}^n_*$ is compact, $\overline{\varphi}$ is a homeomorphism of $\mathbb{R}^n_*$ onto $\overline{\varphi}(\mathbb{R}^n_*)$. However $\mathbb{R}^n_*$ is topologically the sphere S^n, and thus it is not homeomorphic to any proper subset (this is a consequence of the Jordan-Brouwer Separation Theorem (e.g., Greenberg [84, page 79]). Hence $\overline{\varphi}(\mathbb{R}^n_*) = \mathbb{R}^n_*$.

We next turn our attention to determining when the flow is a diffeomorphism of $\mathbb{R}^n$. Recall that a **diffeomorphism** of $\mathbb{R}^n$ is a bijection (one to one and onto) which is $\mathcal{C}^\infty$ and which has an inverse that is also $\mathcal{C}^\infty$. Clearly the hypotheses on the coefficients need to be the intersection of those of Sect. 7 and process Lipschitz.

First we introduce a useful concept, that of *right stochastic exponentials*, which arises naturally in this context. For given n, let Z be an $n \times n$ matrix of given semimartingales. If X is a solution of

$$X_t = I + \int_0^t X_{s-} dZ_s,$$

where X is an $n \times n$ matrix of semimartingales and I is the identity matrix, *then* $X = \mathcal{E}(Z)$, *the (matrix-valued) exponential of* Z. Since the space of $n \times n$ matrices is not commutative, it is also possible to consider *right stochastic integrals*, denoted

$$(Z : H)_t = \int_0^t (dZ_s) H_s,$$

where Z is an $n \times n$ matrix of semimartingales and H is an $n \times n$ matrix of (integrable) predictable processes. If $'$ denotes matrix transpose, then

$$(Z : H) = (H' \cdot Z')',$$

and therefore right stochastic integrals can be defined in terms of stochastic integrals. Elementary results concerning right stochastic integrals are collected in the next theorem. Note that $\int Y_- dZ$ and $[Y, Z]$ denote $n \times n$ matrix-valued processes here.

Theorem 47. *Let* Y, Z *be given* $n \times n$ *matrices of semimartingales,* H *an* $n \times n$ *matrix of locally bounded predictable processes. Then,*

(i) $Y_t Z_t - Y_0 Z_0 = \int_0^t Y_{s-} dZ_s + \int_0^t (dY_s) Z_{s-} + [Y, Z]_t$;
(ii) $[H \cdot Y, Z] = H \cdot [Y, Z]$; *and*
(iii) $[Y, Z : H] = [Y, Z] : H$.

Moreover if F *is an* $n \times n$ *matrix of functional Lipschitz operators, then there exists a unique* $n \times n$ *matrix of* $\mathbb{D}$*-valued processes which is the solution of*

$$X_t = I + \int_0^t (dZ_s) F(X)_{s-}.$$

Proof. The first three identities are easily proved by calculating the entries of the matrices and using the results of Chap. II. Similarly the existence and uniqueness result for the stochastic integral equation is a simple consequence of Theorem 7. □

Theorem 47 allows the definition of the right stochastic exponential.

Definition. The **right stochastic exponential** of an $n \times n$ matrix of semimartingales Z, denoted $\mathcal{E}^R(Z)$, is the (unique) matrix-valued solution of the equation

$$X_t = I + \int_0^t (dZ_s)X_{s-}.$$

We illustrate the relation between left and right stochastic exponentials in the continuous case. The general case is considered in Sect. 10 (see Theorem 63). Note that $\mathcal{E}^R(Z) = \mathcal{E}(Z')'$.

Theorem 48. *Let Z be an $n \times n$ matrix of continuous semimartingales with $Z_0 = 0$. Then $\mathcal{E}(Z)$ and $\mathcal{E}^R(-Z + [Z,Z])$ are inverses; that is, $\mathcal{E}(Z)\mathcal{E}^R(-Z + [Z,Z]) = I$.*

Proof. Let $U = \mathcal{E}(Z)$ and $V = \mathcal{E}^R(-Z + [Z,Z])$. Since $U_0V_0 = I$, it suffices to show that $d(U_tV_t) = 0$, all $t > 0$. Note that

$$\begin{aligned} dV &= (-dZ + d[Z,Z])V, \\ (dU)V &= (UdZ)V, \text{ and} \\ d[U,V] &= Ud[Z,V] = -Ud[Z,Z]V. \end{aligned}$$

Using Theorem 47 and the preceding,

$$\begin{aligned} d(UV) &= UdV + (dU)V + d[U,V] \\ &= U(-dZ + d[Z,Z])V + UdZV - Ud[Z,Z]V \\ &= 0, \end{aligned}$$

and we are done. □

The next theorem is a special case of Theorem 40 (of Sect. 7), but we state it here as a separate theorem for ease of reference.

Theorem 49. *Let $(Z^1, \dots, Z^m)$ be continuous semimartingales with $Z_0^i = 0$, $1 \le i \le m$, and let (f_α^i), $1 \le i \le n$, $1 \le \alpha \le m$, be functions mapping $\mathbb{R}^n$ to $\mathbb{R}$, with locally Lipschitz partial derivatives up to order N, $1 \le N \le \infty$, and bounded first derivatives. Then there exists a solution $X(t,\omega,x)$ to*

$$X_t^i = x_i + \sum_{\alpha=1}^m \int_0^t f_\alpha^i(X_s)dZ_s^\alpha, \quad 1 \le i \le n,$$

such that its flow $\varphi : x \to X(x,t,\omega)$ is N times continuously differentiable on $\mathbb{R}^n$. Moreover the first partial derivatives satisfy the linear equation

$$D_{kt}^i = \delta_k^i + \sum_{\alpha=1}^m \sum_{j=1}^n \int_0^t \frac{\partial f_\alpha^i}{\partial x_j}(X_s)D_{ks}^j dZ_s^\alpha \qquad (1 \le i \le n).$$

where δ_k^j is Kronecker's delta.

Observe that since the first partial derivatives are bounded, the coefficients are globally Lipschitz and it is not necessary to introduce an explosion time. Also, the value $N = \infty$ is included in the statement. The explicit equation for the partial derivatives comes from Theorem 39.

Let D denote the $n \times n$ matrix-valued process

$$D_t = (D^i_{kt})_{1 \le i \le n, 1 \le k \le n}. \tag{$\otimes$}$$

The process D is the right stochastic exponential $\mathcal{E}^R(Y)$, where Y is defined by

$$dY^{i,j}_s = \sum_{\alpha=1}^m \frac{\partial f^i_\alpha}{\partial x_j}(X_s)dZ^\alpha_s.$$

Combining Theorems 48 and 49 and the above observation we have the important following result.

Theorem 50. *With the hypotheses and notation of Theorem 49, the matrix D_t is non-singular for all $t > 0$ and $x \in \mathbb{R}^n$, a.s.*

Theorem 51. *Let $(Z^1, \dots, Z^m)$ be continuous semimartingales and let (f^i_α), $1 \le i \le n$, $1 \le \alpha \le m$, be functions mapping $\mathbb{R}^n$ to $\mathbb{R}$, with partial derivatives of all orders, and bounded first partials. Then the flow of the solution of*

$$X^i_t = x_i + \sum_{\alpha=1}^m \int_0^t f^i_\alpha(X_s)dZ^\alpha_s, \ 1 \le i \le n,$$

is a diffeomorphism from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Proof. Let φ denote the flow of X. Since $(f^i_\alpha)_{1\le i\le n, 1\le\alpha\le m}$ have bounded first partials, they are globally Lipschitz, and hence there are no finite explosions. Moreover since they are C^∞, the flow is C^∞ on $\mathbb{R}^n$ by Theorem 49. The coefficients (f^i_α) are trivially process Lipschitz, hence by Theorem 46 the flow φ is a homeomorphism; in particular it is a bijection of $\mathbb{R}^n$. Finally, the matrix D_t (defined in ($\otimes$) preceding Theorem 50) is non-singular by Theorem 50, thus φ^{-1} is C^∞ by the Inverse Function Theorem. Since φ^{-1} is also C^∞, we conclude φ is a diffeomorphism of $\mathbb{R}^n$. □

9 General Stochastic Exponentials and Linear Equations

Let Z be a given continuous semimartingale with $Z_0 = 0$ and let $\mathcal{E}(Z)_t$ denote the unique solution of the **stochastic exponential equation**

$$X_t = 1 + \int_0^t X_s dZ_s. \tag{$*$}$$

Then $X_t = \mathcal{E}(Z)_t = \exp\{Z_t - \frac{1}{2}[Z,Z]_t\}$ (cf., Theorem 37 of Chap. II). It is of course unusual to have a closed form solution of a stochastic differential equation, and it is therefore especially nice to be able to give an explicit solution of the stochastic exponential equation *when it also has an exogenous driving term.* That is, we want to consider equations of the form

$$X_t = H_t + \int_0^t X_{s-} dZ_s, \qquad (**)$$

where $H \in \mathbb{D}$ (càdlàg and adapted), and Z is a continuous semimartingale. A unique solution of $(**)$ exists by Theorem 7. It is written $\mathcal{E}_H(Z)$.

Theorem 52. *Let H be a semimartingale and let Z be a continuous semimartingale with $Z_0 = 0$. Then the solution $\mathcal{E}_H(Z)$ of equation $(**)$ is given by*

$$\mathcal{E}_H(Z)_t = \mathcal{E}(Z)_t\{H_0 + \int_{0+}^t \mathcal{E}(Z)_s^{-1} d(H_s - [H,Z]_s)\}.$$

Proof. We use the method of "variation of constants." Assume the solution is of the form $X_t = C_t U_t$, where $U_t = \mathcal{E}(Z)_t$, the normal stochastic exponential. The process C is càdlàg while U is continuous. Using integration by parts,

$$\begin{aligned} dX_t &= C_{t-} dU_t + U_t dC_t + d[C,U]_t \\ &= C_{t-} U_t dZ_t + U_t dC_t + U_t d[C,Z]_t \\ &= X_{t-} dZ_t + U_t d\{C_t + [C,Z]_t\}. \end{aligned}$$

If X is the solution of $(**)$, then equating the above with $(**)$ yields

$$dH_t + X_{t-} dZ_t = X_{t-} dZ_t + U_t d\{C_t + [C,Z]_t\}$$

or

$$dH_t = U_t d\{C_t + [C,Z]_t\}.$$

Since U is an exponential it is never zero and $1/U$ is locally bounded. Therefore

$$\frac{1}{U_t} dH_t = dC_t + d[C,Z]_t. \qquad (***)$$

Calculating the quadratic covariation of each side with Z and noting that $[[C,Z],Z] = 0$, we conclude

$$[\frac{1}{U} \cdot H, Z] = [C,Z].$$

Therefore equation $(***)$ becomes

$$\frac{1}{U} dH = dC + \frac{1}{U} d[H,Z],$$

and $C_t = \int_0^t U_s^{-1} d(H_s - [H,Z]_s)$. Recall that $U_t = \mathcal{E}(Z)_t$ and $X_t = C_t U_t$, and the theorem is proved. $\square$

Since $\mathcal{E}(Z)_t^{-1} = 1/\mathcal{E}(Z)_t$ appears in the formula for $\mathcal{E}_H(Z)$, it is worthwhile to note that (for Z a continuous semimartingale)

$$d\left(\frac{1}{\mathcal{E}(Z)}\right) = \frac{dZ - d[Z,Z]}{\mathcal{E}(Z)}$$

and also

$$\frac{1}{\mathcal{E}(Z)} = \mathcal{E}(-Z + [Z,Z]).$$

A more complicated formula for $\mathcal{E}_H(Z)$ exists when Z is not continuous (see Yoeurp-Yor [236]). The next theorem generalizes Theorem 52 to the case where H is not necessarily a semimartingale.

Theorem 53. *Let H be càdlàg, adapted (i.e., $H \in \mathbb{D}$), and let Z be a continuous semimartingale with $Z_0 = 0$. Let $X_t = \mathcal{E}_H(Z)_t$ be the solution of*

$$X_t = H_t + \int_0^t X_{s-} dZ_s.$$

Then $X_t = \mathcal{E}_H(Z)_t$ is given by

$$X_t = H_t + \mathcal{E}(Z)_t \int_0^t \mathcal{E}(Z)_s^{-1}(H_{s-} dZ_s - H_{s-} d[Z,Z]_s).$$

Proof. Let $Y_t = X_t - H_t$. Then Y satisfies

$$\begin{aligned} Y_t &= \int_0^t H_{s-} dZ_s + \int_0^t Y_{s-} dZ_s \\ &= K_t + \int_0^t Y_{s-} dZ_s, \end{aligned}$$

where K is the semimartingale $H_- \cdot Z$. By Theorem 52,

$$Y_t = \mathcal{E}(Z)_t \{K_0 + \int_{0+}^t \mathcal{E}(Z)_s^{-1} d(K_s - [K,Z]_s)\}$$

and since $K_0 = 0$ and $[K,Z]_t = \int_0^t H_{s-} d[Z,Z]_s$,

$$Y_t = \mathcal{E}(Z)_t \int_0^t \mathcal{E}(Z)_s^{-1}(H_{s-} dZ_s - H_{s-} d[Z,Z]_s),$$

from which the result follows. □

Theorem 54 uses the formula of Theorem 52 to give a pretty result on the comparison of solutions of stochastic differential equations.

Lemma. Suppose that F is functional Lipschitz such that if $X_t(\omega) = 0$, then $F(X)_{t-}(\omega) > 0$ for continuous processes X. Let C be a continuous increasing process and let X be the solution of

$$X_t = x_0 + \int_{0+}^{t} F(X)_{s-} dC_s,$$

with $x_0 > 0$. Then $P\{\exists\, t > 0 : X_t \leq 0\} = 0$.

Proof. Let $T = \inf\{t > 0 : X_t = 0\}$. Since $X_0 \geq 0$ and X is continuous, $X_s \geq 0$ for all $s < T$ on $\{T < \infty\}$. The hypotheses then imply that $F(X)_{T-} > 0$ on $\{T < \infty\}$, which is a contradiction. □

Comment. In the previous lemma if one allows $x_0 = 0$, then it is necessary to add the hypothesis that C be strictly increasing at 0. One then obtains the same conclusion.

Theorem 54 (Comparison Theorem). *Let $(Z^\alpha)_{1 \leq \alpha \leq m}$ be continuous semimartingales with $Z_0^\alpha = 0$, and let F_α be process Lipschitz. Let A be a continuous, adapted process with increasing paths, strictly increasing at $t = 0$. Let G and H be process Lipschitz functionals such that $G(X)_{t-} > H(X)_{t-}$ for any continuous semimartingale X. Finally, let X and Y be the unique solutions of*

$$X_t = x_0 + \int_{0+}^{t} G(X)_{s-} dA_s + \int_0^t F(X)_{s-} dZ_s,$$
$$Y_t = y_0 + \int_{0+}^{t} H(Y)_{s-} dA_s + \int_0^t F(Y)_{s-} dZ_s$$

where $x_0 \geq y_0$ and F and Z are written in vector notation. Then $P\{\exists\, t > 0 : X_t \leq Y_t\} = 0$.

Proof. Let

$$U_t = X_t - Y_t,$$
$$N_t = \int_0^t \{F(X)_{s-} - F(Y)_{s-}\}(X_s - Y_s)^{-1} 1_{\{X_s \neq Y_s\}} dZ_s, \text{ and}$$
$$C_t = x_0 - y_0 + \int_{0+}^{t} \{G(X)_{s-} - H(Y)_{s-}\} dA_s.$$

Then $U_t = C_t + \int_0^t U_{s-} dN_s$, and by Theorem 52

$$U_t = \mathcal{E}(N)_t \{(x_0 - y_0) + \int_{0+}^{t} \mathcal{E}(N)_s^{-1} dC_s\}.$$

Next set

$$V_t = \frac{1}{\mathcal{E}(N)_t} U_t,$$

and define the operator K on continuous processes W by

$$K(W) = G(W\mathcal{E}(N) + Y) - H(Y).$$

Note that since G and H are process Lipschitz, if $W_t = 0$ then $G(W\mathcal{E}(N) + Y)_{t-} = G(Y)_{t-}$. Therefore K has the property that $W_t(w) = 0$ implies that $K(W)_{t-} > 0$. Note further that $K(V) = G(U + Y) - H(Y) = G(X) - H(Y)$.

Next observe that

$$\begin{aligned} V_t &= \frac{1}{\mathcal{E}(N)_t} U_t \\ &= x_0 - y_0 + \int_{0+}^t \mathcal{E}(N)_s^{-1} dC_s \\ &= x_0 - y_0 + \int_{0+}^t \mathcal{E}(N)_s^{-1} \{G(X)_{s-} - H(Y)_{s-}\} dA_s \\ &= x_0 - y_0 + \int_{0+}^t K(V)_{s-} \mathcal{E}(N)_s^{-1} dA_s \\ &= x_0 - y_0 + \int_0^t K(V)_{s-} dD_s, \end{aligned}$$

where $D_t = \int_{0+}^t \mathcal{E}(N)_s^{-1} dA_s$. Then D is a continuous, adapted, increasing process which is strictly increasing at zero. Since V satisfies an equation of the type given in the lemma, we conclude that a.s. $V_t > 0$ for all t. Since $\mathcal{E}(N)_t^{-1}$ is strictly positive (and finite) for all $t \geq 0$, we conclude $U_t > 0$ for all $t > 0$, hence $X_t > Y_t$ for all $t > 0$. □

Comment. If $x_0 > y_0$ (i.e., $x_0 = y_0$ is not allowed), then the hypothesis that A is strictly increasing at 0 can be dropped.

The theory of flows can be used to generalize the formula of Theorem 52. In particular, the homeomorphism property is used to prove Theorem 55.

Consider the system of linear equations given by

$$X_t = H_t + \sum_{j=1}^m \int_0^t A_{s-}^j X_{s-} dZ_s^j \tag{$*4$}$$

where $H = (H^i)$, $1 \leq i \leq n$ is a vector of n semimartingales, X takes values in $\mathbb{R}^n$, and A^j is an $n \times n$ matrix of adapted, càdlàg processes. The processes Z^j, $1 \leq j \leq m$, are given, continuous semimartingales which are zero at zero.

Define the operators F_j on $\mathbb{D}^n$ by

$$F(X)_t = A_t^j X_t$$

where A^j is the $n \times n$ matrix specified above. The operators F_j are essentially process Lipschitz. (The Lipschitz processes can be taken to be $\|A_t^j\|$ which is càdlàg, not càglàd, but this is unimportant since one takes $F(X)_{t-}$ in the equation.)

Before examining equation (*4), consider the simpler system ($1 \leq i, k \leq n$),

$$U_t^{i,k} = \delta_k^i + \sum_{j=1}^m \int_0^t \sum_{\ell=1}^n (A_{i,\ell}^j)_{s-} U_{s-}^{\ell,k} dZ_s^j$$

where

$$\delta_k^i = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases}$$

Letting I denote the $n \times n$ identity matrix and writing the preceding in matrix notation yields

$$U_t = I + \sum_{j=1}^m \int_0^t A_{s-}^j U_{s-} dZ_s^j, \tag{*5}$$

where U takes its values in the space of $n \times n$ matrices of adapted processes in $\mathbb{D}$.

Theorem 55. *Let A^j, $1 \leq j \leq m$, be $n \times n$ matrices of càdlàg, adapted processes, and let U be the solution of (*5). Let X_t^x be the solution of (*4) where $H_t = x$, $x \in \mathbb{R}^n$. Then $X_t^x = U_t x$ and for almost all ω, for all t and x the matrix $U_t(\omega)$ is invertible.*

Proof. Note that U_t is an $n \times n$ matrix for each (t, ω) and $x \in \mathbb{R}^n$, so that $U_t x$ is in $\mathbb{R}^n$. If $X_t^x = U_t x$, then since the coefficients are process Lipschitz we can apply Theorem 46 (which says that the flow is a homeomorphism of $\mathbb{R}^n$) to obtain the invertibility of $U_t(\omega)$.

Note that U is also a right stochastic exponential. Indeed, $U = \mathcal{E}^R(V)$, where $V_t = \int_0^t \sum_{j=1}^m A_{s-}^j dZ_s^j$, and therefore the invertibility also follows from Theorem 48.

Thus we need to show only that $X_t^x = U_t x$. Since $U_t x$ solves (*4) with $H_t = x$, we have $U_t x = X_t^x$ a.s. for each x. Note that a.s. the function $x \mapsto U(\omega)x$ is continuous from $\mathbb{R}^n$ into the subspace of $\mathcal{D}^n$ consisting of continuous functions; in particular $(t, x) \mapsto U_t(\omega)x$ is continuous. Also as shown in the proof of Theorem 46, $(x, t) \mapsto X_t^x$ is continuous in x and right continuous in t. Since $U_t x = X_t^x$ a.s. for each fixed x and t, the continuity permits the removal of the dependence of the exceptional set on x and t. □

Let U^{-1} denote the $n \times n$ matrix-valued process with continuous trajectories a.s. defined by $(U^{-1})_t(\omega) = (U_t(\omega))^{-1}$.

Recall equation (*4)

$$X_t = H_t + \sum_{j=1}^m \int_0^t A_{s-}^j X_{s-} dZ_s^j, \tag{*4}$$

where H is a column vector of n semimartingales and $Z_0^j = 0$. Let $[H, Z^j]$ denote the column vector of n components, the ith one of which is $[H^i, Z^j]$.

Theorem 56. *Let H be a column vector of n semimartingales, Z^j $(1 \leq j \leq m)$ be continuous semimartingales with $Z_0^j = 0$, and let A^j, $1 \leq j \leq m$ be $n \times n$ matrices of processes in $\mathbb{D}$. Let U be the solution of equation (*5). Then the solution X^H of (*4) is given by*

$$X_t^H = U_t H_0 + U_t \int_{0+}^t U_s^{-1}(dH_s - \sum_{j=1}^m A_{s-}^j d[H, Z^j]_s)$$

Proof. Write X^H as the matrix product UY. Recall that U^{-1} exists by Theorem 48, hence $Y = U^{-1}X^H$ is a semimartingale, that we need to find explicitly. Using matrix notation throughout, we have

$$d(UY) = dH + \sum_{j=1}^m A_-^j X_- dZ^j.$$

Integration by parts yields (recall that U is continuous)

$$(dU)Y_- + U(dY) + d[U, Y] = dH + \sum_{j=1}^m A_-^j U_- Y_- dZ^j,$$

by replacing X with UY on the right side above. However U satisfies (*5) and therefore

$$(dU)Y_- = \sum_{j=1}^m A_-^j U_- Y_- dZ^j,$$

and combining this with the preceding gives

$$U(dY) + d[U, Y] = dH,$$

or equivalently

$$dY = U^{-1} dH - U^{-1} d[U, Y]. \tag{*6}$$

Taking the quadratic covariation of the preceding equation with Z, we have

$$d[Y, Z^j] = U^{-1} d[H, Z^j],$$

since $[U^{-1}d[U, Y], Z^j] = 0$, $1 \leq j \leq m$. However since U satisfies (*5),

$$\begin{aligned} d[U, Y] &= \sum_{j=1}^m A_-^j U_- d[Y, Z^j] = \sum_{j=1}^m A_-^j U U^{-1} d[H, Z^j] \\ &= \sum_{j=1}^m A_-^j d[H, Z^j], \end{aligned}$$

since U equals U_-. Substitute the above expression for $d[U,Y]$ into (*6) and we obtain

$$dY = U^{-1}(dH - \sum_{j=1}^{m} A^j_- d[H, Z^j]),$$

and since $X^H = UY$, the theorem is proved. □

10 Flows as Diffeomorphisms: The General Case

In this section we study the same equations as in Sect. 8, namely

$$X_t^i = x_i + \sum_{\alpha=1}^{m} \int_0^t F_\alpha^i(X)_{s-} dZ_s^\alpha, \quad 1 \le i \le n, \tag{$*$}$$

except that the semimartingales $(Z^\alpha)_{1\le\alpha\le m}$ *are no longer assumed to be continuous.* For simplicity we still assume that $\mathbf{Z}_0 = 0$. In the general case it is not always true that the flows of solutions are diffeomorphisms of $\mathbb{R}^n$, as the following example shows.

Example. Consider the exponential equation in $\mathbb{R}$,

$$X_t = x + \int_0^t X_{s-} dZ_s.$$

Let Z be a semimartingale, $Z_0 = 0$, such that Z has a jump of size -1 at a stopping time T, $T > 0$ a.s. Then all trajectories, starting at any initial value x, become zero at T and stay there after T, as is trivially seen by the closed form of the solution with initial condition x:

$$X_t = x \ \exp\{Z_t - \frac{1}{2}[Z,Z]_t^c\} \prod_{0<s\le t} (1 + \Delta Z_s)e^{-\Delta Z_s}.$$

Therefore, injectivity of the flow fails, and the flow cannot be a diffeomorphism of $\mathbb{R}$.

We examine both the injectivity of the flow and when it is a diffeomorphism of $\mathbb{R}^n$. Recall the hypotheses of Sect. 7, to which we add one, denoted (H3).

Hypothesis (H1). *Z^α are given semimartingales with $Z_0^\alpha = 0$, $1 \le \alpha \le m$.*

Hypothesis (H2). *$f_\alpha^i : \mathbb{R}^n \to \mathbb{R}$ are given functions, $1 \le i \le n$, $1 \le \alpha \le m$, and $f(x)$ denotes the $n \times m$ matrix $(f_\alpha^i(x))$.*

The system of equations

$$X_t^i = x^i + \sum_{\alpha=1}^{m} \int_0^t f_\alpha^i(X_{s-}) dZ_s^\alpha, \quad 1 \le i \le n \tag{$*$}$$

may also be written

$$X_t = x + \int_0^t f(X_{s-})dZ_s \tag{$*$}$$

where X_t and x are column vectors in $\mathbb{R}^n$, $f(X_{s-})$ is an $n \times m$ matrix, and Z is a column vector of m semimartingales.

Hypothesis (H3). f is $\mathcal{C}^\infty$ and has bounded derivatives of all orders.

Note that by Theorem 40 of Sect. 7, under (H3) the flow is $\mathcal{C}^\infty$. The key to studying the injectivity (and diffeomorphism properties) is an analysis of the jumps of the semimartingale driving terms.

Choose an $\varepsilon > 0$, the actual size of which is yet to be determined. For $(Z^\alpha)_{1\le\alpha\le m}$ we can find stopping times $0 = T_0 < T_1 < T_2 < \cdots$ tending a.s. to ∞ such that

$$Z^{\alpha,j} = (Z^\alpha)^{T_j-} - (Z^\alpha)^{T_{j-1}}$$

have an $\underline{\underline{H}}^\infty$ norm less than ε (cf., Theorem 5). Note that by Theorem 1, $[Z^{\alpha,j}, Z^{\alpha,j}]_\infty^{1/2} < \varepsilon$ as well, hence the jumps of each $Z^{\alpha,j}$ are smaller than ε. Therefore all of the "large" jumps of $Z^{\alpha,j}$ occur only at the times (T_j), $j \ge 1$.

Let $X_t^j(x)$ denote the solution of $(*)$ driven by the semimartingales $Z^{\alpha,j}$. Outside of the interval (T_{j-1}, T_j) the solution is

$$X_t^j(x) = \begin{cases} x, & \text{if } t \le T_{j-1}, \\ X_{T_j-}^j(x), & \text{if } t \ge T_j. \end{cases}$$

Next define the **linkage operators**

$$H^j(x) = x + f(x)\Delta Z_{T_j},$$

using vector and matrix notation. We have the following obvious result.

Theorem 57. *The solution X of $(*)$ is equal to, for $T_n \le t < T_{n+1}$,*

$$X_t(x) = X_t^{n+1}(x_n+),$$

where

$$\begin{aligned} x_0+ &= x & & \\ x_1- &= X_{T_1-}^1(x), & x_1+ &= H^1(x_1-) \\ x_2- &= X_{T_2-}^2(x_1+), & x_2+ &= H^2(x_2-) \\ &\ \vdots & &\ \vdots \\ x_n- &= X_{T_n-}^n(x_{(n-1)}+), & x_n+ &= H^n(x_n-). \end{aligned}$$

Theorem 58. *The flow $\varphi : x \to X_t(x,\omega)$ of the solution X of $(*)$ is a diffeomorphism if the collections of functions*

$$x \mapsto X_t^j(x,\omega) \qquad x \mapsto H^j(x,\omega)$$

are diffeomorphisms.

Proof. By Theorem 57, the solution $(*)$ can be constructed by composition of functions of the types given in the theorem. Since the composition of diffeomorphisms is a diffeomorphism, the theorem is proved. □

We begin by studying the functions $x \to X^j_{T_j}(x,\omega)$ and $x \to X^{n+1}_t(x,\omega)$. Note that by our construction and choice of the times T_j, *we need only to consider the case where $Z = Z^j$ has a norm in $\underline{\underline{H}}^\infty$ smaller than ε.*

The following classical result, due to Hadamard, underlies our analysis.

Theorem 59 (Hadamard's Theorem). *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be C^∞. Suppose*

(i) $\lim_{\|x\|\to\infty} \|g(x)\| = \infty$, and
(ii) the Jacobian matrix $g'(x)$ is an isomorphism of $\mathbb{R}^n$ for all x.

Then g is a diffeomorphism of $\mathbb{R}^n$.

Proof. By the Inverse Function Theorem the function g is a local diffeomorphism, and hence it suffices to show it is a bijection of $\mathbb{R}^n$.

To show that g is onto (i.e., a surjection), first note that $g(\mathbb{R}^n)$ is open and non-empty. It thus suffices to show that $g(\mathbb{R}^n)$ is a closed subset of $\mathbb{R}^n$, since $\mathbb{R}^n$ itself is the only nonempty subset of $\mathbb{R}^n$ that is open and closed. Let $(x_i)_{i\geq 1}$ be a sequence of points in $\mathbb{R}^n$ such that $\lim_{i\to\infty} g(x_i) = y \in \mathbb{R}^n$. We will show that $y \in g(\mathbb{R}^n)$. Let $x_i = t_i v_i$, where $t_i > 0$ and $\|v_i\| = 1$. By choosing a subsequence if necessary we may assume that v_i converges to $v \in S^n$, the unit sphere, as i tends to ∞. Next observe that the sequence $(t_i)_{i\geq 1}$ must be bounded by condition (i) in the theorem: for if not, then $t_i = \|x_i\|$ tends to ∞ along a subsequence and then $\|g(x_{i_k})\|$ tends to ∞ by (i), which contradicts that $\lim_{i\to\infty} g(x_i) = y$. Since $(t_i)_{i\geq 1}$ is bounded we may assume $\lim_{i\to\infty} t_i = t_0 \in \mathbb{R}^n$ again by taking a subsequence if necessary. Then $\lim_{i\to\infty} x_i = t_0 v$, and by the continuity of g we have $y = \lim_{i\to\infty} g(x_i) = g(t_0 v)$.

To show g is injective (i.e., one-to-one), we first note that g is a local homeomorphism, and moreover g is finite-to-one. Indeed, if there exists an infinite sequence $(x_n)_{n\geq 1}$ such that $g(x_n) = y_0$, all n, for some y_0, then by condition (i) the sequence must be bounded in norm and therefore have a cluster point. By taking a subsequence if necessary we can assume that x_n tends to $\hat{x}$ (the cluster point), where $g(x_n) = y_0$, all n. By the continuity of g we have $g(\hat{x}) = y_0$ as well. This then violates the condition that g is a local homeomorphism, and we conclude that g is finite-to-one.

Since g is a finite-to-one surjective homeomorphism, it is a covering map.[16] However since $\mathbb{R}^n$ is simply connected the only covering space of $\mathbb{R}^n$ is $\mathbb{R}^n$ (the fundamental group of $\mathbb{R}^n$ is trivial). Therefore the fibers $g^{-1}(x)$ for $x \in \mathbb{R}^n$ each consist of one point, and g is injective. □

The next step is to show that the functions $x \mapsto X^j_{T_j}(x,\omega)$ and $x \mapsto X^{n+1}_t(x,\omega)$ of Theorem 58 satisfy the two conditions of Theorem 59 and are

[16] For the algebraic topology used here, the reader can consult, for example, Munkries [183, Chapter 8].

thus diffeomorphisms. This is done in Theorems 62 and 64. First we give a result on weak injectivity which is closely related to Theorem 41.

Theorem 60. *Let Z^α be semimartingales, $1 \le \alpha \le m$ with $Z_0^\alpha = 0$, and let F be an $n \times m$ matrix of process Lipschitz operators with non-random Lipschitz constant K. Let $H^i \in \mathbb{D}$, $1 \le i \le n$ (càdlàg, adapted). If $\sum_{\alpha=1}^m \|Z^\alpha\|_{\underline{\underline{H}}^\infty} < \varepsilon$, for $\varepsilon > 0$ sufficiently small, then the flow of the solution of*

$$X_t = x + H_t + \int_0^t F(X)_{s-} dZ_s$$

is weakly injective.[17]

Proof. Let $x, y \in \mathbb{R}^n$, and let X^x, X^y denote the solutions of the above equation with initial conditions x, y, respectively. Let $u = x - y$, $U = X^x - X^y$, and $V = F(X^x)_- - F(X^y)_-$. Then $V \in \mathbb{L}$ and $|V| \le K|U_-|$. Also,

$$U_t = u + \int_0^t V_s dZ_s.$$

Therefore $\Delta U_s = \sum_\alpha V_s^\alpha \Delta Z_s^\alpha$ and moreover (using the Einstein convention to leave the summations implicit)

$$\begin{aligned} \|\Delta U_s\| &\le \|V_s^\alpha\| \|\Delta Z_s^\alpha\| \\ &\le C\|U_{s-}\|\varepsilon \\ &< \frac{1}{2}\|U_{s-}\| \end{aligned}$$

if ε is small enough. Consequently $\|U_s\| \ge \frac{1}{2}\|U_{s-}\|$. Define $T = \inf\{t > 0 : U_{t-} = 0\}$. Then $U_{t-} \ne 0$ on $[0, T)$ and the above implies $U_t \ne 0$ on $[0, T)$ as well. Using Itô's formula for $f(x) = \log\|x\|$, as in the proof of Theorem 41 we have

$$\begin{aligned} &\log\|U_t\| - \log\|u\| \qquad (**) \\ &= \int_0^t \frac{U_{s-}^i V_s^{i,\alpha}}{\|U_{s-}\|^2} dZ_s^\alpha + \frac{1}{2}\int_0^t \frac{V_s^{i,\alpha} V_s^{i,\beta}}{\|U_{s-}\|^2} d[Z^\alpha, Z^\beta]_s^c \\ &\quad - \int_0^t \frac{U_{s-}^i U_{s-}^j V_s^{i,\alpha} V_s^{j,\beta}}{\|U_{s-}\|^4} d[Z^\alpha, Z^\beta]_s^c \\ &\quad + \sum_{0<s\le t} \{\log\|U_s\| - \log\|U_{s-}\| - \frac{U_{s-}^i}{\|U_{s-}\|^2} \Delta U_s^i\}. \end{aligned}$$

For s fixed, let

[17] We are using vector and matrix notation, and the Einstein convention on sums. The Einstein convention is used throughout this section.

$$J_s = \{\log \|U_s\| - \log \|U_{s-}\| - \frac{U^i_{s-}}{\|U_{s-}\|^2}\Delta U^i_s\},$$

so that the last sum on the right side of equation (**) can be written $\sum_{0<s\leq t} J_s$. By Taylor's Theorem

$$|J_s| \leq C\|\Delta U_s\|^2 \sup_{\substack{i,j \\ x\in I_s}} |\frac{\partial^2 f}{\partial x_i \partial x_j}(x)|$$

where $f(x) = \log \|x\|$, and I_s denotes the segment with extremities U_s and U_{s-}. Since

$$\sup_{\substack{i,j \\ x\in I_s}} |\frac{\partial^2 f}{\partial x_i \partial x_j}(x)| \leq \sup_{x\in I_s} \left(\frac{1}{\|x\|^2}\right),$$

and since $\|\Delta U_s\| \leq \frac{1}{2}\|U_{s-}\|$, we deduce

$$\sup_{x\in I_s} \left(\frac{1}{\|x\|^2}\right) \leq \frac{4}{\|U_{s-}\|^2}$$

which in turn implies

$$|J_s| \leq C\|U_{s-}\|^2\|\Delta Z_s\|^2 \frac{4}{\|U_{s-}\|^2}.$$

Therefore

$$\sum_{0<s\leq t} |J_s| \leq C \sum_{0<s\leq t} \|\Delta Z_s\|^2 \leq C\sum_\alpha [Z^\alpha, Z^\alpha]_t < \infty$$

on $[0,T)$. Returning to (**), as t increases to T, the left side tends to $-\infty$ on $\{T<\infty\}$ and the right side remains finite. Therefore $P(T<\infty)=0$, and U and U_- never vanish, which proves the theorem. □

Theorem 61. *Let $(Z^\alpha)_{1\leq\alpha\leq m}$ be semimartingales, $Z^\alpha_0 = 0$, F an $n\times m$ matrix of process Lipschitz operators with a non-random Lipschitz constant, and $H^i \in \mathbb{D}$, $1\leq i\leq n$. If $\sum_{\alpha=1}^m \|Z^\alpha\|_{\underline{H}^\infty} < \varepsilon$ for $\varepsilon>0$ sufficiently small, then for $r\in\mathbb{R}$ there exist uniformly locally bounded predictable processes $H^\alpha(x,y)$ and $K^{\alpha,\beta}(x,y)$, which depend on r, such that*

$$\|X^x - X^y\|^r = \|x-y\|^r \mathcal{E}(\Lambda_r(x,y))$$

where X^x is the solution of

$$X_t = x + H_t + \int_0^t F(X)_{s-}dZ_s.$$

The semimartingale Λ_r is given by

$$\Lambda_r(x,y)_t = \int_0^t H_s^\alpha(x,y)dZ_s^\alpha + \int_0^t K_s^{\alpha,\beta}(x,y)d[Z^\alpha, Z^\beta]_s^c + J_t$$

where $J_t = \sum_{0<s\leq t} A_s$, *and where* A_s *is an adapted process such that* $|A_s| \leq C_r(\Delta Z_s^\alpha)^2$.

Proof. Fix $x, y \in \mathbb{R}^n$, $x \neq y$, and let $U = X^x - X^y$, $V = F(X^x)_- - F(X^y)_-$. By Theorem 60 U is never zero. As in the proof of Theorem 43, by Itô's formula,

$$\begin{aligned}\|U_y\|^r &= \|x-y\|^r + \int_0^t r\|U_{s-}\|^{r-2}U_{s-}^i dU_s^i \\ &+ \frac{1}{2}\int_0^t r\{(r-2)\|U_{s-}\|^{r-4}U_{s-}^i U_{s-}^j + \delta_j^i\|U_{s-}\|^{r-2}\}d[U^i, U^j]_s^c \\ &+ \sum_{0<s\leq t}\{\|U_s\|^r - \|U_{s-}\|^r - r\|U_{s-}\|^{r-2}U_{s-}^i\Delta U_s^i\}.\end{aligned}$$

Let L_s denote the summands of the last sum on the right side of the above equation. If $g(x) = \|x\|^r$, then

$$|L_s| \leq C\|\Delta U_s\|^2 \sup_{\substack{i,j \\ x\in I_s}} |\frac{\partial^2 g}{\partial x_i \partial x_j}|,$$

where I_s is the segment with boundary U_s and U_{s-}. However,

$$|\frac{\partial^2 g}{\partial x_i \partial x_j}| = r|(r-2)\|x\|^{r-4}x_i x_j + \delta_j^i\|x\|^{r-2}|$$

which is less than $C_r\|x\|^{r-2}$. However as in the proof of Theorem 60, we have $\|\Delta U_s\| \leq \frac{1}{2}\|U_{s-}\|$, which implies (for all r positive or negative) that $\|x\|^r \leq C_r\|U_{s-}\|^r$ for all x between U_{s-} and U_s. Hence (the constant C changes)

$$\begin{aligned}|L_s| &\leq C\|\Delta U_s\|^2\|U_{s-}\|^{r-2} \\ &\leq C\|U_{s-}\|^r\|\Delta Z_s\|^2.\end{aligned}$$

Therefore, let $A_s = \|U_{s-}\|^{-r}L_s$, and we have

$$\sum_{0<s\leq t}|A_s| \leq \sum_{0<s\leq t} C\|\Delta Z_s\|^2 \leq C\sum_\alpha [Z^\alpha, Z^\alpha]_t,$$

an absolutely convergent series with a bound independent of (x, y). To complete the proof it suffices to take

$$H_s^\alpha(x,y) = r\|U_{s-}\|^{-2}(U_{s-}, V_s^\alpha)1_{\{U_{s-}\neq 0\}}$$

and

$$K_s^{\alpha,\beta}(x,y) = \frac{1}{2}r\{(r-2)\|U_{s-}\|^4(U_{s-},V_s^\alpha)(U_{s-},V_s^\beta) + \|U_{s-}\|^{-2}(V_s^\alpha,V_s^\beta)\}1_{\{U_{s-}\neq 0\}},$$

as in the proof of Theorem 43. Note that $1_{\{U_{s-}\neq 0\}}$ is indistinguishable from the zero process by weak injectivity (Theorem 60). These choices for H^α and $K^{\alpha,\beta}$ are easily seen to work by observing that $U_t^i = \sum_{\alpha=1}^m \int_0^t V_s^{i,\alpha} dZ_s^\alpha$, and the preceding allows us to conclude that

$$\|U_t\|^r = \|x-y\|^r + \int_0^t \|U_{s-}\|^r dA_r(x,y)_s,$$

and the result follows. □

Theorem 62. *Let $(Z^\alpha)_{1\leq\alpha\leq m}$ be semimartingales, $Z_0^\alpha = 0$, F an $n\times m$ matrix of process Lipschitz operators with a non-random Lipschitz constant, and $H^i \in \mathbb{D}$, $1\leq i\leq n$. Let $X = X(t,\omega,x)$ be the solution of*

$$X_t = x + \int_0^t F(X)_{s-}dZ_s. \tag{$*$}$$

If $\sum_{\alpha=1}^m \|Z^\alpha\|_{\underline{\underline{H}}^\infty} < \varepsilon$ for $\varepsilon > 0$ sufficiently small, then for each $N<\infty$ and almost all ω

$$\lim_{\|x\|\to\infty} \inf_{s\leq N} \|X(s,\omega,x)\| = \infty.$$

Proof. The proof is essentially the same as that of Theorem 45, so we only sketch it. For $x\neq 0$ let $Y^x = \|X^x - X^0\|^{-1}$, which is well-defined by weak injectivity (Theorem 60). Then

$$|Y^x| = \|x\|^{-1}\mathcal{E}(A_{-1}(x,0))$$

$$\begin{aligned}|Y^x - Y^y| &\leq \|X^x - X^y\|\|X^x - X^0\|^{-1}\|X^y - X^0\|^{-1}\\ &= \|x-y\|\|x\|^{-1}\|y\|^{-1}\mathcal{E}(A_1(x,y))\mathcal{E}(A_{-1}(x,0))\mathcal{E}(A_{-1}(y,0))\end{aligned}$$

by Theorem 61, where $A_r(x,y)$ is as defined in Theorem 61. Set $Y^\infty = 0$.

Since $\|Z\|_{\underline{\underline{H}}^\infty} < \varepsilon$ each Z^α has jumps bounded by ε, and the process J_t defined in Theorem 61 also has jumps bounded by $C_r\varepsilon^2$. Therefore we can stop the processes $A_r(x,y)$ at an appropriately chosen sequence of stopping times $(T_\ell)_{\ell\geq 1}$ increasing to ∞ a.s. such that each $A_r(x,y)\in\mathcal{S}(\varepsilon)$ for a given ε, and for each ℓ, uniformly in (x,y). However if Z is a semimartingale in $\mathcal{S}(\varepsilon)$, then since $\mathcal{E}(A_r(x,y))$ satisfies the equation $U_t = 1 + \int_{0+}^t U_{s-}dA_r(x,y)_s$, by Lemma 2 of Sect. 3 of this chapter we have

$$\|\mathcal{E}(A_r(x,y))\|_{\underline{\underline{S}}^p} \leq C(p,z) < \infty,$$

where $C(p,z)$ is a constant depending on p and $z = \|\Lambda_r(x,y)\|_{\underline{\underline{H}}^\infty} \le k_\ell$, the bound for ℓ, provided of course that ε is sufficiently small. We conclude that for these T_ℓ there exist constants C_ℓ such that

$$\|(Y^x - Y^y)^{T_\ell}\|_{\underline{\underline{S}}^p} \le C_\ell d(x,y),$$

where $p > n$, and where d is the distance on $\mathbb{R}^n \setminus \{0\}$ given by $d(x,y) = \frac{\|x-y\|}{\|x\|\|y\|}$.

Set

$$\hat{Y}^x = \begin{cases} Y^{x\|x\|^{-2}}, & 0 < \|x\| < \infty, \\ Y^\infty = 0, & \|x\| = 0. \end{cases}$$

Then $\|(\hat{Y}^x - \hat{Y}^y)^{T_\ell}\|^p_{\underline{\underline{S}}^p} \le C^p_\ell \|x-y\|^p$ on $\mathbb{R}^n$, and by Kolmogorov's Lemma (Theorem 72 of Chap. IV) we conclude that $\lim_{\|x\|\to 0} \hat{Y}^x$ exists and it is zero. Since $(\hat{Y}^x)^{-1} = \|X^{x\|x\|^{-2}} - X^0\|$, the result follows. □

If φ is the flow of the solution of $(*)$, Theorem 62 shows that

$$\lim_{\|x\|\to\infty} \|\varphi(x)\| = +\infty,$$

and the first condition in Hadamard's Theorem (Theorem 59) is satisfied. Theorem 63 allows us to determine when the second condition in Hadamard's Theorem is also satisfied (see Theorem 64), but it has an independent interest. First, however, some preliminaries are needed.

For given n, let Z be an $n \times n$ matrix of given semimartingales. Recall that $X = \mathcal{E}(Z)$ denotes *the (matrix-valued) exponential of* Z, and that $\mathcal{E}^R(Z)$ denotes *the (matrix-valued) right stochastic exponential of* Z, which was defined in Sect. 8, following Theorem 47.

Recall that in Theorem 48 we showed that if Z is an $n \times n$ matrix of *continuous* semimartingales with $Z_0 = 0$, then $\mathcal{E}(Z)\mathcal{E}^R(-Z + [Z,Z]) = I$, or equivalently $\mathcal{E}(-Z + [Z,Z])\mathcal{E}^R(Z) = I$. The general case is more delicate.

Theorem 63. *Let Z be an $n \times n$ matrix of semimartingales with $Z_0 = 0$. Suppose that $W_t = -Z_t + [Z,Z]^c_t + \sum_{0<s\le t}(I + \Delta Z_s)^{-1}(\Delta Z_s)^2$ is a well-defined semimartingale. Then*

$$\mathcal{E}(W)_t \mathcal{E}^R(Z)_t = I$$

for all $t \ge 0$.

Proof. Let $U = \mathcal{E}(W)$, $V = \mathcal{E}^R(Z)$, and $J_t = \sum_{0<s\le t}(I + \Delta Z_s)^{-1}(\Delta Z_s)^2$. Then,

$$\begin{aligned} dU &= U_-(-dZ + d[Z,Z]^c + dJ), \\ dV &= (dZ)V_-, \end{aligned}$$

and therefore

$$\begin{aligned}(dU)V_- &= U_-(-dZ + d[Z,Z]^c + dJ)V_-,\\ U_-dV &= U_-(dZ)V_-,\\ d[U,V] &= U_-d[W,Z]V_-\\ &= -U_-d[Z,Z]V_- + U_-d[J,Z]V_-,\end{aligned}$$

since $d[Z,[Z,Z]^c] = 0$. By Theorem 47

$$d(UV) = U_-dV + (dU)V_- + d[U,V];$$

using the above calculations several terms cancel, yielding

$$d(UV) = U_-d[Z,Z]^cV_- + U_-(dJ)V_- - U_-d[Z,Z]V_- + U_-d[J,Z]V_-.$$

Since $[Z,Z]_t = [Z,Z]_t^c + \sum_{0<s\le t}(\Delta Z_s)^2$, and

$$[J,Z]_t = \sum_{0<s\le t} \Delta J_s \Delta Z_s = \sum_{0<s\le t} (\Delta Z_s)^3 (I + \Delta Z_s)^{-1},$$

the preceding becomes

$$\begin{aligned}d(UV) = -U_-d(\textstyle\sum(\Delta Z_s)^2)V_- &+ U_-d(\textstyle\sum (\Delta Z_s)^2(I+\Delta Z_s)^{-1})V_-\\ &+ U_-d(\textstyle\sum (\Delta Z_s)^3(I+\Delta Z_s)^{-1})V_-.\end{aligned}$$

Since

$$(\Delta Z_s)^2 = ((\Delta Z_s)^2 + (\Delta Z_s)^3)(I + \Delta Z_s)^{-1},$$

the above equation implies $d(UV)_t = 0$, all $t \ge 0$. However $U_0V_0 = I$, and therefore $U_tV_t = I$, all $t \ge 0$. □

Corollary. Let Z be a square matrix of semimartingales. If $\|Z\|_{\underline{\underline{H}}^\infty} < \varepsilon$ for $\varepsilon > 0$ sufficiently small, then $\mathcal{E}^R(Z)_t$ is invertible for all $t \ge 0$.

Proof. If $\|Z\|_{\underline{\underline{H}}^\infty} < \varepsilon$ then the jumps of Z are bounded by $\varepsilon > 0$ and therefore the process W of Theorem 63 is a well-defined semimartingale. □

Theorem 64. *Let $(Z^\alpha)_{1\le\alpha\le m}$ be semimartingales with $Z_0 = 0$, and let f be a matrix of coefficients satisfying Hypotheses (H2) and (H3). Let X be the unique solution of*

$$X_t = x + \int_0^t f(X_{s-})dZ_s. \qquad (*)$$

The Jacobian matrix $D_k^i(t,\omega,x) = \frac{\partial}{\partial x_k}X^i(t,\omega,x)$ is invertible for each $t \ge 0$ provided $\|Z\|_{\underline{\underline{H}}^\infty} < \varepsilon$ for sufficiently small $\varepsilon > 0$.

Proof. By Theorem 39 (in Sect. 7) the Jacobian matrix D satisfies the right stochastic exponential equation

$$D_t = I + \int_0^t \left(\frac{\partial f_\alpha^i}{\partial x_k}(X_{s-})dZ_s^\alpha\right)D_{s-},$$

and the matrix semimartingale differential $\frac{\partial f_\alpha^i}{\partial x_k}(X_{s-})dZ_s^\alpha$ satisfies the hypotheses of the corollary of Theorem 63, whence the result. □

Before stating the principal result of this section, we need to define two subsets of $\mathbb{R}^m$; recall that under Hypotheses (H1), (H2), and (H3), that $Z = (Z^\alpha)_{1\le\alpha\le m}$ is a given m-tuple of semimartingales and that $f(x) = (f_\alpha^i(x))$ is an $n \times m$ matrix of C^∞ functions. Let

$$\begin{aligned}\mathcal{D} &= \{z \in \mathbb{R}^m : H(x) = x + f(x)z \text{ is a diffeomorphism of } \mathbb{R}^n\}\\ \mathcal{I} &= \{z \in \mathbb{R}^m : H(x) = x + f(x)z \text{ is injective in } \mathbb{R}^n\}.\end{aligned}$$

Clearly $\mathcal{D} \subset \mathcal{I}$.

Theorem 65. *Let Z and f be as given in Hypotheses (H1), (H2), and (H3), and let X be the solution of*

$$X_t = x + \int_0^t f(X_{s-})dZ_s. \qquad (*)$$

The flow of X is a.s. a diffeomorphism of $\mathbb{R}^n$ (resp. trajectories of X from different initial points a.s. never meet) for all t if and only if all the jumps of Z belong to $\mathcal{D}$ (resp. all the jumps of Z belong to $\mathcal{I}$).

Proof. Recall the processes $X_{T_j}^j(x,\omega)$ and $X_t^{n+1}(x,\omega)$ defined in Theorem 57, and the linkage operator $H^j(x) = x + f(x)\Delta Z_{T_j}$, defined immediately preceding Theorem 57. By hypothesis the linkage operators $H^j(x)$ are clearly diffeomorphisms of $\mathbb{R}^n$ (resp. injective), and by Theorems 62 and 64, Hadamard's conditions are satisfied (Theorem 59), and therefore the functions $x \mapsto X_{T_j}^j(x,\omega)$ and $x \mapsto X_t^{n+1}(x,\omega)$ are diffeomorphisms of $\mathbb{R}^n$ if $\varepsilon > 0$ is taken small enough in the definition of the stopping times $(T_j)_{j\ge1}$, which it is always possible to do. Therefore by Theorem 58 the flow $\varphi : x \to X_t(x,\omega)$ is a.s. a diffeomorphism of $\mathbb{R}^n$ for each $t > 0$.

The necessity is perhaps the more surprising part of the theorem. First observe that by Hadamard's Theorem (Theorem 59) *the set $\mathcal{D}$ contains a neighborhood of the origin.* Indeed, if z is small enough and x is large enough then $\|f(x)z\| \le \|x\|/2$ since f is Lipschitz, which implies that $\|H(x)\| \ge \|x\|/2$ and thus condition (i) of Hadamard's Theorem is satisfied. On the other hand $H'(x) = I + f'(x)Z$ is invertible for all x for $\|z\|$ small enough because $f'(x)$ is bounded; therefore condition (ii) of Hadamard's Theorem is satisfied. Since $f(x)$ is C^∞ (Hypothesis (H3)), we conclude that $\mathcal{D}$ contains a neighborhood of the origin.

To prove necessity, set

$$\Gamma_1 = \{\omega : \exists\, s > 0 \text{ with } \Delta Z_s(\omega) \in \mathcal{D}^c\},$$
$$\Gamma_2 = \{\omega : \exists\, s > 0 \text{ with } \Delta Z_s(\omega) \in \mathcal{I}^c\}.$$

Suppose $P(\Gamma_1) > 0$. Since $\mathcal{D}$ contains a neighborhood of the origin, there exists an $\varepsilon > 0$ such that all the jumps of Z less than ε are in $\mathcal{D}$. We also take ε so small that all the functions $x \mapsto X^x_{T_i}(x)$ are diffeomorphisms as soon as the linkage operators H^k are, all $k \leq i$.

Since the jumps of Z smaller than ε are in $\mathcal{D}$, the jumps of Z that are in $\mathcal{D}^c$ must take place at the times T_i. Let

$$\Lambda^j = \{\omega : \Delta Z_{T_i} \in \mathcal{D}, \text{ all } i < j, \text{ and } \Delta Z_{T_j} \in \mathcal{D}^c\}.$$

Since $P(\Gamma_1) > 0$, there must exist a j such that $P(\Lambda^j) > 0$. Then for $\omega \in \Lambda^j$, $x \mapsto X_{T_j-}(x, \omega)$ is a diffeomorphism, but $H^j(x, \omega)$ is not a diffeomorphism. Let $\omega_0 \in \Lambda_j$ and $t_0 = T_j(\omega_0)$. Then $x \mapsto X_{t_0}(x, \omega_0)$ is not a diffeomorphism, and therefore

$$P\{\omega : \exists\, t \text{ such that } x \to X_t(x, \omega) \text{ is not a diffeomorphism}\} > 0,$$

and we are done. The proof of the necessity of the jumps belonging to $\mathcal{I}$ to have injectivity is analogous. □

Corollary. Let Z and f be as given in Hypotheses (H1), (H2), and (H3), and let X be the solution of

$$X_t = x + \int_0^t f(X_{s-})dZ_s.$$

Then different trajectories of X can meet only at the jumps of Z.

Proof. We saw in the proof of Theorem 65 that two trajectories can intersect only at the times T_j that slice the semimartingales Z^α into pieces of $\underline{\underline{H}}^\infty$ norm less than ε. If the Z^α do not jump at T_{j_0} for some j_0, however, and paths of X intersect there, one need only slightly alter the construction of T_{j_0} (cf., the proof of Theorem 5, where the times T_j were constructed), so that T_{j_0} is not included in another sequence that ε-slices $(Z^\alpha)_{1\leq\alpha\leq m}$, to achieve a contradiction. (Note that if, however, $(Z^\alpha)_{1\leq\alpha\leq m}$ has a large jump at T_{j_0}, then it cannot be altered.) □

11 Eclectic Useful Results on Stochastic Differential Equations

We begin this collection of mostly technical results concerning stochastic differential equations with some useful moment estimates. And we begin the moment estimates with preliminary estimates for stochastic integrals. The first result is trivial, and the second is almost as simple.

Lemma. For any predictable (matrix-valued) process H and for any $p > 1$, and for $0 \leq t \leq 1$, we have

$$E\{\sup_{s \leq t} \left| \int_0^s H_s ds \right|^p \} \leq \int_0^t E\{|H_s|^p\} ds.$$

Proof. Since $0 \leq t \leq 1$, we have that ds on $[0,1]$ is a probability measure, and since $f(x) = x^p$ is convex for $p > 1$, we can use Jensen's inequality. The result then follows from Fubini's Theorem. □

Lemma. For any predictable (matrix-valued) process H and multidimensional Brownian motion B, and for any $p \geq 2$, and for $0 \leq t \leq 1$, there exists a constant c_p depending only on p such that

$$E\{\sup_{s \leq t} \left| \int_0^s H_s dB_s \right|^p \} \leq c_p \int_0^t E\{|H_s|^p\} ds.$$

Proof. The proof for the case $p = 2$ is simply Doob's inequality; we give it for the one dimensional case.

$$E\{\sup_{s \leq t} (\int_0^s H_s dB_s)^2\} \leq 4E\{[H \cdot B, H \cdot B]_t\} = 4E\{\int_0^t H_s^2 ds\} = 4 \int_0^t E\{H_s^2\} ds,$$

where the last equality is by Fubini's Theorem. For the case $p > 2$, we use Burkholder's inequality (see Theorem 48 of Chap. IV,

$$E\{\sup_{s \leq t} \left| \int_0^s H_s dB_s \right|^p \} \leq c_p E\{[H \cdot B, H \cdot B]_t^{p/2}\} \leq c_p \int_0^t E\{|H_s|^p\} ds$$

where the last inequality follows from Jensen's inequality (recall that $p \geq 2$ so that $f(x) = x^{p/2}$ is a convex function, and since $0 \leq t \leq 1$ we have used that ds on $[0,1]$ is a probability measure) and Fubini's Theorem. □

Theorem 66. *Let Z be a d-dimensional Lévy process with $E\{\| Z_t \|_{L^p}\} < \infty$ for $0 \leq t \leq 1$, H a predictable (matrix-valued) process and $p \geq 2$. Then there exists a finite constant K_p such that for $0 \leq t \leq 1$ we have*

$$E\{\sup_{s \leq t} \left| \int_0^s H_s dZ_s \right|^p \} \leq K_p \int_0^t E\{|H_s|^p\} ds.$$

Proof. We give the proof for the one dimensional case.[18] Since the Lévy process Z can be decomposed $Z_t = bt + \sigma B_t + M_t$, where M is a purely discontinuous martingale (that is, M is orthogonal in the L^2 sense to the stable subspace generated by continuous L^2 martingales), it is enough to prove the inequality separately for $Z_t = bt$, and $Z_t = \sigma B_t$, and $Z_t = M_t$. The preceding two lemmas prove the first two cases, so we give the proof of only the third.

[18] An alternative (and simpler) proof using random measures can be found in [105].

In the computations below, the constants C_p and the functions $K_p(\cdot)$ vary from line to line. Choose the rational number k such that $2^k \le p < 2^{k+1}$. Applying the Burkholder-Gundy inequalities[19] for $p \ge 2$ we have

$$E\{\left|\int_0^t H_s dM_s\right|^p\} \le (C_p)^p E\{\left|\int_0^t |H_s|^2 d[M,M]_s\right|\}^{p/2}.$$

Set

$$\alpha_M := E\{[M,M]_1\} = E\{\sum_{s\le 1}|\Delta M_s|^2\} = \int |x|^2 \nu_M(dx) < \infty.$$

Since $[M,M]$ is also a Lévy process, we have that $[M,M]_t - \alpha_M t$ is also a martingale. Therefore the inequality above becomes

$$\begin{aligned} E\{\left|\int_0^t H_s dM_s\right|^p\} \le & C_p E\{\left|\int_0^t |H_s|^2 d([M,M]_s - \alpha_M s)\right|^{p/2}\} \\ & + K_p(t)\alpha_M^{p/2} E\{\left|\int_0^t |H_s|^2 ds\right|^{p/2}\}. \end{aligned}$$

We apply a Burkholder-Gundy inequality again to the first term on the right side above to obtain

$$\begin{aligned} E\{\left|\int_0^t H_s dM_s\right|^p\} \le & C_p E\{(\sum_{s\le t}|H_s \Delta M_s|^4)^{p/4}\} \\ & + K_p(t)[\int |x|^2 \nu_M(dx)]^{p/2} E\{\int_0^t |H_s|^p ds\}. \end{aligned}$$

We continue recursively to get

$$\begin{aligned} E\{\left|\int_0^t H_s dM_s\right|^p\} \le & C_p E\{(\sum_{s\le t}|H_s \Delta M_s|^{2^{k+1}})^{p/2^{k+1}}\} \\ & + K_p(t)(\sum_{i=1}^k [\int |x|^{2^i} \nu_M(dx)]^{p2^{-i}}) E\{\int_0^t |H_s|^p ds\}. \end{aligned}$$

Next we use the fact that, for any sequence a such that $\|a\|_{l^q}$ is finite, $\|a\|_{l^2} \le \|a\|_{l^q}$ for $1 \le q \le 2$. As $1 \le p2^{-k} < 2$ we get

$$\begin{aligned} [\sum_{s\le t}|H_s \Delta M_s|^{2^{k+1}}]^{p/2^{k+1}} &= [\sum_{s\le t}(|H_s\Delta M_s|^{2^k})^2]^{\frac{1}{2}\frac{p}{2^k}} \\ &\le \sum_{s\le t}|H_s \Delta M_s|^p \end{aligned}$$

whence

[19] The Davis inequality is for the (important) case $p = 1$ only.

$$E\{\{\sum_{s\leq t}|H_s\Delta M_s|^{2^{k+1}}\}^{p/2^{k+1}}\}\leq E\{\sum_{s\leq t}|H_s\Delta M_s|^p\}.$$

Note that $\sum_{s\leq t}|\Delta M_s|^p$ is an increasing, adapted, càdlàg process, and its (non-random) compensator is $A_t = t\int|x|^p\nu_M(dx)$, which is finite by hypothesis. Since $|H|^p$ is a predictable process,

$$\int_0^t |H_s|^p d(\sum_{r\leq s}|\Delta M_r|^p - A_s)$$

is a martingale with zero expectation. Therefore we have

$$E\{|\int_0^t H_s dM_s|^p\} \leq [C_p\int|x|^p\nu_M(dx) + K_p(t)\sum_{i=1}^k(\int|x|^{2^i}\nu_M(dx))^{p2^{-i}}] \cdot E\{\int_0^t |H_s|^p ds\}$$

It remains to show that, for any $1\leq i\leq k$,

$$(\int|x|^{2^i}\nu_M(dx))^{p2^{-i}}\leq(\int|x|^2\nu_M(dx))^{p/2}+\int|x|^p\nu_M(dx).$$

Let $\lambda_M := \int|x|^2\nu_M(dx)$, so that

$$\mu_M(dx) := \frac{1}{\lambda_M}|x|^2\nu_M(dx)$$

is a probability measure. Denote 2^i by q. One has to show

$$\lambda_M^{p/q}(\int|x|^{q-2}\mu_M(dx))^{p/q}\leq\lambda_M^{p/2}+\lambda_M\int|x|^{p-2}\mu_M(dx).$$

If

$$(\int|x|^{q-2}\mu_M(dx))^{p/q}\leq\lambda_M^{p/2-p/q}$$

the preceding inequality is obvious. On the other hand, if

$$\lambda_M\leq(\int|x|^{q-2}\mu_M(dx))^{2/(q-2)}$$

then it is sufficient to prove that

$$\lambda_M^{p/q-1}(\int|x|^{q-2}\mu_M(dx))^{p/q}\leq\int|x|^{p-2}\mu_M(dx).$$

But the bound on λ_M and Jensen's inequality give the result. □

Theorem 66 can be used to prove many different results concerning solutions of stochastic differential equations driven by Lévy processes. We give two examples. (These types of results are true more extensively (for example in n dimensions) with similar proofs; see for example [105].) Our first example is a moment estimate. As a convenience, we continue to work on the time interval $[0, 1]$.

Theorem 67. *Let σ be continuously differentiable with a bounded derivative, and let Z be a Lévy process with Lévy measure ν such that $\int_{|x|>1} |x|^p \nu(dx) < \infty$, which is equivalent to $E\{|Z_t|^p\} < \infty$ for $p \geq 2$ and for all t, $0 \leq t \leq 1$. Let X^x denote the unique solution of the stochastic differential equation*

$$dX_t = \sigma(X_{t-})dZ_t, \quad X_0 = x. \tag{$*$}$$

Then we have

$$E\{\sup_s |X_s^x|^p\} \leq K(1 + |x|^p)$$

where K is of course a positive constant.

Before proving this theorem we need to recall Gronwall's inequality.[20]

Theorem 68 (Gronwall's Inequality). *Let α be a function from $\mathbb{R}_+$ to itself, and suppose*

$$\alpha(s) \leq c + k \int_0^s \alpha(r)dr < \infty$$

for $0 \leq s \leq t$. Then $\alpha(t) \leq ce^{kt}$. Moreover if $c = 0$ then α vanishes identically.

The proof is simple and usually taught in elementary differential equations courses.

Proof of Theorem 67. This is a consequence of Theorem 66 and Gronwall's inequality. □

The next corollary follows easily from Theorem 67.

Theorem 69. *Let σ and Z be as in Theorem 67. Let g be continuous and for $q > 0$ define*

$$\| g \|_q = \inf\{a > 0 : |g(x)| \leq a(1 + |x|^q)\}.$$

For $q = 0$ we let $\| g \|_0$ be the L^∞ norm. If P_t denotes the transition semigroup of the Markov process X^x of equation ($*$), then we have for $q \geq 0$

$$\| P_t g \|_q \leq K \| g \|_q .$$

[20] Extensions of this classic version of Gronwall's inequality are in the exercises for this chapter. See also [182].

We now turn to our second example, which consists of moment estimates on the flows of the solution X^x of equation $(*)$. We first establish some notation. For a Lévy process Z with Lévy measure ν and with finite p-th moment for $p \geq 2$, we decompose it as $Z_t = bt + cB_t + M_t$ where B is a standard Brownian motion and M is a martingale independent from B and L^2 orthogonal to all continuous martingales. We then define

$$\eta_p = |b| + |c| + \int (|y|^2 1_{\{|y|\leq 1\}} + |y|^p 1_{\{|y|>1\}})\nu(dy)$$

which is finite since by assumption $E\{|Z_t|^p\} < \infty$ and $p \geq 2$. Here we consider the multidimensional case: both Z and the solution X^x of equation $(*)$ are assumed to be d-dimensional, where $d \geq 1$.[21] We let $X_s^{x,(k)}$ represent a k-th derivative of the flow of X^x.

Theorem 70. *For $k \geq 1$ assume that σ is differentiable k times and that all partial derivatives of σ of order greater than or equal to one are bounded. Assume also that Z has moments of order kp with $p \geq 2$. With the notation defined right before this theorem, we have that there exists a constant $K(k, p, \sigma, \eta_p)$ such that for $0 \leq s \leq 1$,*

$$E\{\sup_s |X_s^{x,(k)}|^p\} \leq K.$$

Proof. We recall from Theorem 39 that X^x together with $X^{x,(k)}$ constitute the unique solution of the system of equations

$$\begin{aligned} X_t^i &= x_i + \sum_{\alpha=1}^m \int_0^t \sigma_\alpha^i(X_{s-})dZ_s^\alpha \\ D_{kt}^i &= \delta_k^i + \sum_{\alpha=1}^m \sum_{j=1}^n \int_0^t \frac{\partial \sigma_\alpha^i}{\partial x_j}(X_{s-})D_{ks-}^j dZ_s^\alpha, \end{aligned} \tag{D}$$

as long as $k \geq 2$, and with a slightly simpler formulation if $k = 1$. (Note that this is the same equation (D) as on page 305.) We can write the second equation above in a slightly more abstract way

$$\begin{aligned} X_t^{x,(k)} &= \int_0^t \nabla\sigma(X_{s-}^x)X_{s-}^{x,(k)}dZ_s \\ &\quad + \sum_{i=2}^k \int_0^t \nabla^i\sigma(X_{s-}^x)F_{k,i}(X_{s-}^{x,(1)}, \ldots, X_{s-}^{x,(k-i+1)})dZ_s, \end{aligned}$$

if $k \geq 2$; and if $k = 1$ the equation is the simpler

[21] One could also allow the dimensions of Z and X^x to be different, but we do not address that case here.

$$X_t^{x,(1)} = I_d + \int_0^t \nabla\sigma(X_{s-}^x)X_{s-}^{x,(1)}dZ_s.$$

We next observe that the components of $F_{k,i}(x^{(1)}, \ldots, x^{(k-i+1)})$ are sums of terms of the form

$$\prod_{j=1}^{k-i+1} \prod_{r=1}^{\alpha_j} x^{(j),l_r}, \quad \text{where } \sum_j j\alpha_j = k, \tag{$**$}$$

and where $x^{(j),l}$ is the l-th component of $x^{(j)} \in \mathbb{R}^{d^j}$, and an "empty" product equals 1.

We thus want now to prove that

$$E\{\sup_t \left| \int_0^t \nabla^i\sigma(X_{s-}^x)F_{N,i}(X_{s-}^{x,(1)}, \ldots, X_{s-}^{x,(k-i+1)})dZ_s \right|^p\} \leq K$$

for all $i = 2, \ldots, N$ and for some constant $K = K(k, p, \sigma, \eta_{kp})$ (in the remainder of the proof $K = K(k, p, \sigma, \eta_{kp})$ varies from line to line). And of course, it is enough to prove that if G is any monomial as in $(**)$, then

$$E\{\sup_t \left| \int_0^t \nabla^i\sigma(X_{s-}^x)G(X_{s-}^{x,(1)}, \ldots, X_{s-}^{x,(k-i+1)})dZ_s \right|^p\} \leq K.$$

We will use Theorem 66 and the fact that $\nabla^i\sigma$ is bounded. Note that Theorem 66 implies that

$$E\{\sup_{s\leq t} | \int_0^s H_s dZ_s|^{p'}\} \leq K(p, \eta_p) \int_0^t E\{|H_s|^{p'}\}ds$$

for $2 \leq p' \leq p$, where the constant $K(p, \eta_p)$ depends only on p, η_p, and the dimensions of H and Z. For this we see that the left side of the previous equation is smaller than

$$K\, E\{\prod_{j=1}^{k+1-i} \sup_s |X_s^{x,(j)}|^{p\alpha_j}\} \leq K \prod_{j=1}^{k+1-i} (E\{\sup_s |X_s^{x,(j)}|^{kp/j}\})^{j\alpha_j/k}$$

by Hölder's inequality, since $\sum_j j\alpha_j = k$. The recurrence assumption yields that each expectation above is smaller than some constant $K(p, k, \sigma, \eta_{kp})$, and we obtain the required inequality. □

We now turn to the positivity of solutions of stochastic differential equations. This issue arises often in applications (for example in mathematical finance) when one wants to model a dynamic phenomenon where for reasons relating to the application, the solution should always remain positive. We begin with a very simple result, already in widespread use.

Theorem 71. *Let Z be a continuous semimartingale, σ a continuous function, and suppose there exists a (path-by-path) unique, non-exploding solution to the equation*

$$X_t = X_0 + \int_0^t \sigma(X_s)X_s dZ_s$$

with $X_0 > 0$ almost surely. Let $T = \inf\{t > 0 : X_t = 0\}$. Then $P(T < \infty) = 0$.

In words, if the coefficient of the equation is of the form $x\sigma(x)$ with $\sigma(x)$ continuous on $[0, \infty)$, then the solution X stays strictly positive in finite time if it begins with a strictly positive initial condition. We also remark that in the applied literature the equation is often written in the form

$$\frac{dX_t}{X_t} = \sigma(X_t)dZ_t.$$

Note that if $\sigma(x)$ is any function, one can then write $\sigma(x) = x(\frac{\sigma(x)}{x})$, but we then need the new function $\frac{\sigma(x)}{x}$ to be continuous, well-defined for all nonnegative x, and we also need for a unique solution to exist that *has no explosions*. This of course amounts to a restriction on the function σ. Since $x\sigma(x)$ being Lipschitz continuous is a sufficient condition to have a unique and non-exploding solution, we can require in turn for σ to be both bounded and Lipschitz, which of course implies that $x\sigma(x)$ is itself Lipschitz.

Proof of Theorem 71. Define $T_n = \inf\{t > 0 : X_t = 1/n \text{ or } X_t = X_0 \vee n\}$, and note that $P(T_n > 0) = 1$ because $P(X_0 > 0) = 1$ and Z is continuous. Using Itô's formula up to time T_n we have

$$\ln(|X_{T_n}|) = \ln(X_0) + \int_0^{T_n} \sigma(X_s)dZ_s - \frac{1}{2}\int_0^{T_n} \sigma(X_s)^2 d[Z, Z]_s,$$

and since σ is assumed continuous, it is bounded on the compact set $[1/n, n]$, say by a constant c. Since the stopping times $T_n \uparrow T$ we see that on the event $\{T < \infty\}$ the left side of the above equation tends to ∞ while the right side remains finite, a contradiction. Therefore $P(T < \infty) = 0$. □

Theorem 71 has an analogue when Z is no longer assumed to be continuous.

Theorem 72. *Let Z be a semimartingale and let σ be a bounded, continuous function such that there exists a unique non-exploding solution of*

$$X_t = X_0 + \int_0^t \sigma(X_{s-})X_{s-}dZ_s$$

with $X_0 > 0$ almost surely. If in addition we have $|\Delta Z_s| \leq \frac{1-\varepsilon}{\|\sigma\|_\infty}$ almost surely for all $s \geq 0$ for some ε with $0 < \varepsilon < 1$, then the solution X stays positive a.s. for all $t > 0$.

Proof. $\| \sigma \|_\infty$ refers to the L^∞ norm of σ. The proof is similar to the proof of Theorem 71. Using Itô's formula gives

$$\ln(|X_t|) = \ln(X_0) + \int_0^t \sigma(X_{s-})dZ_s - \frac{1}{2}\int_0^t \sigma(X_{s-})^2 d[Z,Z]_s^c$$
$$+ \sum_{s\le t}\{\ln|X_s| - \ln|X_{s-}| - \sigma(X_{s-})\Delta Z_s\},$$

and let us consider the last term, the summation. We have

$$\sum_{s\le t}\{\ln|X_s| - \ln|X_{s-}| - \sigma(X_{s-})\Delta Z_s\}$$
$$= \frac{1}{2}\sum_{s\le t}\frac{1}{(X_{s-} + H_s X_{s-}\sigma(X_{s-})\Delta Z_s)^2} X_{s-}^2\sigma(X_{s-})^2(\Delta Z_s)^2$$
$$= \frac{1}{2}\sum_{s\le t}\frac{1}{(1 + H_s\sigma(X_{s-})\Delta Z_s)^2}\sigma(X_{s-})^2(\Delta Z_s)^2$$

where $|H_s| \le 1$ by the Mean Value Theorem, applied ω-by-ω, which implies the above is

$$\ge \frac{1}{2}\sum_{s\le t}\frac{1}{\varepsilon^2}\sigma(X_{s-})^2(\Delta Z_s)^2,$$

which is a convergent series for each finite time t. The only ways X can be zero or negative is to cross 0 while continuous, which cannot happen by an argument analogous to the proof of Theorem 71, or to jump to 0 or across into negative values. Assume X has not yet reached $(-\infty, 0]$ at time $s-$. Then we have

$$|\Delta X_s| \le X_{s-}|\sigma(X_{s-})\Delta Z_s| \le X_{s-}(1-\varepsilon) < X_{s-},$$

which implies that X cannot cross into $(-\infty, 0]$ by jumping. Hence we have that X must be positive on $[0, \infty)$. □

We now turn to something completely different. In Sect. 6, Theorem 32, we showed that an equation of the form

$$X_t^i = X_0 + \sum_{j=1}^d \int_0^t f_j^i(X_{s-})dZ_s^j \qquad (*)$$

with Z a vector of independent Lévy processes, had a (unique) solution which was strong Markov. We now show what amounts to a converse, which is perhaps surprising. That is, if one wants a solution of a stochastic differential equation to be time homogeneous Markov, one essentially must have that the driving semimartingale is a Lévy process!

Theorem 73. *Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space with Z a càdlàg semimartingale defined on the space. Let f be a Borel function which is never 0 and is such that for every $x \in \mathbb{R}$ the equation*

$$dX_t = f(X_{t-})dZ_t, \quad X_0 = x$$

has unique (strong) solution, which we denote X^x. If each of the processes X^x is a time homogeneous Markov process with the same transition semigroup, then Z is a Lévy process.

For the proof of this theorem, we assume the reader is familiar with the abstract theory of Markov processes, as can be found (for example) in [215].

Proof. Let Ω' be the function space of càdlàg functions defined on $\mathbb{R}_+$ with $X_t(\omega) = \omega(t)$, the projection operator. Let $\mathbb{F}'$ be the canonical filtration and let (θ'_t) be the shift operators. If P'_x denotes the law of X^x, then we have that $(\Omega', \mathbb{F}', \theta'_t, X', P'_x)$ is a Markov process in the Dynkin (or Blumenthal-Getoor) sense. Since f is never zero, we can write

$$Z_t = Z_0 + \int_0^t f(X^x_{s-})^{-1} dX^x_s.$$

We can therefore define on Ω', relative to each law P'_x, the stochastic integral

$$Z'_t = \int_0^t f(X'_{s-})^{-1} dX'_s,$$

and we have that the law of Z' under P'_x is the law of the process $Z - Z_0$.

However Z' is also an additive functional, and hence the Markov property implies

$$\begin{aligned} E'_x\{g(Z'_{t+s} - Z'_t)|\mathcal{F}'\} &= E'_x\{g(Z'_s) \circ \theta'_t|\mathcal{F}'_t\} \\ &= E'_{X'_t}\{g(Z'_s)\} = E\{g(Z_s - Z_0)\}, \end{aligned}$$

where we have also used that the law of Z' under P'_x is the law of the process $Z - Z_0$. This in turn implies that $Z'_{t+s} - Z'_t$ is P'_x independent of $\mathcal{F}'_t$, and hence also independent from Z'_r for all $r \leq t$, and it also implies that the law of $Z'_{t+s} - Z'_t$ is the same as the law of $Z_s - Z_0$. Using the identity of the laws of Z' under P'_x with that of the process $Z - Z_0$ once again, the result follows. □

Bibliographic Notes

The extension of the $\underline{\underline{H}}^p$ norm from martingales to semimartingales was implicit in Protter [197] and first formally proposed by Emery [64]. A comprehensive account of this important norm for semimartingales can be found in

Dellacherie-Meyer [46]. Emery's inequalities (Theorem 3) were first established in Emery [64], and later extended by Meyer [174].

Existence and uniqueness of solutions of stochastic differential equations driven by general semimartingales was first established by Doléans-Dade [51] and Protter [198], building on the result for continuous semimartingales in Protter [197]. Before this Kazamaki [123] published a preliminary result, and of course the literature on stochastic differential equations driven by Brownian motion and Lebesgue measure, as well as Poisson processes, was extensive. See, for example, the book of Gihman-Skorohod [81] in this regard. These results were improved and simplified by Doléans-Dade-Meyer [54] and Emery [65]; our approach is inspired by Emery [65]. Métiver-Pellaumail [161] have an alternative approach. See also Métivier [158]. Other treatments can be found in Doléans-Dade [52] and Jacod [103]. The approach of L. Schwartz [214] is nicely presented in [44].

The stability theory is due to Protter [199], Emery [65], and also to Métivier-Pellaumail [162]. The semimartingale topology is due to Emery [66] and Métivier-Pellaumail [162], while a pedagogic treatment is in Dellacherie-Meyer [46].

The generalization of Fisk-Stratonovich integrals to semimartingales is due to Meyer [171]. The treatment here of Fisk-Stratonovich differential equations is new. The idea of quadratic variation is due to Wiener [230]. Theorem 18, which is a random Itô's formula, appears in this form for the first time. It has an antecedent in Doss-Lenglart [58], and for a very general version (containing some quite interesting consequences), see Sznitman [222]. Theorem 19 generalizes a result of Meyer [171], and Theorem 22 extends a result of Doss-Lenglart [58]. Theorem 24 and its corollary is from Itô [99]. Theorem 25 is inspired by the work of Doss [57] (see also Ikeda-Watanabe [92] and Sussman [221]). The treatment of approximations of the Fisk-Stratonovich integrals was inspired by Yor [238]. For an interesting application see Rootzen [211]. For more examples of solutions of stochastic differential equations with formulas, see [128].

The results of Sect. 6 are taken from Protter [196] and Çinlar-Jacod-Protter-Sharpe [34]. A comprehensive pedagogic treatment when the Markov solutions are diffusions can be found in either Stroock-Varadhan [220] or Rogers-Williams [209, 210].

Work on flows of stochastic differential equations goes back to 1961 and the work of Blagovescenskii-Freidlin [18] who considered the Brownian case. For recent work on flows of stochastic differential equations, see Kunita [131, 133], Ikeda-Watanabe [92] and the references therein. There are also flows results for the Brownian case in Gihman-Skorohod [81], but they are L^2 rather than almost sure results. Much of our treatment is inspired by the work of Meyer [177] and that of Uppman [224, 225] for the continuous case, however results are taken from other articles as well. For example, the example following Theorem 38 is due to Leandre [140], while the proof of Theorem 41, the non-confluence of solutions in the continuous case, is due to Emery [68]; an alternative proof is in Uppman [225]. For the general (right continuous) case, we follow the work of Leandre [141]. A similar result was obtained by Fujiwara-Kunita [78].

Our presentation on moment estimates follows [205] and [105], although it dates back to [16]. A perhaps simpler treatment using random measures can be found in [105]. See also [106]. The results on when solutions of stochastic differential equations are always positive is well known, at least in the Brownian case, but a reference is not easily found in the literature. The general case may be new. Theorem 73 is from [108].

Exercises for Chapter V

Exercise 1. Let X be a solution of the stochastic differential equation

$$dX_t = X_t^2 dB_t + X_t^3 dt$$

with initial condition $X_0 \neq 0$. Show that X is defined on $[0,T)$, where $T = \inf\{t > 0 : B_t = 1/X_0\}$. Moreover show that X is the unique solution on $[0,T)$. Finally show that X explodes at T.

Exercise 2. Let $\sigma(x) = \sqrt{|x|}$. Note that σ is Lipschitz continuous everywhere except at $x = 0$. Let X be a solution of the equation

$$dX_t = \sigma(X_t) dB_t + \frac{1}{2} dt$$

with initial condition X_0. Find an explicit solution of this equation and determine where it is defined.

Exercise 3. Let X be the unique solution of the equation

$$dX_t = \sigma(X_t) dB_t + b(X_t) dt$$

with X_0 given, and σ and b both Lipschitz continuous. Let the generator of the (time homogeneous) Markov process X be given by

$$Af(x) = \lim_{t \to 0, t > 0} \frac{P_t f(x) - f(x)}{t}$$

where $P_t f(x)$ equals the transition function[22] $P_s(X_0, f)$ when $X_0 = x$ almost surely. A is called the **infinitesimal generator** of X. Use Itô's formula to show that if f has two bounded continuous derivatives (written $f \in C_b^2$), then the limit exists and A is defined. In this case f is said to be in the *domain of the generator A*.

[22] This notation is defined in Sect. 5 of Chap. I on page 35.

Exercise 4. (Continuation of Exercise 3.) With the notation and terminology of Exercise 3, show that the infinitesimal generator of X is a linear partial differential operator for functions $f \in C_b^2$ given by

$$Af(x) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 f}{\partial x^2} + b(x)\frac{\partial f}{\partial x}.$$

Exercise 5. Let $f \in C_b^2$, let σ and b be Lipschitz continuous, let X be the unique solution of the stochastic integral equation

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds, \tag{$*$}$$

and let A denote its infinitesimal generator. If $X_0 = x$, where $x \in \mathbb{R}$, prove **Dynkin's expectation formula** for Markov processes in this case:

$$E\{f(X_t) - f(X_0) - \int_0^t Af(X_s)ds\} = 0. \tag{$**$}$$

Exercise 6. As a continuation of Exercise 5, show that

$$(P_t f)(y) - f(y) - \int_0^t (P_s Af)(y)ds = 0$$

for $y \in \mathbb{R}$, and that $\lim_{s\downarrow 0} P_s Af = Af$ for (for example) all f in C^∞ with compact support.

Exercise 7. As in Exercise 5, extend Dynkin's expectation formula by showing that it still holds if one replaces in equation $(**)$ the integration upper limit t with a bounded stopping time T, and use this to verify Dynkin's formula in this case. That is, show that if $f \in C_b^2$ and σ and b are both Lipschitz, and X is a solution of equation $(*)$ of Exercise 5, then

$$f(X_T) - f(X_0) - \int_0^T Af(X_s)ds \text{ is a local martingale.}$$

Exercise 8. Let Z be a Lévy process with Lévy measure ν and let X satisfy the equation

$$X_t = x + \int_0^t f(X_{s-})dZ_s$$

where f is Lipschitz continuous. The process X is taken to be d-dimensional, and Z is n-dimensional, so f takes its values in $\mathbb{R}^d \times \mathbb{R}^n$. We use (column) vector and matrix notation. The initial value is some given $x \in \mathbb{R}^d$. Let P^x denote the corresponding law of X starting at x. Show that if $g \in C^\infty$ with compact support on $\mathbb{R}^d$, then

$$Ag(x) = \nabla g(x)f(x)b + \frac{1}{2}\sum_{i,j=1}^{d}\left(\frac{\partial^2 g}{\partial x^i \partial x^j}(x)\right)(f(x)cf(x)^\star)^{ij}$$
$$+ \int \nu(dy)\left(g(x+f(x)y) - g(x) - \nabla g(x)f(x)\right),$$

where ∇g is a row vector and b is the drift coefficient of Z, $\star$ denotes transpose, and A is of course the infinitesimal generator of X.

Exercise 9. In the setting of Exercise 8 show that

$$g(X_t) - g(X_0) - \int_0^t Ag(X_s)ds \text{ is a local martingale}$$

for each law P^x. Show further that it is a local martingale for initial laws of the form P^μ, where P^μ is given by: for $\Lambda \subset \Omega$ measurable, $P^\mu(\Lambda) = \int_{\mathbb{R}^d} P^x(\Lambda)\mu(dx)$ for any probability law μ on $\mathbb{R}^d$.

Exercise 10. Let Z be an arbitrary semimartingale, and let $T_0 = 0$ and $(T_n)_{n\geq 1}$ denote the increasing sequence of stopping times of the jump times of Z when the corresponding jump size is larger than or equal to 1 in magnitude. Define $\varepsilon_n = \text{sign}(\Delta Z_{T_n})$ with $\varepsilon_0 = 1$. For a semimartingale Y define $U^n(Y)_t$ to be the solution of the exponential equation

$$V_t = 1 + \int_{T_n}^t V_{s-}dY_s$$

for $T_n \leq t < T_{n+1}$. Show that the (unique) solution of the equation

$$X_t = 1 + \int_0^t |X_{s-}|dZ_s$$

is given by $X_t = \sum_{n\geq 0} X_t^n 1_{[T_n,T_{n+1})}(t)$, where $X_t^n = X_{T_n}U^n(\varepsilon_n Z)_t$, for $T_n \leq t < T_{n+1}$.

Exercise 11. Show that if M is an L^2 martingale with $\Delta M_s > -1$, and $\langle M, M\rangle_\infty$ is bounded, then $\mathcal{E}(M)_\infty > 0$ a.s. and it is square integrable. (See [147].)

Exercise 12. Let M be a martingale which is of integrable variation. Suppose also that the compensator of the process $\sum_{s\leq t}|\Delta M_s|$ is bounded. Show that the local martingale $\mathcal{E}(M)$ is also of integrable variation.

Exercise 13. Let Z be a semimartingale with decomposition $Z = L + V$ where L is a local martingale with $\Delta L_s \neq -1$ for all finite s, and V has paths of finite variation on compacts. Let

$$U_t = V_t - \sum_{s\leq t}\frac{\Delta L_s \Delta V_s}{1+\Delta L_s}.$$

Show that $\mathcal{E}(Z) = \mathcal{E}(L)\mathcal{E}(U)$. (See [157].)

***Exercise 14 (Mémin's criterion for exponential martingales).** Let M be a local martingale, let $J_t = \sum_{s\leq t} \Delta M_s 1_{\{|\Delta M_s|\geq \frac{1}{2}\}}$, and let $L_t = J_t - \widetilde{J}_t$ and $N_t = M_t - L_t$. (Both L and N are of course also local martingales.) Let

$$A_t = [M,M]^c + \sum_{0<s\leq t} |\Delta L_s| + \sum_{0<s\leq t} (\Delta N_s)^2$$

and assume that the compensator of A is bounded. Show that this implies that $\mathcal{E}(M)$ is a uniformly integrable martingale. (*Note*: This is related to Kazamaki's and Novikov's criteria, but without the assumption of path continuity. Also see [156].)

Exercise 15 (Gronwall's inequality). . Let C_t be a càdlàg increasing process and suppose $0 \leq A_t \leq \alpha + \int_0^t A_{s-} dC_s$ for $t \geq 0$. Show that $A_t \leq \alpha e^{C_t}$, each $t \geq 0$. (*Note*: The classical Gronwall inequality is usually stated with $C_t = t$. Also see [85].)

The next five exercises outline the *Métivier-Pellaumail method* for showing the existence and uniqueness of solutions of stochastic differential equations.

***Exercise 16.** Let M be an L^2 martingale. Show that for any predictable stopping time T, $E\{(E\{\Delta M_T|\mathcal{F}_{T-}\})^2\} \leq E\{\langle M,M\rangle_{T-}\}$. Show also that for any totally inaccessible stopping time T, we have $E\{(E\{\Delta M_T|\mathcal{F}_{T-}\})^2\} \leq E\{\langle M,M\rangle_T\}$. Conclude that for an arbitrary stopping time T we have

$$E\{(E\{\Delta M_T|\mathcal{F}_{T-}\})^2\} \leq E\{\langle M,M\rangle_{T-}\}.$$

Exercise 17 (Métivier-Pellaumail inequality). Let M be an L^2 martingale and let $M^*_{t-} = \sup_{s<t} |M_s|$. Show that for any stopping time T we have

$$E\{(M^*_{T-})^2\} \leq 4E\{\langle M,M\rangle_{T-} + [M,M]_{T-}\}.$$

(*Hint*: Use Exercise 16).

Exercise 18. Let M be a locally square integrable local martingale. Show that for any stopping time T we have

$$E\{\sup_{s<T}(M_s - M_0)^2\} \leq 4E\{\langle M,M\rangle_{T-} + [M,M]_{T-}\}.$$

Exercise 19 (extended Gronwall's inequality). Let A and C be two increasing processes with A adapted and $E\{A_\infty\} < \infty$, and C such that $0 \leq C_\infty \leq k$ everywhere for some constant k. Suppose that for every stopping time T and for some positive constant α we have

$$E\{A_{T-}\} \leq \alpha + E\{\int_0^{T-} A_{s-} dC_s\}.$$

Then $E\{A_\infty\} \leq \alpha e^k$. (*Hint*: Use Lebesgue's change of time lemma, together with Gronwall's inequality of Exercise 15. Also see [110, page 576].)

Exercise 20. Use Exercises 16 through 19 as needed to show, via a Picard iteration method, that a unique solution exists to the stochastic differential equation

$$dX_t = f(X_{s-})dZ_s$$

where f is Lipschitz continuous and Z is an arbitrary semimartingale.

Exercise 21. Show that there are an infinite number of solutions (on the same probability space) to the equation

$$X_t = 3\int_0^t X_s^{\frac{1}{3}}ds + 3\int_0^t X_s^{\frac{2}{3}}dB_s, \quad X_0 = 0,$$

where B is a standard Brownian motion. (See [121, page 293].)

***Exercise 22.** Let σ satisfy a weakening of the Lipschitz condition of the form $|\sigma(y) - \sigma(x)| \leq \kappa(|x-y|)$ for all x, $y \in \mathbb{R}$, with $\int_0^\varepsilon \frac{1}{\kappa(u)^2}du = \infty$ and κ increasing with $\kappa(0) = 0$. (Such a condition is a variant of a standard condition for uniqueness of solutions in the theory of ordinary differential equations; see for example [35, page 60]. Let b be Lipschitz continuous. Show that there exists a unique solution of the equation

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t$$

where B is standard Brownian motion. (See [232].)

Exercise 23. Let σ and b be Lipschitz continuous and let X be the unique solution of equation $(*)$ of Exercise 5. Show that $2xb(x) + \sigma(x)^2 \leq c + k|x|^2$, where c and k are positive constants. Use Itô's formula on X_t^2 to show the following *moment estimate*: $E\{X_t^2\} \leq (E\{X_0^2\} + ct)e^{kt}$. (See [135].)

***Exercise 24 (Euler method of approximation).** Let σ and b be continuous and such that a unique path-by-path solution X of equation $(*)$ of Exercise 5 exists. Let $(\xi_k)_{k\geq 1}$ be i.i.d. with mean zero and variance τ^2. Let X_0 be independent of $(\xi_k)_{k\geq 1}$ and let X_k be defined inductively by

$$X_{k+1} = X_k + \sigma\frac{\xi_{k+1}}{\sqrt{n}} + \frac{b(X_k)}{n}.$$

Define $X_n(t) = X_{[nt]}$ where $[nt]$ denotes the integer part of nt. Let $B_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k$, $V_n(t) = \frac{[nt]}{n}$, and finally observe that

$$X_n(t) = X(0) + \int_0^t \sigma(X_n(s-))dB_n(s) + \int_0^t b(X_n(s-))dV_n(s).$$

By Donsker's Theorem we know that (B_n, V_n) converges weakly (in distribution) to $(\tau B, V)$, where B is a standard Brownian motion and $V(t) = t$. Show that B_n is a martingale, and that X_n converges weakly to X. (See [136] and [137].)

*__Exercise 25.__ Let Z be a Lévy process which is a square integrable martingale. Let f be continuously differentiable with a bounded derivative and let X^x be the unique solution of

$$X_t^x = x + \int_0^t f(X_{s-}^x)dZ_s.$$

Show that X^x is also a square integrable martingale.

*__Exercise 26.__ In the framework of Exercise 25, let $X_t'^x$ be the solution of

$$X_t'^x = 1 + \int_0^t f'(X_{s-}^x)X_{s-}'^x dZ_s.$$

Show that X'^x is also a square integrable martingale.

**__Exercise 27.__ In the framework of Exercises 25 and 26, for each measurable function g with at most linear growth, set

$$P_t g(x) = E\{g(X_t^x)\} \quad \text{and} \quad Q_t g(x) = E\{g(X_t^x)X_t'^x\}.$$

Assume that f is infinitely differentiable, and that g is twice differentiable, bounded and both of its first two derivatives are bounded. Show that the function $(t,x) \mapsto P_t g(x)$ is twice differentiable in x and once differentiable in t, that all the partial derivatives are continuous in (t,x), and further that

$$\frac{\partial}{\partial x}P_t g(x) = Q_t g'(x).$$

VI

Expansion of Filtrations

1 Introduction

By an expansion of the filtration, we mean that we enlarge the filtration $(\mathcal{F}_t)_{t\geq 0}$ to get another filtration $(\mathcal{H}_t)_{t\geq 0}$ such that the new filtration satisfies the usual hypotheses and $\mathcal{F}_t \subset \mathcal{H}_t$, each $t \geq 0$. There are three questions we wish to address: (1) when does a specific, given semimartingale remain a semimartingale in the enlarged filtration; (2) when do all semimartingales remain semimartingales in the enlarged filtration; (3) what is a new decomposition of the semimartingale for the new filtration.

The subject of the expansion of filtrations began with a seminal paper of K. Itô in 1976 (published in [100] in 1978), when he showed that if B is a standard Brownian motion, then one can expand the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ of B by adding the σ-algebra generated by the random variable B_1 to all $\mathcal{F}_t$ of the filtration, including of course $\mathcal{F}_0$. He showed that B remains a semimartingale for the expanded filtration, he calculated its decomposition explicitly, and he showed that one has the intuitive formula

$$B_1 \int_0^t H_s dB_s = \int_0^t B_1 H_s dB_s$$

where the integral on the left is computed with the original filtration, and the integral on the right is computed using the expanded filtration. Obviously such a result is of interest only for $0 \leq t \leq 1$. We will establish this formula more generally for Lévy processes in Sect. 2.

The second advance for the theory of the expansion of filtrations was the 1978 paper of M. Barlow [7] where he considered the problem that if L is a positive random variable, and one expands the filtration in a minimal way to make L a stopping time, what conditions ensure that semimartingales remain semimartingales for the expanded filtration? This type of question is called *progressive expansion* and it is the topic of Sect. 3.

2 Initial Expansions

Throughout this section we assume given an underlying filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ which satisfies the usual hypotheses. As in previous chapters, for convenience we denote the filtration $(\mathcal{F}_t)_{t\geq 0}$ by the symbol $\mathbb{F}$.

The most elementary result on the expansion of filtrations is due to Jacod and was established in Chap. II (Theorem 5). We recall it here.

Theorem 1 (Jacod's Countable Expansion). *Let $\mathcal{A}$ be a collection of events in $\mathcal{F}$ such that if A_α, $A_\beta \in \mathcal{A}$ then $A_\alpha \cap A_\beta = \phi$, $\alpha \neq \beta$. Let $\mathcal{H}_t$ be the filtration generated by $\mathcal{F}_t$ and $\mathcal{A}$. Then every $((\mathcal{F}_t)_{t\geq 0}, P)$ semimartingale is an $((\mathcal{H}_t)_{t\geq 0}, P)$ semimartingale also.*

We also record a trivial observation as a second elementary theorem.

Theorem 2. *Let X be a semimartingale with decomposition $X = M + A$ and let $\mathcal{G}$ be a σ-algebra independent of the local martingale term M. Let $\mathbb{H}$ denote the filtration obtained by expanding $\mathbb{F}$ with the one σ-algebra $\mathcal{G}$ (that is, $\mathcal{H}_t = \mathcal{F}_t \vee \mathcal{G}$, each $t \geq 0$ and $\mathcal{H}_t = \mathcal{H}_{t+}$). Then X is an $\mathbb{H}$ semimartingale with the same decomposition.*

Proof. Since the local martingale M remains a local martingale under $\mathbb{H}$, the theorem follows. □

We now turn to Lévy processes and an extension of Itô's first theorem. Let Z be a given Lévy process on our underlying space, and define $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ to be the smallest filtration satisfying the usual hypotheses, such that Z_1 is $\mathcal{H}_0$ measurable and $\mathcal{F}_t \subset \mathcal{H}_t$ for all $t \geq 0$.

Theorem 3 (Itô's Theorem for Lévy Processes). *The Lévy process Z is an $\mathbb{H}$ semimartingale. If moreover $E\{|Z_t|\} < \infty$, all $t \geq 0$, then the process*

$$M_t = Z_t - \int_0^{t\wedge 1} \frac{Z_1 - Z_s}{1-s} ds$$

is an $\mathbb{H}$ martingale on $[0, \infty)$.

Proof. We begin by assuming $E\{Z_t^2\} < \infty$, each $t > 0$. Without loss of generality we can further assume $E\{Z_t\} = 0$. Since Z has independent increments, we know Z is an $\mathbb{F}$ martingale. Let $0 \leq s < t \leq 1$ be rationals with $s = j/n$ and $t = k/n$. We set

$$Y_i = Z_{\frac{i+1}{n}} - Z_{\frac{i}{n}}.$$

Then $Z_1 - Z_s = \sum_{i=j}^{n-1} Y_i$ and $Z_t - Z_s = \sum_{i=j}^{k-1} Y_i$. The random variables Y_i are i.i.d. and integrable. Therefore

$$E\{Z_t - Z_s | Z_1 - Z_s\} = E\{\sum_{i=j}^{k-1} Y_i | \sum_{i=j}^{n-1} Y_i\} = \frac{k-j}{n-j} \sum_{i=j}^{n-1} Y_i$$
$$= \frac{t-s}{1-s}(Z_1 - Z_s).$$

The independence of the increments of Z yields $E\{Z_t - Z_s | \mathcal{H}_s\} = E\{Z_t - Z_s | Z_1 - Z_s\}$; therefore $E\{Z_t - Z_s | \mathcal{H}_s\} = \frac{t-s}{1-s}(Z_1 - Z_s)$ for all rationals, $0 \leq s < t \leq 1$. Since Z is an $\mathbb{F}$ martingale, the random variables $(Z_t)_{0 \leq t \leq 1}$ are uniformly integrable, and since the paths of Z are right continuous, we deduce $E\{Z_t - Z_s | \mathcal{H}_s\} = \frac{t-s}{1-s}(Z_1 - Z_s)$ for all reals, $0 \leq s < t \leq 1$. By Fubini's Theorem for conditional expectations the above gives

$$\begin{aligned} E\{M_t - M_s | \mathcal{H}_s\} &= E\{Z_t - Z_s | \mathcal{H}_s\} - \int_s^t \frac{1}{1-u} E\{Z_1 - Z_u | \mathcal{H}_s\} du \\ &= \frac{t-s}{1-s}(Z_1 - Z_s) - \int_s^t \frac{1}{1-u} \frac{1-u}{1-s}(Z_1 - Z_s) du \\ &= 0. \end{aligned}$$

There is a potential problem at $t = 1$ because of the possibility of an explosion. Indeed this is typical of initial enlargements. However if we can show $E\{\int_0^1 \frac{|Z_1 - Z_s|}{1-s} ds\} < \infty$ this will suffice to rule out explosions. By the stationarity and independence of the increments of Z we have $E\{|Z_1 - Z_s|\} \leq E\{(Z_1 - Z_s)^2\}^{\frac{1}{2}} \leq a(1-s)^{\frac{1}{2}}$ for some constant a and for all s, $0 \leq s \leq 1$. Therefore $E\{\int_0^1 \frac{|Z_1 - Z_s|}{1-s} ds\} \leq a \int_0^1 \frac{\sqrt{1-s}}{1-s} ds < \infty$. Note that if $t > 1$ then $\mathcal{F}_t = \mathcal{H}_t$, and it follows that M is a martingale. Since $Z_t = M_t + \int_0^1 \frac{Z_1 - Z_s}{1-s} ds$ we have that Z is a semimartingale.

Next suppose only that $E\{|Z_t|\} < \infty$, $t \geq 0$ instead of Z being in L^2 as we assumed earlier. We define

$$J_t^1 = \sum_{0 < s \leq t} \Delta Z_s 1_{\{\Delta Z_s > 1\}} \quad \text{and} \quad J_t^2 = \sum_{0 < s \leq t} \Delta Z_s 1_{\{\Delta Z_s < -1\}}.$$

(Since Z has càdlàg paths a.s. for each ω there are at most a finite number of jumps bigger than a fixed size on each compact time set; hence each J_t^i is finite a.s.) By the results of Chap. I on Lévy processes we have that $Y_t = Z_t - J_t^1 + J_t^2$ is a Lévy process with bounded jumps, and hence it is square integrable. Additionally, the processes Y, J^1, and J^2 are jointly independent.

Combining these facts with the preceding proof yields

$$N_t = Y_t - \int_0^t \frac{Y_1 - Y_s}{1-s} ds$$

is an $\mathbb{H}'$ martingale, where $\mathbb{H}'$ is the smallest right continuous filtration obtained by expanding $\mathbb{F}$ with (Y_1, J_1^1, J_1^2). Note that $\mathbb{H} \subset \mathbb{H}'$. The same argument, plus the observation that since J^i does not jump at time $t = 1$ and is therefore constant in a (random) neighborhood of 1, which in turn yields that $\int_0^1 \frac{|J_1^i - J_s^i|}{1-s} ds < \infty$ a.s., shows that $(J_t^i - \int_0^t \frac{J_1^i - J_s^i}{1-s} ds)_{t \geq 0}$ is a local martingale for $\mathbb{H}'$. Moreover it is a martingale for $\mathbb{H}'$ as soon as $E\{\int_0^1 \frac{|J_1^i - J_s^i|}{1-s} ds\} < \infty$. But the function $t \mapsto E\{J_t^i\} = a_i t$ for all i by the stationarity of the increments. Hence

$$E\{\int_0^1 \frac{|J_1^i - J_s^i|}{1-s} ds\} = |E\{\int_0^1 \frac{J_1^i - J_s^i}{1-s} ds\}| = |\int_0^1 E\{\frac{J_1^i - J_s^i}{1-s}\} ds|$$
$$= |a_i| \int_0^1 \frac{1-s}{1-s} ds$$
$$= |a_i| < \infty.$$

Since Y, J^1, and J^2 are all independent, we conclude that M is an $\mathbb{H}'$ martingale. Since M is adapted to $\mathbb{H}$, by Stricker's Theorem it is also an $\mathbb{H}$ martingale, and thus Z is an $\mathbb{H}$ semimartingale.

Finally we drop all integrability assumptions on Z. We let

$$J_t^1 = \sum_{0<s\leq t} \Delta Z_s 1_{\{|\Delta Z_s|>1\}} \quad \text{and also} \quad X_t = Z_t - J_t^1.$$

Then X is also a Lévy process, and since X has bounded jumps it is in L^p for all $p \geq 1$, and in particular $E\{X_t^2\} < \infty$, each $t \geq 0$. Let $\mathbb{H}(X_1)$ denote $\mathbb{F}$ expanded by the adjunction of the random variable X_1. Let $\mathbb{K} = \mathbb{H}(X_1) \vee \mathbb{H}(J_1^1)$, the filtration generated by $\mathbb{H}(X_1)$ and $\mathbb{H}(J_1^1)$. Then $\mathbb{H}(Z_1) \subset \mathbb{K}$. But X is a semimartingale on $\mathbb{H}(X_1)$, and since J^1 is independent of X we have by Theorem 2 that X is a $(\mathbb{K}, P)$ semimartingale. Therefore by Stricker's Theorem, X is an $(\mathbb{H}, P)$ semimartingale; and since J_t^1 is a finite variation process adapted to $\mathbb{H}$ we have that Z is an $(\mathbb{H}, P)$ semimartingale as well. □

The most important example of the above theorem is that of Brownian motion, which was Itô's original formula. In this case let $\mathbb{H} = \mathbb{H}(B_1)$, and we have as a special case the $\mathbb{H}$ decomposition of Brownian motion:

$$B_t = (B_t - \int_0^t \frac{B_1 - B_s}{1-s} ds) + \int_0^t \frac{B_1 - B_s}{1-s} ds = \beta_t + \int_0^t \frac{B_1 - B_s}{1-s} ds.$$

Note that the martingale β in the decomposition is continuous and has $[\beta, \beta]_t = t$, which by Lévy's theorem gives that β is also a Brownian motion. A simple calculation gives Itô's original formula, for a process H which is $\mathbb{F}$ predictable:

$$B_1 \int_0^t H_s dB_s = B_1 \int_0^t H_s d\beta_s + B_1 \int_0^t H_s \frac{B_1 - B_s}{1-s} ds$$
$$= \int_0^t B_1 H_s d\beta_s + \int_0^t B_1 H_s \frac{B_1 - B_s}{1-s} ds$$
$$= \int_0^t B_1 H_s dB_s$$

where since the random variable B_1 is $\mathcal{H}_0$ measurable, it can be moved inside the stochastic integral. We can extend this theorem with a simple iteration; we omit the fairly obvious proof.

Corollary. Let Z be a given Lévy process with respect to a filtration $\mathbb{F}$, and let $0 = t_0 < t_1 < \cdots < t_n < \infty$. Let $\mathbb{H}$ denote the smallest filtration satisfying the usual hypotheses containing $\mathbb{F}$ and such that the random variables $Z_{t_1}, \ldots, Z_{t_n}$ are all $\mathcal{H}_0$ measurable. Then Z is an $\mathbb{H}$ semimartingale. If we have a countable sequence $0 = t_0 < t_1 < \cdots < t_n < \cdots$, we let $\tau = \sup_n t_n$, with $\mathbb{H}$ the corresponding filtration. Then Z is an $\mathbb{H}$ semimartingale on $[0, \tau)$.

We next give a general criterion (Theorem 5) to have a local martingale remain a semimartingale in an expanded filtration. (Note that a finite variation process automatically remains one in the larger filtration, so the whole issue is what happens to the local martingales.) We then combine this theorem with a lemma due to Jeulin to show how one can expand the Brownian filtration. Before we begin let us recall that a process X is *locally integrable* if there exist a sequence of stopping times $(T_n)_{n\geq 1}$ increasing to ∞ a.s. such that $E\{|X_{T_n} 1_{\{T_n>0\}}|\} < \infty$ for each n. Of course, if $X_0 = 0$ this reduces to the condition $E\{|X_{T_n}|\} < \infty$ for each n.

Theorem 4. *Let M be an $\mathbb{F}$ local martingale and suppose M is a semimartingale in an expanded filtration $\mathbb{H}$. Then M is a special semimartingale in $\mathbb{H}$.*

Proof. First recall that any local martingale is a special semimartingale. In particular the process $M_t^* = \sup_{s\leq t} |M_s|$ is locally integrable (see Theorem 33 of Chap. III), and this of course remains locally integrable in the expanded filtration $\mathbb{H}$, since stopping times remain stopping times in an expanded filtration. Since M is an $\mathbb{H}$ semimartingale by hypothesis, it is special because M_t^* is locally integrable (see Theorem 34 of Chap. III). □

Theorem 5. *Let M be an $\mathbb{F}$ local martingale, and let H be predictable such that $\int_0^t H_s^2 d[M, M]_s$ is locally integrable. Suppose $\mathbb{H}$ is an expansion of $\mathbb{F}$ such that M is an $\mathbb{H}$ semimartingale. Then M is a special semimartingale in $\mathbb{H}$ and let $M = N + A$ denote its canonical decomposition. The stochastic integral process $(\int_0^t H_s dM_s)_{t\geq 0}$ is an $\mathbb{H}$ semimartingale if and only if the process $(\int_0^t H_s dA_s)_{t\geq 0}$ exists as a path-by-path Lebesgue-Stieltjes integral a.s.*

Proof. First assume that $E\{\int_0^\infty H_s^2 d[M, M]_s\} < \infty$, which implies that $H \cdot M$ (where $H \cdot M$ denotes the stochastic integral process $(\int_0^t H_s dM_s)_{t\geq 0}$) is a square integrable martingale, and hence by Theorem 4 it is a special semimartingale in the $\mathbb{H}$ filtration. Let $M = N + A$ be the canonical $\mathbb{H}$ decomposition of M, and it follows that $H \cdot M = H \cdot N + H \cdot A$ is the canonical $\mathbb{H}$ decomposition of $H \cdot M$. By the lemma following Theorem 23 of Chap. IV we have $E\{\int_0^\infty H_s^2 d[N, N]_s\} \leq E\{\int_0^\infty H_s^2 d[M, M]_s\} < \infty$ and $E\{\int_0^\infty H_s^2 d[A, A]_s\} \leq E\{\int_0^\infty H_s^2 d[M, M]_s\} < \infty$. This allows us to conclude that H is $(\mathcal{H}^2, M)$ integrable, calculated in the $\mathbb{H}$ filtration.

To remove the assumption $E\{\int_0^\infty H_s^2 d[M,M]_s\} < \infty$, we only need to recall that $H \cdot M$ is assumed to be locally square integrable, and thus take M stopped at a stopping time T_n that makes $H \cdot M$ square integrable, and we are reduced to the previous case. □

Theorem 6 (Jeulin's Lemma). *Let R be a positive, measurable stochastic process. Suppose for almost all s, R_s is independent of $\mathcal{F}_s$; and the law (or distribution) of R_s is μ which is independent of s, with $\mu(\{0\}) = 0$ and $\int_0^\infty x\mu(dx) < \infty$. If a is a positive predictable process with $\int_0^t a_s ds < \infty$ a.s. for each t, then the two sets below are equal almost surely:*

$$\{\int_0^\infty R_s a_s ds < \infty\} = \{\int_0^\infty a_s ds < \infty\} \quad a.s.$$

Proof. We first show that $\{\int_0^\infty R_s a_s ds < \infty\} \subset \{\int_0^\infty a_s ds < \infty\}$ a.s. Let A be an event with $P(A) > 0$, and let $J = 1_A$, $J_t = E\{J|\mathcal{F}_t\}$, the càdlàg version of the martingale. Let $j = \inf_t J_t$. Then $j > 0$ on $\{J = 1\}$. We have a.s.

$$E\{1_A R_t|\mathcal{F}_t\} = \int_0^\infty E\{1_A 1_{\{R_t > u\}}|\mathcal{F}_t\} du. \tag{$*$}$$

Consider next

$$E\{1_A 1_{\{R_t>u\}}|\mathcal{F}_t\} = E\{(1_A - 1_{\{R_t \le u\}})^+|\mathcal{F}_t\} \ge (E\{(1_A - 1_{\{R_t\le u\}})|\mathcal{F}_t\})^+$$

by Jensen's inequality, and this is equal to $(E\{1_A|\mathcal{F}_t\} - \mu(0,u])^+$, where μ is the law of R_t. Continuing with $(*)$ we have

$$E\{1_A R_t|\mathcal{F}_t\} \ge \int_0^\infty (E\{1_A|\mathcal{F}_t\} - \mu(0,u])^+ du = \Phi(E\{1_A|\mathcal{F}_t\}) = \Phi(J_t),$$

where $\Phi(x) = \int_0^\infty (x - \mu(0,u])^+ du$. Note that Φ is increasing and continuous on $[0,1]$, and $\Phi > 0$ on $(0,1]$ because $\mu(\{0\}) = 0$.

Choose $A_n = \{\int_0^\infty R_s a_s ds \le n\}$ for the event A in the foregoing. Then

$$\begin{aligned}\infty > nP(A_n) &\ge E\{1_{A_n}\int_0^\infty R_s a_s ds\} = E\{\int_0^\infty 1_{A_n} R_s a_s ds\} \\ &= E\{\int_0^\infty E\{1_{A_n} R_s|\mathcal{F}_s\} a_s ds\} \ge E\{\int_0^\infty \Phi(J_s) a_s ds\},\end{aligned}$$

which implies that $\int_0^\infty \Phi(J_s) a_s ds < \infty$ a.s. But

$$\infty > \int_0^\infty \Phi(J_s) a_s ds \ge \Phi(j) \int_0^\infty a_s ds,$$

and therefore $\int_0^\infty a_s ds < \infty$ a.s. on $\{J = 1\}$. That is, $\int_0^\infty a_s ds < \infty$ a.s. on A_n. Letting n tend to ∞ through the integers gives $\{\int_0^\infty R_s a_s ds < \infty\} \subset$

$\{\int_0^\infty a_s ds < \infty\}$ a.s. We next show the inverse inclusion. Let $T_n = \inf\{t > 0 : \int_0^t a_s ds > n\}$. Then T_n is a stopping time, and $n \geq E\{\int_0^{T_n} a_s ds\}$. Moreover

$$\begin{aligned} E\{\int_0^{T_n} R_s a_s ds\} &= E\{\int_0^{T_n} E\{R_s|\mathcal{F}_{s\wedge T_n}\}a_s ds\} \\ &= E\{\int_0^{T_n} E\{E\{R_s|\mathcal{F}_s\}|\mathcal{F}_{s\wedge T_n}\}a_s ds\} = E\{\int_0^{T_n} E\{R_s\}a_s ds\} \\ &= \alpha E\{\int_0^{T_n} a_s ds\} \leq \alpha n < \infty \end{aligned}$$

where α is the expectation of R_s, which is finite and constant by hypothesis. Therefore $\int_0^{T_n} R_s a_s ds < \infty$ a.s. Let ω be in $\{\int_0^\infty a_s ds < \infty\}$. Then there exists an n (depending on ω) such that $T_n(\omega) = \infty$. Therefore we have the inclusion $\{\int_0^\infty R_s a_s ds < \infty\} \supset \{\int_0^\infty a_s ds < \infty\}$ a.s. and the proof is complete. □

We are now ready to study the Brownian case. The next theorem gives the main result.

Theorem 7. *Let M be a local martingale defined on the standard space of canonical Brownian motion. Let $\mathbb{H}$ be the minimal expanded filtration containing B_1 and satisfying the usual hypotheses. Then M is an $\mathbb{H}$ semimartingale if and only if the integral $\int_0^1 \frac{1}{\sqrt{1-s}}|d[M,B]_s| < \infty$ a.s. In this case $M_t - \int_0^{t\wedge 1} \frac{B_1 - B_s}{1-s} d[M,B]_s$ is an $\mathbb{H}$ local martingale.*

Proof. By the Martingale Representation Theorem we have that every $\mathbb{F}$ local martingale M has a representation $M_t = M_0 + \int_0^t H_s dB_s$, where H is predictable and $\int_0^t H_s^2 ds < \infty$ a.s., each $t > 0$. By Theorem 5 we know that M is an $\mathbb{H}$ semimartingale if and only if $\int_0^t |H_s| \frac{|B_1 - B_s|}{1-s} ds$ is finite a.s., $0 \leq t \leq 1$. We take $a_s = \frac{|H_s|}{\sqrt{1-s}}$, and $R_s = 1_{\{s<1\}} \frac{|B_1 - B_s|}{\sqrt{1-s}}$. Then $\int_0^1 |H_s| \frac{|B_1 - B_s|}{1-s} ds = \int_0^1 a_s R_s ds$, which is finite only if $\int_0^1 a_s ds < \infty$ a.s. by Jeulin's Lemma. Thus it is finite only if $\int_0^1 \frac{|H_s|}{\sqrt{1-s}} ds < \infty$ a.s. But

$$\begin{aligned} \int_0^1 \frac{|H_s|}{\sqrt{1-s}} ds &= \int_0^t \frac{1}{\sqrt{1-s}} |H_s| \, d[B,B]_s = \int_0^t \frac{1}{\sqrt{1-s}} d[|H| \cdot B, B]_s \\ &= \int_0^t \frac{1}{\sqrt{1-s}} |d[H \cdot B, B]_s| = \int_0^t \frac{1}{\sqrt{1-s}} |d[M,B]_s|. \end{aligned}$$

and this completes the proof. □

As an example, let $1/2 < \alpha < 1$, and define

$$H_s = \frac{1}{\sqrt{1-s}}(-\ln(1-s))^{-\alpha} 1_{\{\frac{1}{2}<s<1\}}.$$

Then H is trivially predictable and also $\int_0^1 H_s^2 ds < \infty$. However $\int_0^1 H_s \frac{1}{\sqrt{1-s}} ds$ is divergent. Therefore $M = H \cdot B$ is an $\mathbb{F}$ local martingale which is *not* an $\mathbb{H}$ semimartingale, by Theorem 7, where of course $\mathbb{H} = \mathbb{H}(B_1)$. Thus we conclude that not all $\mathbb{F}$ local martingales (and hence *a fortiori* not all semimartingales) remain semimartingales in the $\mathbb{H}$ filtration.

We now turn to a general criterion that allows the expansion of filtration such that all semimartingales remain semimartingales in the expanded filtration. It is due to Jacod, and it is Theorem 10. The idea is surprisingly simple: recall that for a càdlàg adapted process X to be a semimartingale, if H^n is a sequence of simple predictable processes tending uniformly in (t, ω) to zero, then we must have also that the stochastic integrals $H^n \cdot X$ tend to zero in probability. If we expand the filtration by adding a σ-algebra generated by a random variable L to the $\mathbb{F}$ filtration at time 0 (that is, $\sigma\{L\}$ is added to $\mathcal{F}_0$), then we obtain more simple predictable processes, and it is harder for X to stay a semimartingale. We will find a simple condition on the random variable L which ensures that this condition is not violated. This approach is inherently simpler than trying to show there is a new decomposition in the expanded filtration.

We assume that L is an $(\mathbb{E}, \mathcal{E})$-valued random variable, where E is a *standard Borel space*[1] and $\mathcal{E}$ are its Borel sets, and we let $\mathbb{H}(L)$ denote the smallest filtration satisfying the usual hypotheses and containing both L and the original filtration $\mathbb{F}$. When there is no possibility of confusion, we will write $\mathbb{H}$ in place of $\mathbb{H}(L)$. Note that if $Y \in \mathcal{H}_t^0 = \mathcal{F}_t \vee \sigma\{L\}$, then Y can be written $Y(\omega) = G(\omega, L(\omega))$, where $(\omega, x) \mapsto G(\omega, x)$ is an $\mathcal{F}_t \otimes \mathcal{E}$ measurable function. We next recall two standard theorems from elementary probability theory.

Theorem 8. *Let X^n be a sequence of real-valued random variables. Then X^n converges to 0 in probability if and only if* $\lim_{n\to\infty} E\{\min(1, |X^n|)\} = 0$.

A proof of Theorem 8 can be found in textbooks on probability (see for example [109]).

We write $1 \wedge |X^n|$ for $\min(1, |X^n|)$. Also, given a random variable L, we let $Q_t(\omega, dx)$ denote the **regular conditional distribution of** L with respect to $\mathcal{F}_t$, each $t \geq 0$. That is, for any $\Lambda \in \mathcal{E}$ fixed, $Q_t(\cdot, \Lambda)$ is a version of $E\{1_{\{L\in\Lambda\}} | \mathcal{F}_t\}$, and for any fixed ω, $Q_t(\omega, dx)$ is a probability on $\mathcal{E}$. A second standard elementary result is the following.

Theorem 9. *Let L be a random variable with values in a standard Borel space. Then there exists a regular conditional distribution $Q_t(\omega, dx)$ which is a version of $E\{1_{\{L\in dx\}} | \mathcal{F}_t\}$.*

For a proof of Theorem 9 the reader can see, for example, Breiman [23, page 79].

[1] $(\mathbb{E}, \mathcal{E})$ is a *standard Borel space* if there is a set $\Gamma \in \mathcal{B}$, where $\mathcal{B}$ are the Borel subsets of $\mathbb{R}$, and an injective mapping $\phi : \mathbb{E} \to \Gamma$ such that ϕ is $\mathcal{E}$ measurable and ϕ^{-1} is $\mathcal{B}$ measurable. Note that $(\mathbb{R}^n, \mathcal{B}^n)$ are standard Borel spaces, $1 \leq n \leq \infty$.

Theorem 10 (Jacod's Criterion). *Let L be a random variable with values in a standard Borel space $(\mathbb{E}, \mathcal{E})$, and let $Q_t(\omega, dx)$ denote the regular conditional distribution of L given $\mathcal{F}_t$, each $t \geq 0$. Suppose that for each t there exists a positive σ-finite measure η_t on $(\mathbb{E}, \mathcal{E})$ such that $Q_t(\omega, dx) \ll \eta_t(dx)$ a.s. Then every $\mathbb{F}$ semimartingale X is also an $\mathbb{H}(L)$ semimartingale.*

Proof. Without loss of generality we assume $Q_t(\omega, dx) \ll \eta_t(dx)$ surely. Then by Doob's Theorem on the disintegration of measures there exists an $\mathcal{E} \otimes \mathcal{F}_t$ measurable function $q_t(x, \omega)$ such that $Q_t(\omega, dx) = q_t(x, \omega)\eta_t(dx)$. Moreover since $E\{\int_{\mathbb{E}} Q_t(\cdot, dx)\} = E\{E\{1_{\{Y \in \mathbb{E}\}}|\mathcal{F}_t\}\} = P(Y \in \mathbb{E}) = 1$, we have

$$E\{\int_{\mathbb{E}} Q_t(\cdot, dx)\} = E\{\int_{\mathbb{E}} q_t(x, \omega)\eta_t(dx)\} = \int_{\mathbb{E}} E\{q_t(x, \omega)\}\eta_t(dx) = 1.$$

Hence for almost all x (under $\eta_t(dx)$), we have $q_t(x, \cdot) \in L^1(dP)$.

Let X be an $\mathbb{F}$ semimartingale, and suppose that X is not an $\mathbb{H}(L)$ semimartingale. Then there must exist a $u > 0$ and an $\varepsilon > 0$, and a sequence H^n of simple predictable processes for the $\mathbb{H}$ filtration, tending uniformly to 0 but such that $\inf_n E\{1 \wedge |H^n \cdot X|\} \geq \varepsilon$. Let us suppose that $t_n \leq u$, and

$$H_t^n = \sum_{i=0}^{n-1} J_i^n 1_{(t_i, t_{i+1}]}(t)$$

with $J_i^n \in \mathcal{F}_{t_i} \vee \sigma\{L\}$. Hence J_i^n has the form $g_i(\omega, L(\omega))$, where $(\omega, x) \mapsto g_i(\omega, x)$ is $\mathcal{F}_{t_i} \otimes \mathcal{E}$ measurable. Since H^n is tending uniformly to 0, we can take without loss $|H^n| \leq 1/n$, and thus we can also assume that $|g_i| \leq 1/n$.

We write

$$H_t^{n,x}(\omega) = \sum_{i=0}^{n-1} g_i(\omega, x) 1_{(t_i, t_{i+1}]}(t),$$

and therefore $(x, \omega) \mapsto H_t^{n,x}(\omega)$ and $(x, \omega) \mapsto (H^{n,x} \cdot X)_t(\omega)$ are each $\mathcal{E} \otimes \mathcal{F}_u$ measurable, $0 \leq t \leq u$. Moreover one has clearly $H^n \cdot X = H^{n,L} \cdot X$. Combining the preceding, we have

$$\begin{aligned} E\{1 \wedge |H^n \cdot X_u|\} &= E\{\int_{\mathbb{E}} (1 \wedge |H^{n,x} \cdot X_u|) Q_u(\cdot, dx)\} \\ &= E\{\int_{\mathbb{E}} (1 \wedge |H^{n,x} \cdot X_u|) q_u(\cdot, dx)\eta_u(dx)\} \\ &= \int_{\mathbb{E}} E\{(1 \wedge |H^{n,x} \cdot X_u|) q_u(\cdot, dx)\}\eta_u(dx) \end{aligned}$$

where we have used Fubini's Theorem to obtain the last equality. However the function $h_n(x) = E\{(1 \wedge |H^{n,x} \cdot X_u|) q_u(\cdot, x)\} \leq E\{q_u(\cdot, x)\} \in L^1(d\eta_u)$, and since h_n is non-negative, we have

$$\begin{aligned}\lim_{n\to\infty} E\{1\wedge|H^n\cdot X_u|\} &= \lim_{n\to\infty}\int_{\mathbb{E}} E\{(1\wedge|H^n\cdot X_u|)q_u(\cdot,x)\}\eta_u(dx)\\ &= \int_{\mathbb{E}}\lim_{n\to\infty} E\{(1\wedge|H^n\cdot X_u|)q_u(\cdot,x)\}\eta_u(dx) \qquad (*)\end{aligned}$$

by Lebesgue's Dominated Convergence Theorem. However $q_u(\cdot,x)\in L^1(dP)$ for a.a. x (under $(d\eta_u)$), and if we define $dR = cq_u(\cdot,x)dP$ to be another probability, then convergence in P-probability implies convergence in R-probability, since $R \ll P$. Therefore $\lim_{n\to\infty} E_R\{(1\wedge|H^{n,x}\cdot X_u|)\} = 0$ as well, which implies

$$\begin{aligned}0 &= \lim_{n\to\infty}\frac{1}{c}E_R\{(1\wedge|H^{n,x}\cdot X_u|)\}\\ &= E_P\{(1\wedge|H^{n,x}\cdot X_u|)q_u(\cdot,x)\}\end{aligned}$$

for a.a. x (under $(d\eta_u)$) such that $q_u(\cdot,x)\in L^1(dP)$. Therefore the limit of the integrand in $(*)$ is zero for a.a. x (under $(d\eta_u)$), and we conclude

$$\lim_{n\to\infty} E\{(1\wedge|H^n\cdot X_u|)\} = 0,$$

which is a contradiction. Hence X must be a semimartingale for the filtration $\mathbb{H}^0$, where $\mathcal{H}_t^0 = \mathcal{F}_t\vee\sigma\{L\}$. Let $X = M + A$ be a decomposition of X under $\mathbb{H}^0$. Since $\mathbb{H}^0$ need not be right continuous, the local martingale M need not be right continuous. However if we define $\tilde{M}_t = M_t$ if t is rational; and $\lim_{u\searrow t, u\in\mathbb{Q}}\tilde{M}_u$ if t is not rational; then $\tilde{M}_t$ is a right continuous martingale for the filtration $\mathbb{H}$ where $\mathcal{H}_t = \bigcap_{u>t}\mathcal{H}_u^0$. Letting $\tilde{A}_t = X_t - \tilde{M}_t$, we have that $X_t = \tilde{M}_t + \tilde{A}_t$ is an $\mathbb{H}$ decomposition of X, and thus X is an $\mathbb{H}$ semimartingale. □

A simple but useful refinement of Jacod's Theorem is the following where we are able to replace the family of measures η_t be a single measure η.

Theorem 11. *Let L be a random variable with values in a standard Borel space $(\mathbb{E},\mathcal{E})$, and let $Q_t(\omega,dx)$ denote the regular conditional distribution of L given $\mathcal{F}_t$, each $t\geq 0$. Then there exists for each t a positive σ-finite measure η_t on $(\mathbb{E},\mathcal{E})$ such that $Q_t(\omega,dx)\ll\eta_t(dx)$ a.s. if and only if there exists* **one** *positive σ-finite measure $\eta(dx)$ such that $Q_t(\omega,dx)\ll\eta(dx)$ for all ω, each $t>0$. In this case η can be taken to be the distribution of L.*

Proof. It suffices to show that the existence of η_t for each $t>0$ implies the existence of η with the right properties; we will show that the distribution measure of L is such an η. As in the proof of Theorem 10 let $(x,\omega)\mapsto q_t(x,\omega)$ be $\mathcal{E}\otimes\mathcal{F}_t$ measurable such that $Q_t(\omega,dx) = q_t(x,\omega)\eta_t(dx)$. Let $a_t(x) = E\{q_t(x,\omega)\}$, and define

$$r_t(x,\omega) = \begin{cases}\frac{q_t(x,\omega)}{a_t(x)}, & \text{if } a_t(x)>0,\\ 0, & \text{otherwise.}\end{cases}$$

Note that $a_t(x) = 0$ implies $q_t(x, \cdot) = 0$ a.s. Hence, $q_t(x, \omega) = r_t(x, \omega)a_t(x)$ a.s.; whence $r_t(x, \omega)a_t(x)\eta_t(dx)$ is also a version of $Q_t(\omega, dx)$.

Let η be the law of L. Then for every positive $\mathcal{E}$ measurable function g we have

$$\begin{aligned}\int g(x)\eta(dx) = E\{g(L)\} &= E\{\int_{\mathbb{E}} g(x)Q_t(\cdot, dx)\}\\ &= E\{\int_{\mathbb{E}} g(x)q_t(x,\cdot)\eta_t(dx)\}\\ &= \int_{\mathbb{E}} g(x)E\{q_t(x,\cdot)\}\eta_t(dx)\\ &= \int_{\mathbb{E}} g(x)a_t(x)\eta_t(dx)\end{aligned}$$

from which we conclude that $a_t(x)\eta_t(dx) = \eta(dx)$. Hence, $Q_t(\omega, dx) = r_t(\omega, x)\eta(dx)$, and the theorem is proved. □

We are now able to re-prove some of the previous theorems, which can be seen as corollaries of Theorem 11.

Corollary 1 (Independence). Let L be independent of the filtration $\mathbb{F}$. Then every $\mathbb{F}$ semimartingale is also an $\mathbb{H}(L)$ semimartingale.

Proof. Since L is independent of $\mathcal{F}_t$, $E\{g(L)|\mathcal{F}_t\} = E\{g(L)\}$ for any bounded, Borel function g. Therefore

$$E\{g(L)|\mathcal{F}_t\} = \int_{\mathbb{E}} Q_t(\omega, dx)g(x) = \int_{\mathbb{E}} \eta(dx)g(x) = E\{g(L)\},$$

from which we deduce $Q_t(\omega, dx) = \eta(dx)$, and in particular $Q_t(\omega, dx) \ll \eta(dx)$ a.s., and the result follows from Theorem 11. □

Corollary 2 (Countably-valued random variables). Let L be a random variable taking on only a countable number of values. Then every $\mathbb{F}$ semimartingale is also an $\mathbb{H}(L)$ semimartingale.

Proof. Let L take on the values α_1, α_2, $\alpha_3, \ldots$. The distribution of L is given by $\eta(dx) = \sum_{i=1}^{\infty} P(L = \alpha_i)\varepsilon_{\alpha_i}(dx)$, where $\varepsilon_{\alpha_i}(dx)$ denotes the point mass at α_i. With the notation of Theorem 11, we have that the regular conditional distribution of L given $\mathcal{F}_t$, denoted $Q_t(\omega, dx)$, has density with respect to η given by

$$\sum_j \frac{P(L = \alpha_j|\mathcal{F}_t)}{P(L = \alpha_j)} 1_{\{x=\alpha_j\}}.$$

The result now follows from Theorem 11. □

Corollary 3 (Jacod's Countable Expansion). Let $\mathcal{A} = (A_1, A_2, \dots)$ be a sequence of events such that $A_i \cap A_j = \emptyset$, $i \neq j$, all in $\mathcal{F}$, and such that $\bigcup_{i=1}^{\infty} A_i = \Omega$. Let $\mathbb{H}$ be the filtration generated by $\mathbb{F}$ and $\mathcal{A}$, and satisfying the usual hypotheses. Then every $\mathbb{F}$ semimartingale is an $\mathbb{H}$ semimartingale.

Proof. Define $L = \sum_{i=1}^{\infty} 2^{-i} 1_{A_i}$. Then $\mathbb{H} = \mathbb{H}(L)$ and we need only to apply the preceding corollary. □

Next we consider several examples.

Example (Itô's example). We first consider the original example of Itô, where in the standard Brownian case we expand the natural filtration $\mathbb{F}$ with $\sigma\{B_1\}$. We let $\mathbb{H}$ denote $\mathbb{H}(B_1)$. We have

$$E\{g(B_1)|\mathcal{F}_t\} = E\{g(B_1 - B_t + B_t)|\mathcal{F}_t\} = \int g(x + B_t)\eta_t(dx)$$

where $\eta_t(dx)$ is the law of $B_1 - B_t$ and where we have used that $B_1 - B_t$ is independent of $\mathcal{F}_t$. Note that $\eta_t(dx)$ is a Gaussian distribution with mean 0 and variance $(1-t)$ and thus has a density with respect to Lebesgue measure. Since Lebesgue measure is translation invariant this implies that $Q_t(\omega, dx) \ll dx$ a.s., each $t < 1$. However at time 1 we have $E\{g(B_1)|\mathcal{F}_1\} = g(B_1)$, which yields $Q_1(\omega, dx) = \varepsilon_{\{B_1(\omega)\}}(dx)$, which is a.s. singular with respect to Lebesgue measure. We conclude that any $\mathbb{F}$ semimartingale is also an $\mathbb{H}(B_1)$ semimartingale, for $0 \leq t < 1$, but not necessarily including 1. This agrees with Theorem 7 which implies that there exist local martingales in $\mathbb{F}$ which are not semimartingales in $\mathbb{H}(B_1)$.

Our next example shows how Jacod's criterion can be used to show a somewhat general, yet specific result on the expansion of filtrations.

Example (Gaussian expansions). Let $\mathbb{F}$ again be the standard Brownian filtration satisfying the usual hypotheses, with B a standard Brownian motion. Let $V = \int_0^{\infty} g(s)dB_s$, where $\int_0^{\infty} g(s)^2 ds < \infty$, g a deterministic function. Let $a = \inf\{t > 0 : \int_t^{\infty} g(s)^2 ds = 0\}$. If h is bounded Borel, then as in the previous example

$$\begin{aligned} E\{h(V)|\mathcal{F}_t\} &= E\{h(\int_0^t g(s)dB_s + \int_t^{\infty} g(s)dB_s)|\mathcal{F}_t\} \\ &= \int h(\int_0^t g(s)dB_s + x)\,\eta_t(dx), \end{aligned}$$

where η_t is the law of the Gaussian random variable $\int_t^{\infty} g(s)dB_s$. If $a = \infty$, then η_t is non-degenerate for each t, and η_t of course has a density with respect to Lebesgue measure. Since Lebesgue measure is translation invariant, we conclude that the regular conditional distribution of $Q_t(\omega, dx)$ of V given $\mathcal{F}_t$ also has a density, because

$$Q_t(\omega, h) = E\{h(V)|\mathcal{F}_t\} = \int h(\int_0^t g(s)dB_s + x)\eta_t(dx).$$

Hence by Theorem 10 we conclude that every $\mathbb{F}$ semimartingale is an $\mathbb{H}(V)$ semimartingale.

Example (expansion via the end of a stochastic differential equation). Let B be a standard Brownian motion and let X be the unique solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds$$

where σ and b are Lipschitz. In addition, assume σ and b are chosen so that for h Borel and bounded,

$$E\{h(X_1)|\mathcal{F}_t\} = \int h(x)\pi(1-t, X_t, x)dx$$

where $\pi(1-t, u, x)$ is a deterministic function.[2] Thus $Q_t(\omega, dx) = \pi(1-t, X_t(\omega), x)dx$, and $Q_t(\omega, dx)$ is a.s. absolutely continuous with respect to Lebesgue measure if $t < 1$. Hence if we expand the Brownian filtration $\mathbb{F}$ by initially adding X_1, we have by Theorem 10 that every $\mathbb{F}$ semimartingale is an $\mathbb{H}(X_1)$ semimartingale, for $0 \leq t < 1$.

The mirror of initial expansions is that of **filtration shrinkage**. This has not been studied to any serious extent. We include one result (Theorem 12 below), which can be thought of as a strengthening of Stricker's Theorem, from Chap. II. Recall that if X is a semimartingale for a filtration $\mathbb{H}$, then it is also a semimartingale for any subfiltration $\mathbb{G}$, provided X is adapted to $\mathbb{G}$, by Stricker's Theorem. But what if a subfiltration $\mathbb{F}$ is so small that X is not adapted to it? This is the problem we address. We will deal with the optional projection Z of X onto $\mathbb{F}$.

Definition. Let $H = (H_t)_{t \geq 0}$ be a bounded measurable process. It can be shown that there exists a unique optional process ${}^oH$, also bounded, such that for any stopping time T one has

$$E\{H_T 1_{\{T<\infty\}}\} = E\{{}^oH_T 1_{\{T<\infty\}}\}.$$

The process ${}^oH$ is called **the optional projection of** H.

We remark that ${}^oH_t = H_t$ a.s. for each fixed time t, but the null set depends on t. Therefore were we simply to write as a process $(E\{H_t|\mathcal{F}_t\})_{t \geq 0}$, instead of ${}^oH$, it would not be uniquely determined almost surely. This is why we use the optional projection. The uniqueness follows from Meyer's section theorems, which are not treated in this book, so we ask the reader to accept it on faith.

[2] Sufficient conditions are known for this to be true. These conditions involve X_0 having a nice density, and requirements on the differentiability of the coefficients. See, for example, [190] and [193].

Definition. Let $H = (H_t)_{t\geq 0}$ be a bounded measurable process. It can be shown that there exists a predictable process ${}^pH$, also bounded, such that for any predictable stopping time T one has

$$E\{H_T 1_{\{T<\infty\}}\} = E\{{}^pH_T 1_{\{T<\infty\}}\}.$$

The process ${}^pH$ is called **the predictable projection of** H.

It follows that for the optional projection, for each stopping time T we have

$${}^oH_T = E\{H_T|\mathcal{F}_T\} \text{ a.s. on } \{T<\infty\}$$

whereas for the predictable projection we have that

$${}^pH_T = E\{H_T|\mathcal{F}_{T-}\} \text{ a.s. on } \{T<\infty\}$$

for any *predictable* stopping time T. For fixed times $t > 0$ we have then of course ${}^oH_t = E\{H_t|\mathcal{F}_t\}$ a.s., and one may wonder why we don't simply use the "process" $(E\{H_t|\mathcal{F}_t\})_{t\geq 0}$ instead of the more complicated object ${}^oH$. The reason is that this is defined only almost surely for each t, and we have an uncountable number of null sets; therefore it does not uniquely define a process. (Related to this observation, note in contrast that in some cases $(E\{H_t|\mathcal{F}_t\})_{t\geq 0}$ might exist even when ${}^oH$ does not). We begin our treatment with two simple lemmas.

Lemma. Let $\mathbb{F} \subset \mathbb{G}$, and let X be a $\mathbb{G}$ martingale but not necessarily adapted to $\mathbb{F}$. Let Z denote the optional projection of X onto $\mathbb{F}$. Then Z is an $\mathbb{F}$ martingale.

Proof. Let $s < t$. Then

$$Z_s = E\{X_s|\mathcal{F}_s\} = E\{E\{X_t|\mathcal{G}_s\}|\mathcal{F}_s\} = E\{E\{X_t|\mathcal{G}_t\}|\mathcal{F}_s\} = E\{Z_t|\mathcal{F}_s\}$$

and the result follows. □

The Azéma martingale, a projection of Brownian motion onto a subfiltration to which it is not adapted, is an example of the above lemma. The projection of an increasing process, however, is a submartingale but not an increasing process.

Lemma. Let $\mathbb{F} \subset \mathbb{G}$, and let X be a $\mathbb{G}$ supermartingale but not necessarily adapted to $\mathbb{F}$. Let Z denote the optional projection of X onto $\mathbb{F}$. Then Z is an $\mathbb{F}$ supermartingale.

Proof. Let $s < t$. Then

$$\begin{aligned} Z_s = E\{X_s|\mathcal{F}_s\} \geq E\{E\{X_t|\mathcal{G}_s\}|\mathcal{F}_s\} = E\{X_t|\mathcal{F}_s\} &= E\{E\{X_t|\mathcal{F}_t\}|\mathcal{F}_s\} \\ &= E\{Z_t|\mathcal{F}_s\}, \end{aligned}$$

and the result follows. □

The limitation of the two preceding lemmas is the need to require integrability of the random variables X_t for each $t \geq 0$. We can weaken this condition by a localization procedure.

Definition. We say that a $\mathbb{G}$ semimartingale X starting at 0 is an $\mathbb{F}$ **special, $\mathbb{G}$ semimartingale** if there is a sequence $(T_n)_{n\geq 1}$ of $\mathbb{F}$ stopping times increasing a.s. to ∞, and such that the stopped processes X^{T_n} can be written in the form $X^{T_n} = M^n + A^n$ where M^n is a $\mathbb{G}$ martingale with $M_0^n = 0$ and where A^n has integrable variation over each $[0, t]$, each $t > 0$, and with $A_0 = 0$.

Theorem 12 (Filtration Shrinkage). *Let $\mathbb{G}$ be a given filtration and let $\mathbb{F}$ be a subfiltration of $\mathbb{G}$. Let X be an $\mathbb{F}$ special, $\mathbb{G}$ semimartingale. Then the $\mathbb{F}$ optional projection of X, called Z, exists, and it is a special semimartingale for the $\mathbb{F}$ filtration.*

Proof. Without loss of generality we can assume $T_n \leq n$ for each $n \geq 1$. We set $T_0 = 0$ and let $X^n = X^{T_n} - X^{T_{n-1}}$, and $N^n = M^{T_n} - M^{T_{n-1}}$ with $N^0 = 0$. For each n there are two increasing processes C^n and D^n, each starting at 0, with $X^n = N^n + C^n - D^n$, and moreover we can choose this decomposition such that the following holds:

$$E\{C_\infty^n\} + E\{D_\infty^n\} + E\{\sup_t |N_t|\} < \infty,$$

and where $t \leq T_{n-1}$ implies $C_t^n = D_t^n = N_t^n = 0$, and $t \geq T_n$ implies $C_t^n - C_{T_n}^n = D_t^n - D_{T_n}^n = N_t^n - N_{T_n}^n = 0$. The integrability condition implies that the $\mathbb{F}$ optional projections of C^n, D^n, and N^n all exist and have càdlàg versions. By the previous two lemmas the optional projection of N^n is an $\mathbb{F}$ martingale, and those of C^n and D^n are $\mathbb{F}$ submartingales. Therefore letting ${}^oX^n$, ${}^oN^n$, ${}^oC^n$, and ${}^oD^n$ denote the respective $\mathbb{F}$ optional projections of X^n, N^n, C^n, and D^n, we have that ${}^oX^n = {}^oN^n + {}^oC^n - {}^oD^n$ exists and is a special $\mathbb{F}$ semimartingale.

Since T_{n-1} and T_n are $\mathbb{F}$ stopping times, we have that also the $\mathbb{F}$ optional projections ${}^oN^n$, ${}^oC^n$, and ${}^oD^n$ and hence ${}^oX^n$ are all null over the stochastic interval $[0, T_{n-1}]$ and constant over (T_n, ∞). Then $\sum_{n\geq 1} {}^oX^n$ is a càdlàg version of ${}^oX = Z$ and thus Z is a special $\mathbb{F}$ semimartingale. □

3 Progressive Expansions

We consider the case where we add a random variable gradually to a filtration in order to create a minimal expanded filtration allowing it to be a stopping time. Note that if the initial filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is given by $\mathcal{F}_t = \{\emptyset, \Omega\}$ for all t, then $\mathbb{G} = (\mathcal{G})_{t\geq 0}$ given by $\mathcal{G}_t = \sigma\{L \wedge s; s \leq t\}$ is the smallest expansion of $\mathbb{F}$ making L a stopping time. Note that $\mathcal{G}_t = \sigma\{L \wedge t\}$ as well. Let L be a strictly positive random variable. Let $\Lambda = \{(t, \omega) : t \leq L(\omega)\}$.

Then $L = \sup\{t : (t,\omega) \in \Lambda\}$. In this sense every positive random variable is the end of a random set. Instead however let us begin with a random set $\Lambda \subset \mathbb{R}_+ \times \Omega$ and define L to be the end of the set Λ. That is,

$$L(\omega) = \sup\{t : (t,\omega) \in \Lambda\}$$

where we use the (unusual) convention that $\sup(\emptyset) = 0-$, where $\{0-\}$ is an extra isolated point added to the non-negative reals $[0,\infty]$ and which can be thought of as $0- < 0$. We also define $\mathcal{F}_{0-} = \mathcal{F}_0$. The purpose of $\{0-\}$ is to distinguish between the events $\{\omega : \Lambda(\omega) = \emptyset\}$ and $\{\omega : \Lambda(\omega) = \{0\}\}$, each of which could potentially be added to the expanded filtration.

The smallest filtration expanding $\mathbb{F}$ and making the random variable L a stopping time is $\mathbb{G}^0$ defined by $\mathcal{G}_t^0 = \mathcal{F}_t \vee \sigma\{L \wedge t\}$; but $\mathbb{G}^0$ is not necessarily right continuous. Thus the smallest expanded filtration making L a stopping time and satisfying the usual hypotheses is $\mathbb{G}$ given by $\mathcal{G}_t = \bigcap_{u>t} \mathcal{G}_u^0$. Nevertheless, it turns out that the mathematics is more elegant if we consider expansions slightly more rich than the minimal ones. In order to distinguish progressive expansions from initial ones, we will change the notation for the expanded filtrations. Beginning with our usual filtered probability space satisfying the usual hypotheses $(\Omega, \mathcal{F}, \mathbb{F}, P)$m where of course $\mathbb{F}$ denotes the filtration $(\mathcal{F})_{t \geq 0}$, and a random variable L which is the end of a random set, we define the *expanded filtration* to be $\mathbb{F}^L$ and is given by

$$\Gamma \in \mathcal{F}_t^L \iff \{\Gamma \in \mathcal{F} \text{ and } \exists\, \Gamma_t \in \mathcal{F}_t \,:\, \Gamma \cap \{L > t\} = \Gamma_t \cap \{L > t\}\}.$$

This filtration is easily seen to satisfy the usual hypotheses, and also it makes L into a stopping time. Thus $\mathbb{G} \subset \mathbb{F}^L$. There are two useful key properties the filtration $\mathbb{F}^L$ enjoys.

Lemma. If H is a predictable process for $\mathbb{F}^L$ then there exists a process J which is predictable for $\mathbb{F}$ such that $H = J$ on $[0, L]$. Moreover, if T is any stopping time for $\mathbb{F}^L$ then there exists an $\mathbb{F}$ stopping time S such that $S \wedge L = T \wedge L$ a.s.

Proof. Let Γ be an event in $\mathcal{F}_t^L$. Then events of the form $(t,\infty) \times \Gamma$ form a generating set for $\mathcal{P}(\mathbb{F}^L)$, the predictable sets for $\mathbb{F}^L$. Let $H_s = 1_{\{(t,\infty)\times\Gamma\}}(s)$ and then take J to be $J_s = 1_{(t,\infty)\times\Gamma_t}(s)$. The first result follows by an application of the Monotone Class Theorem. For the stopping time T, note that it suffices to take $H = 1_{[0,T]}$, and let J be the $\mathbb{F}$ predictable process guaranteed by the first half of this lemma, and take $S = \inf\{t : J_t = 0\}$. □

We next define a measure μ^L on $[0,\infty] \times \Omega$ by

$$\mu^L(J) = E\{J_L 1_{\{L>0-\}}\}$$

for any positive, measurable process J. For such a measure μ^L there exists an increasing process $1_{\{t \geq L\}}$ which is null at $0-$ but which can jump at both 0 and $+\infty$. We will denote $A^L = (A^L_t)_{t\geq 0}$, the **(predictable) compensator of $1_{\{t\geq L\}}$ for the filtration $\mathbb{F}$**. Therefore if J is an $\mathbb{F}$ predictable bounded process we have

$$E\{J_L 1_{\{L>0-\}}\} = E\{\int_{[0,\infty]} J_s dA^L_s\}.$$

We now define what will prove to be a process fundamental to our analysis. The process Z defined below was first used in this type of analysis by J. Azéma [3]. Recall that if H is a (bounded, or integrable) $\mathbb{F}^L$ process, then its optional projection ${}^oH$ onto the filtration $\mathbb{F}$ exists. We define

$$Z_t =^o 1_{\{L>t\}}.$$

Note that $1_{\{L>t\}}$ is decreasing, hence by the lemma preceding Theorem 12 we have that Z is an $\mathbb{F}$ supermartingale. We next prove a needed technical result.

Theorem 13. *The set $\{t : 0 \leq t \leq \infty, Z_{t-} = 0\}$ is contained in the set $(L,\infty]$ and is negligible for the measure dA^L.*

Proof. Let $T(\omega) = \inf\{t \geq 0 : Z_t(\omega) = 0,$ or $Z_{t-}(\omega) = 0$ for $t > 0\}$. Then it is a classic result for supermartingales that for almost all ω the function $t \mapsto Z_t(\omega)$ is null on $[T(\omega),\infty]$. (This result is often referred to as "a non-negative supermartingale sticks at zero.") Thus we can write $\{Z = 0\}$ as the stochastic interval $[T,\infty]$, and on $[0,T)$ we have $Z > 0$, $Z_- > 0$. We have $E\{A^L_\infty A^L_T\} = P(T < L) = E\{Z_T 1_{\{T<\infty\}}\} = 0$, hence dA^L is carried by $[0,T]$. Note that since $d1_{\{L\geq t\}}$ is carried by the graph of L, written $[L]$, we have $L \leq T$, and hence we have $Z > 0$, $Z_- > 0$ on $[0,L)$. Next observe that the set $\{Z_- = 0\}$ is predictable, hence $0 = E\{1_{\{Z_{t-}=0\}}d1_{\{L>t\}}\} = E\{1_{\{Z_{t-}=0\}}dA^L_t\}$ and hence $\{Z_- = 0\}$ is negligible for dA^L. Note that this further implies that $P(Z_{L-} > 0) = 1$, and again that $\{Z_- = 0\} \subset (L,\infty]$. □

We can now give a description of martingales for the filtration $\mathbb{F}^L$, as long as we restrict our attention to processes stopped at the time L. What happens after L is more delicate. For an integrable process J we let ${}^pJ$ denote its predictable projection.

Theorem 14. *Let Y be a random variable with $E\{|Y|\} < \infty$. A right continuous version of the martingale $Y_t = E\{Y|\mathcal{F}^L_t\}$ is given by the formula*

$$Y_t = \frac{1}{Z_t}\,{}^o(Y 1_{\{t<L\}}) + Y 1_{\{t\geq L\}}.$$

Moreover the left continuous version Y_- is given by

$$Y_{t-} = \frac{1}{Z_{t-}}{}^p(Y 1_{(0,L]}) + Y 1_{(L,\infty)}.$$

Proof. Let $\mathcal{O}^L$ denote the optional σ-algebra on $\mathbb{R}_+ \times \Omega$, corresponding to the filtration $\mathbb{F}^L$. On $[0, L)$, $\mathcal{O}^L$ coincides with the trace of $\mathcal{O}$ on $[0, L)$. (By $\mathcal{O}$ we mean of course the optional σ-algebra on $\mathbb{R}_+ \times \Omega$ corresponding to the underlying filtration $\mathbb{F}$.) Moreover on $[L, \infty)$, $\mathcal{O}^L$ coincides with the trace of the σ-algebra $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty$ on $[L, \infty)$. The analogous description of $\mathcal{P}^L$ holds, with $[0, L)$ replaced by $(0, L]$, and with $[L, \infty)$ replaced with (L, ∞). It is then simple to check that the formulas give the *bona fide* conditional expectations, and also the right continuity is easily checked on $[0, L)$ and $[L, \infty)$ separately. The second statement follows since $Z_- > 0$ on $(0, L]$ by Theorem 13. □

We now make a simplifying assumption for the rest of this paragraph. This assumption is often satisfied in the cases of interesting examples, and it allows us to avoid having to introduce the dual optional projection of the measure $\varepsilon_L 1_{\{L>0\}}$.

Simplifying assumption to hold for the rest of this paragraph. *We assume L avoids all $\mathbb{F}$ stopping times. That is, $P(L = T) = 0$ for all $\mathbb{F}$ stopping times T.*

Definition. The martingale M^L given by $M_t^L = A_t^L + Z_t$ is called the **fundamental L martingale.**

Note that it is trivial to check that M^L is in fact a martingale, since A^L is the compensator of $1 - Z$. Note also that $M_\infty^L = A_\infty^L$, since $Z_\infty = 0$. Last, note that it is easy to check that M^L is a square integrable martingale.

Theorem 15. *Let X be a square integrable martingale for the $\mathbb{F}$ filtration. Then $(X_{t\wedge L})_{t\geq 0}$ is a semimartingale for the filtration $\mathbb{F}^L$. Moreover $X_{t\wedge L} - \int_0^{t\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s$ is a martingale in the $\mathbb{F}^L$ filtration.*

Proof. Let C be the (non-adapted) increasing process $C_t = 1_{\{t\geq L\}}$. Since C has only one jump at time L we have $E\{X_L\} = E\{\int_0^\infty X_s dC_s\}$. Since X is a martingale it jumps only at stopping times, hence it does not jump at L, and using that A^L is predictable and hence natural we get

$$
\begin{aligned}
E\{X_L\} &= E\{\int_0^\infty X_{s-} dC_s\} = E\{\int_0^\infty X_{s-} dA_s^L\} = E\{X_\infty A_\infty^L\} \\
&= E\{X_\infty M_\infty^L\} = E\{[X, M^L]_\infty\} = E\{\langle X, M^L\rangle_\infty\}. \qquad (*)
\end{aligned}
$$

Suppose that H is a predictable process for $\mathbb{F}^L$, and J is a predictable process for $\mathbb{F}$ which vanishes on $\{Z_- = 0\}$ and is such that $J = H$ on $(0, L]$. We are assured such a process J exists by the lemma preceding Theorem 13. Suppose first that H has the simple form $H = h1_{(t,\infty)}$ for bounded $h \in \mathcal{F}_t^L$. If j is an $\mathcal{F}_t$ random variable equal to h on $\{t < L\}$, then we can take $J = j1_{(t,\infty)}$ and we obtain $H \cdot X_\infty = h(X_L - X_t)1_{\{t<L\}}$. In this way we can define stochastic integrals for non-adapted simple processes. We have then $E\{(H \cdot X)_\infty\} = E\{(J \cdot X)_L\}$, and using our previous calculation, since $J \cdot X$ is another square integrable martingale, we get

$$E\{(H \cdot X)_\infty\} = E\{\langle J \cdot X, M^L\rangle_\infty\} = E\{\int_0^\infty J_s d\langle X, M^L\rangle_s\}.$$

Since $\langle X, M^L\rangle$ is $\mathbb{F}$ predictable, we can replace H by $\frac{H 1_{(0,L]}}{Z_-}$ because it has the same predictable projection on the support of $d\langle X, M^L\rangle$. This yields

$$E\{\int_0^\infty H_s dX_s\} = E\{\int_0^L \frac{H_s}{Z_{s-}} d\langle X, M^L\rangle_s\}. \qquad (**)$$

Last if we take the bounded $\mathbb{F}^L$ predictable process H to be a stochastic interval $[0, T \wedge L]$, where T is an $\mathbb{F}^L$ stopping time, we obtain $E\{X_{T\wedge L} - \int_0^{T\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s\} = 0$, which implies by Theorem 21 of Chap. I that $X_{T\wedge L} - \int_0^{T\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s$ is a martingale. □

We do not need the assumption that X is a square integrable martingale, which we made for convenience. In fact the conclusion of the theorem holds even if X is only assumed to be a local martingale. We get our main result as a corollary to Theorem 15.

Corollary. Let X be a semimartingale for the $\mathbb{F}$ filtration. Then $(X_{t\wedge L})_{t\geq 0}$ is a semimartingale for the filtration $\mathbb{F}^L$.

Proof. If X is a semimartingale then it has a decomposition $X = M + D$. The local martingale term M can be decomposed into $X = V + N$, where V and N are both local martingales, but V has paths of bounded variation on compact time sets, and N has bounded jumps. (This is the Fundamental Theorem of Local Martingales, Theorem 25 of Chap. III.) Clearly V and D remain finite variation processes in the expanded filtration $\mathbb{F}^L$, and since M has bounded jumps it is locally bounded, hence locally square integrable, and since every $\mathbb{F}$ stopping time remains a stopping time for the $\mathbb{F}^L$ filtration, the corollary follows from Theorem 15. □

We need to add a restriction on the random variable L in order to study the evolution of semimartingales in the expanded filtration after the time L.

Definition. A random variable L is called **honest** if for every $t \leq \infty$ there exists an $\mathcal{F}_t$ measurable random variable L_t such that $L = L_t$ on $\{L \leq t\}$.

Note that in particular if L is honest then it is $\mathcal{F}_\infty$ measurable. Also, any stopping time is honest, since then we can take $L = L \wedge t$ which is of course $\mathcal{F}_t$ measurable by the stopping time property.

Example. Let X be a bounded càdlàg adapted processes, and let $X_t^{+*} = \sup_{s\leq t} X_s^+$ and $X^{+*} = \sup_s X_s^+$. Then $L = \inf\{s : X_s^{+*} = X^{+*}\}$, is honest because on the set $\{L \leq t\}$ one has $L = \inf\{s : X_s = X_s^{+*}\}$, which is $\mathcal{F}_t$ measurable.

Theorem 16. *L is an honest time if and only if there exists an optional set $\Lambda \subset [0,\infty] \times \Omega$ such that $L(\omega) = \sup\{t \leq \infty : (t,\omega) \in \Lambda\}$.*

This is often described verbally by saying "L is honest if it is the end of an optional set."

Proof. The end of an optional set is always an honest random variable. Indeed, on $\{L \leq t\}$, the random variable L coincides with the end of the set $\Lambda \cap ([0,t] \times \Omega)$, which is $\mathcal{F}_t$ measurable.

For the converse we suppose L is honest. Let $(L_t)_{t\geq 0}$ be an $\mathbb{F}$ adapted process such that $L = L_t$ on $\{L \leq t\}$. Since we can replace L_t with $L_t \wedge t$ we can assume without loss of generality that $L_t \leq t$. There is also no loss of generality to assume that L_t is increasing with t , since we can further replace L_t with $\sup_{s\leq t} L_s$. Last, it is also no loss, now that it is increasing, to take it right continuous. We thus have that the process $(L_t)_{t\geq 0}$ is optional for the filtration $\mathbb{F}$. Last, L is now the end of the optional set $\{(t,\omega) : L_t(\omega) = t\}$. □

When L is honest we can give a simple and elegant description of the filtration $\mathcal{F}^L$.

Theorem 17. *Let L be an honest time. Define*

$$\mathcal{G}_t = \{\Gamma : \Gamma = (A \cap \{L > t\}) \cup (B \cap \{L \leq t\}) \text{ for some } A, B \in \mathcal{F}_t\}$$

Then $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ constitutes a filtration satisfying the usual hypotheses. Moreover L is a $\mathbb{G}$ stopping time. A process U is predictable for $\mathbb{G}$ if and only if it has a representation of the form

$$U = H1_{[0,L]} + K1_{(L,\infty]}$$

where H and K are $\mathbb{F}$ predictable processes.

Proof. Let $s < t$ and take $H \in \mathcal{G}_s$, of the form $(A \cap \{L > s\}) \cup (B \cap \{L \leq s\})$ with $A, B \in \mathcal{F}_s$. We will show that $H \in \mathcal{G}_t$, which shows that the collection $\mathbb{G}$ is filtering to the right.[3] Since L is an honest time, there must exist $D \in \mathcal{F}_t$ such that $\{L \leq s\} = D \cap \{L \leq t\}$. Therefore

$$H \cap \{L \leq t\} = [(A \cap D^c) \cup (B \cap D)] \cap \{L \leq t\},$$

with $[(A \cap D^c) \cup (B \cap D)] \in \mathcal{F}_t$. The fact that each $\mathcal{G}_t$ is a σ-algebra, and also that $\mathbb{G}$ is right continuous, we leave to the reader. Note that $\{L \leq t\} \in \mathcal{G}_t$ which implies that L is a $\mathbb{G}$ stopping time, as we observed at the start of the proof of Theorem 16. For the last part of the theorem, let $U = (U_t)_{t\geq 0}$ be a càdlàg process adapted to the $\mathbb{G}$ filtration. Let H_t and K_t be $\mathcal{F}_t$ measurable random variables such that

$$U_t = H_t \text{ on } \{L > t\} \text{ and } U_t = K_t \text{ on } \{L \leq t\}.$$

Next set $H_t = \liminf_{s\nearrow t} H_s; K_t = \liminf_{s\nearrow t} J_s$ for all $t > 0$. Suppose that s is a rational number. Take $t = 0$ and we have the equality

[3] That is, if $s < t$, then $\mathcal{G}_s \subset \mathcal{G}_t$.

$$U = H1_{[0,L]} + K1_{(L,\infty]}.$$

One can extend this result to predictable processes by the Monotone Class Theorem. □

Theorem 18. *Let X be a square integrable martingale for the $\mathbb{F}$ filtration. Then X is a semimartingale for the $\mathbb{G}$ filtration if L is an honest time. Moreover X has a $\mathbb{G}$ decomposition*

$$X_t = \{X_t - \int_0^{t\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s + 1_{\{t\geq L\}} \int_L^t \frac{1}{1-Z_{s-}} d\langle X, M^L\rangle_s\}$$
$$+\{\int_0^{t\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s - 1_{\{t\geq L\}} \int_L^t \frac{1}{1-Z_{s-}} d\langle X, M^L\rangle_s\}.$$

Before beginning the proof of the theorem, we establish a lemma we will need in the proof. It is a small extension of the local behavior of the stochastic integral established in Chap. IV.

Lemma (Local behavior of the stochastic integral at random times). Let X and Y be two semimartingales and let $H \in L(X)$ and $J \in L(Y)$. Let U and V be two positive random variables with $U \leq V$. (U and V are not assumed to be stopping times.) Define

$$\Lambda = \{\omega : H_t(\omega) = J_t(\omega) \text{ and } X_t(\omega) - X_U(\omega) = Y_t(\omega) - Y_U(\omega), \text{ for all } t \in [U(\omega), V(\omega))\}.$$

Let $W_t = H \cdot X_t$ and $Z_t = J \cdot Y_t$. Then a.s. on Λ, $W_t(\omega) - W_U(\omega) = Z_t(\omega) - Z_U(\omega)$ for all $t \in [U(\omega), V(\omega))$.

Proof. We know by the Corollary to Theorem 26 of Chap. IV that the conclusion of the lemma is true when U and V are stopping times. Let (u, v) be rationals with $0 < u < v$, and let

$$\Lambda_{uv} = \{\omega : U(\omega) < u < v < V(\omega)\} \cap \Lambda.$$

Then also $W_t - W_u = Z_t - Z_u$, all $t \in [u, v)$, a.s. on Λ_{uv}, since u, v are fixed times and hence *a fortiori* stopping times. Next let

$$A = \left\{\omega : \bigcap_{\substack{u,v\in\mathbb{Q}_+\\ u<v}} \{U(\omega) < u < v < V(\omega), \text{ and } W_t(\omega) - W_u(\omega) = Z_t(\omega) - Z_u(\omega) \text{ for all } t \in [u, v)\}\right\}.$$

The intersection in the definition of A is countable, so null sets do not accumulate. Finally let u decrease to $U(\omega)$ and then v increase to $V(\omega)$, which gives the result. □

Proof of Theorem 18. We first observe that without loss of generality we can assume $X_0 = 0$. Let H be a bounded $\mathbb{F}$ predictable process. We *define* stochastic integrals at the random time L by

$$\int_0^L H_s dX_s = (H \cdot X)_L, \quad \int_L^\infty H_s dX_s = (H \cdot X)_\infty - (H \cdot X)_L$$

where $H \cdot X = (H \cdot X_t)_{t \geq 0}$ is the $\mathbb{F}$ stochastic integral process. That is, we use the usual definition of the stochastic integral, sampling it at the random time L. When H is a simple predictable process, these are reasonable definitions. Moreover we know from the lemma that if H and J are both bounded $\mathbb{F}$ predictable processes and if also $H = J$ on $(L, \infty]$, then $\int_L^\infty H_s dX_s = \int_L^\infty J_s dX_s$ a.s., so the definition is well-defined.

Since X is an $\mathbb{F}$ martingale in L^2 with $X_0 = 0$, we have $E\{\int_0^\infty H_s dX_s\} = 0$. Applying the equalities $(*)$ on page 372 to $H \cdot X$ we have

$$E\{\int_0^L H_s dX_s\} = E\{\int_0^\infty H_s d\langle X, M^L\rangle_s\} = E\{\int_0^L \frac{H_s}{Z_{s-}} d\langle X, M^L\rangle_s\} \qquad (\dagger)$$

where the second equality uses $(**)$ on page 373. Recalling $E\{\int_0^\infty H_s dX_s\} = 0$ and combining this with the above gives (where we have changed the name of the process H to K for clarity slightly later in the proof)

$$E\{\int_L^\infty K_s dX_s\} = -E\{\int_0^\infty K_s d\langle X, M^L\rangle_s\}$$

We next replace K with $K1_{\{Z_- < 1\}}$ to obtain

$$\begin{aligned} E\{\int_L^\infty K_s dX_s\} &= -E\{\int_0^\infty K_s 1_{\{Z_{s-}<1\}} d\langle X, M^L\rangle_s\} \\ &= -E\{\int_L^\infty \frac{K_s}{1 - Z_{s-}} d\langle X, M^L\rangle_s\} \end{aligned} \qquad (\ddagger)$$

because the predictable projection of $\frac{1}{1-Z_{s-}} 1_{(L,\infty)}(s)$ is $1_{\{Z_{s-}<1\}}$. Finally we combine equalities $(\dagger)$ and $(\ddagger)$, and define $U = H1_{[0,L]} + K1_{(L,\infty]}$ to get

$$\begin{aligned} &E\{\int_0^L H_s dX_s + \int_L^\infty K_s dX_s\} \\ &= E\{\int_0^\infty U_s d\{\int_0^{t\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s - 1_{\{t \geq L\}} \int_L^t \frac{1}{1 - Z_{s-}} d\langle X, M^L\rangle_s\} \end{aligned}$$

Up to this point we have not used the hypothesis that L is an honest time. Now we do. Let us take U to be a simple $\mathbb{G}$ predictable process, and using Theorem 17 we have

$$U = H1_{[0,L]} + K1_{(L,\infty]}$$

where H and K are $\mathbb{F}$ predictable processes (not necessarily simple predictable, however). If we further take U bounded by 1 and take the supremum over all such simple predictable processes of the elementary $\mathbb{G}$ stochastic integrals $\int_0^\infty U_s dX_s$, we get that the variation of X, $\mathrm{Var}(X)$ as defined in Chap. III, is less than or equal the expected total variation of

$$d\alpha_t = d\{\int_0^{t\wedge L} \frac{1}{Z_{s-}} d\langle X, M^L\rangle_s + 1_{\{t\geq L\}} \int_L^t \frac{1}{1-Z_{s-}} d\langle X, M^L\rangle_s\}.$$

Thus the process X is a $\mathbb{G}$ quasimartingale, and hence also a semimartingale. We also note that by the preceding we have that $E\{X_T - \alpha_T\} = 0$ for all $\mathbb{G}$ stopping times T, and thus it is a martingale for the $\mathbb{G}$ filtration. □

The next theorem is really a corollary of Theorem 18.

Theorem 19. *Let X be a semimartingale for the $\mathbb{F}$ filtration. Then X is a semimartingale for the $\mathbb{G}$ filtration if L is an honest time.*

Proof. Let $X = M + A$ be a decomposition of X in the $\mathbb{F}$ filtration, where M is a local martingale and A is a finite variation process. Since A remains a finite variation process in the larger $\mathbb{G}$ filtration, we need only concern ourselves with M. By the Fundamental Theorem of Local Martingales (Theorem 25 on page 125), we know that we can write $M = N + V$ where N is a local martingale with bounded jumps, and V is a martingale with paths of finite variation on compacts. Again, V remains a semimartingale in the $\mathbb{G}$ filtration since it is of finite variation, and further N can be locally stopped with $\mathbb{F}$ stopping times T_n tending to ∞ a.s. to be a square integrable martingale for each T_n. Then by Theorem 18 each X^{T_n} is also a $\mathbb{G}$ semimartingale. Since X is locally a $\mathbb{G}$ semimartingale, it is an actual $\mathbb{G}$ semimartingale. □

4 Time Reversal

In this section we apply the results on initial expansions of filtration to some elementary issues regrading time reversal of semimartingales, stochastic integrals, and stochastic differential equations.

Definition. Let $Y = (Y_t)_{0\leq t\leq 1}$ be a càdlàg process. The **time reversal** of Y on $[0,1]$ is defined to be

$$\tilde{Y}^t = \begin{cases} 0, & \text{if } t = 0, \\ Y_{(1-t)-} - Y_{1-}, & \text{if } 0 < t < 1, \\ Y_0 - Y_{1-}, & \text{if } t = 1. \end{cases}$$

Note that time as a superscript denotes reversal. Let $\mathbb{F}$ denote the forward filtration, and correspondingly let $\tilde{\mathbb{F}} = (\mathcal{F}^t)_{0\leq t\leq 1}$ denote a backward filtration. As an example, if B is a standard Brownian motion, $\mathcal{F}^0_t = \sigma\{B_s; s \leq t\}$, and

$\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N}$, where $\mathcal{N}$ are the P-null sets of $\mathcal{F}_1$, while $\mathcal{F}^t = \mathcal{F}_0^t \vee \mathcal{N}$ is the analogous backward filtration for $\tilde{B}$. Our first two theorems are obvious, and we omit the proofs.

Theorem 20. *Let B be a standard Brownian motion on $[0,1]$. Then the time reversal $\tilde{B}$ of B is also a Brownian motion with respect to its natural filtration.*

Theorem 21. *Let Z be a Lévy process on $[0,1]$. Then $\tilde{Z}$ is also a Lévy process, with the same law as $-Z$, with respect to its natural filtration.*

In what follows let us assume that we are given a forward filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le 1}$ and a backward filtration $\tilde{\mathbb{G}} = (\mathcal{G}^t)_{0\le t\le 1}$. We further assume both filtrations satisfy the usual hypotheses.

Definition. A càdlàg process Y is called an $(\mathbb{F}, \tilde{\mathbb{G}})$ **reversible semimartingale** if Y is an $\mathbb{F}$ semimartingale on $[0,1]$ and $\tilde{Y}$ is a $\tilde{\mathbb{G}}$ semimartingale on $[0,1)$.

Note the small lack of symmetry: for the time reversed process we do not include the final time 1. This is due to the occurrence of singularities at the terminal time 1, and including it as a requirement would exclude interesting examples.

Let $\tau_t = (t_0, \ldots, t_k)$ denote a partition of $[0,t]$ with $t_0 = 0$, $t_k = t$, $0 \le t \le 1$, and let

$$S_{\tau_t}(H,Y) = H_0Y_0 + \sum_i (H_{t_{i+1}} - H_{t_i})(Y_{t_{i+1}} - Y_{t_i}).$$

Definition. Let H, Y be two càdlàg processes. The **quadratic covariation** of H and Y, denoted $[H,Y]$, is defined to be

$$\lim_{n\to\infty} S_{\tau_t^n}(H,Y) = [H,Y]_t$$

when this limit exists in probability as $\text{mesh}(\tau_t) \to 0$, and is such that $[H,Y]$ is càdlàg, adapted, and of finite variation a.s.

Theorem 22. *Let Y be an $(\mathbb{F}, \tilde{\mathbb{G}})$ reversible semimartingale, and let H be a càdlàg process such that $H_t \in \mathcal{F}_t$, and $H_t \in \mathcal{G}^{1-t}, 0 \le t \le 1$. Suppose the quadratic variation $[H,Y]$ exists. Then the two processes $[H,Y]$ and $X_t = \int_0^t H_{s-}dY_s$ are $(\mathbb{F}, \tilde{\mathbb{G}})$ reversible semimartingales. Moreover*

$$\tilde{X}_t + \widetilde{[H,Y]}_t = \int_0^t H_{1-s} d\tilde{Y}^s.$$

Proof. Since $H_t \in \mathcal{G}^{1-t}, 0 \le t \le 1$, we have that $H_{1-s} \in \mathcal{G}^{1-(1-s)} = \mathcal{G}^s$, and thus it is adapted and left continuous, so the stochastic integral $\int_0^t H_{1-s}d\tilde{Y}^s$ is well-defined. We choose and fix t, $0 < t < 1$. Let τ be a partition of $[1-t, 1]$, $\tau = (s_1, \ldots, s_n)$, chosen such that $\Delta H_{s_i} = 0$ and $\Delta Y_{s_i} = 0$ a.s.,

$i = 2, \dots, n-1$. Note that one can always choose the partition points $\{s_i\}$ in this manner because

$$E\{\int_0^1 1_{\{|\Delta H_s|>0\}} 1_{\{|\Delta Y_s|>0\}} ds\} = 0,$$

since for each ω, $s \mapsto Y_s(\omega)$ has only countably many jumps, and hence $\int_0^1 1_{\{|\Delta Y_s|>0\}} ds = 0$ a.s. But then $\int_0^1 P(|\Delta Y_s| > 0) ds = 0$, which in turn implies that $P(|\Delta Y_s| > 0) = 0$ for almost all s in $[0,1]$. We now define three new processes:

$$\begin{aligned}
A^\tau &= H_{(1-t)-}\Delta Y_{1-t} + \sum_{i=1}^{n-2} H_{s_i}(Y_{s_{i+1}} - Y_{s_i}) + H_{s_{n-1}}(Y_{1-} - Y_{s_{n-1}}) \\
B^\tau &= -\sum_{i=1}^{n-1} H_{s_{i+1}-}(Y_{s_{i+1}-} - Y_{s_i-}) \\
C^\tau &= H_{1-t}\Delta Y_{1-t} + \sum_{i=1}^{n-2} \{(H_{s_{i+1}} - H_{s_i})(Y_{s_{i+1}} - Y_{s_i})\} \\
&\quad + (H_{1-} - H_{s_{n-1}})(Y_{1-} - Y_{s_{n-1}}).
\end{aligned}$$

Let τ_n be a sequence of partitions of $[0,1]$ with $\lim_{n\to\infty} \mathrm{mesh}(\tau_n) = 0$. Then

$$\begin{aligned}
\lim_{n\to\infty} C^{\tau_n} &= [H,Y]_{1-} - [H,Y]_{1-t} + \Delta H_{1-t}\Delta Y_{1-t} \\
&= [H,Y]_{1-} - [H,Y]_{(1-t)-} \\
&= -\widetilde{[H,Y]}_t
\end{aligned}$$

Since Y is $(\mathbb{F}, \tilde{\mathbb{G}})$ reversible by hypothesis, we know that C^τ is $\tilde{\mathbb{G}}$ adapted. Hence $\widetilde{[H,Y]}$ is $\tilde{\mathbb{G}}$ adapted and moreover since it has paths of finite variation by hypothesis, it is a semimartingale.

Since H is càdlàg we can approximate the stochastic integral with partial sums, and thus

$$\lim_{n\to\infty} A^{\tau_n} = \int_{[1-t,1)} H_{s-} dY_s = X_{1-} - X_{(1-t)-} = -\tilde{X}^t.$$

Since $Y_{s_{i+1}-} - Y_{s_i-} = -(\tilde{Y}^{1-s_i} - \tilde{Y}^{1-s_{i+1}})$, we have

$$\lim_{n\to\infty} B^{\tau_n} = \int_0^t H_{1-s} d\tilde{Y}^s.$$

Combining A^{τ_n}, B^{τ_n}, and C^{τ_n}, we get

$$\begin{aligned}
&A^{\tau_n} + B^{\tau_n} + C^{\tau_n} \\
&= H_{1-t}\Delta Y_{1-t} + \sum_{i=1}^{n-2} H_{s_{i+1}}(Y_{s_{i+1}} - Y_{s_i}) + H_{1-}(Y_{1-} - Y_{s_{n-1}}) \\
&\quad - \sum_{i=1}^{n-1} H_{s_{i+1}}(Y_{s_{i+1}-} - Y_{s_i-}) \\
&= H_{1-t}\Delta Y_{1-t} + \sum_{i=1}^{n-2} H_{s_{i+1}}(\Delta Y_{s_{i+1}} - \Delta Y_{s_i}) - \Delta H_1(Y_{1-} - Y_{s_{n-1}}) - H_1 \Delta Y_{s_{n-1}}.
\end{aligned}$$

Since we chose our partitions τ_n with the property that $\Delta Y_{s_i} = 0$ a.s. for each partition point s_i, the above simplifies to

$$A^{\tau_n} + B^{\tau_n} + C^{\tau_n} = (H_{1-t} - H_{s_i})\Delta Y_{1-t} + \Delta H_1(Y_{s_{n-1}} - Y_{1-}),$$

which tends to 0 since s_1 decreases to $1-t$, and s_{n-1} increases to 1. Therefore

$$0 = \lim_{n\to\infty} A^{\tau_n} + \lim_{n\to\infty} B^{\tau_n} + \lim_{n\to\infty} C^{\tau_n} = -\tilde{X}^t + \int_0^t H_{1-s} d\tilde{Y}^s - \widetilde{[H,Y]}_t.$$

This establishes the desired formula, and since we have already seen that $\widetilde{[H,Y]}_t$ is a semimartingale, we have that X is also a reversible semimartingale as a consequence of the formula. □

We remark that in the case where $\tilde{Y}$ is a $\tilde{\mathbb{G}}$ semimartingale on $[0,1]$ (and not only on $[0,1)$), the same proof shows that $\tilde{X}$ is a $\tilde{\mathbb{G}}$ semimartingale on $[0,1]$.

Example (reversibility for Lévy process integrals). Let Z be a Lévy process on $[0,1]$, and let $\tilde{\mathbb{H}} = (\tilde{\mathcal{H}}_t)_{0\le t\le 1}$ be the filtration generated by $\tilde{Z}$ and expanded by the addition of $\tilde{Z}^1 = Z_0 - Z_{1-}$. Then $\tilde{Z}$ is an $\tilde{\mathbb{H}}$ semimartingale on $[0,1]$ by Theorem 21. Let $H_t = f(Z_t)$ with f such that the quadratic covariation $[H,Z]$ exists. Clearly $H_t = f(Z_t) \in \mathcal{F}_t$. Since

$$\mathcal{H}^{1-t} \supseteq \sigma\{Z_{(1-u)-} - Z_{(1-v)-};\ 0 \le u,v \le 1\} \vee \sigma\{Z_1\},$$

we have that $Z_t = Z_{1-(1-t)} \in \mathcal{H}^{1-t}$, since $Z_t = Z_{t-}$ a.s. Therefore $H_t \in \mathcal{H}^{1-t}$. Using Theorem 22 we conclude that $X_t = \int_0^t f(Z_{s-})dZ_s$ is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale.

Example (reversibility for Brownian integrals). Let B be a standard Brownian motion on $[0,1]$. Let L_t^x denote Brownian local time at the level x and let μ be a signed measure on $\mathbb{R}$. Let $\mathbb{F}$ be the minimal filtration of B satisfying the usual hypotheses, and let $(\tilde{\mathbb{H}}^t)_{0\le t\le 1}$ be the minimal filtration generated by $\tilde{B}$, expanded by adding $\tilde{B}^1 = -B_1$, and satisfying the usual hypotheses. Finally let f be a càdlàg function on $\mathbb{R}$ of finite variation on compacts with primitive F (that is, $F'_+(x) = f(x)$), and define

$$U_t = \int_0^t f(B_s)dB_s + \int_{\mathbb{R}} L_t^x \mu(dx).$$

We will show U is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale.

The hypotheses on f imply that its primitive F is the difference of two convex functions, and therefore $M_t = F(B_t) - F(B_0)$ is an $\mathbb{F}$ semimartingale. Moreover $\tilde{M}^t = F(B_{1-t}) - F(B_1)$, and since we already know that B_{1-t} is an $\tilde{\mathbb{H}}$ semimartingale by Theorem 3, by the convexity of F we conclude that $\tilde{M}$ is also an $\tilde{\mathbb{H}}$ semimartingale. Therefore $\int_0^t f(B_s)dB_s$ is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale as soon as $\frac{1}{2}\int_{\mathbb{R}} L_t^a \eta(da)$ is one, where η is the signed measure 'second derivative' of F. Finally, what we want to show is that $A_t = \frac{1}{2}\int_{\mathbb{R}} L_t^a \mu(da)$ is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale, for any signed measure μ. This will imply that U is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale.

Since A is continuous with paths of finite variation on $[0,1]$, all we really need to show is that $\tilde{A}^t \in \mathcal{H}^t$, each t. But $\tilde{A}^t = A_{1-t} - A_1 = \int_{\mathbb{R}}(L_{1-t}^a - L_1^a)\mu(da)$, and

$$\begin{aligned}
L_{1-t}^a - L_1^a &= \lim_{\varepsilon \to 0}\Big(-\frac{1}{\varepsilon}\int_{1-t}^1 1_{[a,a+\varepsilon)}(B_s)ds\Big) \\
&= \lim_{\varepsilon \to 0}\Big(-\frac{1}{\varepsilon}\int_{1-t}^1 1_{[a-B_1,a-B_1+\varepsilon)}(B_s - B_1)ds\Big) \\
&= \lim_{\varepsilon \to 0}\Big(-\frac{1}{\varepsilon}\int_{1-t}^1 1_{[a-B_1,a-B_1+\varepsilon)}(\tilde{B}^u)du\Big) \\
&= -\Lambda_t^{a-B_1},
\end{aligned}$$

where Λ_t^x is the local time at level x of the standard Brownian motion $\tilde{B}$. Therefore $\tilde{A}^t = -\int_{\mathbb{R}}(\Lambda_t^{a-B_1})\mu(da)$. Since Λ_t^x is $\tilde{\mathbb{H}}$ adapted, and since $B_1 \in \mathcal{H}^0$, we conclude $\tilde{A}$ is $\tilde{\mathbb{H}}$ adapted, and therefore U is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale.

We next consider the time reversal of stochastic differential equations, which is perhaps more interesting than the previous two examples. Let B be a standard Brownian motion and let X be the unique solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds$$

where σ and b are Lipschitz, and moreover, σ and b are chosen so that for h Borel and bounded,

$$E\{h(X_1)|\mathcal{F}_t\} = \int h(x)\pi(1-t, X_t, x)dx$$

where $\pi(1-t, u, x)$ is a deterministic function. As in the example *expansion via the end of a stochastic differential equation* treated earlier on page 367, we know (as a consequence of Theorem 10) that if $\mathbb{F}$ is the natural, completed Brownian filtration, we can expand $\mathbb{F}$ with X_1 to get $\mathbb{H}$: $\mathcal{H}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma\{X_1\}$, and then all $\mathbb{F}$ semimartingales remain $\mathbb{H}$ semimartingales on $[0,1)$.

We fix (t,ω) and define $\phi : \mathbb{R} \to \mathbb{R}$ by $\phi(x) = X(t,\omega,x)$ where $X_t = X(t,\omega,x)$ is the unique solution of

$$X_t = x + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds$$

for $x \in \mathbb{R}$. Recall from Chap. V that ϕ is called the **flow of the solution of the stochastic differential equation**. Again we have seen in Chap. V that under our hypotheses on σ and b the flow ϕ is injective. For $0 < s < t < 1$ define the function $\phi_{s,t}$ to be the flow of the equation

$$X_{s,t} = x + \int_s^t \sigma(X_{s,u})dB_u + \int_s^t b(X_{s,u})du. \tag{$*$}$$

It now follows from the uniqueness of solutions that $X_t = \phi_{s,t}(X_s)$, and in particular $X_1 = \phi_{t,1}(X_t)$, and therefore $\phi_{t,1}^{-1}(X_1) = X_t$, where $\phi_{t,1}^{-1}$ is of course the inverse function of $\phi_{s,t}$. Since the solution $X_{s,t}$ of equation $(*)$ is $\mathcal{G} = \sigma\{B_v - B_u;\ s \leq u,v \leq t\}$ measurable, we have $\phi_{t,1} \in \mathcal{F}^{1-t}$, where $\mathcal{F}^t = \sigma\{\tilde{B}^s; 0 \leq s \leq t\} \vee \mathcal{N}$ where $\mathcal{N}$ are the null sets of $\mathcal{F}$. Let $\mathcal{H}^t = \bigcap_{u>t} \mathcal{F}^u \vee \sigma\{X_1\}$ and we have $\tilde{B}$ is an $\tilde{\mathbb{H}} = (\mathcal{H}^t)_{0\leq t\leq 1}$ semimartingale, and that $X_t = \phi_{t,1}^{-1}(X_1)$ is $\mathcal{H}^{1-t}$ measurable. Therefore $X_t \in \mathcal{F}_t$ and also at the same time $X_t \in \mathcal{H}^{1-t}$. Finally note that since X is a semimartingale and σ is $\mathcal{C}^1$, the quadratic covariation $[\sigma(X), B]$ exists and is of finite variation, and we are in a position to apply Theorem 22.

Theorem 23. *Let B be a standard Brownian motion and let X be the unique solution of the stochastic differential equation*

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds$$

for $0 \leq t \leq 1$, where σ and b are Lipschitz, and moreover, σ and b are chosen so that for h Borel and bounded,

$$E\{h(X_1)|\mathcal{F}_t\} = \int h(x)\pi(1-t,X_t,x)dx$$

where $\pi(1-t,u,x)$ is a deterministic function. Let $\mathbb{H}$ be given by $\mathcal{H}_t = \bigcap_{u>t} \mathcal{F}_u \vee \sigma\{X_1\}$. Then B is an $(\mathbb{F},\tilde{\mathbb{H}})$ reversible semimartingale, and $\hat{X}^t = X_{1-t}$ satisfies the following backward stochastic differential equation

$$Y_t = Z_1 + \int_0^t \sigma(Y_s)d\tilde{B}^s + \int_0^t \sigma'(Y_s)\sigma(Y_s)ds + \int_0^t b(Y_s)ds,$$

for $0 \leq t \leq 1$. In particular, X is an $(\mathbb{F},\tilde{\mathbb{H}})$ reversible semimartingale.

Proof. We first note that B is an $(\mathbb{F}, \tilde{\mathbb{H}})$ reversible semimartingale as we saw in the example on page 367. We have that $[\sigma(X), B] = \int_0^t \sigma'(X_s)\sigma(X_s)ds$, which is clearly of finite variation, since σ' is continuous, and thus $\sigma'(X)$ has paths bounded by a random constant on $[0, 1]$. In the discussion preceding this theorem we further established that $\sigma(X_t) \in \mathcal{H}^{1-t}$, and of course $\sigma(X_t) \in \mathcal{F}_t$. Therefore by Theorem 22 we have

$$\tilde{X}^t = X_{1-t} - X_1 = \int_0^t \sigma(X_{1-s})d\tilde{B}^s + [\widetilde{\sigma(X), B}]_t + \int_{1-t}^1 b(X_s)ds.$$

Observe that $[\widetilde{\sigma(X), B}]_t = \int_{1-t}^1 \sigma'(X_s)\sigma(X_s)ds$. Use the change of variable $u = 1 - s$ in the preceding integral and also in the term $\int_{1-t}^1 b(X_s)ds$ to get

$$X_{1-t} = X_1 + \int_0^t \sigma(X_{1-s})d\tilde{B}^s + \int_0^t \sigma'(X_{1-s})\sigma(X_{1-s})ds + \int_0^t b(X_{1-s})ds,$$

and the proof is complete. □

Bibliographic Notes

The theory of expansion of filtrations began with the work of K. Ito [100] for initial expansions, and with the works of M. Barlow [7] and M. Yor [239] for progressive expansions. Our treatment has benefited from some private notes P. A. Meyer shared with the author, as well as the pedagogic treatment found in [44]. A comprehensive treatment, including most of the early important results, is in the book of Th. Jeulin [114]. Excellent later summaries of main results, including many not covered here, can be found in the lecture notes volume edited by Th. Jeulin and M. Yor [117] and also in the little book by M. Yor [247].

Theorem 1 is due to J. Jacod, but we first learned of it through a paper of P. A. Meyer [175], while Theorem 3 (Itô's Theorem for Lévy processes) was established in [107], with the help of T. Kurtz. Theorems 6 and 7 are of course due to Th. Jeulin (see page 44 of [114]); see also M. Yor [245]. Theorem 10, Jacod's criterion, and Theorem 11 are taken from [104]; Jacod's criterion is the best general result on initial expansions of filtrations. The small result on filtration shrinkage is new and is due to J. Jacod and P. Protter.

Two seminal papers on progressive expansions after M. Barlow's original paper are those of Th. Jeulin and M. Yor [115] and [116]. The idea of using a slightly larger filtration than the minimal one when one progressively expands, dates to M. Yor [240], and the idea to use the process $Z_t =^o 1_{\{L>t\}}$ originates with J. Azéma [3]. The lemma on local behavior of stochastic integrals at random times is due to C. Dellacherie and P. A. Meyer [46], and the idea of the proof of Theorem 18 is ascribed to C. Dellacherie in [44]. The results on time reversal are inspired by the work of J. Jacod and P. Protter [107] and also E. Pardoux [190]. Related results can be found in J. Picard [193]. More recently time reversal has been used to extend Itô's formula for Brownian motion to functions which are more general than just being convex, although we did not present these results in this book. The interested reader can consult for example H. Föllmer, P. Protter and A.N. Shiryaev [77] and also H. Föllmer and P. Protter [76] for the multidimensional case. The case for diffusions is treated by X. Bardina and M. Jolis [6].

Exercises for Chapter VI

Exercise 1. Let $(\Omega, \mathcal{F}, \mathbb{F}, B, P)$ be a standard Brownian motion. Expand $\mathbb{F}$ by the initial addition of $\sigma\{B_1\}$. Let M be the local martingale in the formula

$$B_t = B_0 + M_t + \int_0^{t\wedge 1} \frac{B_1 - B_s}{1-s} ds.$$

Show that the Brownian motion $M = (M_t)_{0\le t\le 1}$ is independent of $\sigma\{B_1\}$.

Exercise 2. Show that the processes J^i defined in the proof of Itô's Theorem for Lévy processes (Theorem 3) are compound Poisson processes. Let N^i denote the Poisson process comprised of the arrival times of J^i, and let $\mathbb{G}^i$ be the natural completed filtration of N^i. Further, show the following three formulae (where we suppress the i superscripts) hold.

(a) $E\{J_t | J_1, \mathcal{G}_t\} = 1_{\{N_1 \ge 1\}} \dfrac{N_t}{N_1} J_1$.
(b) $E\{J_t | J_1, \mathcal{G}_1\} = tJ_1$.
(c) $E\{J_t | J_1\} = tJ_1$.

Exercise 3. In Exercise 1 assume $B_0 = x$ and condition on the event $\{B_1 = y\}$. Show that B is then a **Brownian bridge** beginning at x and ending at y. (See Exercise 9 below for more results concerning filtration expansions and Brownian bridges. Brownian bridges are also discussed in Exercise 24 of Chap. II, and on page 299.)

Exercise 4. Let $(\Omega, \mathcal{F}, \mathbb{F}, Z, P)$ be a standard Poisson process. Expand $\mathbb{F}$ by the initial addition on $\sigma\{Z_1\}$. Let M be the local martingale in the formula

$$Z_t = M_t + \int_0^{t\wedge 1} \frac{Z_1 - Z_s}{1-s} ds.$$

Show that $M = (M_t)_{0\le t\le 1}$ is a time changed compensated Poisson process.

Exercise 5. Let B be standard Brownian motion. Show there exists an $\mathcal{F}_1$ measurable random variable L such that if $\mathbb{G}$ is the filtration $\mathbb{F}$ expanded initially with L, then $\mathcal{G}_t = \mathcal{F}_1$ for $0 \le t \le 1$. Deduce that B is not a semimartingale for the expanded filtration $\mathbb{G}$. Show further that L can be taken to be of the form $L = \int_0^1 f(s) dB_s$ where f is non-random.

Exercises 6 through 9 are linked, with the climax being Exercise 9.

Exercise 6. Let B denote standard Brownian motion on its canonical path space on continuous functions with domain $\mathbb{R}_+$ and range $\mathbb{R}$, with $B_t(\omega) = \omega(t)$. $\mathbb{F}$ is the usual minimal filtration completed (and thus rendered right continuous *a fortiori*). Define $Y_t^U = P\{B_a \in U|\mathcal{F}_t\}$. Show that

$$Y_t^U = \int_U Y_t^u du$$

where

$$Y_t^U = g(u, t, B_t) \quad \text{and} \quad g(u, t, x) = \frac{1}{\sqrt{2\pi(a-t)}} \exp\{\frac{-(u-x)^2}{2(a-t)}\}$$

for $t < a$. Show also that $Y_t^U = 1_{\{B_a \in U\}}$ for $t \geq a$.

Exercise 7. (Continuation of Exercise 6.) Show that $Y_{a-}^u = 0$ except on the null set $\{B_a = u\}$ and infer that $(Y_t^u)_{0 \leq t < a}$ is a positive martingale which is not uniformly integrable on $[0, a]$.

Exercise 8. (Continuation of Exercises 6 and 7.) Show that

$$Y_t^u = Y_0^u + \int_0^t \frac{\partial g}{\partial x}(u, s, B_s) dB_s$$

where

$$\frac{\partial g}{\partial x}(u, s, x) = \frac{u-x}{\sqrt{2\pi(a-t)}} \exp\{\frac{-(u-x)^2}{2(a-t)}\}$$

for $t < a$.

Exercise 9. (Continuation of Exercises 6, 7, and 8.) Show that one can define a probability P_u on $\mathcal{C}([0, a], \mathbb{R})$ such that $P_u(B_a = u) = 1$ with P_u absolutely continuous with respect to P on $\mathcal{F}_t$ for $t < a$, and singular with respect to P on $\mathcal{F}_a$. (*Hint: Sketch of procedure.* Let $\Omega_a = \mathcal{C}([0, a), \mathbb{R})$, be the space of continuous functions mapping $[0, a]$ into $\mathbb{R}$, and define $M_t^u = Y_t^u / Y_t^0$ on $\mathcal{F}_t$ for (of course) $t < a$. Let P_u denote the (unique) probability on $\Omega_a = \mathcal{C}([0, a), \mathbb{R})$ given on $\mathcal{F}_t$ by $dP_u = M_t^u dP$, and confirm that

$$\beta_t = B_t - B_0 - \int \frac{1}{M_s^u} d[B, M^u]_s$$

is a Brownian motion under P_u. Next show that

$$A_t^u = \int \frac{1}{M_s^u} d[B, M^u]_s \quad \text{verifies} \quad dA_t^u = \frac{\frac{\partial g}{\partial x}(u, s, B_t)}{g(u, t, B_t)} = \frac{u - B_t}{a - t}.$$

Finally show that

$$B_t - \frac{a-t}{a}B_0 - \frac{t}{a}u = (a-t)\int_0^t \frac{1}{a-s}d\beta_s$$

under P_u, and thus $(a-t)\int_0^t (a-s)^{-1}d\beta_s$, which is defined only on $[0,a)$, has the same distribution under P_u as the Brownian bridge, whence

$$\lim_{t\to a}(a-t)\int_0^t \frac{1}{a-s}d\beta_s = 0$$

a.s., and deduce the desired result.)

Exercise 10 (expansion by a natural exponential random variable). Let $\mathbb{F}$ be a filtration satisfying the usual hypotheses, and let T be a totally inaccessible stopping time, with $P(T<\infty)=1$. Let $A=(A_t)_{t\geq 0}$ be the compensator of $1_{\{t\geq T\}}$ and let M be the martingale $M_t = 1_{\{t\geq T\}} - A_t$. Note that $A_T = A_\infty$. Let $\mathbb{G}$ be the filtration obtained by initial expansion of $\mathbb{F}$ with $\sigma\{A_\infty\}$. For a bounded, non-random, Borel measurable function f let $M_t^f = \int_0^t f(A_s)dM_s$ which becomes in $\mathbb{G}$,

$$M_t^f = f(A_T)1_{\{t\geq T\}} - \int_0^t f(A_s)dA_s = f(A_T)1_{\{t\geq T\}} - F(A_t),$$

where $F(t) = \int_0^t f(s)ds$, since A is continuous. Show that $E\{f(A_t)\} = E\{F(A_t)\}$, and since f is arbitrary deduce that the distribution of the random variable A_T is exponential of parameter 1.

Exercise 11. Let $\mathbb{F}$, T, A, and $\mathbb{G}$ be as in Exercise 10. Show that if X is an $\mathbb{F}$ martingale with $X_{T-} = X_T$ a.s., then X is also a $\mathbb{G}$ martingale.

Exercise 12. Show that every stopping time for a filtration satisfying the usual hypotheses is the end of a predictable set, and is thus an honest time.

Exercise 13. Let $B=(B_t)_{t\geq 0}$ be a standard three dimensional Brownian motion. Let $L=\sup\{t : \|B_t\|\leq 1\}$, the last exit time from the unit ball. Show that L is an honest time.

Exercise 14. Let M^L be the fundamental L martingale of a progressive expansion of the underlying filtration $\mathbb{F}$ using the non-negative random variable L.

(a) Show that for any $\mathbb{F}$ square integrable martingale X we have $EX_L = EX_\infty M_\infty^L$.
(b) Show that M^L is the only $\mathbb{F}^L$ square integrable martingale with this property for all $\mathbb{F}$ square integrable martingales.

Exercise 15. Give an example of a filtration $\mathbb{F}$ and a non-negative random variable L and an $\mathbb{F}$ semimartingale X such that the post L process $Y_t = X_t - X_{t\wedge L} = X_{t\vee L} - X_t$ is not a semimartingale. (*Hint:* Use the standard Brownian filtrations and let

$$L = \sup\{t < 1 : \int_0^t f(s)dB_s = \frac{1}{2}\int_0^1 f(s)dB_s\}$$

where f is the non-random function of Exercise 5.)

Exercise 16. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space satisfying the usual hypotheses, let L be a non-negative random variable, let $\mathcal{G}_\infty = \mathcal{F}_\infty \vee \sigma\{L\}$, and let $\mathbb{G}$ be given by

$$\mathcal{G}_t = \{\Lambda \in \mathcal{G}_\infty | \exists \Lambda_t \in \mathcal{F}_t, \Lambda \cap \{t < L\} = \Lambda_t \cap \{t < L\}\}.$$

Let X be a $\mathbb{F}$ semimartingale. Give a simple proof (due to M. Yor [239]) that $X_t 1_{\{t<L\}}$ (and hence $X_{t\wedge L}$) is a semimartingale. (*Hint:* Show that it is enough to consider X a uniformly integrable martingale, hence enough to consider X a quasimartingale, and finally enough to consider X a positive supermartingale. Let $Z_t = {}^o 1_{\{L>t\}}$, the $\mathbb{F}$ optional projection of $1_{\{L>t\}}$, which is a supermartingale. Recall from Theorem 13 that $P(Z_{L-} > 0) = 1$. Show that

$$X_t^Z = \frac{X_t}{Z_t} 1_{\{t<L\}}$$

is a $\mathbb{G}$ positive supermartingale. Next show that XZ is an $\mathbb{F}$ special semimartingale, and let $XZ = M + A$ be its canonical decomposition. Observe that

$$X_t 1_{\{t<L\}} = \frac{X_t Z_t}{Z_t} 1_{\{t<L\}} = \frac{M_t + A_t}{Z_t} 1_{\{t<L\}} = M_t^Z + A_t^Z,$$

and argue that M^Z is a $\mathbb{G}$ semimartingale as before, and that A^Z is the product of the two $\mathbb{G}$ semimartingales A and $(Z_t)^{-1} 1_{\{t<L\}}$, and hence it is also a $\mathbb{G}$ semimartingale.)

Exercise 17. Let Z be a Lévy process on $[0, 1]$ and let

$$X_t = X_0 + \int_0^t \sigma(X_{s-})dZ_s + \int_0^t b(X_{s-})ds.$$

Assume that σ and b are smooth, and also assume (which is not always true!) that the flows of X are injective. Find a stochastic differential equation solved by X_{1-t}.

Exercise 18. Let $(\Omega, \mathcal{F}, \mathbb{F}, B, P)$ be a standard Brownian space with B, a standard Brownian motion. Let $\mathbb{Z}$ denote the random zero set of B, which is a closed, perfect, nowhere dense set. Let $\{(L_n, R_n)\}_{n\geq 1}$ denote the random intervals contiguous to $\mathbb{Z}$. Without loss of generality assume that the graphs of L_n, denoted $[L_n]$, are disjoint random sets. P. A. Meyer's conjecture was that one could expand $\mathbb{F}$ to a filtration $\mathbb{G}$ in such a way that B would still be a semimartingale and each L_n would be a $\mathbb{G}$ stopping time. Show this is false (argument due to M. Barlow [8]) by taking $Y_t = |B_t|$ and showing (computed in the $\mathbb{G}$ filtration) that $\int_0^t 1_{\{Y_s>0\}} dY_s = Y_t$, which implies that $\int_0^t 1_{\{Y_s=0\}} dY_s = 0$, which is a contradiction, since as a consequence of Tanaka's formula $\int_0^t 1_{\{Y_s=0\}} dY_s = L_t^0$, the Brownian local time at 0.

References

1. H. Ahn and P. Protter. A remark on stochastic integration. In *Séminaire de Probabilités XXVIII*, volume 1583 of *Lecture Notes in Mathematics*, pages 312–315. Springer-Verlag, Berlin, 1994.
2. L. Arnold. *Stochastic Differential Equations: Theory and Applications.* Wiley, New York, 1973.
3. J. Azéma. Quelques applications de la théorie générale des processus. *Invent. Math.*, 18:293–336, 1972.
4. J. Azéma. Sur les fermés aléatoires. In *Séminaire de Probabilités XIX*, volume 1123 of *Lecture Notes in Mathematics*, pages 397–495. Springer-Verlag, Berlin, 1985.
5. J. Azéma and M. Yor. Etude d'une martingale remarquable. In *Séminaire de Probabilités XXIII*, volume 1372 of *Lecture Notes in Mathematics*, pages 88–130. Springer-Verlag, Berlin, 1989.
6. X. Bardina and M. Jolis. An extension of Itô's formula for elliptic diffusion processes. *Stochastic Processes and their Appl.*, 69:83–109, 1997.
7. M. T. Barlow. Study of a filtration expanded to include an honest time. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 44:307–323, 1978.
8. M. T. Barlow. On the left end points of Brownian excursions. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, page 646. Springer-Verlag, Berlin, 1979.
9. M. T. Barlow, S. D. Jacka, and M. Yor. Inequalities for a pair of processes stopped at a random time. *Proc. London Math. Soc.*, 52:142–172, 1986.
10. R. F. Bass. *Probabilistic Techniques in Analysis.* Springer-Verlag, New York, 1995.
11. R. F. Bass. The Doob-Meyer decomposition revisited. *Canad. Math. Bull.*, 39:138–150, 1996.
12. J. Bertoin. *Lévy Processes.* Cambridge University Press, Cambridge, 1996.
13. K. Bichteler. Stochastic integrators. *Bull. Amer. Math. Soc.*, 1:761–765, 1979.
14. K. Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, 9:49–89, 1981.
15. K. Bichteler. *Stochastic Integration With Jumps.* Cambridge University Press, Cambridge, 2002.
16. K. Bichteler, J.-B. Gravereaux, and J. Jacod. *Malliavin Calculus for Processes With Jumps.* Gordon and Breach, 1987.

17. P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.
18. Yu. N. Blagovescenskii and M. I. Freidlin. Certain properties of diffusion processes depending on a parameter. *Sov. Math.*, 2:633–636, 1961. Originally published in Russian in *Dokl. Akad. Nauk SSSR*, 138:508–511, 1961.
19. R. M. Blumenthal and R. K. Getoor. *Markov Processes and Potential Theory.* Academic Press, New York, 1968.
20. T. Bojdecki. *Teoria General de Procesos e Integracion Estocastica.* Universidad Nacional Autonoma de Mexico, 1988. Segundo edicion.
21. N. Bouleau. Formules de changement de variables. *Ann. Inst. H. Poincaré Probab. Statist.*, 20:133–145, 1984.
22. N. Bouleau and M. Yor. Sur la variation quadratique des temps locaux de certaines semimartingales. *C. R. Acad. Sci. Paris Sér. I Math.*, 292:491–494, 1981.
23. L. Breiman. *Probability.* SIAM, Philadelphia, 1992.
24. P. Brémaud. *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York, 1981.
25. J. L. Bretagnolle. Processus a accroissements indépendants. In *Ecole d'Eté de Probabilités: Processus Stochastiques.*, volume 307 of *Lecture Notes in Mathematics*, pages 1–26. Springer-Verlag, Berlin, 1973.
26. B. Bru and M. Yor. La vie de W. Doeblin et le Pli cacheté 11 668. *La Lettre de L'Académie des sciences*, 2:16–17, 2001.
27. B. Bru and M. Yor. Comments on the life and mathematical legacy of Wolfgang Doeblin. *Finance Stoch.*, 6:3–47, 2002.
28. Mireille Chaleyat-Maurel and T. Jeulin. Caracterisation d'une classe de semimartingales. In *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*, pages 59–109. Springer-Verlag, Berlin, 1979.
29. C. S. Chou. Caracterisation d'une classe de semimartingales. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 250–252. Springer-Verlag, Berlin, 1979.
30. C. S. Chou and P. A. Meyer. Sur la représentation des martingales comme intégrales stochastiques dans les processus ponctuels. In *Séminaire de Probabilités IX*, volume 465 of *Lecture Notes in Mathematics*, pages 226–236. Springer-Verlag, Berlin, 1975.
31. C. S. Chou, P. A. Meyer, and C. Stricker. Sur les intégrales stochastiques de processus prévisibles non bornés. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 128–139. Springer-Verlag, Berlin, 1980.
32. K. L. Chung and R. J. Williams. *Introduction to Stochastic Integration.* Birkhäuser, Boston, second edition, 1990.
33. E. Çinlar. *Introduction to Stochastic Processes.* Prentice Hall, Englewood Cliffs, 1975.
34. E. Çinlar, J. Jacod, P. Protter, and M. Sharpe. Semimartingales and Markov processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 54:161–220, 1980.
35. E. A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill, New York, 1955.
36. P. Courrège. Intégrale stochastique par rapport à une martingale de carré intégrable. In *Séminaire de Théorie du Potentiel dirigé par M. Brelot, G. Choquet et J. Dény, 7é année (1962/1963)*, Paris, 1963. Faculté des Sciences de Paris.

37. K. E. Dambis. On the decomposition of continuous submartingales. *Theory Probab. Appl.*, 10:401–410, 1965.
38. F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.*, 312:215–250, 1998.
39. F. Delbaen and M. Yor. Passport options. Preprint, ETH, 2001.
40. C. Dellacherie. Un exemple de la théorie générale des processus. In *Séminaire de Probabilités IV*, volume 124 of *Lecture Notes in Mathematics*, pages 60–70. Springer-Verlag, Berlin, 1970.
41. C. Dellacherie. *Capacités et processus stochastiques.* Springer-Verlag, Berlin, 1972.
42. C. Dellacherie. Un survol de la théorie de l'intégrale stochastique. *Stochastic Process. Appl.*, 10:115–144, 1980.
43. C. Dellacherie. Mesurabilité des débuts et théorème de section. In *Séminaire de Probabilités XV*, volume 850 of *Lecture Notes in Mathematics*, pages 351–360. Springer-Verlag, Berlin, 1981.
44. C. Dellacherie, B. Maisonneuve, and P. A. Meyer. *Probabilités et Potentiel. Chapitres XVII à XXIV; Processus de Markov (fin); Compléments de Calcul Stochastique.* Hermann, Paris, 1992.
45. C. Dellacherie and P. A. Meyer. *Probabilities and Potential*, volume 29 of *North-Holland Mathematics Studies.* North-Holland, Amsterdam, 1978.
46. C. Dellacherie and P. A. Meyer. *Probabilities and Potential B*, volume 72 of *North-Holland Mathematics Studies.* North-Holland, Amsterdam, 1982.
47. C. Dellacherie, P. A. Meyer, and M. Yor. Sur certaines propriétés des espaces de Banach $\mathcal{H}^1$ et BMO. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 98–113. Springer-Verlag, Berlin, 1978.
48. C. Doléans-Dade. Intégrales stochastiques dépendant d'un paramètre. *Publ. Inst. Stat. Univ. Paris*, 16:23–34, 1967.
49. C. Doléans-Dade. Quelques applications de la formule de changement de variables pour les semimartingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 16:181–194, 1970.
50. C. Doléans-Dade. Une martingale uniformément intégrable mais non localement de carré intégrable. In *Séminaire de Probabilités V*, volume 191 of *Lecture Notes in Mathematics*, pages 138–140. Springer-Verlag, Berlin, 1971.
51. C. Doléans-Dade. On the existence and unicity of solutions of stochastic differential equations. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 36:93–101, 1976.
52. C. Doléans-Dade. Stochastic processes and stochastic differential equations. In *Stochastic Differential Equations: I ciclo 1978, Palazzone della Scuola Normale Superiore, Cortona (Centro Internazionale Matematico Estivo)*, pages 5–75. Liguori Editore, Napoli, 1981.
53. C. Doléans-Dade and P. A. Meyer. Intégrales stochastiques par rapport aux martingales locales. In *Séminaire de Probabilités IV*, volume 124 of *Lecture Notes in Mathematics*, pages 77–107. Springer-Verlag, 1970.
54. C. Doléans-Dade and P. A. Meyer. Equations différentielles stochastiques. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 376–382. Springer-Verlag, Berlin, 1977.
55. J. L. Doob. *Stochastic Processes.* Wiley, 1953.
56. J. L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart.* Springer-Verlag, Berlin, 1984.

57. H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Probab. Statist.*, 13:99–125, 1977.
58. H. Doss and E. Lenglart. Sur l'existence, l'unicité et le comportement asymptotique des solutions d'équations différentielles stochastiques. *Ann. Inst. H. Poincaré Probab. Statist.*, 14:189–214, 1978.
59. L. Dubins and G. Schwarz. On continuous martingales. *Proc. Natl. Acad. Sci. USA*, 53:913–916, 1965.
60. D. Duffie and C. Huang. Multiperiod security markets with differential information. *J. Math. Econom.*, 15:283–303, 1986.
61. D. A. Edwards. A note on stochastic integrators. *Math. Proc. Cambridge Philos. Soc.*, 107:395–400, 1990.
62. N. Eisenbaum. Integration with respect to local time. *Potential Anal.*, 13:303–328, 2000.
63. R. J. Elliott. *Stochastic Calculus and Applications.* Springer-Verlag, Berlin, 1982.
64. M. Emery. Stabilité des solutions des équations différentielles stochastiques; applications aux intégrales multiplicatives stochastiques. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 41:241–262, 1978.
65. M. Emery. Equations différentielles stochastiques lipschitziennes: étude de la stabilité. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 281–293. Springer-Verlag, Berlin, 1979.
66. M. Emery. Une topologie sur l'espace des semimartingales. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 260–280. Springer-Verlag, Berlin, 1979.
67. M. Emery. Compensation de processus à variation finie non localement intégrables. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 152–160. Springer-Verlag, Berlin, 1980.
68. M. Emery. Non confluence des solutions d'une équation stochastique lipschitzienne. In *Séminaire de Probabilités XV*, volume 850 of *Lecture Notes in Mathematics*, pages 587–589. Springer-Verlag, Berlin, 1981.
69. M. Emery. On the Azéma martingales. In *Séminaire de Probabilités XXIII*, volume 1372 of *Lecture Notes in Mathematics*, pages 66–87. Springer-Verlag, Berlin, 1989.
70. M. Emery and M. Yor, editors. *Séminaire de Probabilités 1967–1980: A Selection in Martingale Theory*, volume 1771 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.
71. S. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, New York, 1986.
72. W. Feller. *An Introduction to Probability Theory and Its Applications. Volume II.* Wiley, New York, second edition, 1971.
73. D. L. Fisk. Quasimartingales. *Trans. Amer. Math. Soc.*, 120:369–389, 1965.
74. T. R. Fleming and D. P. Harrington. *Counting Processes and Survival Analysis.* John Wiley, New York, 1991.
75. H. Föllmer. Calcul d'Itô sans probabilités. In *Séminaire de Probabilités XV*, volume 850 of *Lecture Notes in Mathematics*, pages 144–150. Springer-Verlag, Berlin, 1981.
76. H. Föllmer and P. Protter. An extension of Itô's formula in n dimensions. *Probability Theory and Related Fields*, 116:1–20, 2000.
77. H. Föllmer, P. Protter, and A. N. Shiryaev. Quadratic variation and an extension of Itô's formula. *Bernoulli*, 1:149–169, 1995.

78. T. Fujiwara and H. Kunita. Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. *J. Math. Kyoto Univ.*, 25:71–106, 1985.
79. R. K. Getoor and M. Sharpe. Conformal martingales. *Invent. Math.*, 16:271–308, 1972.
80. R. Ghomrasni and G. Peskir. Local time-space calculus and extensions of Itô's formula. Preprint, Aarhus, 2002.
81. I. I. Gihman and A. V. Skorokhod. *Stochastic Differential Equations.* Springer-Verlag, Berlin, 1972.
82. R. D. Gill and S. Johansen. A survey of product-integration with a view towards applications in survival analysis. *Ann. Statist.*, 18:1501–1555, 1990.
83. I. V. Girsanov. On transforming a certain class of stochastic processes by absolutely continuous substitutions of measures. *Theory Probab. Appl.*, 5:285–301, 1960.
84. M. Greenberg. *Lectures on Algebraic Topology.* Benjamin, New York, 1967.
85. T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential eqautions. *Ann. of Math. (2)*, 20:292–296, 1919.
86. A. Gut. An introduction to the theory of asymptotic martingales. In *Amarts and Set Function Processes*, volume 1042 of *Lecture Notes in Mathematics*, pages 1–49. Springer-Verlag, Berlin, 1983.
87. S. W. He, J. G. Wang, and J. A. Yan. *Semimartingale Theory and Stochastic Calculus.* CRC Press, Boca Raton, 1992.
88. T. Hida. *Brownian Motion.* Springer-Verlag, Berlin, 1980.
89. G. A. Hunt. Markoff processes and potentials I. *Illinois J. Math.*, 1:44–93, 1957.
90. G. A. Hunt. Markoff processes and potentials II. *Illinois J. Math.*, 1:316–369, 1957.
91. G. A. Hunt. Markoff processes and potentials III. *Illinois J. Math.*, 2:151–213, 1958.
92. N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam, 1981.
93. K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519–524, 1944.
94. K. Itô. On stochastic integral equations. *Proc. Japan Acad.*, 22:32–35, 1946.
95. K. Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.*, 1:35–47, 1950.
96. K. Itô. Multiple Wiener integral. *J. Math. Soc. Japan*, 3:157–169, 1951.
97. K. Itô. On a formula concerning stochastic differentials. *Nagoya Math. J.*, 3:55–65, 1951.
98. K. Itô. On Stochastic Differential Equations. *Memoirs of the American Mathematical Society*, 4, 1951.
99. K. Itô. Stochastic differentials. *Appl. Math. Optim.*, 1:374–381, 1974.
100. K. Itô. Extension of stochastic integrals. In *Proceedings of International Symposium on Stochastic Differential Equations*, pages 95–109, New York, 1978. Wiley.
101. K. Itô and M. Nisio. On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.*, 4:1–75, 1964.
102. K. Itô and S. Watanabe. Transformation of Markov processes by multiplicative functionals. *Ann. Inst. Fourier*, 15:15–30, 1965.

103. J. Jacod. *Calcul Stochastique et Problèmes de Martingales*, volume 714 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1979.
104. J. Jacod. Grossissement initial, hypothèse (H'), et théorème de Girsanov. In *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*, pages 15–35. Springer-Verlag, Berlin, 1985.
105. J. Jacod, T. G. Kurtz, S. Méléard, and P. Protter. The approximate Euler method for Lévy driven stochastic differential equations. Preprint, 2003.
106. J. Jacod, S. Méléard, and P. Protter. Explicit form and robustness of martingale representations. *Ann. Probab.*, 28:1747–1780, 2000.
107. J. Jacod and P. Protter. Time reversal on Lévy processes. *Ann. Probab.*, 16:620–641, 1988.
108. J. Jacod and P. Protter. Une remarque sur les équations différentielles stochastiques à solutions Markoviennes. In *Séminaire de Probabilités XXV*, volume 1485 of *Lecture Notes in Mathematics*, pages 138–139. Springer-Verlag, Berlin, 1991.
109. J. Jacod and P. Protter. *Probability Essentials*. Springer-Verlag, Heidelberg, second edition, 2003.
110. J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York, second edition, 2003.
111. J. Jacod and M. Yor. Etude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 38:83–125, 1977.
112. Y. Le Jan. Temps d'arrêt stricts et martingales de sauts. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 44:213–225, 1978.
113. S. Janson and M. J. Wichura. Invariance principles for stochastic area and related stochastic integrals. *Stochastic Process. Appl.*, 16:71–84, 1983.
114. T. Jeulin. *Semi-Martingales et Grossissement d'une Filtration*, volume 833 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1980.
115. T. Jeulin and M. Yor. Grossissement d'une filtration et semi-martingales: formules explicites. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 78–97. Springer-Verlag, Berlin, 1978.
116. T. Jeulin and M. Yor. Inégalité de hardy, semimartingales, et faux amis. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 332–359. Springer-Verlag, Berlin, 1979.
117. T. Jeulin and M. Yor. Recapitulatif. In *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*, pages 305–315. Springer-Verlag, Berlin, 1985.
118. G. Johnson and L. L. Helms. Class D supermartingales. *Bull. Amer. Math. Soc.*, 69:59–62, 1963.
119. T. Kailath, A. Segall, and M. Zakai. Fubini-type theorems for stochastic integrals. *Sankhyā Ser. A*, 40:138–143, 1978.
120. G. Kallianpur and C. Striebel. Stochastic differential equations occuring in the estimation of continuous parameter stochastic processes. *Theory Probab. Appl.*, 14:567–594, 1969.
121. I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, second edition, 1991.
122. A. F. Karr. *Point Processes and Their Statistical Inference*. Marcel Dekker, New York, second edition, 1991.

123. N. Kazamaki. Note on a stochastic integral equation. In *Séminaire de Probabilités VI*, volume 258 of *Lecture Notes in Mathematics*, pages 105–108. Springer-Verlag, Berlin, 1972.
124. N. Kazamaki. On a problem of Girsanov. *Tôhoku Math. J. (2)*, 29:597–600, 1977.
125. N. Kazamaki. *Continuous Exponential Martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994.
126. J. F. C. Kingman. An intrinsic description of local time. *J. London Math. Soc.*, 6:725–731, 1973.
127. J. F. C. Kingman and S. J. Taylor. *Introduction to Measure and Probability*. Cambridge University Press, Cambridge, 1966.
128. P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin, 1992.
129. F. B. Knight. Calculating the compensator: method and example. In *Seminar on Stochastic Processes, 1990*, pages 241–252, Boston, 1991. Birkhäuser.
130. P. E. Kopp. *Martingales and Stochastic Integrals*. Cambridge University Press, Cambridge, 1984.
131. H. Kunita. On the decomposition of solutions of stochastic differential equations. In *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 213–255. Springer-Verlag, Berlin, 1981.
132. H. Kunita. Stochastic differential equations and stochastic flow of diffeomorphisms. In *Ecole d'Eté de Probabilités de Saint-Flour XII-1982*, volume 1097 of *Lecture Notes in Mathematics*, pages 143–303. Springer-Verlag, Berlin, 1984.
133. H. Kunita. *Lectures on stochastic flows and applications*, volume 78 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Springer-Verlag, Berlin, 1986.
134. H. Kunita and S. Watanabe. On square integrable martingales. *Nagoya Math. J.*, 30:209–245, 1967.
135. T. G. Kurtz. Lectures on Stochastic Analysis. Preprint, 2001.
136. T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19:1035–1070, 1991.
137. T. G. Kurtz and P. Protter. Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. In *Stochastic Analysis: Liber Amicorum for Moshe Zakai*, pages 331–346. Academic Press, 1991.
138. A. U. Kussmaul. *Stochastic Integration and Generalized Martingales*. Pitman, London, 1977.
139. G. Last and A. Brandt. *Marked Point Processes on the Real Line*. Springer-Verlag, New York, 1995.
140. R. Leandre. Un exemple en théorie des flots stochastiques. In *Séminaire de Probabilités XVII*, volume 986 of *Lecture Notes in Mathematics*, pages 158–161. Springer-Verlag, Berlin, 1983.
141. R. Leandre. Flot d'une équation différentielle stochastique avec semimartingale directrice discontinue. In *Séminaire de Probabilités XIX*, volume 1123 of *Lecture Notes in Mathematics*, pages 271–274. Springer-Verlag, Berlin, 1985.
142. E. Lenglart. Transformation des martingales locales par changement absolument continu de probabilités. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 39:65–70, 1977.
143. E. Lenglart. Une caractérisation des processus prévisibles. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 415–417. Springer-Verlag, Berlin, 1977.

144. E. Lenglart. Appendice à l'exposé précédent: inégalités de semimartingales. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 49–52. Springer-Verlag, Berlin, 1980.
145. E. Lenglart. Semi-martingales et intégrales stochastiques en temps continu. *Rev. CETHEDEC*, 75:91–160, 1983.
146. E. Lenglart, D. Lépingle, and M. Prattle. Présentation unifiée de certaines inégalités de la théorie des martingales. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 26–48. Springer-Verlag, Berlin, 1980.
147. D. Lépingle and J. Mémin. Sur l'intégrabilité uniforme des martingales exponentielles. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 42:175–203, 1978.
148. G. Letta. *Martingales et intégration stochastique.* Scuola Normale Superiore, Pisa, 1984.
149. P. Lévy. *Processus stochastiques et mouvement Brownien.* Gauthier-Villars, Paris, 1948.
150. P. Lévy. Wiener's random function, and other Laplacian random functions. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 171–187, Berkeley, 1951. University of California Press.
151. B. Maisonneuve. Une mise au point sur les martingales continues définies sur un intervalle stochastique. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 435–445. Springer-Verlag, Berlin, 1977.
152. G. Maruyama. On the transition probability functions of the Markov process. *Nat. Sci. Rep. Ochanomizu Univ.*, 5:10–20, 1954.
153. H. P. McKean. *Stochastic Integrals.* Academic Press, New York, 1969.
154. E. J. McShane. *Stochastic Calculus and Stochastic Models.* Academic Press, New York, 1974.
155. E. J. McShane. Stochastic differential equations. *J. Multivariate Anal.*, 5:121–177, 1975.
156. J. Mémin. Décompositions multiplicatives de semimartingales exponentielles et applications. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 35–46. Springer-Verlag, Berlin, 1978.
157. J. Mémin and A. N. Shiryaev. Un critère prévisible pour l'uniforme intégrabilité des semimartingales exponentielles. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 147–161. Springer-Verlag, Berlin, 1979.
158. M. Métivier. *Semimartingales: A Course on Stochastic Processes.* Walter de Gruyter, Berlin and New York, 1982.
159. M. Métivier and J. Pellaumail. On Doléans-Föllmer's measure for quasimartingales. *Illinois J. Math.*, 19:491–504, 1975.
160. M. Métivier and J. Pellaumail. Mesures stochastiques à valeurs dans les espaces L_0. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 40:101–114, 1977.
161. M. Métivier and J. Pellaumail. On a stopped Doob's inequality and general stochastic equations. *Ann. Probab.*, 8:96–114, 1980.
162. M. Métivier and J. Pellaumail. *Stochastic Integration.* Academic Press, New York, 1980.
163. P. A. Meyer. A decomposition theorem for supermartingales. *Illinois J. Math.*, 6:193–205, 1962.
164. P. A. Meyer. Decomposition of supermartingales: the uniqueness theorem. *Illinois J. Math.*, 7:1–17, 1963.

165. P. A. Meyer. *Probability and Potentials.* Blaisdell, Waltham, 1966.
166. P. A. Meyer. Intégrales stochastiques I. In *Séminaire de Probabilités I*, volume 39 of *Lecture Notes in Mathematics*, pages 72–94. Springer-Verlag, Berlin, 1967.
167. P. A. Meyer. Intégrales stochastiques II. In *Séminaire de Probabilités I*, volume 39 of *Lecture Notes in Mathematics*, pages 95–117. Springer-Verlag, Berlin, 1967.
168. P. A. Meyer. Intégrales stochastiques III. In *Séminaire de Probabilités I*, volume 39 of *Lecture Notes in Mathematics*, pages 118–141. Springer-Verlag, Berlin, 1967.
169. P. A. Meyer. Intégrales stochastiques IV. In *Séminaire de Probabilités I*, volume 39 of *Lecture Notes in Mathematics*, pages 142–162. Springer-Verlag, Berlin, 1967.
170. P. A. Meyer. Le dual de "H^1" est "BMO" (cas continu). In *Séminaire de Probabilités VII*, volume 321 of *Lecture Notes in Mathematics*, pages 136–145. Springer-Verlag, Berlin, 1973.
171. P. A. Meyer. Un cours sur les intégrales stochastiques. In *Séminaire de Probabilités X*, volume 511 of *Lecture Notes in Mathematics*, pages 246–400. Springer-Verlag, Berlin, 1976.
172. P. A. Meyer. Le théorème fondamental sur les martingales locales. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 463–464. Springer-Verlag, Berlin, 1977.
173. P. A. Meyer. Sur un théorème de C. Stricker. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 482–489. Springer-Verlag, Berlin, 1977.
174. P. A. Meyer. Inégalités de normes pour les intégrales stochastiques. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 757–762. Springer-Verlag, Berlin, 1978.
175. P. A. Meyer. Sur un théorème de J. Jacod. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 57–60. Springer-Verlag, Berlin, 1978.
176. P. A. Meyer. Caractérisation des semimartingales, d'après Dellacherie. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 620–623. Springer-Verlag, Berlin, 1979.
177. P. A. Meyer. Flot d'une équation différentielle stochastique. In *Séminaire de Probabilités XV*, volume 850 of *Lecture Notes in Mathematics*, pages 103–117. Springer-Verlag, Berlin, 1981.
178. P. A. Meyer. Géométrie différentielle stochastique. In *Colloque en l'honneur de Laurent Schwartz (Ecole Polytechnique, 30 mai-3 juin 1983). Vol. 1*, volume 131 of *Astérisque*, pages 107–114. Société Mathématique de France, Paris, 1985.
179. P. A. Meyer. Construction de solutions d'équation de structure. In *Séminaire de Probabilités XXIII*, volume 1372 of *Lecture Notes in Mathematics*, pages 142–145. Springer-Verlag, Berlin, 1989.
180. T. Mikosch. *Elementary Stochastic Calculus with Finance in View.* World Scientific, Singapore, 1998.
181. P. W. Millar. Stochastic integrals and processes with stationary independent increments. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Vol. III: Probability theory*, pages 307–331, Berkeley, 1972. University of California Press.

182. D. S. Mitrinović, J. E. Pečarić, and A. M. Fink. *Inequalities Involving Functions and Their Integrals and Derivatives.* Kluwer, Dordrecht, 1991.
183. J. R. Munkries. *Topology: A First Course.* Prentice-Hall, Englewood Cliffs, 1975.
184. A. A. Novikov. On an identity for stochastic integrals. *Theory Probab. Appl.*, 17:717–720, 1972.
185. C. O'Cinneide and P. Protter. An elementary approach to naturality, predictability, and the fundamental theorem of local martingales. *Stoch. Models*, 17:449–458, 2001.
186. B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer-Verlag, New York, sixth edition, 2003.
187. S. Orey. *F*-processes. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol. II: Contributions to probability theory. Part 1*, pages 301–313, Berkeley, 1965. University of California Press.
188. Y. Ouknine. Généralisation d'un lemme de S. Nakao et applications. *Stochastics*, 23:149–157, 1988.
189. Y. Ouknine. Temps local du produit et du sup de deux semimartingales. In *Séminaire de Probabilités XXIV*, volume 1426 of *Lecture Notes in Mathematics*, pages 477–479. Springer-Verlag, Berlin, 1990.
190. E. Pardoux. Grossissement d'une filtration et retournement du temps d'une diffusion. In *Séminaire de Probabilités XX*, volume 1204 of *Lecture Notes in Mathematics*, pages 48–55. Springer-Verlag, Berlin, 1986.
191. E. Pardoux. Book Review of *Stochastic Integration and Differential Equations*, first edition, by P. Protter. *Stoch. Stoch. Rep.*, 42:237–240, 1993.
192. J. Pellaumail. *Sur l'intégrale stochastique et la décomposition de Doob-Meyer*, volume 9 of *Astérisque*. Société Mathématique de France, Paris, 1973.
193. J. Picard. Une classe de processus stable par retournement du temps. In *Séminaire de Probabilités XX*, volume 1204 of *Lecture Notes in Mathematics*, pages 56–67. Springer-Verlag, Berlin, 1986.
194. Z. R. Pop-Stojanovic. On McShane's belated stochastic integral. *SIAM J. Appl. Math.*, 22:87–92, 1972.
195. M. H. Protter and C. B. Morrey. *A First Course in Real Analysis.* Springer-Verlag, Berlin, 1977.
196. P. Protter. Markov solutions of stochastic differential equations. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 41:39–58, 1977.
197. P. Protter. On the existence, uniqueness, convergence, and explosions of solutions of systems of stochastic integral equations. *Ann. Probab.*, 5:243–261, 1977.
198. P. Protter. Right-continuous solutions of systems of stochastic integral equations. *J. Multivariate Anal.*, 7:204–214, 1977.
199. P. Protter. H^p stability of solutions of stochastic differential equations. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 44:337–352, 1978.
200. P. Protter. A comparison of stochastic integrals. *Ann. Probab.*, 7:176–189, 1979.
201. P. Protter. Semimartingales and stochastic differential equations. Technical Report 85-25, Purdue University Department of Statistics, 1985.
202. P. Protter. Stochastic integration without tears. *Stochastics*, 16:295–325, 1986.
203. P. Protter. A partial introduction to financial asset pricing theory. *Stochastic Process. Appl.*, 91:169–203, 2001.

204. P. Protter and M. Sharpe. Martingales with given absolute value. *Ann. Probab.*, 7:1056–1058, 1979.
205. P. Protter and D. Talay. The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.*, 25:393–423, 1997.
206. K. M. Rao. On decomposition theorems of Meyer. *Math. Scand.*, 24:66–78, 1969.
207. K. M. Rao. Quasimartingales. *Math. Scand.*, 24:79–92, 1969.
208. D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer-Verlag, New York, third edition, 1999.
209. L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Volume 1: Foundations.* Cambridge University Press, Cambridge, second edition, 2000.
210. L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Volume 2: Itô Calculus.* Cambridge University Press, Cambridge, second edition, 2000.
211. H. Rootzen. Limit distributions for the error in approximations of stochastic integrals. *Ann. Probab.*, 8:241–251, 1980.
212. G. Samorodnitsky and M. S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance.* Chapman and Hall, New York, 1994.
213. L. Schwartz. *Semimartingales and their stochastic calculus on manifolds.* Presses de l'Université de Montréal, Montréal, 1984.
214. L. Schwartz. Convergence de la série de picard pour les eds. In *Séminaire de Probabilités XXIII*, volume 1372 of *Lecture Notes in Mathematics*, pages 343–354. Springer-Verlag, Berlin, 1989.
215. M. Sharpe. *General Theory of Markov Processes.* Academic Press, New York, 1988.
216. J. M. Steele. *Stochastic Calculus and Financial Applications.* Springer-Verlag, New York, 2001.
217. R. L. Stratonovich. A new representation for stochastic integrals. *SIAM J. Control*, 4:362–371, 1966.
218. C. Stricker. Quasimartingales, martingales locales, semimartingales, et filtrations naturelles. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 3:55–64, 1977.
219. C. Stricker and M. Yor. Calcul stochastique dépendant d'un paramètre. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 45:109–134, 1978.
220. D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes.* Springer-Verlag, Berlin, 1979.
221. H. Sussman. On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probab.*, 6:19–41, 1978.
222. A. S. Sznitman. Martingales dépendant d'un paramètre: une formule d'Itô. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 60:41–70, 1982.
223. F. Treves. *Topological Vector Spaces, Distributions and Kernels.* Academic Press, New York, 1967.
224. A. Uppman. Deux applications des semi-martingales exponentielles aux équations différentielles stochastiques. *C. R. Acad. Sci. Paris Sér. A-B*, 290:A661–A664, 1980.
225. A. Uppman. Sur le flot d'une équation différentielle stochastique. In *Séminaire de Probabilités XVI*, volume 920 of *Lecture Notes in Mathematics*, pages 268–284. Springer-Verlag, Berlin, 1982.
226. H. von Weizsäcker and G. Winkler. *Stochastic Integrals.* Vieweg, Braunschweig, 1990.

227. J. Walsh. A diffusion with a discontinuous local time. In *Temps locaux: Exposés du séminaire J. Azéma, M. Yor (1976-1977)*, volume 52-53 of *Astérisque*, pages 37–45. Société Mathématique de France, Paris, 1978.
228. N. Wiener. Differential space. *J. Math. Phys.*, 2:131–174, 1923.
229. N. Wiener. The Dirichlet problem. *J. Math. Phys.*, 3:127–146, 1924.
230. N. Wiener. The quadratic variation of a function and its Fourier coefficients. *J. Math. Phys.*, 3:72–94, 1924.
231. E. Wong. *Stochastic Processes in Information and Dynamical Systems.* McGraw Hill, New York, 1971.
232. T. Yamada. Sur une construction des solutions d'équations différentielles stochastiques dans le cas non-lipschitzien. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 114–131. Springer-Verlag, Berlin, 1978.
233. J. A. Yan. Caractérisation d'une classe d'ensembles convexes de L^1 ou H^1. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 220–222. Springer-Verlag, Berlin, 1980.
234. Ch. Yoeurp. Décomposition des martingales locales et formules exponentielles. In *Séminaire de Probabilités X*, volume 511 of *Lecture Notes in Mathematics*, pages 432–480. Springer-Verlag, Berlin, 1976.
235. Ch. Yoeurp. Solution explicite de l'équation $Z_t = 1+\int_0^t |Z_{s-}| dX_s$. In *Séminaire de Probabilités XIII*, volume 721 of *Lecture Notes in Mathematics*, pages 614–619. Springer-Verlag, Berlin, 1979.
236. Ch. Yoeurp and M. Yor. Espace orthogonal à une semimartingale: applications. Unpublished, 1977.
237. M. Yor. Sur les intégrales stochastiques optionnelles et une suite remarquable de formules exponentielles. In *Séminaire de Probabilités X*, volume 511 of *Lecture Notes in Mathematics*, pages 481–500. Springer-Verlag, Berlin, 1976.
238. M. Yor. Sur quelques approximations d'intégrales stochastiques. In *Séminaire de Probabilités XI*, volume 581 of *Lecture Notes in Mathematics*, pages 518–528. Springer-Verlag, Berlin, 1977.
239. M. Yor. Grossissement d'une filtration et semi-martingales: théorèmes généraux. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 61–69. Springer-Verlag, Berlin, 1978.
240. M. Yor. Grossissement d'une filtration et semi-martingales: théorèmes generaux. In *Séminaire de Probabilités XII*, volume 649 of *Lecture Notes in Mathematics*, pages 61–69. Springer-Verlag, Berlin, 1978.
241. M. Yor. Inégalités entre processus minces et applications. *C. R. Acad. Sci. Paris Sér. A-B*, 286:A799–A801, 1978.
242. M. Yor. Rappels et préliminaires généraux. *Astérisque*, 52-53:17–22, 1978.
243. M. Yor. Sur la continuité des temps locaux associés à certaines semimartingales. In *Temps Locaux*, volume 52 and 53 of *Astérisque*, pages 23–36. Société Mathématique de France, Paris, 1978.
244. M. Yor. Un exemple de processus qui n'est pas une semi-martingale. In *Temps Locaux*, volume 52 and 53 of *Astérisque*, pages 219–222. Société Mathématique de France, Paris, 1978.
245. M. Yor. Application d'un lemme de Jeulin au grossissement de la filtration brownienne. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 189–199. Springer-Verlag, Berlin, 1980.

246. M. Yor. Remarques sur une formule de Paul Lévy. In *Séminaire de Probabilités XIV*, volume 784 of *Lecture Notes in Mathematics*, pages 343–346. Springer-Verlag, Berlin, 1980.
247. M. Yor. *Some Aspects of Brownian Motion. Part II: Some Recent Martingale Problems.* Birkhäuser, Berlin, Germany, 1997.
248. W. Zheng. Une remarque sur une même intégrale stochastique calculée dans deux filtrations. In *Séminaire de Probabilités XVIII*, volume 1059 of *Lecture Notes in Mathematics*, pages 172–178. Springer-Verlag, Berlin, 1984.

Symbol Index

Subject Index