Encyclopedia of Mathematics and its Applications 131

STOCHASTIC CONTROL AND MATHEMATICAL MODELING

Applications in Economics

Hiroaki Morimoto

CAMBRIDGE

CAMBRIDGE

more information – www.cambridge.org/9780521195034

Stochastic Control and Mathematical Modeling

Applications in Economics

This is a concise and elementary introduction to stochastic control and mathematical modeling. This book is designed for researchers in stochastic control theory studying its application in mathematical economics and those in economics who are interested in mathematical theory in control. It is also a good guide for graduate students studying applied mathematics, mathematical economics, and nonlinear PDE theory.

Contents include the basics of analysis and probability, the theory of stochastic differential equations, variational problems, problems in optimal consumption and in optimal stopping, optimal pollution control, and solving the Hamilton–Jacobi–Bellman equations with boundary conditions. Major mathematical requisitions are contained in the preliminary chapters or in the appendix so that readers can proceed without referring to other materials.

Hiroaki Morimoto is a Professor in Mathematics at the Graduate School of Science and Engineering at Ehime University. His research interests include stochastic control, mathematical economics and finance and insurance applications, and the viscosity solution theory.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

The titles below, and earlier volumes in the series, are available from booksellers or from Cambridge University Press at www.cambridge.org.

71 G. E. Andrews, R. Askey and R. Roy *Special Functions*
72 R. Ticciati *Quantum Field Theory for Mathematicians*
73 M. Stern *Semimodular Lattices*
74 I. Lasiecka and R. Triggiani *Control Theory for Partial Differential Equations I*
75 I. Lasiecka and R. Triggiani *Control Theory for Partial Differential Equations II*
76 A. A. Ivanov *Geometry of Sporadic Groups I*
77 A. Schinzel *Polynomials with Special Regard to Reducibility*
78 T. Beth, D. Jungnickel and H. Lenz *Design Theory II, 2nd edn*
79 T. W. Palmer *Banach Algebras and the General Theory of *-Algebras II*
80 O. Stormark *Lie's Structural Approach to PDE Systems*
81 C. F. Dunkl and Y. Xu *Orthogonal Polynomials of Several Variables*
82 J. P. Mayberry *The Foundations of Mathematics in the Theory of Sets*
83 C. Foias et al. *Navier–Stokes Equations and Turbulence*
84 B. Polster and G. F. Steinke *Geometries on Surfaces*
85 R. B. Paris and D. Kaminski *Asymptotics and Mellin–Barnes Integrals*
86 R. J. McEliece *The Theory of Information and Coding, 2nd edn*
87 B. A. Magurn *An Algebraic Introduction to K-Theory*
88 T. Mora *Solving Polynomial Equation Systems I*
89 K. Bichteler *Stochastic Integration with Jumps*
90 M. Lothaire *Algebraic Combinatorics on Words*
91 A. A. Ivanov and S. V. Shpectorov *Geometry of Sporadic Groups II*
92 P. McMullen and E. Schulte *Abstract Regular Polytopes*
93 G. Gierz et al. *Continuous Lattices and Domains*
94 S. R. Finch *Mathematical Constants*
95 Y. Jabri *The Mountain Pass Theorem*
96 G. Gasper and M. Rahman *Basic Hypergeometric Series, 2nd edn*
97 M. C. Pedicchio and W. Tholen (eds.) *Categorical Foundations*
98 M. E. H. Ismail *Classical and Quantum Orthogonal Polynomials in One Variable*
99 T. Mora *Solving Polynomial Equation Systems II*
100 E. Olivieri and M. Eulália Vares *Large Deviations and Metastability*
101 A. Kushner, V. Lychagin and V. Rubtsov *Contact Geometry and Nonlinear Differential Equations*
102 L. W. Beineke and R. J. Wilson (eds.) with P. J. Cameron *Topics in Algebraic Graph Theory*
103 O. J. Staffans *Well-Posed Linear Systems*
104 J. M. Lewis, S. Lakshmivarahan and S. K. Dhall *Dynamic Data Assimilation*
105 M. Lothaire *Applied Combinatorics on Words*
106 A. Markoe *Analytic Tomography*
107 P. A. Martin *Multiple Scattering*
108 R. A. Brualdi *Combinatorial Matrix Classes*
109 J. M. Borwein and J. D. Vanderwerff *Convex Functions*
110 M.-J. Lai and L. L. Schumaker *Spline Functions on Triangulations*
111 R. T. Curtis *Symmetric Generation of Groups*
112 H. Salzmann et al. *The Classical Fields*
113 S. Peszat and J. Zabczyk *Stochastic Partial Differential Equations with Lévy Noise*
114 J. Beck *Combinatorial Games*
115 L. Barreira and Y. Pesin *Nonuniform Hyperbolicity*
116 D. Z. Arov and H. Dym *J-Contractive Matrix Valued Functions and Related Topics*
117 R. Glowinski, J.-L. Lions and J. He *Exact and Approximate Controllability for Distributed Parameter Systems*
118 A. A. Borovkov and K. A. Borovkov *Asymptotic Analysis of Random Walks*
119 M. Deza and M. Dutour Sikirić *Geometry of Chemical Graphs*
120 T. Nishiura *Absolute Measurable Spaces*
121 M. Prest Purity *Spectra and Localisation*
122 S. Khrushchev *Orthogonal Polynomials and Continued Fractions*
123 H. Nagamochi and T. Ibaraki *Algorithmic Aspects of Graph Connectivity*
124 F. W. King Hilbert *Transforms I*
125 F. W. King Hilbert *Transforms II*
126 O. Calin and D.-C. Chang *Sub-Riemannian Geometry*
127 M. Grabisch et al. *Aggregation Functions*
128 L. W. Beineke and R. J. Wilson (eds.) with J. L. Gross and T. W. Tucker *Topics in Topological Graph Theory*
129 J. Berstel, D. Perrin and C. Reutenauer *Codes and Automata*
130 T. G. Faticoni *Modules over Endomorphism Rings*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Stochastic Control and Mathematical Modeling

Applications in Economics

HIROAKI MORIMOTO

Ehime University, Matsuyama, Japan

CAMBRIDGE
UNIVERSITY PRESS

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

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521195034

First published 2010

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Morimoto, Hiraoki.
Stochastic control and mathematical modeling : applications in economics /
Hiroaki Morimoto.
p. cm. – (Encyclopedia of mathematics and its applications)
Includes bibliographical references and index.
ISBN 978-0-521-19503-4 (hardback)
1. Stochastic control theory. 2. Optimal stopping (Mathematical statistics)
3. Stochastic differential equations. I. Title. II. Series.
QA402.37.M67 2010
629.8′312 – dc22 2009042538

ISBN 978-0-521-19503-4 Hardback

To My Teacher M. Nisio

Contents

Preface

The purpose of this book is to provide a fundamental description of stochastic control theory and its applications to dynamic optimization in economics. Its content is suitable particularly for graduate students and scientists in applied mathematics, economics, and engineering fields.

A stochastic control problem poses the question: what is the optimal magnitude of a choice variable at each time in a dynamical system under uncertainty? In stochastic control theory, the state variables and control variables, respectively, describe the random phenomena of dynamics and inputs. The state variable in the problem evolves according to stochastic differential equations (SDE) with control variables. By steering of such control variables, we aim to optimize some performance criteria as expressed by the objective functional. Stochastic control can be viewed as a problem of decision making in maximization or minimization. This subject has created a great deal of mathematics as well as a large variety of applications in economics, mathematical finance, and engineering.

This book provides the basic elements of stochastic differential equations and stochastic control theory in a simple and self-contained way. In particular, a key to the stochastic control problem is the dynamic programming principle (DPP), which leads to the notion of viscosity solutions of Hamilton–Jacobi–Bellman (HJB) equations. The study of viscosity solutions, originated by M. Crandall and P. L. Lions in the 1980s, provides a useful tool for dealing with the lack of smoothness of the value functions in stochastic control. The main idea used to solve this maximization problem is summarized as follows:

(a) We formulate the problem and define the supremum of the objective functional over the class of all control variables, which is called the *value function*.
(b) We verify that the DPP holds for the value function.
(c) By the DPP, the value function can be viewed as a unique viscosity solution of the HJB equation associated with this problem.
(d) The uniform ellipticity and the uniqueness of viscosity solutions show the existence of a unique classical solution to the boundary value problem of

the HJB equation. This gives the smoothness of the viscosity solution of the HJB equation.

(e) We seek a candidate of optimal control by using the HJB equation. By using Itô's formula, we show the optimality.

This book is divided into three parts: Part I − Stochastic Calculus and Optimal Control Theory; Part II − Applications to Mathematical Models in Economics; and Part III − a collection of appendices providing background materials.

Part I consists of Chapters 1–5. In Chapter 1, we present the elements of stochastic calculus and SDEs, and in Chapter 2, we present the formulation of the weak solutions of SDEs, the concept of regular conditional probability, the Yamada–Watanabe theorem on weak and strong solutions, and the Markov property of a solution of SDE.

In Chapter 3, we introduce the DPP to issue (b). The verification of the DPP is rather difficult compared to the deterministic case. The Yamada–Watanabe theorem in Chapter 2 makes its proof exact. The supremum of (a) is taken over all systems in the weak sense.

Chapter 4 provides the theory of viscosity solutions of the HJB equations for (c). Using Ishii's lemma, we show the uniqueness results on viscosity solutions.

Chapter 5 is devoted to the boundary value problem of the HJB equations for (d) in the classical sense. Section 5.4 explains how to apply (a)–(e) in stochastic control.

Part II consists of Chapters 6–12. Here we present diverse applications of stochastic control theory to the mathematical models in economics. In Chapters 6–10, we take the state variables in these models as the remaining stock of a resource, the labor supply, and the price of the stock. The criteria in the maximization procedure are often given by the utility function of consumption rates as the control variables. Along (a)–(e), an optimal control is shown to exist.

Chapters 11 and 12 deal with the linear and nonlinear variational inequalities, instead of the HJB equations, which are associated with the stopping time problem. The variational inequality is analyzed by the viscosity solutions approach for optimality.

Part III consists of Appendices A–H. These provide some background material for understanding stochastic control theory as quickly as possible.

The prerequisites for this book are basic probability theory and functional analysis (see e.g., R. B. Ash [2], H. L. Royden [139], and A. Friedman [69]). See M. I. Kamien and N. L. Schwartz [80], A. C. Chiang [33], A. K. Dixit and R. S. Pindyck [46], L. Ljungqvist and T. J. Sargent [107], and R. S. Merton [114], for economics references.

Acknowledgments

I have greatly profited from the valuable comments and thoughtful suggestions of many of my colleagues, friends, and teachers, in particular: A. Bensoussan,

N. El Karoui, W. H. Fleming, H. Ishii, Y. Ishikawa, K. Kamizono, N. Kazamaki, S. Koike, H. Kunita, J. L. Menaldi, H. Nagai, M. Okada, S. Sakaguchi, Ł. Stettner, T. Tsuchikura, J. Zabczyk, and X. Y. Zhou.

I would like to thank I. Karatzas very much for reading the earlier drafts and his encouragement to publish this book. Thanks also go to K. Kawaguchi for many discussions on the mathematical treatment of economics and for his knowledge of the economic problems. I wish to thank T. Adachi, H. Kaise, and C. Liu for their careful reading of the manuscript and assistance at various stages.

H. Morimoto
Matsuyama
January 2009

Part I

Stochastic Calculus and Optimal Control Theory

1

Foundations of Stochastic Calculus

We are concerned here with a stochastic differential equation,

$$\begin{aligned} dX(t) &= b(X(t))dt + \sigma(X(t))dB(t), \quad t \geq 0, \\ X(0) &= x \in \mathbf{R}^N, \end{aligned}$$

in N-dimensional Euclidean space $\mathbf{R}^N$. Here b, σ are Lipschitz functions, called the *drift term* and the *diffusion term*, respectively, and $\{B(t)\}$ is a standard Brownian motion equation defined on a probability space $(\Omega, \mathcal{F}, P)$. This equation describes the evolution of a finite-dimensional dynamical system perturbed by noise, which is formally given by $dB(t)/dt$. In economic applications, the stochastic process $\{X(t)\}$ is interpreted as the labor supply, the price of stocks, or the price of capital at time $t \geq 0$. We present a reasonable definition of the second term with uncertainty and basic elements of calculus on the stochastic differential equation, called *stochastic calculus*.

A. Bensoussan [16], I. Karatzas and S. E. Shreve [87], N. Ikeda and S. Watanabe [75], I. Gihman and A. Skorohod [72], A. Friedman [68], B. Øksendal [132], D. Revuz and M. Yor [134], R. S. Liptzser and A. N. Shiryayev [106] are basic references for this chapter.

1.1 Review of Probability

1.1.1 Random Variables

Definition 1.1.1. *A triple* $(\Omega, \mathcal{F}, P)$ *is a* probability space *if the following assertions hold:*

(a) Ω *is a set.*

(b) $\mathcal{F}$ *is a* σ*-algebra, that is,* $\mathcal{F}$ *is a collection of subsets of* Ω *such that*

  (i) $\Omega,\ \phi \in \mathcal{F}$,

  (ii) *if* $A \in \mathcal{F}$*, then* $A^c := \Omega \setminus A \in \mathcal{F}$,

  (iii) *if* $A_n \in \mathcal{F}, n = 1, 2, \ldots,$ *then* $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$.

(c) *P is a probability measure, that is, a map $P : \mathcal{F} \to [0, 1]$, such that*
 (i) $P(\Omega) = 1$,
 (ii) *if $A_n \in \mathcal{F}$, $n = 1, 2, \dots$, disjoint, then $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$.*

Definition 1.1.2. *A probability space $(\Omega, \mathcal{F}, P)$ is complete if $A \in \mathcal{F}$ has $P(A) = 0$ and $B \subset A$, then $B \in \mathcal{F}$ (and, of course, $P(B) = 0$), that is, $\mathcal{F}$ contains all P-null sets.*

Remark 1.1.3. *Any probability space $(\Omega, \mathcal{F}, P)$ can be made complete by the completion of measures due to Carathéodory. We also refer to the proof of the Daniell Theorem, Theorem 2.1 in Chapter 2.*

Definition 1.1.4. *For any collection $\mathcal{G}$ of subsets of Ω, we define a smallest σ-algebra $\sigma(\mathcal{G})$ containing $\mathcal{G}$ by*

$$\sigma(\mathcal{G}) = \cap\{\mathcal{F} : \mathcal{G} \subset \mathcal{F},\ \mathcal{F}\ \sigma\text{-algebra of } \Omega\},$$

which is the σ-algebra generated by $\mathcal{G}$.

Example 1.1.5. *On the set of real numbers $\mathbf{R}$, we take $\mathcal{G} = \{$open intervals$\}$ and denote by $\mathcal{B}(\mathbf{R})$ the σ-algebra $\sigma(\mathcal{G})$ generated by $\mathcal{G}$, which is the Borel σ-algebra on $\mathbf{R}$.*

Definition 1.1.6. *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space.*

(a) *A map $X : \Omega \to \mathbf{R}$ is a random variable if*

$$X^{-1}(B) := \{\omega : X(\omega) \in B\} \in \mathcal{F}, \qquad \text{for any } B \in \mathcal{B}(\mathbf{R}).$$

(b) *For any random variable X, we define the σ-algebra $\sigma(X)$ generated by X as follows:*

$$\sigma(X) = \sigma(\mathcal{G}) = \mathcal{G}, \quad \mathcal{G} := \{X^{-1}(B) : B \in \mathcal{B}(\mathbf{R})\} \subset \mathcal{F}.$$

Proposition 1.1.7. *Let X, Y be two random variables. Then Y is $\sigma(X)$ measurable if and only if there exists a Borel measurable function $g : \mathbf{R} \to \mathbf{R}$ such that*

$$Y(\omega) = g(X(\omega)), \qquad \text{for all } \omega \in \Omega.$$

Proof. Since $Y = Y^+ - Y^-$, we will show the "only if" part when $Y \geq 0$.

(1) Suppose that Y is a simple random variable. Then Y is of the form:

$$Y(\omega) = \sum_{i=1}^{n} y_i 1_{F_i}(\omega),$$

where $y_i \geq 0$, $F_i \in \sigma(X)$ and the F_i are pairwise disjoint. By definition, there exists $D_i \in \mathcal{B}(\mathbf{R})$, for each i, such that $F_i = X^{-1}(D_i)$. Clearly, the D_i are pairwise disjoint. Define

$$g(y) = \begin{cases} y_i, & y \in D_i, \\ 0, & y \notin \bigcup_{i=1}^{n} D_i. \end{cases}$$

Then

$$Y(\omega) = \sum_{i=1}^{n} y_i 1_{\{X^{-1}(D_i)\}}(\omega) = \sum_{i=1}^{n} y_i 1_{D_i}(X(\omega)) = g(X(\omega)).$$

(2) In the general case, there exists a sequence of simple random variables Y_n converging to Y. Let g_n be the corresponding sequence of measurable functions such that $Y_n = g_n(X)$. Define

$$g(y) = \liminf_{n\to\infty} g_n(y).$$

Then g is $\mathcal{B}(\mathbf{R})$ measurable and

$$Y(\omega) = \liminf_{n} Y_n(\omega) = \liminf_{n} g_n(X(\omega)) = g(X(\omega)), \quad \omega \in \Omega. \quad \blacksquare$$

1.1.2 Expectation, Conditional Expectation

Definition 1.1.8. *Let X be a random variable. The quantity*

$$E[X] = \int_{\Omega} X(\omega) dP(\omega)$$

is the expectation of X, where $E[X^+]$ or $E[X^-]$ is finite.

Definition 1.1.9. *Let X, Y be two random variables on a complete probability space $(\Omega, \mathcal{F}, P)$.*

(a) *The expression $X = Y$ will indicate that $X = Y$ a.s., that is, $P(X \neq Y) = 0$.*

(b) *For $1 \leq p < \infty$, the norm $\|X\|_p$ of X is defined by*

$$\|X\|_p = (E[|X|^p])^{1/p}.$$

(c) *If $p = \infty$, then*

$$\|X\|_\infty = ess\ sup|X| = \inf\{\sup_{\omega\notin N}|X(\omega)| : N \in \mathcal{F}, P(N) = 0\}.$$

(d) *The L^p spaces are defined by*

$$L^p = L^p(\Omega) = \{X : random\ variable, \|X\|_p < \infty\}.$$

Proposition 1.1.10.

(i) *$L^p(\Omega)$ is a Banach space, that is, a complete normed linear space, for $1 \leq p \leq \infty$.*

(ii) *$L^2(\Omega)$ is a Hilbert space, that is, a complete inner product space, with inner product $(X, Y) = E[X \cdot Y], \quad X, Y \in L^2(\Omega)$.*

For the proof, see A. Friedman [69, chapter 3].

Definition 1.1.11. *Let X_n, $n = 1, 2, \ldots$, and X be random variables.*

(a) *$X_n \to X$ a.s. if $P(X_n \to X$ as $n \to \infty) = 1$.*

(b) $X_n \to X$ *in probability if* $P(|X_n - X| \geq \varepsilon) \to 0$ *as* $n \to \infty$*, for any* $\varepsilon > 0$.
(c) $X_n \to X$ *in* L^p *if* $\|X_n - X\|_p \to 0$ *as* $n \to \infty$.

Proposition 1.1.12. *Let* $X_n, n = 1, 2, \dots,$ *and* X *be random variables.*

(i) *If* $X_n \to X$ *a.s., then* $X_n \to X$ *in probability.*
(ii) *If* $X_n \to X$ *in* L^p $(p \geq 1)$*, then* $X_n \to X$ *in probability.*
(iii) $X_n \to X$ *in probability if and only if* $E[\frac{|X_n - X|}{1+|X_n - X|}] \to 0$.
(iv) *Let* $\varepsilon_n \geq 0$ *and* $\sum_{n=1}^{\infty} \varepsilon_n < \infty$*. If* $\sum_{n=1}^{\infty} P(|X_{n+1} - X_n| \geq \varepsilon_n) < \infty$*, then* X_n *converges a.s.*
(v) *If* $X_n \to X$ *in probability, then it contains a subsequence* $\{X_{n_k}\}$ *such that* $X_{n_k} \to X$ *a.s.*

For the proof, see A. Friedman [69, chapter 2].

Definition 1.1.13. *A family* $\{X_n : n \in \mathbf{N}\}$ *of random variables* X_n *on* $(\Omega, \mathcal{F}, P)$ *is uniformly integrable if*

$$\lim_{a \to \infty} \sup_n \int_{\{|X_n| \geq a\}} |X_n| dP = 0.$$

Proposition 1.1.14. *Assume that one of the following assertions is satisfied:*

(i) $E[\sup_n |X_n|] < \infty$,
(ii) $\sup_n E[|X_n|^p] < \infty$, *for some* $p > 1$.

Then $\{X_n\}$ *is uniformly integrable.*

Proof.

(1) We set $Y = \sup_n |X_n|$. Then, by (i),

$$P(Y \geq c) \leq \frac{1}{c} E[Y] \to 0 \quad \text{as } c \to \infty.$$

Thus

$$\sup_n \int_{\{|X_n| \geq c\}} |X_n| dP \leq \int_{\{Y \geq c\}} Y dP \to 0 \quad \text{as } c \to \infty.$$

(2) By Chebyshev's inequality,

$$\sup_n P(|X_n| \geq c) \leq \frac{1}{c^p} \sup_n E[|X_n|^p].$$

Thus, by (ii) and Hölder's inequality,

$$\begin{aligned}\sup_n \int_{\{|X_n| \geq c\}} |X_n| dP &\leq \sup_n (E[|X_n|^p])^{1/p} (E[1_{\{|X_n| \geq c\}}])^{1/q} \\ &\leq \sup_n E[|X_n|^p] (\frac{1}{c})^{p/q} \to 0 \quad \text{as } c \to \infty,\end{aligned}$$

where $1/p + 1/q = 1$. ■

Proposition 1.1.15. *Let $\{X_n\}$ be a sequence of integrable random variables such that $X_n \to X$ a.s. Then $\{X_n\}$ is uniformly integrable if and only if*

$$\lim_{n\to\infty} E[|X_n - X|] = 0.$$

Proof. Let $Y_n = X_n - X$.

(1) Suppose that $\{X_n\}$ is uniformly integrable. Since

$$\begin{aligned} E[|X_n|] &= E[|X_n|1_{\{|X_n|\geq a\}}] + E[|X_n|1_{\{|X_n|<a\}}] \\ &\leq \sup_n E[|X_n|1_{\{|X_n|\geq a\}}] + aP(|X_n| < a), \qquad \text{for any } a > 0, \end{aligned}$$

we have $\sup_n E[|X_n|] < \infty$, taking sufficiently large $a > 0$. By Fatou's lemma

$$E[|X|] \leq \liminf_{n\to\infty} E[|X_n|] < \infty.$$

Also, by Chebyshev's inequality,

$$\begin{aligned} E[|Y_n|] &\leq E[|X_n|1_{\{|Y_n|\geq a\}} + |X|1_{\{|Y_n|\geq a\}} + |Y_n|1_{\{|Y_n|<a\}}] \\ &\leq E[|X_n|1_{\{|X_n|\geq c\}}] + 2cP(|Y_n| \geq a) + E[|X|1_{\{|X|\geq c\}}] \\ &\quad + E[|Y_n|1_{\{|Y_n|<a\}}] \\ &\leq \sup_n E[|X_n|1_{\{|X_n|\geq c\}}] + \frac{2c}{a}\sup_n E[|X_n| + |X|] + E[|X|1_{\{|X|\geq c\}}] \\ &\quad + (1+a)E[\frac{|Y_n|}{1+|Y_n|}]. \end{aligned}$$

Letting $n \to \infty$, $a \to \infty$, and then $c \to \infty$, we get $\limsup_{n\to\infty} E[|Y_n|] = 0$.

(2) Conversely, it is easy to see that

$$P(|Y_n| \geq a) \leq \frac{1}{a}E[|Y_n|] \to 0 \qquad \text{as } a \to \infty$$

and

$$\begin{aligned} E[|Y_n|1_{\{|Y_n|\geq a\}}] &= E[(|Y_n| - |Y_n| \wedge m)1_{\{|Y_n|\geq a\}}] + E[(|Y_n| \wedge m)1_{\{|Y_n|\geq a\}}] \\ &\leq E[|Y_n| - |Y_n| \wedge m] + mP(|Y_n| \geq a). \end{aligned}$$

Hence, letting $a \to \infty$ and then $m \to \infty$, we get

$$\limsup_{a\to\infty} E[|Y_n|1_{\{|Y_n|\geq a\}}] = 0, \qquad \text{for each } n.$$

Next, for any $k \in \mathbf{N}$,

$$\begin{aligned} \sup_n E[|Y_n|1_{\{|Y_n|\geq a\}}] &\leq \sup_{n\leq k} E[|Y_n|1_{\{|Y_n|\geq a\}}] + \sup_{n>k} E[|Y_n|1_{\{|Y_n|\geq a\}}] \\ &\leq \sum_{n=1}^{k} E[|Y_n|1_{\{|Y_n|\geq a\}}] + \sup_{n>k} E[|Y_n|]. \end{aligned}$$

Letting $a \to \infty$ and then $k \to \infty$, we have $\lim_{a\to\infty} \sup_n E[|Y_n| 1_{\{|Y_n| \geq a\}}] = 0$. By the same calculation as shown in (1), we have

$$\begin{aligned}\sup_n E[|X_n| 1_{\{|X_n|\geq c\}}] &\leq \sup_n E[|Y_n| 1_{\{|X_n|\geq c\}} + |X| 1_{\{|X_n|\geq c\}}] \\ &\leq \sup_n E[|Y_n| 1_{\{|Y_n|\geq a\}}] + \frac{2a}{c} \sup_n E[|X_n|] \\ &\quad + E[|X| 1_{\{|X|\geq a\}}].\end{aligned}$$

Letting $c \to \infty$ and then $a \to \infty$, we obtain uniform integrability. ∎

Definition 1.1.16. *Let $X \in L^1(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$, that is, $\mathcal{G} \subset \mathcal{F}$ is a σ-algebra. A random variable $Y \in L^1(\Omega, \mathcal{G}, P)$ is the conditional expectation of X given $\mathcal{G}$ if*

$$\int_A Y dP = \int_A X dP, \quad \textit{for all } A \in \mathcal{G}.$$

We write $Y = E[X|\mathcal{G}]$.

Proposition 1.1.17. *Let $X \in L^1(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ a sub-σ-algebra of $\mathcal{F}$. Then the conditional expectation $Y \in L^1(\Omega, \mathcal{G}, P)$ of X given $\mathcal{G}$ exists uniquely.*

Proof. Without loss of generality, we may assume $X \geq 0$. Define

$$\mu(A) = \int_A X dP, \quad \text{for } A \in \mathcal{G}.$$

Then μ is a finite measure, absolutely continuous with respect to P. By the Radon–Nikodým Theorem (cf. A. Friedman [69]), there exists, uniquely, $Y \in L^1(\Omega, \mathcal{G}, P)$, $Y \geq 0$, such that $\mu(A) = \int_A Y dP$, for $A \in \mathcal{G}$. ∎

Remark 1.1.18. *We recall that $L^2(\Omega, \mathcal{G}, P)$ is a closed subspace of the Hilbert space $L^2(\Omega, \mathcal{F}, P)$. If $X \in L^2(\Omega, \mathcal{F}, P)$, then $E[X|\mathcal{G}]$ coincides with the orthogonal projection $\hat{X}$ of X to $L^2(\Omega, \mathcal{G}, P)$, that is,*

$$E[|X - E[X|\mathcal{G}]|^2] = \min\{E[|X - Y|^2] : Y \in L^2(\Omega, \mathcal{G}, P)\}.$$

Proposition 1.1.19. *Let $X_n, X \in L^1(\Omega, \mathcal{F}, P)$, $n = 1, 2, \ldots,$ and $\mathcal{H}, \mathcal{G}$ be two sub-σ-algebras of $\mathcal{F}$. Then, the following assertions hold:*

(i) $E[E[X|\mathcal{G}]] = E[X]$.
(ii) $E[X|\mathcal{G}] = X$ *a.s. if X is $\mathcal{G}$-measurable.*
(iii) $E[aX_1 + bX_2|\mathcal{G}] = aE[X_1|\mathcal{G}] + bE[X_2|\mathcal{G}]$ *a.s.*, $\quad a, b \in \mathbf{R}$.
(iv) $E[X|\mathcal{G}] \geq 0$ *a.s. if $X \geq 0$.*
(v) $E[X_n|\mathcal{G}] \nearrow E[X|\mathcal{G}]$ *a.s. if $X_n \nearrow X$.*
(vi) $E[\liminf_{n\to\infty} X_n|\mathcal{G}] \leq \liminf_{n\to\infty} E[X_n|\mathcal{G}]$ *a.s. if $X_n \geq 0$.*
(vii) $E[X_n|\mathcal{G}] \to E[X|\mathcal{G}]$ *a.s. if $X_n \to X$ a.s. and* $\sup_n |X_n| \in L^1$.
(viii) $f(E[X|\mathcal{G}]) \leq E[f(X)|\mathcal{G}]$ *a.s. if $f : \mathbf{R} \to \mathbf{R}$ is convex, and $f(X) \in L^1(\Omega)$.*
(ix) $E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}]$ *a.s. if $\mathcal{H} \subset \mathcal{G}$.*

(x) $E[ZX|\mathcal{G}] = ZE[X|\mathcal{G}]$ *a.s. if* $X \in L^p(\Omega, \mathcal{F}, P)$, $Z \in L^q(\Omega, \mathcal{G}, P)$, *for* $p = 1$, $q = \infty$ *or* $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$.
(xi) $E[X|\mathcal{G}] = E[X]$ *a.s. if* X *and* 1_A *are independent for any* $A \in \mathcal{G}$.

Proof. The proof is obtained by integrating over arbitrary sets $A \in \mathcal{G}$ and by using several properties of the integrals. In particular, (v) is immediately from the monotone convergence theorem. For (vi), we apply (v) to $Y_n = \inf_{m \geq n} X_m$, and (vii) follows from (vi). ■

1.2 Stochastic Processes

1.2.1 General Notations

Definition 1.2.1. *A quadruple* $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ *is a stochastic basis if the following assertions hold:*

(a) $(\Omega, \mathcal{F}, P)$ *is a complete probability space.*
(b) $\{\mathcal{F}_t\}_{t\geq 0}$ *is a filtration, that is, a nondecreasing family of sub-*σ*-algebra of* $\mathcal{F}$:

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \text{ for } 0 \leq s < t < \infty.$$

(c) *The filtration* $\{\mathcal{F}_t\}_{t\geq 0}$ *satisfies the following ("usual") conditions*:
  (i) $\{\mathcal{F}_t\}_{t\geq 0}$ *is right-continuous:* $\mathcal{F}_t = \mathcal{F}_{t+} := \cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}$, *for all* $t \geq 0$.
  (ii) $\mathcal{F}_0$ *contains all* P*-null sets in* $\mathcal{F}$.

Definition 1.2.2. *An* N*-dimensional stochastic process* $X = \{X_t\}_{t\geq 0}$ *on a complete probability space* $(\Omega, \mathcal{F}, P)$ *is a collection of* $\mathbf{R}^N$*-valued random variables* $X(t, \omega)$, $\omega \in \Omega$. *For fixed* $\omega \in \Omega$, *the set* $\{X(t, \omega) : t \geq 0\}$ *is a path of* X.

Definition 1.2.3. *Let* $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ *be a stochastic basis, and let* $\mathcal{B}(\mathbf{R})$, $\mathcal{B}(\mathbf{R}^N)$, $\mathcal{B}([0, t])$ *be Borel* σ*-algebra on* $\mathbf{R}$, $\mathbf{R}^N$, $[0, t]$.

(a) *A stochastic process* $\{X_t\}_{t\geq 0}$ *is* $\{\mathcal{F}_t\}_{t\geq 0}$*-adapted* (*with respect to the filtration* $\{\mathcal{F}_t\}_{t\geq 0}$) *if* X_t *is* $\mathcal{F}_t$*-measurable for all* $t \geq 0$.
Such a stochastic process will be denoted by $\{(X_t, \mathcal{F}_t)\}_{t\geq 0}$ *or* $(X_t, \mathcal{F}_t)$, $\{(X(t), \mathcal{F}_t)\}$.
(b) *A stochastic process* $\{X_t\}_{t\geq 0}$ *is measurable if, for all* $A \in \mathcal{B}(\mathbf{R}^N)$,

$$\{(t, \omega) : X_t(\omega) \in A\} \in \mathcal{B}([0, \infty)) \otimes \mathcal{F}.$$

(c) *A measurable adapted stochastic process* $\{X_t\}_{t\geq 0}$ *is progressively measurable if, for each* $t \geq 0$ *and* $A \in \mathcal{B}(\mathbf{R}^N)$,

$$\{(s, \omega) : 0 \leq s \leq t,\ X_s(\omega) \in A\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t.$$

(d) *A stochastic process* $\{X_t\}_{t\geq 0}$ *is continuous* (*right-continuous*) *if there is* $\Omega_0 \in \mathcal{F}$ *with* $P(\Omega_0) = 1$ *such that* $t \to X_t(\omega)$ *is continuous* (*right-continuous*) *for every* $\omega \in \Omega_0$.

(e) *Two stochastic processes X and Y are indistinguishable if there is $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$ such that $X_t(\omega) = Y_t(\omega)$ for all $t \geq 0$ and $\omega \in \Omega_0$. The expression $X = Y$ will indicate that X and Y are indistinguishable.*
(f) *Y is a modification of X if $P(X_t = Y_t) = 1$ for all $t \geq 0$.*

Proposition 1.2.4. *Let $\{(X_t, \mathcal{F}_t)\}_{t\geq 0}$ be a stochastic process on a stochastic basis $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$.*

(i) *If $X = \{X_t\}$ is a right-continuous modification of 0, then $X = 0$.*
(ii) *If $X = \{X_t\}$ is right-continuous, then X is progressively measurable.*

Proof. Let $\Omega_0 = \{\omega : t \to X_t(\omega)$ right-continuous$\}$ and $P(\Omega_0) = 1$.

(1) Since $\{X_t\}$ is a modification of 0, $P(X_r \neq 0) = 0$ for each $r \in \mathbf{Q}_+$, and

$$P(\bigcup_{r\in\mathbf{Q}_+} \{\omega : X_r(\omega) \neq 0\}) = 0.$$

Define $\Omega' = (\cap_{r\in\mathbf{Q}_+}\{\omega : X_r(\omega) = 0\}) \cap \Omega_0$. It is obvious that $P(\Omega') = 1$ and $X_t(\omega) = 0$, for all $t \geq 0$ and $\omega \in \Omega'$.

(2) Taking into account the indistinguishable process of $\{X_t\}$, we may consider that $X(t, \omega) = 0$ if $\omega \notin \Omega_0$. Fix $t \geq 0$, and let $\delta = \{t_0 = 0 < t_1 < t_2 < \cdots < t_n = t\}$ be a partition of $[0, t]$, with $t_k = kt/2^n$, $k = 0, 1, \ldots, 2^n$. Define

$$X^\delta(s, \omega) = \begin{cases} X(0, \omega)1_{\{0\}}(s) + \sum_{k=1}^{2^n} X(t_k, \omega)1_{(t_{k-1},t_k]}(s) & \text{if } \omega \in \Omega_0, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, the map $(s, \omega) \to X^\delta(s, \omega)$ is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$-measurable. Letting $n \to \infty$, by right continuity, we see that $X^\delta(s, \omega) \to X(s, \omega)$ for all $s \in [0, t]$ and $\omega \in \Omega$. Thus the map $(s, \omega) \to X(s, \omega)$ is measurable for this σ-algebra, so $\{X_t\}$ is progressively measurable. ■

Remark 1.2.5.

(a) *Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration such that $\mathcal{F}_0$ contains all $P-$null sets. If $X = Y$ and X is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$, then Y is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$.*
(b) *For any filtration $\{\mathcal{G}_t\}_{t\geq 0}$ we can obtain the right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$ defined by $\mathcal{F}_t = \mathcal{G}_{t+} := \bigcap_{\varepsilon>0} \mathcal{G}_{t+\varepsilon}$.*
(c) *The filtration in connection with the stochastic process $\{X_t\}_{t\geq 0}$ is the σ-algebra*

$\sigma(X_s, s \leq t)$ generated by $\{X_s, s \leq t\}$, where

$$\sigma(X_s, s \leq t) := \sigma(\mathcal{G}_t), \quad \mathcal{G}_t = \{X_s^{-1}(A) : A \in \mathcal{B}(\mathbf{R}^N), s \leq t\}.$$

(d) *Furthermore, a stochastic basis $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t^X\})$ is obtained by setting*

$$\mathcal{F}_t^X = \mathcal{H}_{t+}, \quad \mathcal{H}_t = \sigma(X_s, s \leq t) \vee \mathcal{N} := \sigma(\sigma(X_s, s \leq t) \cup \mathcal{N}),$$

where $\mathcal{N}$ is the collection of P null sets.

1.2.2 Brownian Motion

Definition 1.2.6.

(a) *The real-valued stochastic process* $\{B_t\}_{t\geq 0}$ *is a (one-dimensional standard) Brownian motion if*
  (i) $\{B_t\}_{t\geq 0}$ *is continuous,* $B_0 = 0$ *a.s.,*
  (ii) $B_t - B_s$ *is independent of* $\mathcal{F}_s^B$,
  (iii) $P(B_t - B_s \in A) = \int_A \frac{1}{\sqrt{2\pi(t-s)}} \exp\{-\frac{x^2}{2(t-s)}\}dx$, *for* $t > s$, $A \in \mathcal{B}(\mathbf{R})$.

(b) *The* N*-dimemsional stochastic process* $B(t) = (B_1(t), B_2(t), \ldots, B_N(t))$ *is a (standard)* N*-dimensional Brownian motion if the* N*-components* $B_i(t)$ *are independent one-dimensional standard Brownian motions.*

Proposition 1.2.7. *A continuous stochastic process* $\{B_t\}_{t\geq 0}$ *with* $B_0 = 0$ *is an* N*-dimensional Brownian motion if and only if*

$$E[e^{i(\xi, B_t - B_s)}|\mathcal{F}_s^B] = e^{-\frac{|\xi|^2}{2}(t-s)}, \quad t > s, \quad \xi \in \mathbf{R}^N,$$

where $i = \sqrt{-1}$, $(\ ,\ )$ *denotes the inner product of* $\mathbf{R}^N$, *and* $|\xi| = (\xi, \xi)^{1/2}$.

Proof. For $t > s$, the random variable $Y := B_t - B_s$ is normally distributed with mean 0 and covariance $(t - s)I$ (I : identity), that is to say, the characteristic function of Y is given by

$$E[e^{i(\xi, Y)}] = e^{-\frac{|\xi|^2}{2}(t-s)}.$$

Let Y be independent of $\mathcal{F}_s^B$. Then it is easy to see that

$$E[e^{i(\xi, Y)}|\mathcal{F}_s^B] = E[e^{i(\xi, Y)}].$$

Conversely,

$$E[e^{i(\xi, Y)}e^{i(\eta, Z)}] = E[E[e^{i(\xi, Y)}|\mathcal{F}_s^B]e^{i(\eta, Z)}] = E[e^{i(\xi, Y)}]E[e^{i(\eta, Z)}],$$

for any $\eta \in \mathbf{R}^N$ and $\mathcal{F}_s^B$-measurable random variable Z. Thus we conclude that Y and Z are independent. ■

Remark 1.2.8. *The existence of a Brownian motion can be shown by introducing the probability measure* P, *called a Wiener measure, on the space* $C([0, \infty) : \mathbf{R}^N)$ *of* $\mathbf{R}^N$*-valued contiunuous functions on* $[0, \infty)$. *See I. Karatzas and S. E. Shreve [87] and K. Ito and H. P. Mackean [75] for details. The remarkable properties of the Brownian motion are as follows:*

(a) *The Brownian motion* $\{B_t\}$ *is not differentiable a.s.,*
(b) *The total variation of* $\{B_t\}$ *on* $[0, T]$ *is infinite a.s.*

1.3 Stopping Times

Definition 1.3.1. *Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ be a stochastic basis. A map $\tau : \Omega \to [0, \infty]$ is a stopping time if*

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t, \quad \textit{for all } t \geq 0.$$

For any stopping time τ, the σ-algebra $\mathcal{F}_\tau$ is defined by

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, \textit{ for all } t \geq 0\}.$$

Proposition 1.3.2. *Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ be a stochastic basis. Let $\{X_t\}_{t\geq 0}$ be a continuous, $\{\mathcal{F}_t\}_{t\geq 0}$-adapted, $\mathbf{R}^N$-valued process and A an open or closed subset of $\mathbf{R}^N$. Then*

$$\tau(\omega) = \inf\{t \geq 0 : X_t(\omega) \in A\}, \quad \textit{with understanding } \inf\{\phi\} = \infty,$$

is a stopping time.

Proof. Let $\Omega_0 = \{t \to X_t(\omega) \text{ continuous}\}$ and $P(\Omega_0) = 1$.

(1) Let A be open. If $\tau(\omega) < t$ and $\omega \in \Omega_0$, then there is $s < t$ such that $X_s(\omega) \in A$. Taking $r_n \in \mathbf{Q}_+$ such that $r_n \searrow s$, we have $X_{r_n}(\omega) \in A$, for some $r_n < t$. Hence,

$$\{\tau < t\} \cap \Omega_0 = \bigcup_{r<t,\ r\in\mathbf{Q}_+} (\{X_r \in A\} \cap \Omega_0) \in \mathcal{F}_t.$$

Thus, by the completeness of $\mathcal{F}_t$,

$$\{\tau \leq t\} = \bigcap_{n\in\mathbf{N}} \{\tau < t + 1/n\} \in \mathcal{F}_{t+} = \mathcal{F}_t, \quad \text{for all } t \geq 0.$$

(2) Let A be closed. For each $n \in \mathbf{N}$, the set $A_n := \{x : d(x, A) < 1/n\}$ is open, where $d(x, A) = \inf\{|x - y| : y \in A\}$. Let $X_{r_m}(\omega) \in A_n,\ \omega \in \Omega_0$, and $r_m \to s \leq t$. Then

$$d(X_{r_m}(\omega), A) \leq d(X_{r_m}(\omega), A_n) + \frac{1}{n}.$$

Passing to the limit, we get $d(X_s(\omega), A) = 0$. This implies $X_s(\omega) \in A$, and then $\tau(\omega) \leq s$. Hence

$$\{\tau \leq t\} \cap \Omega_0 = \bigcap_{n\in\mathbf{N}} \bigcup_{r\in([0,t)\cap\mathbf{Q})\cup\{t\}} (\{X_r \in A_n\} \cap \Omega_0) \in \mathcal{F}_t.$$

By the same argument as above, we have $\{\tau \leq t\} \in \mathcal{F}_t$, for all $t \geq 0$. ■

Proposition 1.3.3. *Let $\tau, \sigma,$ and $\tau_n, n = 1, 2, \ldots,$ be stopping times. Then the following assertions hold:*

(i) *$\tau \wedge \sigma, \tau \vee \sigma,$ and $\tau + a$ $(a \in \mathbf{R}_+)$ are stopping times.*
(ii) *τ is $\mathcal{F}_\tau$-measurable.*
(iii) *If $\tau \leq \sigma$, then $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.*
(iv) *If $A \in \mathcal{F}_\tau$, then $A \cap \{\tau \leq \sigma\} \in \mathcal{F}_\sigma$.*
(v) *$\mathcal{F}_\tau \cap \mathcal{F}_\sigma = \mathcal{F}_{\tau\wedge\sigma}$.*

(vi) $\{\tau < \sigma\}, \{\tau = \sigma\}, \{\tau > \sigma\} \in \mathcal{F}_{\tau\wedge\sigma}$.
(vii) $\sup_n \tau_n$ *and* $\inf_n \tau_n$ *are stopping times.*

For the proof, see I. Karatzas and S. E. Shreve [87, section 1.1.2, pp. 6–11].

Definition 1.3.4. *Let $X = \{X_t\}_{t\geq 0}$ be a measurable process on a stochastic basis $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$, and let τ be a stopping time. We define the random variable X_τ by $X_\tau(\omega) = X(\tau(\omega), \omega)$ on $\{\tau < \infty\}$, and $X_\tau(\omega) = X_\infty(\omega)$ on $\{\tau = \infty\}$ if $X_\infty(\omega)$ is defined for all $\omega \in \Omega$.*

Proposition 1.3.5. *Let $X = \{X_t\}_{t\geq 0}$ be an adapted process on a stochastic basis $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$, and let τ be a stopping time. If X is progressively measurable, then X_τ is $\mathcal{F}_\tau$-measurable and the stopped process $X^\tau = \{X_{t\wedge\tau}\}_{t\geq 0}$ at τ is progressively measurable.*

Proof.

(1) For fixed $t \geq 0$, it is easy to see that $\tau \wedge t$ is an $\mathcal{F}_t$-measurable random variable. Hence the map $(s, \omega) \in [0, t] \times \Omega \to H(s, \omega) := (\tau(\omega) \wedge s, \omega) \in [0, t] \times \Omega$ is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ measurable. Thus the map $(s, \omega) \to X(\tau(\omega) \wedge s, \omega) = X \circ H(s, \omega)$ is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ measurable, so $\{X_{\tau\wedge t}\}$ is progressively measurable.
(2) Let $B \in \mathcal{B}(\mathbf{R}^N)$. Taking $s = t$ in (1), the map $\omega \in \Omega \to X \circ H(t, \omega)$ is $\mathcal{F}_t$-measurable. Hence

$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\tau\wedge t} \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t, \qquad \text{for all } t \geq 0.$$

This implies that $\{X_\tau \in B\} \in \mathcal{F}_\tau$, that is, X_τ is $\mathcal{F}_\tau$ measurable. ■

1.4 Martingales

Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ be a stochastic basis.

Definition 1.4.1. *A stochastic process $X = \{(X_t, \mathcal{F}_t)\}_{t\geq 0}$ is a martingale (supermartingale submartingale) if $E[|X_t|] < \infty$, for all $t \geq 0$, and*

$$E[X_t|\mathcal{F}_s] = X_s, \quad P\text{-}a.s., \quad \text{for any } s \leq t,$$

(if $E[X_t|\mathcal{F}_s] \leq X_s$ or $E[X_t|\mathcal{F}_s] \geq X_s$).

Proposition 1.4.2. *Let $\{B_t\}$ is a one-dimensional Brownian motion. Then*

(i) $(B_t, \mathcal{F}_t^B)$ *is a martingale,*
(ii) $(B_t^2 - t, \mathcal{F}_t^B)$ *is a martingale.*

Proof. Let $t \geq s$.

(1) $E[B_t - B_s|\mathcal{F}_s^B] = E[B_t - B_s] = 0, \quad$ a.s.

(2) By Proposition 1.1.19 (x), we have

$$\begin{aligned} E[B_t^2 - B_s^2|\mathcal{F}_s^B] &= E[\{(B_t - B_s) + B_s\}^2 - B_s^2|\mathcal{F}_s] \\ &= E[(B_t - B_s)^2 + 2(B_t - B_s)B_s|\mathcal{F}_s^B] \\ &= E[(B_t - B_s)^2] = t - s, \quad \text{a.s.} \end{aligned}$$

■

Theorem 1.4.3 (Doob's maximal inequality). *Let* $\{(X_t, \mathcal{F}_t)\}_{t\geq 0}$ *be a right-continuous stochastic process and* $0 < T < \infty$, $\lambda > 0$. *Then we have*

(i) *If* $\{X_t\}$ *is a nonnegative submartingale, then*

$$P(\sup_{0\leq t\leq T} X_t \geq \lambda) \leq \frac{1}{\lambda} E[X_T].$$

(ii) *If* $\{X_t\}$ *is a nonnegative submartingale or a martingale, then*

$$E[\sup_{0\leq t\leq T} |X_t|^p] \leq (\frac{p}{p-1})^p E[|X_T|^p], \quad \textit{for } 1 < p < \infty.$$

Proof. Let D be a countable dense subset of $[0, T]$ defined by $[0, T) \cap \mathbf{Q}$, and let $D_n = \{t_i \in D : t_0 = 0 \leq t_1 \leq \cdots \leq t_n, i = 0, 1, \ldots, n\}$ be a sequence of subsets of D such that $D_n \nearrow \bigcup_n D_n = D$ as $n \to \infty$. By right continuity, we observe that

$$\sup_{0\leq t\leq T} X_t = \sup_{t\in D} X_t = \lim_{n\to\infty} \max_{t\in D_n} X_t, \quad \text{a.s.,}$$

and it is $\mathcal{F}_T$-measurable. Furthermore,

$$\begin{aligned} P(\sup_{t\in D} X_t \geq \lambda) &= \lim_{m\to\infty} P(\sup_{t\in D} X_t > \lambda - \frac{1}{m}) \\ &= \lim_{m\to\infty} \lim_{n\to\infty} P(\max_{t\in D_n} X_s > \lambda - \frac{1}{m}). \end{aligned}$$

(1) We claim that

$$P(\max_{t\in D_n} X_t > \lambda) \leq \frac{1}{\lambda} E[X_T].$$

Put $Y_i = X_{t_i}$ and $\mathcal{G}_i = \mathcal{F}_{t_i}$, for $i = 0, 1, \ldots, n$. Let us define

$$\tau = \min\{i \leq n : Y_i > \lambda\}, \ = n \quad \text{if } \{\ \} = \phi.$$

Then τ is a $\mathcal{G}_i$-stopping time such that $\tau \leq n$ a.s. Since $\{Y_i\}$ is a $\mathcal{G}_i$-submartingale,

$$\begin{aligned} E[Y_\tau] &= \sum_{i=0}^n E[Y_i 1_{\{\tau=i\}}] \leq \sum_{i=0}^n E[E[Y_n|\mathcal{G}_i]1_{\{\tau=i\}}] \\ &= \sum_{i=0}^n E[Y_n 1_{\{\tau=i\}}] = E[Y_n] \leq E[X_T]. \end{aligned}$$

We set $Y_n^* = \max_{i \le n} Y_i = \max_{t \in D_n} X_t$ and note that $Y_\tau > \lambda$ on $\{Y_n^* > \lambda\}$ and $\{\tau = n\}$ on $\{Y_n^* \le \lambda\}$. Therefore,

$$\begin{aligned} E[Y_\tau] &= E[Y_\tau 1_{\{Y_n^* > \lambda\}}] + E[Y_\tau 1_{\{Y_n^* \le \lambda\}}] \\ &\ge \lambda P(Y_n^* > \lambda) + E[Y_n 1_{\{Y_n^* \le \lambda\}}], \end{aligned}$$

which implies

$$\lambda P(\max_{t \in D_n} X_t > \lambda) \le E[Y_n] - E[Y_n 1_{\{Y_n^* \le \lambda\}}] = E[Y_n 1_{\{Y_n^* > \lambda\}}] \le E[X_T].$$

(2) Let $E[|X_T|^p] < \infty$. For any $k > 0$,

$$E[(Y_n^* \wedge k)^p] = \int_0^\infty x^p dP(Y_n^* \wedge k \le x) = p \int_0^\infty x^{p-1} P(Y_n^* \wedge k > x) dx.$$

Letting $k \to \infty$, by (1), Fubini's theorem and Hölder's inequality, we have

$$\begin{aligned} E[(Y_n^*)^p] &= p \int_0^\infty x^{p-1} P(Y_n^* > x) dx \le p \int_0^\infty x^{p-2} E[Y_n 1_{\{Y_n^* > x\}}] dx \\ &= pE[Y_n \int_0^{Y_n^*} x^{p-2} dx] = \frac{p}{p-1} E[Y_n (Y_n^*)^{p-1}] \\ &= qE[Y_n (Y_n^*)^{p-1}] \le qE[Y_n^p]^{1/p} E[(Y_n^*)^{(p-1)q}]^{1/q} \\ &= qE[Y_n^p]^{1/p} E[(Y_n^*)^p]^{1/q}, \end{aligned}$$

where $1/p + 1/q = 1$. Put $y = E[(Y_n^*)^p]$. Then y satisfies $y^q \le ay$ with $q > 1$, for some $a > 0$, and thus y is finite. Therefore,

$$E[(Y_n^*)^p]^{1/p} \le qE[Y_n^p]^{1/p}.$$

By Fatou's lemma and Proposition 1.1.19 (viii), we deduce

$$E[\sup_{t \in D} (X_t)^p]^{1/p} \le \liminf_{n \to \infty} E[(Y_n^*)^p]^{1/p} \le \liminf_{n \to \infty} qE[Y_n^p]^{1/p} \le qE[X_T^p]^{1/p}.$$

When $\{X_t\}$ is a martingale, we apply this inequality to a nonnegative submartingale $\{|X_t|\}$ to obtain the result. ∎

Theorem 1.4.4 (Optional sampling theorem)**.** *Let $(X_t, \mathcal{F}_t)$ be a right-continuous martingale (or submartingale) and let τ, σ be two bounded stopping times. Then*

$$E[X_\tau | \mathcal{F}_\sigma] = X_{\tau \wedge \sigma}, \quad P\text{-}a.s. \quad (or\ E[X_\tau | \mathcal{F}_\sigma] \ge X_{\tau \wedge \sigma}, \quad P\text{-}a.s.)$$

Proof. We will show only the case when $\{X_t\}$ is a submartingale.

(1) Suppose that $\sigma \le \tau \le T$ a.s. and σ, τ take values in a finite subset $\delta_n := \{t_i : t_0 = 0 < t_1 < \cdots < t_n, i = 0, 1, \ldots, n\}$ of $[0, T]$. Put $Y_i = X_{t_i}$ for each i. Let $A \in \mathcal{F}_\sigma$ and $A_i = A \cap \{\sigma = t_i\}$. For $j \ge i$, $A_i \cap \{\tau \ge t_{j+1}\} \in \mathcal{F}_{t_j}$. Hence, by the submartingale property,

$$\begin{aligned} E[Y_j 1_{A_i \cap \{\tau \ge t_j\}}] &= E[Y_j 1_{A_i \cap \{\tau = t_j\}}] + E[Y_j 1_{A_i \cap \{\tau \ge t_{j+1}\}}] \\ &\le E[X_\tau 1_{A_i \cap \{\tau = t_j\}}] + E[Y_{j+1} 1_{A_i \cap \{\tau \ge t_{j+1}\}}]. \end{aligned}$$

Taking the summation over j with $t_{n+1} > T$, we have

$$E[Y_i 1_{A_i \cap \{\tau \ge t_i\}}] \le \sum_{j=i}^{n} E[X_\tau 1_{A_i \cap \{\tau = t_j\}}] = E[X_\tau 1_{A_i \cap \{\tau \ge t_i\}}],$$

and then

$$E[X_\sigma 1_{A_i}] = E[X_\sigma 1_{A_i \cap \{\tau \ge \sigma\}}] \le E[X_\tau 1_{A_i \cap \{\tau \ge \sigma\}}] = E[X_\tau 1_{A_i}].$$

Taking the summation over i, we deduce $E[X_\sigma 1_A] \le E[X_\tau 1_A]$, that is, $X_\sigma \le E[X_\tau | \mathcal{F}_\sigma]$ a.s.

(2) Suppose $\sigma \le \tau \le T$ a.s. We take a finite set δ_n for each n such that $t_i = iT/2^n$, $i = 0, 1, \ldots, 2^n$. Define

$$\rho_n = \sum_{i=1}^{2^n} t_i 1_{\{t_{i-1} \le \rho < t_i\}} + T 1_{\{\rho = T\}} \quad \text{for} \quad \rho = \tau \ \text{ or } \ \sigma.$$

Then ρ_n is a stopping time taking values in δ_n, $\rho_n \searrow \rho$, and $\tau_n \ge \sigma_n$. By right continuity, $\lim_{n\to\infty} X_{\tau_n} = X_\tau$ and $\lim_{n\to\infty} X_{\sigma_n} = X_\sigma$ hold a.s. We apply (1) to obtain

$$E[X_{\tau_n} | \mathcal{F}_{\sigma_n}] \ge X_{\sigma_n}, \quad \text{a.s.},$$

and

$$E[X_{\tau_n} 1_A] \ge E[X_{\sigma_n} 1_A] \quad \text{for } A \in \mathcal{F}_\sigma.$$

By (3) below, Proposition 1.1.15 is applicable and so $\lim_{n\to\infty} X_{\rho_n} = X_\rho \in L^1(\Omega)$. Passing to the limit, we deduce $E[X_\tau 1_A] \ge E[X_\sigma 1_A]$.

(3) We claim that $\{X_{\rho_n}\}$ is uniformly integrable. By Proposition 1.1.19 (viii), $\{|X_t|\}$ is a nonnegative submartingale. By Doob's maximal inequality,

$$P(\sup_n |X_{\rho_n}| \ge c) \le P(\sup_{0 \le t \le T} |X_t| \ge c) \le \frac{1}{c} E[|X_T|] \to 0 \quad \text{as } c \to \infty.$$

Thus, by (1)

$$\begin{aligned} \sup_n E[|X_{\rho_n}| 1_{\{|X_{\rho_n}| \ge c\}}] &\le \sup_n E[|X_T| 1_{\{|X_{\rho_n}| \ge c\}}] \\ &\le E[|X_T| 1_{\{\sup_n |X_{\rho_n}| \ge c\}}] \to 0 \quad \text{as } c \to \infty. \end{aligned}$$

(4) In the general case, by (2) and Proposition 1.3.3 (vi), we have

$$\begin{aligned} E[X_\tau | \mathcal{F}_\sigma] 1_{\{\tau \ge \sigma\}} &= E[X_{\tau \vee \sigma} 1_{\{\tau \ge \sigma\}} | \mathcal{F}_\sigma] = E[X_{\tau \vee \sigma} | \mathcal{F}_\sigma] 1_{\{\tau \ge \sigma\}} \\ &\ge X_\sigma 1_{\{\tau \ge \sigma\}} = X_{\tau \wedge \sigma} 1_{\{\tau \ge \sigma\}}. \end{aligned}$$

Furthermore,

$$E[X_\tau | \mathcal{F}_\sigma] 1_{\{\tau < \sigma\}} = E[X_{\tau \wedge \sigma} 1_{\{\tau < \sigma\}} | \mathcal{F}_\sigma] = X_{\tau \wedge \sigma} 1_{\{\tau < \sigma\}}.$$

By addition, we obtain the desired inequality. ∎

Remark 1.4.5. *Let $\{X_t\}$ be a right-continuous submartingale such that $E[\sup_{t\geq 0}|X_t|] < \infty$. Then the optional sampling theorem holds by the uniform integrability of $\{X_{\rho\wedge t}\}$ for $\rho = \tau$ or σ if τ, σ are two stopping times such that $P(\tau, \sigma < \infty) = 1$. See I. Karatzas and S. E. Shreve [87] for weaker conditions and the submartingale convergence theorem.*

1.5 Stochastic Integrals

1.5.1 Definition of the Stochastic Integral

Let $B = \{(B_t, \mathcal{F}_t)\}_{t\geq 0}$ be a one-dimensional Brownian motion on a stochastic basis $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$, that is to say, $\{(B_t, \mathcal{F}_t)\}_{t\geq 0}$ is a continuous martingale with $B_0 = 0$ such that

$$\{(B_t^2 - t, \mathcal{F}_t)\}_{t\geq 0} \text{ is a martingale.}$$

We note by Theorem 1.6.2 below, due to Lévy, that $\{B_t\}_{t\geq 0}$ is a Brownian motion. We will construct the stochastic integral, denoted by $I(f) = \int_0^T f_t dB_t$, for appropriate stochastic processes $\{(f_t, \mathcal{F}_t)\}_{t\geq 0}$ and $T > 0$.

Definition 1.5.1.

(a) *For $p \geq 1$, let $\mathcal{M}_p(0, T)$ be the space of progressively measurable processes $f = \{(f_t, \mathcal{F}_t)\}$ such that*

$$E[\int_0^T |f_t|^p dt] < \infty, \quad \textit{for fixed } T > 0.$$

(b) *Let $\mathcal{M}_0(0, T)$ be the space of step processes $\phi \in \mathcal{M}_2(0, T)$, that is, there is a partition $\{t_0 = 0 < t_1 < t_2 < \cdots < t_n = T\}$ such that*

$$\phi_t = \phi_{(0)} 1_{\{0\}}(t) + \sum_{i=1}^n \phi_{(i)} 1_{(t_{i-1}, t_i]}(t), \quad \textit{for some } \phi_{(i)} \in L^\infty(\Omega, \mathcal{F}_{t_{i-1}}, P).$$

(c) *We define the stochastic integral $I(\phi)$ for $\phi \in \mathcal{M}_0(0, T)$ by*

$$I(\phi) = I_T(\phi) = \sum_{i=1}^n \phi_{(i)}(B_{t_i} - B_{t_{i-1}}).$$

Lemma 1.5.2. *The space $\mathcal{M}_0(0, T)$ is dense in $\mathcal{M}_2(0, T)$.*

Proof. Let $f = \{f_t\} \in \mathcal{M}_2(0, T)$, and we will show that there exists a sequence $\phi^{(n)} \in \mathcal{M}_0(0, T)$ such that

$$E[\int_0^T |f_t - \phi_t^{(n)}|^2 dt] \to 0 \quad \text{as } n \to \infty.$$

(1) Suppose that $\{f_t\}$ is continuous and bounded. Define

$$\phi_t^{(n)} = f_{t_0} 1_{\{t=0\}} + \sum_{k=1}^{n} f_{t_{k-1}} 1_{(t_{k-1},t_k]}(t), \qquad \text{for } t_k = kT/n,\ k = 0, 1, \ldots, n.$$

Then $\{\phi^{(n)}\} \in \mathcal{M}_0(0, T)$ and $\phi_t^{(n)} \to f_t$ for each t a.s. By the dominated convergence theorem, we have $\lim_{n\to\infty} E[\int_0^T |f_t - \phi_t^{(n)}|^2 dt] = 0$.

(2) Let $\{f_t\}$ be progressively measurable and bounded. Put $h(t, \omega) = \int_0^t f(s, \omega)ds$. By progressive measurability, $\{h_t\}$ is a measurable process, and the random variable $h(t, \omega)$ is $\mathcal{F}_t$-measurable for each $t \in [0, T]$. Define

$$h^\varepsilon(t, \omega) = \frac{1}{\varepsilon} \int_{(t-\varepsilon)^+}^{t} f(s, \omega)ds = \frac{h(t, \omega) - h((t-\varepsilon)^+, \omega)}{\varepsilon}, \qquad \text{for } \varepsilon > 0.$$

By Lebesgue's Theorem, there exists a derivative $h'(t, \omega)$ for almost all $t \in [0, T]$ a.s., and

$$\lim_{\varepsilon\to 0} |h'(t, \omega) - h^\varepsilon(t, \omega)| = \lim_{\varepsilon\to 0} |f(t, \omega) - h^\varepsilon(t, \omega)| = 0, \quad \text{a.s.}$$

Hence

$$\lim_{\varepsilon\to 0} \int_0^T |h^\varepsilon(t, \omega) - f(t, \omega)|^2 dt = 0, \quad \text{a.s.},$$

and

$$\lim_{\varepsilon\to 0} E[\int_0^T |h^\varepsilon(t, \omega) - f(t, \omega)|^2 dt] = 0.$$

(3) In the general case, we define the bounded progressively measurable process $f^{(m)} \in \mathcal{M}_2(0, T)$ by

$$f_t^{(m)} = f_t 1_{\{|f_t|\le m\}}, \qquad \text{for each } m \in \mathbf{N}.$$

By the dominated convergence theorem, we have

$$E[\int_0^T |f_t - f_t^{(m)}|^2 dt] \to 0 \quad \text{as } m \to \infty.$$

Applying (2) to obtain a sequence of continuous bounded processes $h^{(m,\varepsilon)}$ approximating $f^{(m)}$, and then (1) to $h^{(m,\varepsilon)}$, we obtain the assertion. ■

Remark 1.5.3. *Lebesgue's Theorem states that if $f \in L^1[a, b]$, then*

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)dt = f(x) \quad \textit{for almost all } x \in [a, b].$$

For the proof, see H. L. Royden [139, chapter 5, Theorems 2 and 9].

Theorem 1.5.4. *Let $f \in \mathcal{M}_2(0, T)$ and let $\phi^{(m)} \in \mathcal{M}_0(0, T)$, $m = 1, 2, \ldots,$ be such that*

$$E[\int_0^T |f_t - \phi_t^{(m)}|^2 dt] \to 0 \quad \textit{as } m \to \infty.$$

Then $I(\phi^{(m)})$ converges to some $I(f)$ in $L^2(\Omega, \mathcal{F}_T, P)$, independent of a choice of $\phi^{(m)}$.

Proof.

(1) We claim that for $\phi \in \mathcal{M}_0(0, T)$,

$$E[|I(\phi)|^2] = E[\int_0^T |\phi_t|^2 dt].$$

For brevity, we set $\Delta_i B = B_{t_i} - B_{t_{i-1}}$ and $\Delta_i t = t_i - t_{i-1}$. Then

$$|I(\phi)|^2 = \sum_{i=0}^n \phi_{(i)}^2 (\Delta_i B)^2 + 2 \sum_{i<j} \phi_{(i)} \phi_{(j)} (\Delta_i B)(\Delta_j B).$$

In view of Proposition 1.4.2

$$E[\phi_{(i)}^2 (\Delta_i B)^2 | \mathcal{F}_{t_{i-1}}] = \phi_{(i)}^2 E[(\Delta_i B)^2 | \mathcal{F}_{t_{i-1}}] = \phi_{(i)}^2 \ \Delta_i t, \quad \text{a.s.,}$$

and

$$E[\phi_{(i)} \phi_{(j)} (\Delta_i B)(\Delta_j B) | \mathcal{F}_{t_{j-1}}] = \phi_{(i)} \phi_{(j)} (\Delta_i B) E[\Delta_j B | \mathcal{F}_{t_{j-1}}] = 0, \quad \text{a.s.,} \quad i < j.$$

Therefore,

$$E[|I(\phi)|^2] = \sum_{i=1}^n E[\phi_{(i)}^2] \ \Delta_j t.$$

On the other hand,

$$|\phi_t|^2 = \phi_{(0)}^2 1_{\{0\}}(t) + \sum_{i=1}^n \phi_{(i)}^2 1_{(t_{i-1}, t_i]}(t).$$

This implies that

$$E[\int_0^T |\phi_t|^2 dt] = E[\sum_{i=1}^n \phi_{(i)}^2 \ \Delta_i t] = E[|I(\phi)|^2],$$

as required.

(2) Let $f \in \mathcal{M}_2(0, T)$. It is clear that

$$E[\int_0^T |\phi_t^{(m)} - \phi_t^{(n)}|^2 dt]^{1/2} \leq E[\int_0^T |\phi_t^{(m)} - f_t|^2 dt]^{1/2} + E[\int_0^T |f_t - \phi_t^{(n)}|^2 dt]^{1/2} \to 0 \quad \text{as } n, m \to \infty.$$

Hence, by (1)

$$E[|I(\phi^{(m)}) - I(\phi^{(n)})|^2] = E[\int_0^T |\phi_t^{(m)} - \phi_t^{(n)}|^2 dt] \to 0 \quad \text{as } n, m \to \infty.$$