Applications of
Mathematics

21

Philip Protter

Stochastic Integration and Differential Equations

A New Approach

随机积分和微分方程 [英]

Springer-Verlag
World Publishing Corp

Philip Protter

Stochastic Integration and Differential Equations

A New Approach

Springer-Verlag
World Publishing Corp

Philip Protter
Mathematics and Statistics Departments
Purdue University
West Lafayette, IN 47907
USA

Managing Editor

A. V. Balakrishnan
Systems Science Department
University of California
Los Angeles, CA 90024
USA

Mathematics Subject Classification (1980): 60H05, 60H20, 60G44

ISBN 3-540-50996-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-50996-8 Springer-Verlag New York Berlin Heidelberg
ISBN 7 − 5062 − 1100 − 9

Reprinted by World Publishing Corporation, Beijing, 1992
for distribution and sale in The People's Republic of China only

Preface

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 [2] provided an outline for just such a pedagogic approach. I developed this into a series of lectures (Protter [6]), using the work of K. Bichteler [2], E. Lenglart [3] and P. Protter [7], 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 [1], Dellacherie-Meyer [1, 2], or Ethier-Kurtz [1]). 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 [1], Kopp [1], Métivier [1], Rogers-Williams [1] and to a much lesser extent Letta [1] are examples. The books of McKean [1], Chung-Williams [1], and Karatzas-Shreve [1] 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. 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 me with 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. We reintroduce Meyer's original notion of a process being natural, allowing for less abstract and more intuitive proofs. However in what is essentially an optional last section (Sect. 8) we give a simple proof that a process with paths of integrable variation is natural if and only if it is predictable, since natural processes are referred to as predictable in the literature.

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. These more general integrands allow us to give a presentation of the theory 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. Throughout Chap. V we have tried to achieve a balance between maximum generality and the simplicity of the proofs.

CHAPTER 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* $(\mathcal{F}_t)_{0 \le t \le \infty}$. By a filtration we mean a family of σ-algebras $(\mathcal{F}_t)_{0 \le t \le \infty}$ that is increasing: $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \le t$.

Definition. A filtered complete probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{0 \le t \le \infty})$ is said to satisfy the **usual hypotheses** if

(i) $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$;
(ii) $\mathcal{F}_t = \cap_{u>t} \mathcal{F}_u$, all t, $0 \le t < \infty$; that is, the filtration $(\mathcal{F}_t)_{0 \le t \le \infty}$ 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 $\{T \le t\} \in \mathcal{F}_t$, every t, $0 \le t \le \infty$.

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

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

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

A **stochastic process** X on $(\Omega, \mathcal{F}, P)$ is a collection of random variables $(X_t)_{0 \le t < \infty}$. 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. (The nonsensical word *càdlàg* is an acronym from the French "continu à droite, limites à gauche".)

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. □

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 $\{T < t\} \in \mathcal{F}_t$, $0 \leq t < \infty$. 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. □

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

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

is a stopping time.

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

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

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

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

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

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

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

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

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

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,*

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

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

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 $\varphi : \{T \leq t\} \mapsto [0,\infty) \times \Omega$ 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 $(\{T \leq t\}, \mathcal{F}_t \cap \{T \leq t\})$ into $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ are the Borel sets of $\mathbb{R}$. Therefore

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

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 $S \leq T$ 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.s. ω, $t \to \Delta X_t = 0$ 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 $\{t : |\Delta X_t| > \frac{1}{n}\}$ must be finite for each n, since $t_0 < \infty$. Using Theorem 4 we define stopping times for each n inductively as follows:

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

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

---

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

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 $\Delta X_T 1_{\{T<\infty\}} = \Delta Y_T 1_{\{T<\infty\}}$ a.s. for each stopping time T, 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 [1] or Dellacherie-Meyer [1].

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 $\mathcal{H}$)
(iii) If $(f_n)_{n\geq 1} \subset \mathcal{H}$ and $0 \leq f_1 \leq f_2 \leq \cdots \leq f_n \leq \ldots$ 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 [1, p. 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 [1, p. 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 [2], or Ethier-Kurtz [1]. Also, recall that we will always assume as given a filtered, complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le\infty}, P)$, where the filtration $(\mathcal{F}_t)_{0\le t\le\infty}$ is assumed to be right continuous.

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

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

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

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

A random variable Y closing a martingale is not necessarily unique. We give a 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 $t \to E\{X_t\}$ is right continuous if and only if there exists a unique 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 $t \to E\{X_t\}$ is constant, and hence it has a right continuous modification.

Corollary 1. *If $X = (X_t)_{0\le t<\infty}$ is a martingale then there exists a unique modification Y of X which is càdlàg.*

Since all martingales have right continuous modifications, *we will always assume that we are taking the right continuous version*, without any special mention. Note that it follows from Corollary 1 that a right continuous martingale is càdlàg.

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

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

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

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

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

(i) $(U_\alpha)_{\alpha\in A}$ is uniformly integrable.
(ii) $\sup_{\alpha\in A} E\{|U_\alpha|\} < \infty$, and whatever $\epsilon > 0$ there exists $\delta > 0$ such that $\Lambda \in \mathcal{F}$, $P(\Lambda) \le \delta$, imply $E\{|U_\alpha 1_\Lambda|\} < \epsilon$.
(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)⇒(ii) and (iii)⇒(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 $(X_t)_{t\ge 0}$ is uniformly integrable if and only if $Y = \lim_{t\to\infty} X_t$ exists a.s., $E\{|Y|\} < \infty$, and $(X_t)_{0\le t\le\infty}$ 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 $(X_n)_{n\le 0}$ refers to a process indexed by the nonpositive integers: $\cdots X_{-2}, X_{-1}, X_0$.

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

Theorem 14 (Backwards Convergence). *Let $(X_n)_{n\le 0}$ 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 $\|{}^{\pi_t}Y\|_{L^2} \le \|Y\|_{L^2}$, by part (iii) of Theorem 11 we have that X is uniformly integrable (take $G(x) = x^2$). □

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

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 $S \le T$ a.s. Then X_S and X_T are integrable and*

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

Theorem 16 has a similar version for supermartingales, but we will not have need of it. See Dellacherie-Meyer [2].

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

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

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

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

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