## Contents

1 Preliminaries

1.1 Functional Analysis in Hilbert Spaces ...................................................... 3
1.2 Measure Theory .............................................................................................. 5
  1.2.1 $\pi$-Classes, Algebras and $\sigma$-Algebras of Sets .................................. 5
  1.2.2 Measures .................................................................................................. 7
  1.2.3 Measurable Functions and Convergence ................................................... 8
  1.2.4 Integration Theory .................................................................................... 10
  1.2.5 Lebesgue Spaces ....................................................................................... 14
  1.2.6 The Radon-Nikodym Theorem .................................................................. 16
  1.2.7 The Fubini-Tonelli Theorem .................................................................... 18
1.3 Product $\sigma$-Algebras on Function Spaces .................................................. 21
  1.3.1 $\sigma$-Algebras Generated by Functions ................................................... 21
  1.3.2 $\sigma$-algebras on Function Spaces ........................................................... 24
1.4 Probability Theory .......................................................................................... 30
  1.4.1 Probability Spaces, Random Variables, and Expectation .......................... 30
  1.4.2 Independence ............................................................................................ 33
  1.4.3 Characteristic Functions .......................................................................... 34
  1.4.4 The Gaussian Distribution ....................................................................... 36
  1.4.5 Conditional Expectation .......................................................................... 37
1.5 Monotone Class Theorems .............................................................................. 42
1.6 Problems ......................................................................................................... 49

2 Discrete-Parameter Martingales

2.1 Filtrations and Stopping Times ....................................................................... 54
2.2 Martingales; Basic Definitions and Properties ............................................... 59
2.3 The Optional Sampling Theorem ..................................................................... 61
2.4 The Basic Martingale Inequalities .................................................................. 68
2.5 The Martingale Convergence Theorem .......................................................... 71
2.6 Uniform Integrability ....................................................................................... 79
2.7 Uniformly Integrable Supermartingales .............................................. 84
2.8 Problems ......................................................................................... 90

3 Elements of Continuous-Parameter Stochastic Processes 94
3.1 Basic Definitions ............................................................................ 94
3.2 The Wiener Process ....................................................................... 103
3.3 Continuous-Parameter Stopping Times ......................................... 108
3.4 Fine-Tuning a Continuous-Parameter Raw Filtration...................... 115
3.5 Problems ......................................................................................... 124

4 Continuous Parameter-Martingales 128
4.1 Basic Definitions ............................................................................ 128
4.2 Sample Path Properties .................................................................. 132
4.3 Martingale Inequalities .................................................................. 142
4.4 Martingale Convergence Theorems ............................................. 145
4.5 The Optional Sampling Theorem .................................................. 150
4.6 Local Martingales ........................................................................... 153
4.7 The Quadratic Variation Process .................................................. 157
4.8 Problems ......................................................................................... 177

5 Stochastic Integration 183
5.1 Lebesgue-Stieltjes Integration ...................................................... 183
5.2 Random Lebesgue-Stieltjes Integration ....................................... 196
5.3 Stochastic Integration .................................................................... 201
5.4 Itô’s Change-of-Variables Formula .............................................. 227
5.5 Exponential Semimartingales ....................................................... 238
5.6 Levy’s Characterization of a Wiener Process .............................. 246
5.7 Changes of Measure and the Girsanov Transformation .............. 248
5.8 Problems ......................................................................................... 261

6 Representation of Martingales as Stochastic Integrals 269
6.1 Wiener Filtrations and Stochastic Integrals .................................. 270

A Glossary of Notation 282
A.1 Spaces of Stochastic Processes .................................................... 282
A.2 Miscellaneous Symbols ................................................................. 283
Chapter 1

Preliminaries

In this chapter we sketch the basic mathematical essentials with which readers should be familiar in order to use these notes. These essentials are mostly quite modest, comprising only the simplest aspects of functional analysis in Hilbert spaces, measure theory, and probability, and are summarized in Sections 1.1, 1.2, and 1.4. These summaries are not a thorough introduction to the afore-mentioned topics, our purpose being merely to recall without proofs some of the main concepts and results for later reference, and to bring the otherwise murky issue of “prerequisites” into a sharper focus. Two other sections in this chapter deserve comment, namely Section 1.3 on arbitrary products of measurable spaces, and Section 1.5 on monotone class theorems. These deal with measure-theoretic issues which are nevertheless often omitted from standard introductions to measure theory, and with which we do not assume familiarity on the part of readers. In view of their importance in later chapters we have given a rather detailed and self-contained introduction to these ideas, in contrast to the rather brief summaries elsewhere in the chapter.

For well prepared readers who are thoroughly familiar with the ideas of the present chapter, it is possible and probably advisable to begin at Chapter 2, returning to the present chapter only for occasional reference. On the other hand, readers who perceive deficiencies in their mathematical equipment in the course of perusing this chapter will hopefully gain a clearer idea on how to remedy their background knowledge. More extensive background, at a level which is ample for all of our requirements, can be found in (Ash [1], Chaps. 1 - 6), (Chow and Teicher [3], Chaps. 1 - 7), and (Williams [31], Chaps. 1 - 9).

1.1 Functional Analysis in Hilbert Spaces

For completeness we recall the notion of a real vector space: A non-empty set $L$ is a real vector space when the following axioms hold:

(1) For each pair of elements $x, y \in L$ there is a unique element in $L$, denoted by $x + y$, called the vector sum of $x$ and $y$, such that: (i) $x + y = y + x, \forall x, y \in L$; (ii) $(x + y) + z = x + (y + z), \forall x, y, z \in L$;
(iii) there is a unique element \( \theta \in \mathbb{L} \), called the zero element, such that \( x + \theta = \theta + x = x \), \( \forall x \in \mathbb{L} \).

(2) Corresponding to any \( \alpha \in \mathbb{IR} \) and \( x \in \mathbb{L} \) is a unique element \( \alpha x \in \mathbb{L} \), called the scalar product of \( \alpha \) and \( x \), such that: (iv) \( \alpha(\beta x) = (\alpha\beta)x \), \( \forall \alpha, \beta \in \mathbb{IR} \), \( \forall x \in \mathbb{L} \); (v) \( 0x = \theta \) and \( 1x = x \), \( \forall x \in \mathbb{L} \); (vi) \( (\alpha + \beta)x = \alpha x + \beta x \) and \( \alpha(x + y) = \alpha x + \alpha y \), \( \forall \alpha, \beta \in \mathbb{IR} \), \( \forall x, y \in \mathbb{L} \).

In a real vector space \( \mathbb{L} \), the element \( (-1)x \), \( x \in \mathbb{L} \), is called the negative of \( x \) and is denoted by \( -x \). Clearly, \( x + (-1)x = 1x + (-1)x = (1 - 1)x = \theta \), thus \( x + (-x) = \theta \).

A subset \( \mathbb{M} \) of a real vector space \( \mathbb{L} \) is called a vector subspace when \( \alpha x + \beta y \in \mathbb{M} \), \( \forall x, y \in \mathbb{M} \), \( \forall \alpha, \beta \in \mathbb{IR} \).

Given a real vector space \( \mathbb{L} \) with zero element \( \theta \), a mapping \( x \to \|x\| : \mathbb{L} \to [0, \infty) \), called a norm on \( \mathbb{L} \) when the following conditions hold: (a) \( \|\theta\| = 0 \); (b) \( \|\alpha x\| = |\alpha|\|x\| \), \( \forall \alpha \in \mathbb{IR} \), \( \forall x \in \mathbb{L} \); (c) \( \|x + y\| \leq \|x\| + \|y\| \), \( \forall x, y \in \mathbb{L} \); (d) If \( \|x\| = 0 \) then \( x = \theta \). A mapping \( x \to \|x\| : \mathbb{L} \to [0, \infty) \) for which (a), (b) and (c) hold, but (d) is not necessarily true, is called a pseudo-norm on \( \mathbb{L} \). A real vector space on which a norm is defined is called a normed vector space.

In the context of a normed vector space \( \mathbb{L} \), a sequence \( \{x_n, \; n = 1, 2, \ldots\} \) in \( \mathbb{L} \) is a Cauchy sequence when

\[
\lim_{n \to \infty} \sup_{m \geq n} \|x_m - x_n\| = 0.
\]

A normed vector space \( \mathbb{L} \) is called a Banach space when each Cauchy sequence in \( \mathbb{L} \) converges to some member of \( \mathbb{L} \). One can obtain a rather special class of Banach spaces, called Hilbert spaces, when the norm is defined by means of a postulated inner product: An inner product on a real vector space \( \mathbb{L} \) is a real-valued function \( (x, y) \), defined for each pair \( x, y \in \mathbb{L} \) such that: (a) \( (x, y) = (y, x) \) and \( (\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) \) \( \forall x, y, z \in \mathbb{L} \), \( \forall \alpha, \beta \in \mathbb{IR} \); (b) \( (x, x) \geq 0 \), \( \forall x \in \mathbb{L} \) and \( (x, x) = 0 \) if and only if \( x = \theta \).

A real vector space on which an inner product is defined is called an inner product space. We can define a norm on an inner product space \( \mathbb{L} \) by \( \|x\| \overset{\triangle}{=} (x, x)^{1/2} \), \( \forall x \in \mathbb{L} \), and we call the inner product space a Hilbert space when the resulting normed vector space is a Banach space. Two elements \( x \) and \( y \) of a Hilbert space are called orthogonal when \( (x, y) = 0 \), and we write \( x \perp y \) to denote orthogonality of \( x \) and \( y \). Also, if \( A \) is some subset of a Hilbert space \( \mathbb{L} \) and \( x \in \mathbb{L} \) then we write \( x \perp A \) to denote \( x \perp y \), \( \forall y \in A \).

We recall two fundamental theorems about Hilbert spaces for later use:

**Theorem 1.1.1 (Frechet)** Suppose \( \mathbb{M} \) is a closed vector subspace of a Hilbert space \( \mathbb{L} \). Then each \( x \in \mathbb{L} \) has the unique representation \( x = y + z \) where \( y \in \mathbb{M} \), \( z \perp \mathbb{M} \), and \( \|x - y\| = \inf_{v \in \mathbb{M}} \|x - v\| \).

We call the element \( y \) in this theorem the orthogonal projection of \( x \) onto \( \mathbb{M} \). Next, we recall the notion of a bounded linear functional: A mapping \( \phi : \mathbb{L} \to \mathbb{IR} \) is called a linear functional on a normed vector space \( \mathbb{L} \) when \( \phi(\alpha x + \beta y) = \alpha \phi(x) + \beta \phi(y) \), \( \forall \alpha, \beta \in \mathbb{IR} \), \( \forall x, y \in \mathbb{L} \). A linear functional \( \phi : \mathbb{L} \to \mathbb{IR} \) is said to be a bounded linear functional on \( \mathbb{L} \) when there is some constant \( B \in [0, \infty) \) such that \( |\phi(x)| \leq B\|x\| \), \( \forall x \in \mathbb{L} \).
Theorem 1.1.2 (Riesz) Suppose $\phi$ is a bounded linear functional on a Hilbert space $L$. Then there is a unique element $y \in L$ such that $\phi(x) = (x, y)$, $\forall x \in L$.

1.2 Measure Theory

1.2.1 $\pi$-Classes, Algebras and $\sigma$-Algebras of Sets

Given a set $E$, we call a designated collection $C$ of subsets of $E$ a $\pi$-class over $E$ if $A \cap B \in C$ whenever $A$ and $B$ are members of $C$. Specializing this concept, we say that a designated collection $C$ of subsets of $E$ is an algebra over $E$ when $C$ is a $\pi$-class over $E$ with the additional properties that $E \in C$, and $A^c \in C$ whenever $A \in C$. Specializing still further, we say that a designated collection $C$ of subsets of $E$ is a $\sigma$-algebra over $E$ when it is an algebra over $E$ and has the following additional property: $\bigcap_{1 \leq n < \infty} A_n \in C$ whenever $\{A_n, n = 1, 2, \ldots\}$ is a sequence of members of $C$. In these definitions the set $E$ is usually called the universal set of the $\pi$-class, algebra, or $\sigma$-algebra in question. We normally use the generic symbol $\mathcal{S}$ to denote a $\sigma$-algebra, and a pair $(E, \mathcal{S})$ comprising a set $E$ and a $\sigma$-algebra $\mathcal{S}$ over $E$, is called a measurable space. If $H$ and $G$ are given collections of subsets of $E$, then we write $H \subset G$ to indicate that $A \in H$ implies $A \in G$. In the special case where $H$ and $G$ are $\sigma$-algebras over $E$, we say that $H$ is a sub $\sigma$-algebra of $G$ when $H \subset G$.

Remark 1.2.1 The following observations follow easily from the definition of a $\sigma$-algebra:

(a) Suppose that $E$ is a given set. Then the collection $\{\emptyset, E\}$ is obviously a $\sigma$-algebra over $E$, as is the collection of all possible subsets of $E$, which is usually denoted by $\mathcal{P}(E)$ and called the power set of $E$. Clearly $\{\emptyset\}$ and $\mathcal{P}(E)$ are, respectively, the smallest and largest possible $\sigma$-algebras which one can define over the set $E$.

(b) Suppose that $\mathfrak{C}$ is a given collection of $\sigma$-algebras over the set $E$, that is, each $\mathcal{S} \in \mathfrak{C}$ is itself a $\sigma$-algebra over $E$, and let $\bigcap \{\mathcal{S} : \mathcal{S} \in \mathfrak{C}\}$ denote the collection of those subsets of $E$ which are members of each and every $\mathcal{S} \in \mathfrak{C}$, namely

$$\bigcap \{\mathcal{S} : \mathcal{S} \in \mathfrak{C}\} \triangleq \{A \subset E : A \in \mathcal{S} \text{ for each } \mathcal{S} \in \mathfrak{C}\}.$$ 

One trivially verifies that $\bigcap \{\mathcal{S} : \mathcal{S} \in \mathfrak{C}\}$ is itself a $\sigma$-algebra over $E$.

(c) The truly interesting $\sigma$-algebras over the set $E$ occur somewhere between the extreme cases $\{\emptyset, E\}$ and $\mathcal{P}(E)$, and are typically constructed in the following way: suppose that $\mathcal{C}$ is a given collection of subsets of the set $E$, and define

$$\mathfrak{C} \triangleq \{\mathcal{S} \subset \mathcal{P}(E) : \mathcal{S} \text{ is a } \sigma\text{-algebra over } E, \text{ and } \mathcal{C} \subset \mathcal{S}\}.$$ 

Clearly $\mathfrak{C}$ is a non-empty collection of $\sigma$-algebras over $E$, since an obvious member of $\mathfrak{C}$ is the power set $\mathcal{P}(E)$. One trivially observes that (i) $\mathcal{C} \subset \bigcap \{\mathcal{S} : \mathcal{S} \in \mathfrak{C}\}$, and (ii) if $\mathcal{O}$ is any $\sigma$-algebra over $E$ such
that \( \mathcal{C} \subset \mathcal{O} \), then \( \mathcal{O} \in \mathcal{C} \), and hence \( \bigcap \{ S : S \in \mathcal{C} \} \subset \mathcal{O} \). It follows that \( \bigcap \{ S : S \in \mathcal{C} \} \) is the \textit{smallest possible} \( \sigma \)-algebra over \( E \) which includes every set in \( \mathcal{C} \) among its members. We usually denote this \( \sigma \)-algebra by the simpler notation \( \sigma \{ \mathcal{C} \} \), and call it the \( \sigma \)-algebra \textit{generated by} the given collection of sets \( \mathcal{C} \).

\textbf{Example 1.2.2} \hspace{1em} (a) Suppose \( E \triangleq \mathbb{R} \equiv [-\infty, \infty] \) and let \( \mathcal{C} \subset \mathcal{P}(E) \) comprise the sets \(-\infty, \{ \infty \}\) and \((-\infty, a], \forall \ a \in \mathbb{R} \). Then \( \sigma \{ \mathcal{C} \} \) is called the \textbf{Borel \( \sigma \)-algebra} over \( \mathbb{R} \), and is denoted by \( \mathcal{B}(\mathbb{R}) \).

For any \( I \in \mathcal{B}(\mathbb{R}) \) we put \( \mathcal{B}(I) \equiv \{ A \cap I : A \in \mathcal{B}(\mathbb{R}) \} \). Thus \( (I, \mathcal{B}(I)) \) is a measurable space and \( \mathcal{B}(I) \) is called the Borel \( \sigma \)-algebra over \( I \). The particular case of \( \mathcal{B}(\mathbb{R}) \), resulting from \( I \triangleq \mathbb{R} \), is called the \textbf{Borel \( \sigma \)-algebra over the real line}. \( \mathcal{B}(\mathbb{R}) \) can also be synthesized in several other ways. To mention just one, if we let \( \mathcal{C} \) denote the collection of all open subsets in \( \mathbb{R} \), then it is easily checked that \( \mathcal{B}(\mathbb{R}) = \sigma \{ \mathcal{C} \} \).

(b) More generally, suppose \( E \triangleq \mathbb{R}^d \) for some integer \( d \geq 1 \), and let \( \mathcal{C} \) be the collection of all open sets in \( \mathbb{R}^d \). The \( \sigma \)-algebra \( \sigma \{ \mathcal{C} \} \) is called the \textbf{Borel \( \sigma \)-algebra} over \( \mathbb{R}^d \) and is denoted by \( \mathcal{B}(\mathbb{R}^d) \). There are several other ways of arriving at the Borel \( \sigma \)-algebra over \( \mathbb{R}^d \). For example, for each \( x = (x^1, x^2, \ldots, x^d) \) in \( \mathbb{R}^d \), let \((-\infty, x] \) denote the semi-infinite cell given by \( \{(\xi^1, \xi^2, \ldots, \xi^d) \in \mathbb{R}^d : \xi^i \leq x^i \ \forall \ i = 1, 2, \ldots, d\} \).

Put \( \mathcal{D} \equiv \{(-\infty, x] : x \in \mathbb{R}^d \} \). It is easy to verify that \( \mathcal{B}(\mathbb{R}^d) = \sigma \{ \mathcal{D} \} \).

Suppose that one is given sets \( A_n \subset E \) for each \( n = 1, 2, \ldots \). The \textbf{limit superior} of this sequence of sets, denoted by \( \limsup_{n \to \infty} A_n \), is defined to be the set of all points \( x \in E \) which belong to \( A_n \) for \textit{infinitely many} values of \( n \). Thus \( \limsup_{n \to \infty} A_n \) is the set of all \( x \in E \) with the property that, for each positive integer \( N \), there is some integer \( n(x) \geq N \) such that \( x \in A_{n(x)} \). On the other hand, the \textbf{limit inferior} of the sequence of sets \( \{A_n, \ n = 1, 2, \ldots\} \), which is denoted by \( \liminf A_n \), is defined to be the set of all points \( x \in E \) which belong to \( A_n \) for all except finitely many values of \( n \). Thus \( \liminf A_n \) is the set of all \( x \in E \) with the property that there is some positive integer \( n(x) \) such that \( x \in A_n \) for all \( n \geq n(x) \). One easily sees that

\begin{equation}
\limsup_{n \to \infty} A_n = \bigcap_{1 \leq m \leq \infty} \left( \bigcup_{m \leq n < \infty} A_n \right), \quad \liminf_{n \to \infty} A_n = \bigcup_{1 \leq m \leq \infty} \left( \bigcap_{m \leq n < \infty} A_n \right),
\end{equation}

and

\begin{equation}
\liminf_{n \to \infty} A_n \subset \limsup_{n \to \infty} A_n.
\end{equation}

We say that the sequence of sets \( \{A_n, \ n = 1, 2, \ldots\} \) \textbf{converges to a limit} when set equality holds in (1.2), denote the common value of the limit inferior and limit superior in (1.2) by \( \lim_{n \to \infty} A_n \), and call this set the \textbf{limit} of the sequence. A sequence of sets \( \{A_n, \ n = 1, 2, \ldots\} \) is called \textbf{increasing} when \( A_n \subset A_{n+1} \), \( \forall n = 1, 2, \ldots \), and \textbf{decreasing} when \( A_{n+1} \subset A_n \), \( \forall n = 1, 2, \ldots \). Sequences of sets which are either increasing or decreasing will be called \textbf{monotone sequences} of sets. It is easy to verify that, if \( \{A_n, \ n = 1, 2, \ldots\} \) is a monotone sequence of sets then it converges to a limit which is given by

\[ \bigcup_{1 \leq n < \infty} A_n \quad \text{or} \quad \bigcap_{1 \leq n < \infty} A_n. \]
depending on whether the sequence is increasing or decreasing. Finally, it is clear that if \( \mathcal{S} \) is a given \( \sigma \)-algebra over \( E \) and \( A_n \in \mathcal{S} \) for all \( n = 1, 2, \ldots \) then the limit superior, limit inferior and limit (when the latter exists) of the sequence are also members of \( \mathcal{S} \).

### 1.2.2 Measures

Suppose that \( \mathcal{A} \) is an algebra over a set \( E \). Then a mapping \( \mu : \mathcal{A} \to [0, \infty] \) is called a measure on \( \mathcal{A} \) when \( \mu(\emptyset) = 0 \) and

\[
\mu\left( \bigcup_{1 \leq n < \infty} A_n \right) = \sum_{1 \leq n < \infty} \mu(A_n)
\]

for each sequence of sets \( \{A_n, \ n = 1, 2, \ldots\} \) in \( \mathcal{A} \) such that \( A_m \cap A_n = \emptyset \) whenever \( m \neq n \), and \( \bigcup_{1 \leq n < \infty} A_n \in \mathcal{A} \). A measure \( \mu \) on \( \mathcal{A} \) is called a finite measure when \( \mu(E) < \infty \), and, more generally, is said to be a \( \sigma \)-finite measure when there is some sequence \( \{E_n, \ n = 1, 2, \ldots\} \) of sets in \( \mathcal{A} \) such that \( E = \bigcup_{1 \leq n < \infty} E_n \) and \( \mu(E_n) < \infty \), \( \forall \ n = 1, 2, \ldots \) The great majority of measures that one comes across in practice are either finite or \( \sigma \)-finite measures. We are usually interested in measures which are defined on \( \sigma \)-algebras, rather than just algebras. One obtains such measures by starting with a measure on an algebra and then using the following extension theorem:

**Theorem 1.2.3 (Caratheodory-Hahn)** Suppose that \( \mathcal{A} \) is an algebra over a set \( E \) and \( \mu \) is a \( \sigma \)-finite measure on \( \mathcal{A} \). Then there exists a unique \( \sigma \)-finite measure \( \bar{\mu} : \sigma(\mathcal{A}) \to [0, \infty] \) such that \( \bar{\mu}(A) = \mu(A), \forall A \in \mathcal{A} \).

Uniqueness in Theorem 1.2.3 is closely related to the following result:

**Theorem 1.2.4** Suppose that \( \mathcal{C} \) is a \( \pi \)-class over a set \( E \) and put \( \mathcal{S} \doteq \sigma(\mathcal{C}) \). If \( \mu_1 \) and \( \mu_2 \) are finite measures on \( \mathcal{S} \), such that \( \mu_1(E) = \mu_2(E) \) and \( \mu_1(A) = \mu_2(A), \forall A \in \mathcal{C} \), then \( \mu_1(A) = \mu_2(A), \forall A \in \mathcal{S} \).

By far the most important measure arising in applications is Lebesgue’s measure on Euclidean space. Let \( \mathcal{C} \) denote the collection of all subsets of \( \mathbb{R} \) having any of the possible forms \( (-\infty, a], (b, \infty) \) or \( (a, b] \), \( a, b \in \mathbb{R} \). One associates a length \( b - a \) with any set \( (a, b] \), and a length of \( +\infty \) with sets of the form \( (-\infty, a] \) or \( (b, \infty) \). Now, let \( \mathcal{A} \) be the collection of all possible finite unions of disjoint sets taken from \( \mathcal{C} \). It is easy to check that \( \mathcal{A} \) is an algebra over \( \mathbb{R} \) and \( \sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}) \). If a set \( A \in \mathcal{A} \) has the form \( A = \bigcup_{i=1}^{n} C_i \) for some disjoint sets \( C_i \in \mathcal{C} \), then we assign a length or **Lebesgue measure** to \( A \) by

\[
\lambda(A) \doteq \sum_{i=1}^{n} \text{length of } C_i.
\]

In this way we get a mapping \( \lambda : \mathcal{A} \to [0, \infty] \). An application of the Heine-Borel covering theorem shows that \( \lambda \) is a measure on \( \mathcal{A} \), and it is of course obvious that this measure is \( \sigma \)-finite. Thus Theorem 1.2.3 ensures there is a unique measure \( \overline{\lambda} \) which extends Lebesgue measure from \( \mathcal{A} \) to \( \mathcal{B}(\mathbb{R}) \). In this way one gets Lebesgue measure on the Borel \( \sigma \)-algebra over \( \mathbb{R} \). Usually we omit the overbar, and also
denote the extension of $\lambda$ from $\mathcal{A}$ to $\mathcal{B}(\mathbb{R})$ by $\lambda$. An exactly analogous construction can be used to get Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$, which we shall again denote by $\lambda$.

A triple $(E, \mathcal{S}, \mu)$ comprising a set $E$, a $\sigma$-algebra $\mathcal{S}$ over $E$, and a measure $\mu$ on $\mathcal{S}$ is called a **measure space**. This measure space is called $\sigma$-finite or finite, according as the measure $\mu$ on $\mathcal{S}$ is $\sigma$-finite or finite. The next result lists the most basic properties of measure spaces:

**Theorem 1.2.5** Suppose that $(E, \mathcal{S}, \mu)$ is a measure space, and $\{A_n, n = 1, 2, \ldots\}$ is a sequence of sets in $\mathcal{S}$. Then:

(a) $\mu(\lim \inf_{n \to \infty} A_n) \leq \lim \inf_{n \to \infty} \mu(A_n)$;

(b) $\mu(\lim \sup_{n \to \infty} A_n) \geq \lim \sup_{n \to \infty} \mu(A_n)$ provided that

(1.3) $\mu\left( \bigcup_{1 \leq n < \infty} A_n \right) < \infty$;

(c) If $A_n \subset A_{n+1}$, $\forall n = 1, 2, \ldots$ then $\lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n)$;

(d) If $A_{n+1} \subset A_n$, $\forall n = 1, 2, \ldots$ and $\mu(A_1) < \infty$, then $\lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n)$;

(e) If $\lim_{n \to \infty} A_n$ exists and (1.3) holds then $\lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n)$;

(f) If $\sum_{1 \leq n < \infty} \mu(A_n) < \infty$ then $\mu(\lim \sup_{n \to \infty} A_n) = 0$.

Theorem 1.2.5(f) is usually called the **Borel-Cantelli theorem** in the case where $\mu(E) = 1$.

### 1.2.3 Measurable Functions and Convergence

Suppose that $(E, \mathcal{S})$ and $(F, \mathcal{B})$ are two measurable spaces and $f : E \to F$ is a given mapping. We say that the function $f$ is $\mathcal{S}/\mathcal{B}$-measurable when $f^{-1}[B] \in \mathcal{S}$ for each $B \in \mathcal{B}$. Here $f^{-1}[B] \triangleq \{x \in E : f(x) \in B\}$ when $B \subset F$. If we have a third measurable space $(H, \mathcal{G})$ and $g : F \to H$ is a $\mathcal{B}/\mathcal{G}$-measurable mapping, then clearly the mapping $g \circ f : E \to H$ defined by $(g \circ f)(x) \triangleq g(f(x))$, $\forall x \in E$, is $\mathcal{S}/\mathcal{G}$-measurable.

We shall mostly be interested in mappings $f$ on a measurable space $(E, \mathcal{S})$ which take their values in range sets such as $\mathbb{R}^d$, $[0, \infty)$ or $[0, \infty]$. In each of these cases it is understood that the $\sigma$-algebra in the range set is the associated Borel $\sigma$-algebra, and if $f$ is measurable in the sense indicated above we usually omit reference to the Borel $\sigma$-algebra and say that $f$ is $\mathcal{S}$-measurable. Finally, if the $\sigma$-algebra $\mathcal{S}$ is clear from the context, then we abbreviate the terminology still further and say that $f$ is measurable.

The next propositions list the simplest properties of measurable functions taking values in $\mathbb{R}$. Here, and elsewhere in these notes, we use the following notation: if $\{f_n, n = 1, 2, \ldots\}$ is a sequence of functions on set $E$ taking values in $\mathbb{R}$ then $\sup_{n \geq 1} f_n$ denotes the mapping on $E$ whose value at each $x \in E$ is given by $\sup_{n \geq 1} f_n(x)$. A similar convention is understood for the infimum, limit supremum and limit infimum of the functions $f_n$. Likewise, if $f$ and $g$ are mappings on $E$ with values in $\mathbb{R}$, we use $f + g$ to denote the mapping on $E$ whose value at each $x \in E$ is given by $f(x) + g(x)$. Similar
conventions are used for the difference, product and quotient of two real-valued functions. Also, we write \( f \wedge g \) and \( f \vee g \) to indicate the functions on \( E \) whose values at each \( x \) are given by \( \min\{f(x), g(x)\} \) and \( \max\{f(x), g(x)\} \) respectively.

**Proposition 1.2.6** Suppose that \( \{f_n, n = 1, 2, \ldots\} \) is a sequence of \( \mathbb{R} \)-valued measurable functions on a measurable space \((E, S)\). Then \( \sup_{n \geq 1} f_n, \inf_{n \geq 1} f_n, \limsup_{n \to \infty} f_n, \) and \( \liminf_{n \to \infty} f_n \) are measurable functions. If \( f \) and \( g \) are \( \mathbb{R} \)-valued measurable functions on \((E, S)\) then so are \( f \vee g \) and \( f \wedge g \).

**Proposition 1.2.7** Suppose that \( f \) and \( g \) are \( \mathbb{R} \)-valued measurable functions on a measurable space \((E, S)\), and \( \alpha \) and \( \beta \) are real numbers. Then \( \alpha f + \beta g \) and \( fg \) are measurable functions. The preceding statement continues to hold when \( f \) and \( g \) are \([0, \infty]\)-valued measurable functions and \( \alpha, \beta \in [0, \infty] \).

**Proposition 1.2.8** Suppose that \( f \) is an \( \mathbb{R} \)-valued measurable function on a measurable space \((E, S)\), and define the mapping \( g : E \to \mathbb{R} \) by

\[
g(x) \triangleq \begin{cases} 
1/f(x), & \text{when } f(x) \neq 0 \\
0, & \text{when } f(x) = 0.
\end{cases}
\]

Then \( g \) is a measurable function.

If \( A \subset E \) then we use \( I_A \) to denote the **indicator function** of \( A \), namely the mapping \( I_A : E \to \mathbb{R} \) given by \( I_A(x) \triangleq 1, \forall x \in A \), and \( I_A(x) \triangleq 0, \forall x \notin A \). If \((E, S)\) is a measure space and \( A \in S \) then \( I_A \) is clearly \( S \)-measurable. A mapping \( f \) from a set \( E \) into \( \mathbb{R} \) is called a **simple function** when \( f \) assumes only finitely many values. If \( f \) is a \( S \)-measurable simple function on the measure space \((E, S)\) then it is readily apparent that \( f \) can be represented (non-uniquely) in the form

\[
(1.4) \quad f(x) = \sum_{i=1}^{N} a_i I_{A_i}(x)
\]

for some positive integer \( N \), numbers \( a_i \in \mathbb{R} \), and sets \( A_i \in S \).

**Proposition 1.2.9** Suppose that \( f \) is a measurable mapping on a measure space \((E, S)\) and takes values in \([0, \infty]\). Then there exists some sequence \( \{f_n, n = 1, 2, \ldots\} \) of \( S \)-measurable simple functions, such that

(a) \( 0 \leq f_n(x) \leq f_{n+1}(x) \leq f(x), \forall n = 1, 2, \ldots, \forall x \in E \);
(b) \( \lim_{n \to \infty} f_n(x) = f(x), \forall x \in E \).

Now suppose that \((E, S, \mu)\) is a measure space and \( \{f_n, n = 1, 2, \ldots\} \) is a sequence of measurable mappings on \( E \) with values in \( \overline{\mathbb{R}} \). It follows from Proposition 1.2.6 that the set

\[
C \triangleq \{x \in E : \limsup_{n \to \infty} f_n(x) = \liminf_{n \to \infty} f_n(x)\}
\]

is measurable.
is a member of $\mathcal{S}$, and likewise, if $f : E \to \mathbb{R}$ is a $\mathcal{S}$-measurable mapping, then the set

$$D \triangleq \{ x \in C : \lim_{n \to \infty} f_n(x) = f(x) \}$$

is also a member of $\mathcal{S}$. We say that the sequence $\{f_n, n = 1, 2, \ldots\}$ converges $\mu$-almost everywhere ($\mu$-a.e.) when $\mu(C^c) = 0$, and converges $\mu$-a.e. to a limit $f$ when $\mu(D^c) = 0$.

An alternative notion of convergence, closely related to convergence $\mu$-a.e., can be formulated as follows: suppose that $\{f_n, n = 1, 2, \ldots\}$ is a sequence of $\mathbb{R}$-valued measurable functions on a measure space $(E, \mathcal{S}, \mu)$. We say that this sequence converges in $\mu$-measure to a measurable $\mathbb{R}$-valued function $f$ on $(E, \mathcal{S}, \mu)$ when $\lim_{n \to \infty} \mu(\{|f_n - f| \geq \epsilon\}) = 0$ for each $\epsilon > 0$. Notice that our restriction that the $f_n$ be real-valued avoids the possibility of the undefined combination $\infty - \infty$ in this definition.

**Theorem 1.2.10** Suppose $\{f_n, n = 1, 2, \ldots\}$ is a sequence of $\mathbb{R}$-valued measurable functions on the measure space $(E, \mathcal{S}, \mu)$ converging in $\mu$-measure to both $f$ and $g$. Then $\mu\{f \neq g\} = 0$.

The following theorem elucidates the relationship between the two notions of convergence we have just formulated:

**Theorem 1.2.11** Suppose that $\{f_n, n = 1, 2, \ldots\}$ is a sequence of $\mathbb{R}$-valued measurable functions on the measure space $(E, \mathcal{S}, \mu)$, and let $f$ be some $\mathbb{R}$-valued measurable function defined on $(E, \mathcal{S}, \mu)$.

(a) If $\mu(E) < \infty$ and $\{f_n, n = 1, 2, \ldots\}$ converges $\mu$-a.e. to $f$ then $\{f_n, n = 1, 2, \ldots\}$ converges in $\mu$-measure to $f$;

(b) If $\{f_n, n = 1, 2, \ldots\}$ converges in $\mu$-measure to $f$ then there is a subsequence $\{f_{n_r}, r = 1, 2, \ldots\}$ of $\{f_n, n = 1, 2, \ldots\}$ which converges to $f$ $\mu$-a.e.

### 1.2.4 Integration Theory

Suppose that $f$ is a measurable simple function on a measure space $(E, \mathcal{S}, \mu)$, with a representation given by (1.4). We define the **Lebesgue integral** of $f$ on $E$ with respect to the measure $\mu$ as

$$\int_E f d\mu \triangleq \sum_{i=1}^{N} a_i \mu(A_i).$$

It is easy (although tedious) to show that when $f$ in (1.4) is written in some alternative representation

$$f(x) = \sum_{j=1}^{M} b_j I_{B_j}(x)$$

for $b_j \in \mathbb{R}$ and $B_j \in \mathcal{S}$, then

$$\sum_{i=1}^{N} a_i \mu(A_i) = \sum_{j=1}^{M} b_j \mu(B_j),$$

10
and hence the Lebesgue integral of a measurable simple function is properly defined. Next, suppose that $f$ is a measurable function on $(E, S, \mu)$ taking values in $[0, \infty]$. Then we define the **Lebesgue integral** of $f$ on $E$ with respect to $\mu$ as

$$
\int_E f \, d\mu \triangleq \sup_g \left\{ \int_E g \, d\mu \right\}
$$

where the supremum on the right hand side is taken over all measurable simple functions $g$ such that $0 \leq g(x) \leq f(x), \forall x \in E$. Notice that the integral we have just defined assumes its values in $[0, \infty]$ and in particular can take the value $+\infty$. Next, for $f$ as above and $A \in S$, we define the Lebesgue integral of $f$ over $A$ with respect to $\mu$ by

$$
\int_A f \, d\mu \triangleq \int_E f I_A \, d\mu.
$$

It follows from the preceding definition that, if $f$ and $g$ are measurable functions on $(E, S, \mu)$ such that $0 \leq f \leq g$, then

$$
\int_E f \, d\mu \leq \int_E g \, d\mu.
$$

The issue of interchanging limits and integrals is considered in the next result, which is the familiar Monotone Convergence Theorem:

**Theorem 1.2.12 (Levi)** Suppose that $\{f_n, \, n = 1, 2, \ldots\}$ is a sequence of $[0, \infty]$-valued measurable functions on a measure space $(E, S, \mu)$. If $f_n(x) \leq f_{n+1}(x), \forall n = 1, 2, \ldots, \forall x \in E$, then

$$
\lim_{n \to \infty} \int_E f_n \, d\mu = \int_E \lim_{n \to \infty} f_n \, d\mu.
$$

One immediate consequence of this theorem is the following:

**Corollary 1.2.13** Suppose $f$ is a $[0, \infty]$-valued measurable function on a measure space $(E, S, \mu)$. Define the mapping $\nu : S \to [0, \infty]$ by

$$
\nu(A) \triangleq \int_A f \, d\mu, \quad \forall A \in S.
$$

Then $\nu$ is a measure on $(E, S)$.

**Remark 1.2.14** Suppose $(E, S, \mu)$ is a measure space and $\mathcal{H} \subset S$ is a sub-$\sigma$-algebra of $S$. We call the mapping

$$
\mu_H : \mathcal{H} \to [0, \infty]
$$

defined by $\mu_H(A) \triangleq \mu(A), \forall A \in \mathcal{H}$, the **restriction** of $\mu$ to $\mathcal{H}$. Clearly $\mu_H$ is a measure on $(E, \mathcal{H})$, so that we have a another measure space $(E, \mathcal{H}, \mu_H)$ besides $(E, S, \mu)$. If $f$ is a $[0, \infty]$-valued $\mathcal{H}$-measurable function on $(E, \mathcal{H}, \mu_H)$ then the Lebesgue integral of $f$ on $E$ with respect to $\mu_H$ is defined. But, since
$\mathcal{H} \subset \mathcal{S}$, $f$ is also a $\mathcal{S}$-measurable function on $(E, \mathcal{S}, \mu)$, hence the Lebesgue integral of $f$ on $E$ with respect to $\mu$ is defined as well. It follows at once from Theorem 1.2.12 and Proposition 1.2.9 that

$$\int_A f \, d\mu = \int_A f \, d\mu_H, \quad \forall \ A \in \mathcal{H},$$

when $f : E \to [0, \infty]$ is $\mathcal{H}$-measurable.

Closely related to Theorem 1.2.12 is the following:

**Theorem 1.2.15 (Fatou)** Suppose that $\{f_n, n = 1, 2, \ldots\}$ is a sequence of $[0, \infty]$-valued measurable functions on a measure space $(E, \mathcal{S}, \mu)$. Then

$$\liminf_{n \to \infty} \int_E f_n \, d\mu \geq \int_E \liminf_{n \to \infty} f_n \, d\mu.$$

The next result shows that integration of non-negative mappings is linear in a partial sense:

**Theorem 1.2.16** Suppose $f$ and $g$ are $[0, \infty]$-valued measurable mappings on a measurable space $(E, \mathcal{S}, \mu)$ and $\alpha, \beta \in [0, \infty]$. Then

$$\int_E \alpha f + \beta g \, d\mu = \alpha \int_E f \, d\mu + \beta \int_E g \, d\mu. \quad (1.5)$$

Next, we extend the notion of Lebesgue integral to the case of functions which take both positive and negative values. Here we must be careful to avoid the undefined combination $\infty - \infty$. Suppose that $f$ is an $\mathcal{R}$-valued measurable function on a measure space $(E, \mathcal{S}, \mu)$. We define the positive and negative parts of $f$ by the functions $f_+(x) \overset{\triangle}{=} \max\{f(x), 0\}$ and $f_-(x) \overset{\triangle}{=} \max\{-f(x), 0\}$, and observe that $f(x) = f_+(x) - f_-(x), \forall \ x \in E$. The function $f$ is said to be $\mu$-integrable when it is $\mathcal{S}$-measurable, and

$$\int_E f_+ \, d\mu < \infty \quad \text{and} \quad \int_E f_- \, d\mu < \infty, \quad (1.6)$$

in which case the **Lebesgue integral** of $f$ on $E$ with respect to $\mu$ is defined by

$$\int_E f \, d\mu \overset{\triangle}{=} \int_E f_+ \, d\mu - \int_E f_- \, d\mu.$$

There are several things to note about this definition. Firstly, if one or both of the conditions in (1.6) fails to hold then the Lebesgue integral of $f$ is undefined. It follows that whereas the integral of a non-negative measurable mapping is always defined and may assume the value $+\infty$, the Lebesgue integral of a measurable function taking both positive and negative values is defined only when (1.6) holds and is then $\mathcal{R}$-valued (or finitely valued). Finally, it is readily seen from Theorem 1.2.16 that $f$ is $\mu$-integrable if and only if

$$\int_E |f| \, d\mu < \infty.$$
Theorem 1.2.17 Suppose that $f$ is a $[0, \infty]$-valued function on a measurespace $(E, \mathcal{S}, \mu)$ and $\nu$ is the measure on $(E, \mathcal{S})$ in Corollary 1.2.13.

(a) If $g$ is a $[0, \infty]$-valued measurable function on $(E, \mathcal{S})$ then

\begin{equation}
\int_A g \, d\nu = \int_A g \, d\mu, \quad \forall A \in \mathcal{S}.
\end{equation}

(b) If $g$ is an $\mathbb{R}$-valued measurable function on $(E, \mathcal{S})$ then $g$ is $\nu$-integrable if and only if $gf$ is $\mu$-integrable in which case (1.7) holds.

The next result shows that Lebesgue integration of functions taking positive and negative values is linear, provided we impose enough restrictions to ensure that the functions concerned are integrable and to avoid the undefined combination $\infty - \infty$:

Theorem 1.2.18 Suppose $f$ and $g$ are $\mathbb{R}$-valued measurable mappings on a measurable space $(E, \mathcal{S}, \mu)$ and $\alpha, \beta \in \mathbb{R}$. If $f$ and $g$ are $\mu$-integrable then $\alpha f + \beta g$ is $\mu$-integrable and (1.5) holds.

We next consider a change-of-variables lemma for Lebesgue integration. Suppose that $(E, \mathcal{S})$ and $(F, \mathcal{B})$ are measure spaces, $f : E \to F$ is a $\mathcal{S}/\mathcal{B}$-measurable mapping, and $g : F \to \mathbb{R}$ is a $\mathcal{B}$-measurable mapping. As observed in Section 1.2.3, the mapping $g \circ f : E \to \mathbb{R}$, defined by $(g \circ f)(x) \triangleq g(f(x)), \forall x \in E$, is $\mathcal{S}$-measurable. If $\mu$ is a measure on $\mathcal{S}$, define the mapping $f^\star \mu : \mathcal{B} \to [0, \infty]$ by $(f^\star \mu)(A) \triangleq \mu(f^{-1}[A]), \forall A \in \mathcal{B}$. Clearly $f^\star \mu$ is a measure on $\mathcal{B}$, usually called the image measure of $\mu$ through $f$.

Lemma 1.2.19 Suppose $(E, \mathcal{S}, \mu)$ is a measure space, $(F, \mathcal{B})$ is a measurable space, and $f : E \to F$ is a $\mathcal{S}/\mathcal{B}$-measurable mapping. We have the following:

(a) Suppose that $g$ is a $[0, \infty]$-valued measurable function on $(F, \mathcal{B})$. Then

\begin{equation}
\int_F g \, d(f^\star \mu) = \int_E g \circ f \, d\mu.
\end{equation}

(b) Suppose that $g$ is an $\mathbb{R}$-valued measurable function on $(F, \mathcal{B})$. Then $g$ is $f^\star \mu$-integrable if and only if $g \circ f$ is $\mu$-integrable, in which case (1.8) holds.

The next theorem formulates a “uniform continuity” property for Lebesgue integrals:

Theorem 1.2.20 Suppose that $f$ is a $\mu$-integrable function on a measure space $(E, \mathcal{S}, \mu)$. Then, corresponding to each $\epsilon > 0$, there is some $\delta(\epsilon) > 0$ such that

\[ \int_A |f| \, d\mu < \epsilon \]

for each $A \in \mathcal{S}$ such that $\mu(A) < \delta(\epsilon)$.

We now look at the issue of interchanging limits and integrals when integrating functions which can take both positive and negative values. The next result is the familiar Dominated Convergence Theorem:
Theorem 1.2.21 (Lebesgue) Suppose that \( \{f_n, n = 1, 2, \ldots \} \) is a sequence of \( \mathbb{R} \)-valued measurable functions on \((E, S, \mu)\) converging \( \mu \)-a.e. to a \( \mathbb{R} \)-valued measurable function \( f \). If there exists some measurable \( \mu \)-integrable function \( g \) such that \(|f_n| \leq |g| \) \( \mu \)-a.e. for each \( n = 1, 2, \ldots \), then \( f \) is \( \mu \)-integrable, and
\[
\lim_{n \to \infty} \int_E |f_n - f| d\mu = 0.
\]

One useful consequence of this theorem is:

Corollary 1.2.22 (Scheffé) Suppose that \( \{f_n, n = 1, 2, \ldots \} \) are \( \mu \)-integrable \( \mathbb{R} \)-valued functions on a measure space \((E, S, \mu)\) converging \( \mu \)-a.e. to some \( \mathbb{R} \)-valued \( \mu \)-integrable limit \( f \). Then
\[
\lim_{n \to \infty} \int_E |f_n - f| d\mu = 0 \quad \text{if and only if} \quad \lim_{n \to \infty} \int_E |f_n| d\mu = \int_E |f| d\mu.
\]

Despite the tremendous usefulness of the dominated convergence theorem there are applications where it is difficult to verify the existence of a “dominating” function \( g \). Accordingly, in Chapter 2, we shall develop an alternative result which will allow us to interchange limits and integration even when no dominating function is clearly apparent (see Section 2.6, especially Theorem 2.6.10).

Theorem 1.2.23 (Markov inequality) Suppose that \( f \) is a \([0, \infty]\)-valued measurable function on a measure space \((E, S, \mu)\). Then
\[
\mu\{f \geq \lambda\} \leq \frac{1}{\lambda} \int_{[f \geq \lambda]} f d\mu \leq \frac{1}{\lambda} \int_E f d\mu, \quad \forall \lambda \in (0, \infty).
\]

Suppose that \( f \) and \( g \) are measurable functions on \((E, S, \mu)\). Then clearly \( \{f \neq g\} \) is a member of \( S \); \( f \) and \( g \) are said to be equal \( \mu \)-a.e. when \( \mu\{f \neq g\} = 0 \), in which case we write \( f = g \ \mu \)-a.e. By the same token \( f \) is said to be non-negative \( \mu \)-a.e. when \( \mu\{f < 0\} = 0 \), in which case we write \( f \geq 0 \ \mu \)-a.e. The next result summarises the pertinent facts about relations involving functions which hold only \( \mu \)-a.e.

Theorem 1.2.24 Suppose that \( f \) and \( g \) are measurable \( \mathbb{R} \)-valued mappings defined on \((E, S, \mu)\).
(a) If \( f = g \ \mu \)-a.e. and \( f \) is \( \mu \)-integrable, then \( g \) is \( \mu \)-integrable and \( \int_E f d\mu = \int_E g d\mu \).
(b) If \( \int_A f d\mu \geq 0 \) for each \( A \in S \), then \( f \geq 0 \ \mu \)-a.e. and if \( \int_A f d\mu = 0 \) for each \( A \in S \), then \( f = 0 \ \mu \)-a.e.
(c) If \( f \) takes values in \([0, \infty]\) and \( \int_E f d\mu < \infty \), then \( \{f = +\infty\} \in S \) and \( \mu\{f = +\infty\} = 0 \).
(d) Suppose \( f \) takes values in \([0, \infty]\). Then \( f = 0 \ \mu \)-a.e. if and only if \( \int_E f d\mu = 0 \).

1.2.5 Lebesgue Spaces

Suppose that \((E, S, \mu)\) is a given measure space and \( p \in (0, \infty) \) is a constant. We denote by \( L^p(E, S, \mu) \) the set of all \( \mathbb{R} \)-valued \( S \)-measurable functions such that \( |f|^p \) is \( \mu \)-integrable, and write
\[
\|f\|_p \triangleq \left\{ \int_E |f|^p d\mu \right\}^{\frac{1}{p}}.
\]

(1.10)
Theorem 1.2.25 (Hölder inequality) Suppose that \( p \) and \( q \) are constants in \((1, \infty)\) with \((1/p) + (1/q) = 1\). If \( f \in L^p(E, \mathcal{S}, \mu) \) and \( g \in L^q(E, \mathcal{S}, \mu) \), then \( fg \in L^1(E, \mathcal{S}, \mu) \) and

\[
\|fg\|_1 \leq \|f\|_p \|g\|_q.
\]

The numbers \( p \) and \( q \) in Theorem 1.2.25 are called conjugate exponents. The special case of Hölder’s inequality in which the conjugate exponents are equal, namely \( p = q = 2 \), is especially important in probability theory and goes by the name of the Cauchy–Schwarz inequality.

Theorem 1.2.26 (Minkowski inequality) Suppose that \( p \in [1, \infty) \) is a constant and \( f \) and \( g \) are members of \( L^p(E, \mathcal{S}, \mu) \). Then \( f + g \) belong to \( L^p(E, \mathcal{S}, \mu) \), and

\[
\|f + g\|_p \leq \|f\|_p + \|g\|_p.
\]

In view of Minkowski’s inequality, it follows that \( L^p(E, \mathcal{S}, \mu) \) is a real vector space with vector addition and scalar multiplication defined in the usual pointwise sense, the zero element being the function \( f(x) = 0, \forall \ x \in E \). We want to use (1.10) and Theorem 1.2.26 to define a norm on \( L^p(E, \mathcal{S}, \mu) \) for each \( p \in [1, \infty) \), but since the condition \( \|f\|_p = 0 \) ensures only that \( f = 0 \) \( \mu \)-a.e. and not that \( f \) is the zero element of \( L^p(E, \mathcal{S}, \mu) \), we see that \( \| \cdot \|_p \) defined by (1.10) is only a pseudo-norm on \( L^p(E, \mathcal{S}, \mu) \) (see Section 1.1). In order for \( \| \cdot \|_p \) to serve as a norm, it is necessary to make some adjustments in our interpretation of the vector space structure in \( L^p(E, \mathcal{S}, \mu) \). Define

\[
[f] \overset{\Delta}{=} \{ f' \in L^p(E, \mathcal{S}, \mu) : \|f - f'\|_p = 0 \}
\]

for each \( f \in L^p(E, \mathcal{S}, \mu) \), and then put

\[
L^p_*(E, \mathcal{S}, \mu) \overset{\Delta}{=} \{ [f] : f \in L^p(E, \mathcal{S}, \mu) \}.
\]

If we write \( f \leftrightarrow g \) to indicate \( g \in [f] \) then it is readily apparent that \( \leftrightarrow \) is an equivalence relation on \( L^p(E, \mathcal{S}, \mu) \) and \([f]\) is the equivalence class generated by \( f \). In particular, for \( \alpha, \beta \in \mathbb{R} \) and \( f, g \in L^p(E, \mathcal{S}, \mu) \), we have

\[
[\alpha f + \beta g] = [\alpha f' + \beta g']
\]

for all \( f' \in [f], g' \in [g] \), and hence we can define vector addition and scalar multiplication in \( L^p_*(E, \mathcal{S}, \mu) \) by

\[
\alpha[f] + \beta[g] \overset{\Delta}{=} [\alpha f' + \beta g']
\]

for arbitrary \( f' \in [f] \) and \( g' \in [g] \). This turns \( L^p_*(E, \mathcal{S}, \mu) \) into a real vector space with zero element \([0]\), and if we put \( \|[f]\|_p \overset{\Delta}{=} \|f\|_p \) then \( L^p_*(E, \mathcal{S}, \mu) \) becomes a normed vector space. What the preceding construction really achieves is to justify our view of \( L^p(E, \mathcal{S}, \mu) \), for each \( p \in [1, \infty) \), as a normed vector space with a norm defined by (1.10), provided that we weaken the sense of equality in \( L^p(E, \mathcal{S}, \mu) \) and regard two elements \( f \) and \( g \) in this space as identical when \( \|f - g\|_p = 0 \). We shall henceforth always interpret \( L^p(E, \mathcal{S}, \mu) \), usually called the Lebesgue space of order \( p \), in this sense. As the next theorem shows this normed vector space is actually complete:
Theorem 1.2.27 Suppose that \((E, \mathcal{S}, \mu)\) is a measure space. Then \(L^p(E, \mathcal{S}, \mu)\) is a Banach space for each \(p \in [1, \infty)\).

In the special case where \(p = 2\) we define the inner product of two functions \(f\) and \(g\) in \(L^2(E, \mathcal{S}, \mu)\), written as \((f, g)\), by
\[
(f, g) \triangleq \int_E fg \, d\mu,
\]
where it is clear from the Cauchy-Schwarz inequality that the integral on the right side of (1.11) exists. With this notion of inner product, and in view of Theorem 1.2.27, the normed vector space \(L^2(E, \mathcal{S}, \mu)\) becomes a Hilbert space.

### 1.2.6 The Radon-Nikodym Theorem

**Definition 1.2.28** Suppose that \(\nu\) and \(\mu\) are two measures on the measurable space \((E, \mathcal{S})\). Let \(\mathcal{N}_\nu\) denote the collection of all \(\nu\)-null sets of \(\mathcal{S}\), namely
\[
\mathcal{N}_\nu \triangleq \{A \in \mathcal{S} : \nu(A) = 0\},
\]
and define \(\mathcal{N}_\mu\), the collection of all \(\mu\)-null sets of \(\mathcal{S}\), in the same way. We say that \(\nu\) is absolutely continuous with respect to \(\mu\) on \(\mathcal{S}\) when \(\mathcal{N}_\mu \subset \mathcal{N}_\nu\). Equivalently, \(\nu\) is absolutely continuous with respect to \(\mu\) on \(\mathcal{S}\) whenever \(A \in \mathcal{S}\) and \(\mu(A) = 0\) implies that \(\nu(A) = 0\). We use the notation \(\nu \ll \mu[\mathcal{S}]\) to indicate that \(\nu\) is absolutely continuous with respect to \(\mu\) on \(\mathcal{S}\).

**Theorem 1.2.29** Suppose that \(\nu\) and \(\mu\) are finite measures on \((E, \mathcal{S})\). Then \(\nu \ll \mu[\mathcal{S}]\) if and only if the following holds: For each \(\epsilon \in (0, \infty)\) there exists some \(\delta(\epsilon) \in (0, \infty)\) such that \(A \in \mathcal{S}\) and \(\mu(A) < \delta(\epsilon)\) imply \(\nu(A) < \epsilon\).

Suppose \(f\) is a \([0, \infty]\)-valued measurable function on the measure space \((E, \mathcal{S}, \mu)\), and define \(\nu : \mathcal{S} \to [0, \infty] \) by
\[
\nu(A) \triangleq \int_A f \, d\mu, \quad \forall A \in \mathcal{S}.
\]
We see at once from Corollary 1.2.13 that \(\nu\) is a measure on \((E, \mathcal{S})\) with \(\nu \ll \mu[\mathcal{S}]\). The next result is a partial converse to this statement and asserts that if \(\nu \ll \mu[\mathcal{S}]\), and \(\nu\), \(\mu\), are \(\sigma\)-finite, then a relationship between \(\nu\) and \(\mu\) of the form (1.12) necessarily holds:

**Theorem 1.2.30 (Radon-Nikodym)** Suppose that \(\nu\) and \(\mu\) are \(\sigma\)-finite measures on a measurable space \((E, \mathcal{S})\) with \(\nu \ll \mu[\mathcal{S}]\). Then we have the following:

(a) There exists some \([0, \infty]\)-valued measurable function \(f\) on \((E, \mathcal{S})\) such that (1.12) holds;
(b) For any \([0, \infty]\)-valued measurable function \(\tilde{f}\) on \((E, \mathcal{S})\) such that \(\nu(A) = \int_A \tilde{f} \, d\mu, \forall A \in \mathcal{S}\), we have that \(\mu\{f \neq \tilde{f}\} = 0\).
Remark 1.2.31 In Theorem 1.2.30 we clearly have
\[ \int_E f d\mu < \infty, \]
if and only if \( \nu \) is a finite measure \( (\nu(E) < \infty) \).

Remark 1.2.32 The \([0, \infty]\)-valued and \( S \)-measurable mapping \( f \) determined by Theorem 1.2.30(a) is called a **Radon-Nikodym derivative** of \( \nu \) with respect to \( \mu \) relative to the \( \sigma \)-algebra \( S \). Theorem 1.2.30(b) shows that this mapping is uniquely determined to within \( \mu \)-null sets. We shall use the notation \( \frac{d\nu}{d\mu} \) to indicate an arbitrary but fixed choice of the Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \) relative to the \( \sigma \)-algebra \( S \). The arbitrary character of this choice allows us a certain amount of leeway in choosing Radon-Nikodym derivatives with nice properties. In particular, when \( \nu(E) < \infty \), and hence \( \int_E f d\mu < \infty \), we have \( \mu\{f = +\infty\} = 0 \). Since we prefer our functions to be \( \mathbb{R} \)-valued rather than \( \mathbb{R} \)-valued wherever possible, when \( \nu(E) < \infty \) we shall always take \( \frac{d\nu}{d\mu} \) to be the \([0, \infty]\)-valued \( S \)-measurable mapping defined on \( E \) by
\[
\frac{d\nu}{d\mu}(x) \triangleq \begin{cases} 
  f(x) & \text{when } f(x) \in [0, \infty) \\
  0 & \text{when } f(x) = +\infty.
\end{cases}
\]

Remark 1.2.33 If \( \nu \ll \mu[S] \) and \( g : (E, S) \to [0, \infty] \) is a measurable mapping, then, in view of Theorem 1.2.30 and Theorem 1.2.17(a), we see that
\[
\int_A g d\nu = \int_A g \left( \frac{d\nu}{d\mu} \right) d\mu, \quad \forall A \in S;
\]
and, if \( g : (E, S) \to \mathbb{R} \) is measurable, then, from Theorem 1.2.30 and Theorem 1.2.17(b), it follows that \( g \) is \( \nu \)-integrable if and only if \( g(d\nu/d\mu) \) is \( \mu \)-integrable, in which case (1.13) holds.

Remark 1.2.34 Suppose that \( \nu \) and \( \mu \) are finite measures on the measure space \((E, S)\), with \( \nu \ll \mu[S] \), and suppose that \( \mathcal{H} \subset S \) is a sub-\( \sigma \)-algebra of \( S \). Let \( \nu_H \) and \( \mu_H \) be the restrictions to \( \mathcal{H} \) of \( \nu \) and \( \mu \) respectively (see Remark 1.2.14). Clearly \( \nu_H \) and \( \mu_H \) are finite measures on \((E, \mathcal{H})\), and \( \nu_H \ll \mu_H[\mathcal{H}] \).

Thus, identifying \((E, S)\), \( \nu \) and \( \mu \) in Theorem 1.2.30 with \((E, \mathcal{H})\), \( \nu_H \) and \( \mu_H \) in the present remark, we see that there is some \( \mathcal{H} \)-measurable function \( f_H : E \to [0, \infty] \) such that
\[
\nu(A) = \int_A f_H d\mu, \quad \forall A \in \mathcal{H},
\]
and that, if \( \bar{f}_H : E \to [0, \infty] \) is some \( \mathcal{H} \)-measurable function such that (1.14) holds with \( \bar{f}_H \) in place of \( f_H \), then \( \mu\{f_H \neq \bar{f}_H\} = 0 \). We shall use the notation
\[
\frac{d\nu}{d\mu}\bigg\vert_{\mathcal{H}}
\]
to indicate an arbitrary but fixed choice of a \([0, \infty]\)-valued and \( \mathcal{H} \)-measurable function \( f_H \) for which (1.14) holds, and call this function a **Radon-Nikodym derivative** of \( \nu \) with respect to \( \mu \) relative
to the \( \sigma \)-algebra \( \mathcal{H} \). Identifying \((E, \mathcal{S})\) and \( f \) in Theorem 1.2.17 with \((E, \mathcal{H})\) and \( f_H \) in the present remark, we see that, for any \( \mathcal{H} \)-measurable \( g : E \to \overline{\mathbb{R}} \), one has

\[
\int_E |g|d\nu < \infty \quad \text{if and only if} \quad \int_E |f_H g|d\mu < \infty,
\]

in which case

\[
\int_A g d\nu = \int_A g f_H d\mu, \quad \forall \ A \in \mathcal{H}.
\]

**Definition 1.2.35** Suppose that \( \nu \) and \( \mu \) are measures on a measurable space \((E, \mathcal{S})\). If \( \nu \ll \mu \mathcal{S} \) and \( \mu \ll \nu \mathcal{S} \) (i.e. \( N_\nu = N_\mu \)) then \( \nu \) and \( \mu \) are said to be **equivalent measures** on \( \mathcal{S} \), and we write \( \nu \equiv \mu \mathcal{S} \) to denote equivalence of \( \nu \) and \( \mu \) on \( \mathcal{S} \).

**Remark 1.2.36** Suppose that \( \mu \) is a measure on the measurable space \((E, \mathcal{S})\). We call a \([0, \infty)\)-valued measurable function \( f \) on \((E, \mathcal{S})\) a \( \mu \)-strictly positive function when \( f > 0 \) \( \mu \)\-a.e. i.e. \( \mu \{x \in E : f(x) = 0\} = 0 \). When \( f \) is \( \mu \)-strictly positive on \((E, \mathcal{S})\), we shall slightly abuse notation and use \( f^{-1} \) to denote the mapping on \( E \) defined by \( f^{-1}(x) \triangleq 1/f(x) \) when \( f(x) > 0 \) and \( f^{-1}(x) \triangleq 0 \) when \( f(x) = 0 \). Clearly, \( f^{-1} \) is \([0, \infty)\)-valued and measurable on \((E, \mathcal{S})\), with

\[\mu \{f^{-1}f \neq 1\} = 0.\]

The next result is just a simple consequence of Theorem 1.2.24(d):

**Proposition 1.2.37** Suppose that \( \nu \) and \( \mu \) are \( \sigma \)-finite measures on \((E, \mathcal{S})\) with \( \nu \ll \mu \mathcal{S} \), and let \( f \triangleq d\nu/d\mu \) be a Radon-Nikodym derivative of \( \nu \) with respect to \( \mu \) relative to the \( \sigma \)-algebra \( \mathcal{S} \). Then the following hold:

(a) \( f \) is \( \nu \)-strictly positive;
(b) \( f \) is \( \mu \)-strictly positive if and only if \( \nu \equiv \mu \mathcal{S} \);
(c) If \( \nu \equiv \mu \mathcal{S} \) then \( f^{-1} \) is a Radon-Nikodym derivative of \( \mu \) with respect to \( \nu \) relative to the \( \sigma \)-algebra \( \mathcal{S} \).

### 1.2.7 The Fubini-Tonelli Theorem

Suppose that \( A_1 \) and \( A_2 \) are given sets. We define the **ordered product** of \( A_1 \) with \( A_2 \), denoted by \( A_1 \otimes A_2 \), as the set of ordered pairs

\[A_1 \otimes A_2 \triangleq \{(x_{1}, x_{2}) : x_{i} \in A_{i} \quad i = 1, 2\}.
\]

Now suppose that \((E_1, \mathcal{S}_1)\) and \((E_2, \mathcal{S}_2)\) are two measurable spaces. If \( A_1 \in \mathcal{S}_1 \) and \( A_2 \in \mathcal{S}_2 \) then we call \( A_1 \otimes A_2 \) a **measurable rectangle** with sides \( A_1 \) and \( A_2 \) in the product set \( E_1 \otimes E_2 \). If we let \( \mathcal{C} \) be the collection of all such measurable rectangles, namely

\[\mathcal{C} \triangleq \{A_1 \otimes A_2 : A_i \in \mathcal{S}_i \quad i = 1, 2\},\]
then we call the generated $\sigma$-algebra $\sigma\{C\}$ over $E_1 \otimes E_2$ the ordered product of the $\sigma$-algebras $S_1$ and $S_2$. This $\sigma$-algebra is denoted by $S_1 \otimes S_2$ and the resulting measurable space $(E_1 \otimes E_2, S_1 \otimes S_2)$ is called the ordered product of the measurable spaces $(E_1, S_1)$ and $(E_2, S_2)$.

Lemma 1.2.38 Suppose that $(E_1, S_1)$ and $(E_2, S_2)$ are measurable spaces and let $f$ be an $\overline{\mathbb{R}}$-valued measurable function on the product measurable space $(E_1 \otimes E_2, S_1 \otimes S_2)$. Then, for each $x_1 \in E_1$ and $x_2 \in E_2$, the mappings $f(\cdot, x_2)$ and $f(x_1, \cdot)$ are measurable on $(E_1, S_1)$ and $(E_2, S_2)$ respectively.

Next, suppose that $(E_1, S_1, \mu_1)$ and $(E_2, S_2, \mu_2)$ are two measure spaces. We want to define a measure on the ordered product $(E_1 \otimes E_2, S_1 \otimes S_2)$ which is in some sense a product of the measures $\mu_1$ and $\mu_2$. For this purpose we need the following lemma:

Lemma 1.2.39 Suppose that $(E_1, S_1, \mu_1)$ and $(E_2, S_2, \mu_2)$ are $\sigma$-finite measure spaces. Then there is a unique measure $\mu$ on the product measure space $(E_1 \otimes E_2, S_1 \otimes S_2)$ such that

\[
\mu(A_1 \otimes A_2) = \mu_1(A_1)\mu_2(A_2)
\]

for all $A_i \in S_i$, $i = 1, 2$.

The unique measure in Lemma 1.2.39 is called the ordered product of the measures $\mu_1$ and $\mu_2$ and is denoted by $\mu_1 \otimes \mu_2$. One should note in passing that the hypothesis that the measure spaces in Lemma 1.2.39 be $\sigma$-finite is essential. In fact, without this condition there may exist several measures $\mu$ on $(E_1 \otimes E_2, S_1 \otimes S_2)$ for which (1.17) holds.

We have thus seen how to use two $\sigma$-finite measure spaces $(E_1, S_1, \mu_1)$ and $(E_2, S_2, \mu_2)$ to construct a unique product measure space $(E_1 \otimes E_2, S_1 \otimes S_2, \mu_1 \otimes \mu_2)$. Now we are ready for the main issue of this section, namely integration of mappings defined on a product of measure spaces. Suppose that $f$ is a $[0, \infty]$-valued $S_1 \otimes S_2$-measurable function on the product of two $\sigma$-finite measure spaces $(E_1, S_1, \mu_1)$ and $(E_2, S_2, \mu_2)$. Then, of course, the Lebesgue integral of $f$ on $E_1 \otimes E_2$ with respect to the product measure $\mu_1 \otimes \mu_2$ is defined, and will be denoted by

\[
\int_{E_1 \otimes E_2} f \, d\mu_1 \otimes \mu_2 \quad \text{or} \quad \int_{E_1 \otimes E_2} f(x_1, x_2) \, d\mu_1 \otimes \mu_2(x_1, x_2).
\]

In view of Lemma 1.2.38 the mapping $f(\cdot, x_2)$ has a Lebesgue integral on $E_1$ with respect to measure $\mu_1$ for each $x_2 \in E_2$. We denote this integral by

\[
\int_{E_1} f(x_1, x_2) \, d\mu_1(x_1)
\]

and observe that we now have a $[0, \infty]$-valued function on $E_2$ which we shall indicate by

\[
\int_{E_1} f(x_1, \cdot) \, d\mu_1(x_1) \quad \text{or} \quad x_2 \rightarrow \int_{E_1} f(x_1, x_2) \, d\mu_1(x_1).
\]
Identical observations can be made, and an analogous notation used, for the case where we fix \( x_1 \in E_1 \) and integrate over \( E_2 \) with respect to \( \mu_2 \). It is natural to ask about the measurability of the mapping in (1.19) and the relationship between the preceding integrals. The answers to these questions are given by the following:

**Theorem 1.2.40 (Tonelli)** Suppose that \( f \) is a \([0, \infty]\)-valued \( S_1 \otimes S_2 \)-measurable mapping defined on the ordered product of two \( \sigma \)-finite measure spaces \((E_1, S_1, \mu_1)\) and \((E_2, S_2, \mu_2)\). Then the mappings

\[
\int_{E_1} f(x_1, \cdot) \, d\mu_1(x_1) \quad \text{and} \quad \int_{E_2} f(\cdot, x_2) \, d\mu_2(x_2)
\]

are measurable on \((E_2, S_2, \mu_2)\) and \((E_1, S_1, \mu_1)\) respectively, and

\[
\int_{E_1 \otimes E_2} f(x_1, x_2) \, d\mu_1(x_1) \otimes d\mu_2(x_2) = \int_{E_1} \left\{ \int_{E_2} f(x_1, x_2) \, d\mu_2(x_2) \right\} \, d\mu_1(x_1)
\]

\[
= \int_{E_2} \left\{ \int_{E_1} f(x_1, x_2) \, d\mu_1(x_1) \right\} \, d\mu_2(x_2).
\]

We next consider the case where \( f \) is a measurable \( \mathbb{R} \)-valued mapping on the product to the two \( \sigma \)-finite measure spaces \((E_1, S_1, \mu_1)\) and \((E_2, S_2, \mu_2)\). From Theorem 1.2.40 we see that \( f \) is \( \mu_1 \otimes \mu_2 \)-integrable on \( E_1 \otimes E_2 \) if and only if

\[
\int_{E_1} \left\{ \int_{E_2} |f(x_1, x_2)| \, d\mu_2(x_2) \right\} \, d\mu_1(x_1) < \infty
\]

if and only if

\[
\int_{E_2} \left\{ \int_{E_1} |f(x_1, x_2)| \, d\mu_1(x_1) \right\} \, d\mu_2(x_2) < \infty.
\]

Define

\[
N_1 \triangleq \left\{ x_1 \in E_1 : \int_{E_2} |f(x_1, x_2)| \, d\mu_2(x_2) = +\infty \right\}
\]

and

\[
N_2 \triangleq \left\{ x_2 \in E_2 : \int_{E_1} |f(x_1, x_2)| \, d\mu_1(x_1) = +\infty \right\}.
\]

It follows at once from Theorem 1.2.40 that \( N_i \in S_i, \ i = 1, 2 \), and we see from the definition of Lebesgue integrals for functions taking both positive and negative values that \( \int_{E_i} f(x_1, x_2) \, d\mu_1(x_1) \) is defined (and \( \mathbb{R} \)-valued) if and only if \( x_2 \notin N_2 \) (an obviously analogous statement holds for integration over \( E_2 \) with respect to \( \mu_2 \)). Define \( f_1 : E_1 \to \mathbb{R} \) by

\[
f_1(x_1) \triangleq \begin{cases} 
\int_{E_2} f(x_1, x_2) \, d\mu_2(x_2) & \text{if } x_1 \notin N_1 \\
0 & \text{if } x_1 \in N_1
\end{cases}
\]

and let \( f_2 \) be the analogous mapping defined over \( E_2 \). The next theorem ensures that the sets \( N_i \) are of measure zero (i.e. negligible) when \( f \) is \( \mu_1 \otimes \mu_2 \)-integrable, and provides an analogue of Theorem 1.2.40 when \( f \) can take both positive and negative values.
Theorem 1.2.41 (Fubini) Suppose that $f$ is an $\mathbb{R}$-valued $S_1 \otimes S_2$-measurable mapping defined on the product of two $\sigma$-finite measure spaces $(E_1, \mathcal{S}_1, \mu_1)$ and $(E_2, \mathcal{S}_2, \mu_2)$. If $f$ is $\mu_1 \otimes \mu_2$-integrable then $\mu_1(N_1) = \mu_2(N_2) = 0$, $f_i$ is $\mu_i$-integrable for $i = 1, 2$, and

$$\int_{E_1} f_1(x_1) d\mu_1(x_1) = \int_{E_2} f_2(x_2) d\mu_2(x_2) = \int_{E_1 \otimes E_2} f(x_1, x_2) d\mu_1 \otimes \mu_2(x_1, x_2)$$

where $f_1(\cdot)$ is defined by (1.25) and $f_2(\cdot)$ is analogously defined.

The pair of results in Theorems 1.2.40 and 1.2.41 are usually referred to as the Fubini-Tonelli theorem. Since the sets $N_i$ in (1.23) and (1.24) have measure zero, and hence contribute nothing to the integrals in question, we usually write the first two integrals in (1.26) in iterated form, namely

$$\int_{E_1} \left\{ \int_{E_2} f(x_1, x_2) d\mu_2(x_2) \right\} d\mu_1(x_1) = \int_{E_2} \left\{ \int_{E_1} f(x_1, x_2) d\mu_1(x_2) \right\} d\mu_2(x_2)$$

$$= \int_{E_1 \otimes E_2} f(x_1, x_2) d\mu_1 \otimes \mu_2(x_1, x_2),$$

it being understood that the “inner” integrals in braces are put equal to zero (or indeed any fixed value) when $x_1 \in N_1$ and $x_2 \in N_2$.

1.3 Product $\sigma$-Algebras on Function Spaces

A notion which is essential to the modern theory of stochastic processes is that of an arbitrary product of measurable spaces. Our main goal in this section is to formulate this concept. The idea of an arbitrary product of measurable spaces is simple enough, although frequently not included in basic introductions to measure theory and probability. Accordingly, in contrast to the rather brief summaries of the preceding sections, we shall give a rather complete and self-contained introduction to this topic.

1.3.1 $\sigma$-Algebras Generated by Functions

In measure theory one is usually given a fixed $\sigma$-algebra with respect to which measurability is specified and on which all measures are defined. In probability theory, on the other hand, one often wants to deal with several different $\sigma$-algebras over a common underlying set, typically the sample space in question. There is a definite technique for constructing such $\sigma$-algebras, and our first task is to set forth the main elements of this construction.

Suppose that $(F, \mathcal{B})$ is a given measurable space and $f$ is a given mapping from some set $E$ into $F$. We emphasize that there is, as yet, no $\sigma$-algebra over $E$ and hence we cannot speak of the measurability of $f$. Now define a collection of subsets of $E$ by

$$\sigma\{f\} \triangleq \{f^{-1}(B) : B \in \mathcal{B}\}.$$
One trivially verifies that $\sigma\{f\}$ is a $\sigma$-algebra over $E$, that $f$ is $\sigma\{f\}/\mathcal{B}$-measurable, and that, for any $\sigma$-algebra $\mathcal{S}$ over $E$ such that $f$ is $\mathcal{S}/\mathcal{B}$-measurable, we have $\sigma\{f\} \subset \mathcal{S}$. In short, the $\sigma$-algebra $\sigma\{f\}$ defined in (1.28) is the smallest among all $\sigma$-algebras $\mathcal{S}$ over $E$ such that $f$ is $\mathcal{S}/\mathcal{B}$-measurable. It is called the $\sigma$-algebra generated by the mapping $f$. The next result establishes that $f$ “completely determines” functions on $E$ which are $\sigma\{f\}$-measurable:

**Theorem 1.3.1 (Doob)** Suppose that $E$ is a set, $(F, \mathcal{B})$ is a measure space, and $f : E \to F$ is a given mapping. If a mapping $Z : E \to \overline{\mathbb{R}}$ is $\sigma\{f\}$-measurable then there exists some $\mathcal{B}$-measurable mapping $\Psi : F \to \overline{\mathbb{R}}$ such that $Z(x) = \Psi(f(x)), \forall x \in E$.

Proof: Observe that, when $A \in \sigma\{f\}$ then $A = f^{-1}(B)$ for some $B \in \mathcal{B}$, whence $I_A(x) = I_{f^{-1}(B)}(x) = I_B(f(x))$ for all $x \in E$. Now suppose that $Z$ is some simple $\sigma\{f\}$-measurable function from $E$ into $\overline{\mathbb{R}}$. Then, by the preceding observation,

$$Z(x) = \sum_{i=1}^{n} \alpha_i I_{A_i}(x) \equiv \sum_{i=1}^{n} \alpha_i I_{B_i}(f(x)), \quad \forall x \in E,$$

where $\alpha_i \in \mathbb{R}$, $A_i \in \sigma\{f\}$, and $B_i \in \mathcal{B}$ are such that $A_i = f^{-1}(B_i), i = 1, 2, \ldots n$. If we define $\Psi : F \to \overline{\mathbb{R}}$ by

$$\Psi(\xi) \triangleq \sum_{i=1}^{n} \alpha_i I_{B_i}(\xi), \quad \forall \xi \in F,$$

then it follows that $\Psi$ is a $\mathcal{B}$-measurable simple function, and $Z(x) = \Psi(f(x)), \forall x \in E$. The theorem has thus been established when $Z$ is a simple $\sigma\{f\}$-measurable function. Next, suppose that $Z$ is a $[0, \infty]$-valued, $\sigma\{f\}$-measurable, function on $E$. By Proposition 1.2.9 there is a sequence \{$Z_n, n = 1, 2, \ldots$\} of simple $\sigma\{f\}$-measurable functions on $E$ such that \{$Z_n(x), n = 1, 2, \ldots$\} converges (monotonically) to $Z(x)$ for each $x$, and we have also just shown that $Z_n \equiv \Psi_n(f), \forall n = 1, 2, \ldots$, where $\Psi_n$ is some $\mathcal{B}$-measurable simple function from $F$ into $\overline{\mathbb{R}}$. Let

$$C \triangleq \{\xi \in F : \liminf_{n \to \infty} \Psi_n(\xi) = \limsup_{n \to \infty} \Psi_n(\xi)\},$$

and define $\Psi : F \to \overline{\mathbb{R}}$ as follows: $\Psi(\xi) \triangleq \lim_{n \to \infty} \Psi_n(\xi), \forall \xi \in C$ and $\Psi(\xi) \triangleq 0, \forall \xi \notin C$. From Proposition 1.2.6 we observe that $C \in \mathcal{B}$ and $\Psi$ is $\mathcal{B}$-measurable. Since

$$\lim_{n \to \infty} \Psi_n(f(x)) = \lim_{n \to \infty} Z_n(x) = Z(x), \quad \forall x \in E,$$

it follows that $f(x) \in C, \forall x \in E$. Thus, from the definition of $\Psi$, we get

$$\Psi(f(x)) = \lim_{n \to \infty} \Psi_n(f(x)) = Z(x), \quad \forall x \in E.$$

This establishes the theorem when $Z$ is non-negative valued and $\sigma\{f\}$-measurable. Finally, for the case where $Z$ is a general $\sigma\{f\}$-measurable function taking both positive and negative values, apply the preceding result to the positive and negative parts of $Z$. $\blacksquare$
Remark 1.3.2 We next want to introduce the notion of a $\sigma$-algebra generated by more than one mapping on a set $E$. To this end it is useful to establish the following notation: suppose that $\mathcal{G}_\lambda$ is a $\sigma$-algebra over some set $E$ for each $\lambda$ in some “indexing” set $\Lambda$. We shall write $\sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$ to denote the $\sigma$-algebra over $E$ generated by the collection of sets

$$
\mathcal{C} \triangleq \{A \subset E : A \in \mathcal{G}_\lambda \text{ for some } \lambda \in \Lambda\} \equiv \bigcup_{\lambda \in \Lambda} \mathcal{G}_\lambda
$$

(recall Remark 1.2.1). Then we see that $\mathcal{G}_\lambda \subset \sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$ for each $\lambda \in \Lambda$ and that, for any $\sigma$-algebra $\mathcal{S}$ over $E$ such that $\mathcal{G}_\lambda \subset \mathcal{S}, \forall \lambda \in \Lambda$, we have $\sigma\{\mathcal{G}_\lambda : \lambda \in \Lambda\} \subset \mathcal{S}$. Thus, $\sigma\{\mathcal{G}_\lambda : \lambda \in \Lambda\}$ is the smallest or minimal $\sigma$-algebra over $E$ which includes the sets in $\mathcal{G}_\lambda$ for each $\lambda \in \Lambda$.

Now suppose we are given a collection of measurable spaces $\{(F_\lambda, \mathcal{B}_\lambda), \lambda \in \Lambda\}$ where $\Lambda$ is an indexing set, and that, for each $\lambda \in \Lambda$, we have a mapping $f_\lambda$ defined on some set $E$ and assuming values in $F_\lambda$. Let $\mathcal{G}$ be the $\sigma$-algebra over $E$ defined by

$$
\mathcal{G} \triangleq \sigma\{f_\lambda, \lambda \in \Lambda\}.
$$

Clearly the mapping $f_\lambda$ is $\mathcal{G}/\mathcal{B}_\lambda$-measurable for each $\lambda \in \Lambda$, and, in view of Remark 1.3.2, for any $\sigma$-algebra $\mathcal{S}$ over $E$ such that $f_\lambda$ is $\mathcal{S}/\mathcal{B}_\lambda$-measurable for each $\lambda \in \Lambda$, we have $\mathcal{G} \subset \mathcal{S}$. Thus $\mathcal{G}$ is the smallest among all $\sigma$-algebras $\mathcal{S}$ over $E$ with the property that $f_\lambda$ is $\mathcal{S}/\mathcal{B}_\lambda$-measurable for each $\lambda \in \Lambda$. We call $\mathcal{G}$ the $\sigma$-algebra generated by the set of mappings $\{f_\lambda, \lambda \in \Lambda\}$. It is usual to slightly abuse notation and write $\sigma\{f_\lambda, \lambda \in \Lambda\}$ for this $\sigma$-algebra (strictly speaking we should use the right hand side of (1.30)).

We next look at the structure of $\sigma\{f_\lambda, \lambda \in \Lambda\}$ in the case where $\Lambda$ is uncountably infinite. To this end we need the following result about uncountably infinite collections of $\sigma$-algebras:

**Proposition 1.3.3** Suppose $\Lambda$ is an uncountably infinite indexing set and $\mathcal{G}_\lambda$ is a $\sigma$-algebra over some set $E$ for each $\lambda \in \Lambda$. Then, for each $A \in \sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$, there is some countable subset $\Xi \subset \Lambda$, generally depending on the set $A$, such that $A \in \sigma\{\mathcal{G}_\lambda, \lambda \in \Xi\}$.

**Proof:** Define

$$
\mathcal{H} \triangleq \{A \in \sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\} : \exists \text{ countable set } \Xi \subset \Lambda \text{ s.t. } A \in \sigma\{\mathcal{G}_\lambda, \lambda \in \Xi\}\}.
$$

Clearly $E \in \mathcal{H}$, and $A \in \mathcal{H}$ implies that $A^c \in \mathcal{H}$. Finally, since a countable union of countable subsets of $\Lambda$ is also a countable subset of $\Lambda$, one easily checks that the union of a sequence of sets in $\mathcal{H}$ is a member of $\mathcal{H}$. Thus $\mathcal{H}$ is a $\sigma$-algebra over $E$. Moreover, it is clear that $\mathcal{G}_\lambda \subset \mathcal{H}$ for each $\lambda \in \Lambda$, and hence $\mathcal{H} = \sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$, as required. \[\square\]

**Corollary 1.3.4** Suppose $\Lambda$ is an uncountably infinite indexing set and $\mathcal{G}_\lambda$ is a $\sigma$-algebra over some set $E$ for each $\lambda \in \Lambda$. If a mapping $Z : E \to \overline{\mathbb{R}}$ is $\sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$-measurable then there is a countable subset $\Xi \subset \Lambda$, generally depending on the mapping $Z$, such that $Z$ is $\sigma\{\mathcal{G}_\lambda, \lambda \in \Xi\}$-measurable.
Proof: The assertion is an immediate consequence of Proposition 1.3.3 in the case where $Z = I_A$ for some $A \in \sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$, and thus also holds when $Z$ is a $\sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$-measurable simple function. Now suppose that $Z$ is a $[0, \infty]$-valued and $\sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$-measurable function. By Propositions 1.2.6 and 1.2.9, together with the fact that a countable union of countable subsets of $\Lambda$ is also a countable subset of $\Lambda$, one readily checks that $Z$ is $\{\mathcal{G}_\lambda, \lambda \in \Xi\}$-measurable for some countable set $\Xi \subset \Lambda$. Finally, when $Z$ is an $\mathbb{R}$-valued, $\sigma\{\mathcal{G}_\lambda, \lambda \in \Lambda\}$-measurable mapping on $E$, we just apply this result to the positive and negative parts of $Z$.

For later applications we re-state Corollary 1.3.4 as follows:

**Corollary 1.3.5** Suppose $\Lambda$ is an uncountably infinite indexing set, $\{(F_\lambda, \mathcal{B}_\lambda), \lambda \in \Lambda\}$ is a collection of measurable spaces, and $f_\lambda : E \to F_\lambda$ is a given mapping on some set $E$ for each $\lambda \in \Lambda$. Then a mapping $Z : E \to \mathbb{R}$ is $\sigma\{f_\lambda, \lambda \in \Lambda\}$-measurable if and only if there is a countable set $\Xi \subset \Lambda$, generally depending on the mapping $Z$, such that $Z$ is $\sigma\{f_\lambda, \lambda \in \Xi\}$-measurable.

**Remark 1.3.6** In applications where one comes across the $\sigma$-algebra $\sigma\{f_\lambda, \lambda \in \Lambda\}$ it is often the case that the mappings $f_\lambda$ take values in a common Euclidean space $\mathbb{R}^d$, so that we identify $\left((F_\lambda, \mathcal{B}_\lambda) \triangleq (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\right)$ for each $\lambda \in \Lambda$. If, in this case, we expand the $\mathbb{R}^d$-valued mappings $f_\lambda$ as vectors of $\mathbb{R}$-valued functions, namely $f_\lambda \triangleq (f^1_\lambda, f^2_\lambda, \ldots, f^d_\lambda)$, then it is easily verified that

$$\sigma\{f_\lambda : \lambda \in \Lambda\} = \sigma\{f^i_\lambda : \lambda \in \Lambda, i = 1, 2, \ldots d\}.$$

### 1.3.2 $\sigma$-algebras on Function Spaces

We are now ready to formulate the concept of an arbitrary product of a given measurable space. Our first definition deals with a simpler notion, namely an arbitrary product of a given set:

**Definition 1.3.7** Suppose that $F$ and $\Lambda$ are given sets. We shall write $F^\Lambda$ or $\bigotimes_{\lambda \in \Lambda} F$ to denote the set of all mappings defined on $\Lambda$ with values in $F$. We call this set of mappings the $\Lambda$-fold product of $F$.

Since $F^\Lambda$ is a collection of mappings, it is also called a **function space**.

**Example 1.3.8** (a) Suppose $F \triangleq \mathbb{R}$ and $\Lambda \triangleq \{1, 2, \ldots, n\}$ for some integer $n \geq 1$. If we interpret a member $x = (x^1, x^2, \ldots, x^n)$ of $\mathbb{R}^n$ as a mapping from the domain $\{1, 2, \ldots, n\}$ into $\mathbb{R}$, then clearly $F^\Lambda = \mathbb{R}^n$.

(b) Suppose $F \triangleq \mathbb{R}$ and $\Lambda \triangleq \{0, 1, 2, \ldots\}$. Then $F^\Lambda$ denotes the set of all $\mathbb{R}$-valued mappings on the domain $\{0, 1, 2, \ldots\}$. Equivalently, we can regard $F^\Lambda$ as the set of all infinite sequences of real numbers which are indexed by the non-negative integers.

(c) Suppose $F \triangleq \mathbb{R}^d$ and $\Lambda \triangleq \{0, 1, 2, \ldots\}$. Then $F^\Lambda$ denotes the set of all $\mathbb{R}^d$-valued mappings defined on the domain $\{0, 1, 2, \ldots\}$.
(d) Suppose $F \triangleq \mathbb{R}^d$ and $\Lambda \triangleq [0, \infty)$. Then $F^\Lambda$ denotes the set of all $\mathbb{R}^d$-valued mappings defined on the domain $[0, \infty)$.

If $\psi \in F^\Lambda$ and $\lambda \in \Lambda$, then $\psi(\lambda)$ denotes the “value” in $F$ assumed by the mapping $\psi$ when “evaluated” at $\lambda$. It is useful to formalize this notion in the following way:

**Definition 1.3.9** Suppose that $\Lambda$ and $F$ are sets. For each $\lambda \in \Lambda$ define the mapping $\mathfrak{x}_\lambda : F^\Lambda \rightarrow F$ as follows:

$$\mathfrak{x}_\lambda(\psi) \triangleq \psi(\lambda), \ \forall \psi \in F^\Lambda.$$  

The mapping $\mathfrak{x}_\lambda$ in (1.31) is called the **evaluation mapping** on $F^\Lambda$ at $\lambda$.

Next, suppose that $(F, \mathcal{B})$ is a measurable space. Having formulated the notion of the $\Lambda$-fold product $F^\Lambda$, we now want to use $\mathcal{B}$ to construct a $\sigma$-algebra over $F^\Lambda$.

**Definition 1.3.10** Suppose that $\Lambda$ is a set and $(F, \mathcal{B})$ is a measurable space. We shall use $\mathcal{B}^\Lambda$ or $\bigotimes_{\lambda \in \Lambda} \mathcal{B}$ to denote the $\sigma$-algebra over $F^\Lambda$ generated by the mappings

$$\psi \rightarrow \mathfrak{x}_\lambda(\psi) : F^\Lambda \rightarrow F$$

for each $\lambda \in \Lambda$ (recall Section 1.3.1). Thus,

$$\mathcal{B}^\Lambda \triangleq \sigma\{\mathfrak{x}_\lambda, \ \lambda \in \Lambda\}.$$  

This $\sigma$-algebra is called the **$\Lambda$-fold product** of the $\sigma$-algebra $\mathcal{B}$.

**Remark 1.3.11** Given a set $\Lambda$ and a measurable space $(F, \mathcal{B})$, we see that $\mathcal{B}^\Lambda$ is the smallest among all $\sigma$-algebras $\mathcal{S}$ over $F^\Lambda$ with the property that the evaluation mappings $\mathfrak{x}_\lambda : F^\Lambda \rightarrow F$ are $\mathcal{S}/\mathcal{B}$-measurable for each $\lambda \in \Lambda$. Definitions 1.3.7 and 1.3.10 yield a new measurable space $(F^\Lambda, \mathcal{B}^\Lambda)$ called the **$\Lambda$-fold product** of the measurable space $(F, \mathcal{B})$.

**Remark 1.3.12** We see from Definition 1.3.10 that $\mathcal{B}^\Lambda = \sigma\{\mathcal{D}\}$ where

$$\mathcal{D} \triangleq \{\mathfrak{x}_\lambda^{-1}(\Gamma) : \ \lambda \in \Lambda, \ \Gamma \in \mathcal{B}\}.$$  

The set $\mathfrak{x}_\lambda^{-1}(\Gamma) \equiv \{\psi \in F^\Lambda : \psi(\lambda) \in \Gamma\}$, where $\Gamma \in \mathcal{B}$, is called a **strip** in $F^\Lambda$. Thus, $\mathcal{B}^\Lambda$ is the smallest $\sigma$-algebra over $F^\Lambda$ which includes the strips $\mathfrak{x}_\lambda^{-1}(\Gamma)$ for every $\lambda \in \Lambda$ and $\Gamma \in \mathcal{B}$. Closely related to strips are rectangles, which are defined as follows: let $n$ be some positive integer, and fix some $\lambda_i \in \Lambda$ and $\Gamma_i \in \mathcal{B}$, $i = 1, 2, \ldots, n$. A subset of $F^\Lambda$ with the form

$$\text{rect}(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma_1, \ldots, \Gamma_n) \triangleq \bigcap_{i=1}^{n} \mathfrak{x}_{\lambda_i}^{-1}(\Gamma_i)$$

is called a **finite-dimensional rectangle** in $F^\Lambda$. Let $\mathcal{R}$ be the collection of all finite-dimensional rectangles in $F^\Lambda$. Clearly a strip is always a finite-dimensional rectangle, thus $\mathcal{D} \subset \mathcal{R}$, and therefore $\mathcal{B}^\Lambda \subset \sigma\{\mathcal{R}\}$. On the other hand, since a finite-dimensional rectangle is given by the intersection of finitely many strips, we see that $\mathcal{R} \subset \mathcal{B}^\Lambda$, thus $\sigma\{\mathcal{R}\} \subset \mathcal{B}^\Lambda$. It follows that $\mathcal{B}^\Lambda = \sigma\{\mathcal{R}\}$. 

25
Remark 1.3.13 In the case where \((F, B)\) is a measurable space and \(\Lambda \triangleq \{1, 2, \ldots, n\}\) for some positive integer \(n\), we shall usually write \(F^n\) and \(B^n\) respectively for the \(\Lambda\)-fold products of \(F\) and \(B\) (in place of the more correct but rather cumbersome \(F^{(1,2,\ldots,n)}\) and \(B^{(1,2,\ldots,n)}\)), and we shall call \(F^n, B^n\) and \((F^n, B^n)\) the \(n\)-fold products of \(F, B\) and \((F, B)\) respectively. If \(\Gamma_i \in B\), \(i = 1, 2, \ldots, n\), then we write

\[
\bigotimes_{i=1}^{n} \Gamma_i \triangleq \{ \psi = (\psi^1, \psi^2, \ldots, \psi^n) \in F^n : \psi^i \in \Gamma_i, \ i = 1, 2, \ldots, n \}
\]

for the rectangle with sides \(\Gamma_i, \ i = 1, 2, \ldots, n\). If \(R\) denotes the collection of all such rectangular subsets of \(F^n\) then, by Remark 1.3.12, we see that \(\sigma\{R\} = B^n\).

Example 1.3.14 Suppose \(\Lambda \triangleq \{1, 2, \ldots, n\}\) and \((F, B) \triangleq (IR, B(IR))\). Then it is a simple exercise to verify that \((F^n, B^n) = (IR^n, B(IR^n))\).

Remark 1.3.15 Strips and finite-dimensional rectangles are particularly simple subsets of \(F^\Lambda\), and are the building blocks for the \(\sigma\)-algebra \(B^\Lambda\), as is seen from Remark 1.3.12. There is also another useful class of subsets of \(F^\Lambda\), rather more general than strips or rectangles, which we can characterize as follows: Let \(n\) be a positive integer, fix some \(\lambda_i \in \Lambda, \forall i = 1, 2, \ldots, n\), and let \(\Gamma \in B^n\). A subset of \(F^\Lambda\) with the form

\[
cyl(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma) \triangleq \{ \psi \in F^\Lambda : (\mathcal{X}_{\lambda_1}(\psi), \mathcal{X}_{\lambda_2}(\psi), \ldots, \mathcal{X}_{\lambda_n}(\psi)) \in \Gamma \}
\]

is called a finite-dimensional \textbf{cylinder set} in \(F^\Lambda\), and the set \(\Gamma\) is called a \textbf{base} of the cylinder set. For obvious reasons the cylinder set in (1.34) is said to have an \(n\)-\textbf{dimensional base}. Clearly, a finite-dimensional rectangle having the form given by (1.32) is a finite-dimensional cylinder set with a base which is given by (1.33).

Proposition 1.3.16 Suppose that \((F, B)\) is a measurable space and \(\Lambda\) is a set. Let \(C\) denote the collection of all finite-dimensional cylinder sets in \(F^\Lambda\). Then:

(a) \(C\) is an algebra over \(F^\Lambda\);

(b) \(\sigma\{C\} = B^\Lambda\).

Proof: (a) Showing that \(C\) is an algebra over \(F^\Lambda\) involves only elementary set manipulations and is left to the reader.

(b) Since the collection of finite-dimensional rectangles in \(F^\Lambda\) is included within \(C\), we see from Remark 1.3.12 that \(B^\Lambda \subset \sigma\{C\}\). To establish the opposite set inclusion fix some positive integer \(n\), some \(\lambda_i \in \Lambda, \ i = 1, 2, \ldots, n\), and put

\[
\mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \triangleq \{ \Gamma \in B^n : cyl(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma) \in B^\Lambda \}.
\]

Now it is clear from (1.34) that, for any \(\Gamma \in B^n\), we have

\[
[cyl(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma)]^c = cyl(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma^c),
\]

26
and, for any sequence \( \{ \Gamma_k, \ k = 1, 2, \ldots \} \) of sets in \( \mathcal{B}^n \),

\[
\text{cyl}(\lambda_1, \lambda_2, \ldots, \lambda_n; \bigcup_{1 \leq k < \infty} \Gamma_k) = \bigcup_{1 \leq k < \infty} \text{cyl}(\lambda_1, \lambda_2, \ldots, \lambda_n; \Gamma_k).
\]

From these observations it follows that \( \mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \) is a \( \sigma \)-algebra over \( F^n \). Moreover, since \( \mathcal{B}^\Lambda \) includes all finite-dimensional rectangles with the form (1.32), it follows that \( \mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \) includes all rectangles in \( F^n \) with the form (1.33), for \( \Gamma_i \in \mathcal{B}, \ i = 1, 2, \ldots, n \). Thus, in view of Remark 1.3.13, we see that \( \mathcal{B}^n \subset \mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \), and therefore, since (1.35) shows that \( \mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \) is included within \( \mathcal{B}^n \), we have \( \mathcal{B}^n = \mathcal{A}(\lambda_1, \lambda_2, \ldots, \lambda_n) \). Thus, in view of the arbitrary choice of the \( n \)-fold sequence \( \{ \lambda_1, \lambda_2, \ldots, \lambda_n \} \), it is clear that \( \mathcal{C} \subset \mathcal{B}^\Lambda \), so that \( \sigma\{\mathcal{C}\} \subset \mathcal{B}^\Lambda \) as required.

**Remark 1.3.17** We are now going to generalize Theorem 1.3.1 to the case of arbitrarily many mappings from a set \( E \) into a measure space \( (F, \mathcal{B}) \). To this end the following notation is useful: when \( \Lambda \) is the set of non-negative integers, namely \( \Lambda \triangleq \{0, 1, 2, \ldots\} \), then we shall write \( F^\infty \) and \( \mathcal{B}^\infty \) for the \( \Lambda \)-fold products \( F^\Lambda \) and \( \mathcal{B}^\Lambda \) respectively. Moreover, we shall frequently write elements \( x \in F^\infty \) in expanded form as follows,

\[
(x_0, x_1, x_2, \ldots), \ x_i \in F,
\]

and, if \( \Psi \) is a function on \( F^\infty \), then we shall write the value \( \Psi(x) \) at some \( x \in F^\infty \) as

\[
\Psi(x_0, x_1, x_2, \ldots).
\]

**Remark 1.3.18** Suppose that \( E \) is some set, \( (F, \mathcal{B}) \) is a measure space, and \( f_n : E \to F, \ \forall \ n = 0, 1, 2, \ldots, \) is a given sequence of mappings. In accordance with Section 1.3.1 we can define the \( \sigma \)-algebras

\[
\sigma\{f_n\} \triangleq \{f_n^{-1}(\Gamma) : \ \Gamma \in \mathcal{B}\}, \ \forall \ n = 0, 1, 2, \ldots
\]

and

\[
\sigma\{f_n, \ n = 0, 1, 2, \ldots\} \triangleq \sigma\{\sigma\{f_n\}, \ n = 0, 1, 2, \ldots\}.
\]

over the set \( E \). Also, define the mapping \( f : E \to F^\infty \) by

\[
f(x) \triangleq (f_0(x), f_1(x), f_2(x), \ldots), \ \forall \ x \in E.
\]

Then we can define another \( \sigma \)-algebra over \( E \), namely

\[
\sigma\{f\} \triangleq \{f^{-1}(A) : \ A \in \mathcal{B}^\infty\}.
\]

The next proposition shows that the \( \sigma \)-algebras in (1.37) and (1.38) are identical:

**Proposition 1.3.19** With reference to Remark 1.3.18, we have \( \sigma\{f\} = \sigma\{f_n, \ n = 0, 1, 2, \ldots\} \).
Proof: To show that $\sigma\{f\} \subset \sigma\{f_n, n = 0, 1, 2, \ldots\}$, let $\mathcal{R}$ be the collection of all finite-dimensional measurable rectangles in $\mathcal{B}^\infty$, namely sets of the form

$$A = \{\psi \in F^\infty : \chi_{r_i}(\psi) \in \Gamma_i, \ \forall i = 1, 2, \ldots, m\}$$

for some positive integer $m$, non-negative integers $r_i$ and sets $\Gamma_i \in \mathcal{B}$, $i = 1, 2, \ldots, m$. Clearly

$$f^{-1}(A) = \{x \in E : \chi_{r_i}(f(x)) \in \Gamma_i, \ \forall i = 1, 2, \ldots, m\}$$

(1.39)

$$= \bigcap_{i=1}^{m}\{x \in E : f_{r_i}(x) \in \Gamma_i\}.$$  

Thus $f^{-1}(A) \in \sigma\{f_n, n = 0, 1, 2, \ldots\}$ for all $A \in \mathcal{R}$. Next, define

$$A \triangleq \{A \subset F^\infty : f^{-1}(A) \in \sigma\{f_n, n = 0, 1, 2, \ldots\}\}.$$

(1.40)

Routine set manipulations show that $\mathcal{A}$ is a $\sigma$-algebra over $F^\infty$, and we have seen that $\mathcal{R} \subset \mathcal{A}$. In view of Remark 1.3.12, we get $\mathcal{B}^\infty = \sigma\{\mathcal{R}\} \subset \mathcal{A}$, and thus, from (1.40),

$$\{f^{-1}(A) : A \in \mathcal{B}^\infty\} \subset \sigma\{f_n, n = 0, 1, 2, \ldots\}.$$

Equivalently, from (1.38), we have $\sigma\{f\} \subset \sigma\{f_n, n = 0, 1, 2, \ldots\}$. Establishing the opposite set inclusion involves only routine set manipulations and is left to the reader.

The next result generalizes Theorem 1.3.1 to the case of countably many mappings:

**Theorem 1.3.20** Suppose that $E$ is a set, $(F, \mathcal{B})$ is a measure space, and $f_n : E \to F$, $n = 0, 1, 2, \ldots$, is a given sequence of mappings. If a mapping $Z : E \to \mathbb{R}$ is $\sigma\{f_n, n = 0, 1, 2, \ldots\}$-measurable, then there exists some $\mathcal{B}^\infty$-measurable mapping $\Psi : F^\infty \to \mathbb{R}$, such that

$$Z(x) = \Psi(f_0(x), f_1(x), f_2(x), \ldots), \ \forall x \in E.$$  

(1.41)

Proof: Define the mapping $f : E \to F^\infty$ by

$$f(x) \triangleq (f_0(x), f_1(x), f_2(x), \ldots), \ \forall x \in E.$$  

In view of Proposition 1.3.19, it follows that $Z$ is $\sigma\{f\}$-measurable. Applying Theorem 1.3.1, but with $(F, \mathcal{B})$ there interpreted as $(F^\infty, \mathcal{B}^\infty)$, we see that there is some $\mathcal{B}^\infty$-measurable mapping $\Psi : F^\infty \to \mathbb{R}$ such that $Z(x) = \Psi(f(x))$, $\forall x \in E$.  

For completeness we state the following special case of Theorem 1.3.20, which follows by an obviously similar argument:
Theorem 1.3.21 Suppose that $E$ is a set, $(F, \mathcal{B})$ is a measurable space, and $f_i : E \to F$, $i = 1, 2, \ldots, n$, are mappings. If a mapping $Z : E \to \overline{F}$ is $\sigma\{f_1, f_2, \ldots, f_n\}$-measurable, then there is some $\mathcal{B}^\sigma$-measurable mapping $\Psi : F^n \to \overline{F}$, such that

$$Z(x) = \Psi(f_1(x), f_2(x), \ldots, f_n(x)), \quad \forall x \in E.$$ 

Next, we consider the case where there are uncountably infinitely many mappings on $E$ into $(F, \mathcal{B})$:

Theorem 1.3.22 Suppose that $E$ is a set, $(F, \mathcal{B})$ is a measurable space, and $f_\lambda : E \to F$ is a mapping for each $\lambda$ in some uncountably infinite indexing set $\Lambda$. If a mapping $Z : E \to \overline{F}$ is $\sigma\{f_\lambda, \lambda \in \Lambda\}$-measurable, then there exists some sequence $\{\lambda_0, \lambda_1, \lambda_2, \ldots\} \subset \Lambda$, and some $\mathcal{B}^\infty$-measurable mapping $\Psi : F^\infty \to \overline{F}$, such that

$$(1.42) \quad Z(x) = \Psi(f_{\lambda_0}(x), f_{\lambda_1}(x), f_{\lambda_2}(x), \ldots), \quad \forall x \in E.$$ 

Proof: In view of Corollary 1.3.5 there is some countable set $\Xi \subset \Lambda$ such that $Z$ is $\sigma\{f_\lambda, \lambda \in \Xi\}$-measurable. Let the sequence $\{\lambda_0, \lambda_1, \lambda_2, \ldots\}$ be some arbitrarily chosen but fixed enumeration of the countable set $\Xi$. The existence of a measurable mapping $\Psi : (F^\infty, \mathcal{B}^\infty) \to \overline{F}$ such that (1.42) holds, is now an immediate consequence of Proposition 1.3.20 applied to the sequence of mappings $\{f_{\lambda_n}, n = 0, 1, 2, \ldots\}$. □

Remark 1.3.23 Theorem 1.3.22 really asserts that a mapping which is $\sigma\{f_\lambda, \lambda \in \Lambda\}$-measurable, for $\Lambda$ an uncountably infinite set, is completely determined by only countably many of the functions $f_\lambda$.

We can use the preceding results to get a useful characterization of the sets in $\mathcal{B}^{\Lambda}$ when $\Lambda$ is uncountably infinite:

Theorem 1.3.24 Suppose that $\Lambda$ is an uncountably infinite set and $(F, \mathcal{B})$ is a measure space. If $A \in \mathcal{B}^{\Lambda}$ then there is some sequence $\{\lambda_0, \lambda_1, \lambda_2, \ldots\} \subset \Lambda$ and some $\Gamma \in \mathcal{B}^\infty$ such that

$$(1.43) \quad A = \{\psi \in F^{\Lambda} : (\mathcal{X}_{\lambda_0}(\psi), \mathcal{X}_{\lambda_1}(\psi), \mathcal{X}_{\lambda_2}(\psi), \ldots) \in \Gamma\}.$$ 

Proof: By Definition 1.3.10 we have $\mathcal{B}^{\Lambda} = \sigma\{\mathcal{X}_\lambda, \lambda \in \Lambda\}$. Thus, if $A \in \mathcal{B}^{\Lambda}$, then, interpreting $E$ and $f_\lambda$ in Proposition 1.3.22 as $F^{\Lambda}$ and $\mathcal{X}_\lambda$ respectively, we get some sequence $\{\lambda_0, \lambda_1, \lambda_2, \ldots\} \subset \Lambda$ and some measurable mapping $\Psi : (F^\infty, \mathcal{B}^\infty) \to \overline{F}$ such that

$$(1.44) \quad I_A(\psi) = \Psi(\mathcal{X}_{\lambda_0}(\psi), \mathcal{X}_{\lambda_1}(\psi), \mathcal{X}_{\lambda_2}(\psi), \ldots), \quad \forall \psi \in F^{\Lambda}.$$ 

Next, fix some $x \overset{\Delta}{=} (x_0, x_1, x_2, \ldots) \in F^\infty$, and let $\psi$ be any member of $F^{\Lambda}$ such that $\psi(\lambda_i) = x_i$, $\forall i = 0, 1, 2, \ldots$ Then either $\psi \in A$, in which case it follows from (1.44) that $\Psi(x) = 1$, or $x \notin A$, in which case we similarly obtain $\Psi(x) = 0$. Thus, $\Psi$ is a two-valued mapping assuming values in $\{0, 1\}$, and (1.43) follows upon defining $\Gamma \overset{\Delta}{=} \{x \in F^\infty : \Psi(x) = 1\}$. □

Remark 1.3.25 One can interpret Theorem 1.3.24 as saying that, even when $\Lambda$ is uncountably infinite, each set $A$ in the $\Lambda$-fold product $\mathcal{B}^{\Lambda}$ is really just a type of cylinder set, rather similar in form to (1.34), but with a base whose dimension is countably infinite.
1.4 Probability Theory

1.4.1 Probability Spaces, Random Variables, and Expectation

A probability space is a finite measure space \((E, S, \mu)\) with \(\mu(E) = 1\). To connote the specifically probabilistic aspect of these spaces it customary to assign them the distinctive notation \((\Omega, \mathcal{F}, P)\), in which the set \(\Omega\) is called the sample space, members of the \(\sigma\)-algebra \(\mathcal{F}\) are dubbed events, and the measure \(P\) is the probability measure on the sample space. An \(\mathbb{R}\)-valued \(\mathcal{F}/\mathcal{B}(\mathbb{R})\)-measurable mapping \(X\) defined on a probability space \((\Omega, \mathcal{F}, P)\) is called a random variable. Extending this terminology, an \(\mathbb{R}^d\)-valued mapping \(X\) on a probability space \((\Omega, \mathcal{F}, P)\) is called an \(\mathbb{R}^d\)-valued random variable or random vector when \(X\) is \(\mathcal{F}/\mathcal{B}(\mathbb{R}^d)\)-measurable. If an \(\mathbb{R}^d\)-valued mapping \(X\) on a probability space \((\Omega, \mathcal{F}, P)\) is expanded as a vector of \(\mathbb{R}\)-valued mappings, namely \(X = (X^1, X^2, \ldots X^d)^T\), then it is readily seen that \(X\) is a random vector if and only if each \(X^i\) is a random variable.

Suppose \(X = (X^1, X^2, \ldots X^d)\) is some \(\mathbb{R}^d\)-valued random variable defined on a probability space \((\Omega, \mathcal{F}, P)\). Define the mapping \(\mu_X : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]\) by \(\mu_X(A) \triangleq P(X^{-1}(A))\). In the terminology of §1.2.4, \(\mu_X\) is the image measure of \(P\) through \(X\). In the present probabilistic context we prefer to call \(\mu_X\) the distribution measure generated by \(X\). For each \(x = (x^1, x^2, \ldots x^d)\) in \(\mathbb{R}^d\) we use the notation \(\{X \leq x\}\) to indicate the event \(\{\omega : X^i(\omega) \leq x^i, \ \forall \ i = 1, 2, \ldots d\}\). The mapping \(F_X : \mathbb{R}^d \to [0, 1]\) defined by \(F_X(x) \triangleq P(\{X \leq x\})\), \(\forall \ x \in \mathbb{R}^d\), is called the cumulative distribution function of \(X\). It is clear that \(F_X(x) = \mu_X((-\infty, x])\), \(\forall \ x \in \mathbb{R}^d\). The collection of sets \(D \triangleq \{(-\infty, x] : x \in \mathbb{R}^d\}\) is a \(\pi\)-class over \(\mathbb{R}^d\), and in view of Example 1.2.2(d) we see that \(\mathcal{B}(\mathbb{R}^d) = \sigma(D)\). We then have the following immediate consequence of Theorem 1.2.4:

**Proposition 1.4.1** Suppose \(X\) and \(\tilde{X}\) are \(\mathbb{R}^d\)-valued random variables on probability spaces \((\Omega, \mathcal{F}, P)\) and \((\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})\) respectively, with identical distribution functions: \(F_X(x) = F_{\tilde{X}}(x)\), \(\forall \ x \in \mathbb{R}^d\). Then \(\mu_X(A) = \mu_{\tilde{X}}(A)\), \(\forall \ A \in \mathcal{B}(\mathbb{R}^d)\).

If there exists a \([0, \infty]\)-valued measurable mapping \(f_X\) on \((\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))\) such that

\[
F_X(x) = \int_{(-\infty, x]} f_X d\lambda, \quad \forall \ x \in \mathbb{R}^d,
\]

then \(X\) is said to have a density function \(f_X\) (here \(\lambda\) denotes Lebesgue measure on \(\mathcal{B}(\mathbb{R}^d)\)). It is now easily verified that \(\mu_X \ll \lambda[\mathcal{B}(\mathbb{R}^d)]\) and

\[
f_X = \frac{d\mu_X}{d\lambda}, \quad \lambda - \text{a.e.}
\]

Suppose that \(X\) is random variable on some probability space \((\Omega, \mathcal{F}, P)\). The expectation of \(X\), denoted by \(EX\) or \(E[X]\), is defined to be the Lebesgue integral of \(X\) on \(\Omega\) with respect to the measure \(P\):

\[
EX \triangleq \int_{\Omega} X dP.
\]
The conditions governing existence of this integral are exactly those that pertain to general Lebesgue integrals: if $X$ assumes values in $[0, \infty)$ then $EX$ is always defined and may take the value $+\infty$. On the other hand, when $X$ is $\mathbb{R}$-valued, then $EX$ is defined if and only if $E|X| < \infty$ in which case $EX$ is a real number. A random variable $X$ is called integrable when $E|X| < \infty$, and square integrable when $E[X^2] < \infty$. In accordance with the notation established in §1.2.5, for each $p \in (0, \infty)$ we use $L^p(\Omega, \mathcal{F}, P)$ to denote the collection of all random variables $X$ on the probability space $(\Omega, \mathcal{F}, P)$ such that $|X|^p$ is $P$-integrable; when $p \in [1, \infty)$ then $L^p(\Omega, \mathcal{F}, P)$ is a Banach space with norm $\|X\|_p \triangleq \left\{E[|X|^p]\right\}^{1/p}$ (Theorem 1.2.27). A random variable $X$ on $(\Omega, \mathcal{F}, P)$ is called bounded when there is some constant $B \in (0, \infty)$ such that $|X(\omega)| \leq B$, $\forall \omega \in \Omega$. A bounded random variable clearly belongs to $L^p(\Omega, \mathcal{F}, P)$ for each $p \in (0, \infty)$.

If $X$ is a random variable on probability space $(\Omega, \mathcal{F}, P)$ and $A \in \mathcal{F}$, then we use the notation $E[X; A]$ or $E[X1_A]$ to denote the expectation of $X$ over $A$:

$$E[X; A] \triangleq E[X1_A] \equiv \int_A X \, dP,$$

assuming that the integral in question is defined.

If $X$ is a square integrable random variable then we define its variance, written $\text{Var}(X)$, as

$$\text{Var}(X) \triangleq E[(X - EX)^2],$$

and, if $X$ and $Y$ are two square integrable random variables on a common probability space, we define the covariance if $X$ with $Y$, written $\text{Cov}(X, Y)$, as

$$\text{Cov}(X, Y) \triangleq E[(X - EX)(Y - EY)].$$

An immediate consequence of these definitions, together with the trivial identity

$$xy = \frac{1}{4}\{(x + y)^2 - (x - y)^2\}, \quad x, y \in \mathbb{R}^d,$$

is that

$$(1.46) \quad \text{Cov}(X, Y) = \frac{1}{4}\{\text{Var}(X + Y) - \text{Var}(X - Y)\}$$

for two square integrable random variables $X$ and $Y$ on the same probability space. A structural relationship analogous to (1.46) will show up in Chapter 4 when we look at the quadratic and co-quadratic variation processes of square integrable martingales.

Some of the preceding ideas extend to a multidimensional setting. An $\mathbb{R}^d$-valued random vector $X \triangleq (X^1, X^2, \ldots X^d)$ on a probability space $(\Omega, \mathcal{F}, P)$ is called integrable when each $X^i$ is integrable, and is called square integrable when each $X^i$ is square integrable. If $X$ is integrable then we use $EX$ or $E[X]$ to denote the member of $\mathbb{R}^d$ whose $i$-th scalar element is $EX^i$. If $X \triangleq (X^1, X^2, \ldots X^d)$ and $Y \triangleq (Y^1, Y^2, \ldots Y^d)$ are square integrable random vectors on a probability space $(\Omega, \mathcal{F}, P)$ we define...
the covariance of $X$ with $Y$, written $\text{Cov}(X, Y)$, to be the $d$ by $d$ matrix whose $(i, j)$-th scalar element is $\text{Cov}(X^i, Y^j)$, and we define the variance of $X$, written $\text{Var}(X)$, by $\text{Var}(X) \triangleq \text{Cov}(X, X)$.

Suppose that $X$ is an $\mathbb{R}^d$-valued random vector on $(\Omega, \mathcal{F}, P)$, and $g : \mathbb{R}^d \to [0, \infty)$ is $\mathcal{B}(\mathbb{R}^d)$-measurable. In view of Lemma 1.2.19 it follows that

$$\text{Cov}(g(X), Y) = \int_{\mathbb{R}^d} g d\mu_X.$$

On the other hand, if $g : \mathbb{R}^d \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R}^d)$-measurable, then $E|g(X)| < \infty$ if and only if $g$ is $\mu_X$ integrable, in which case (1.47) holds. If, moreover, $X$ has a density function $f_X$, and $g : \mathbb{R}^d \to [0, \infty)$ is $\mathcal{B}(\mathbb{R}^d)$-measurable, then we see from (1.47) and Theorem 1.2.17 that

$$E[g(X)] = \int_{\mathbb{R}^d} g f_X d\lambda,$$

while, if $g$ takes values in $\mathbb{R}$, then $g f_X$ is $\lambda$-integrable if and only if $g$ is $\mu_X$-integrable, in which case (1.48) continues to hold.

The basic inequalities of Hölder and Minkowski continue to hold in the spaces of random variables $L^p(\Omega, \mathcal{F}, P)$ subject to the restrictions on $p$ in Theorems 1.2.25 and 1.2.26, but the fact that $P(\Omega) = 1$ allows us to formulate yet another essential inequality. For this purpose we need the notion of a convex function: a function $c$ from an open interval $I \subset \mathbb{R}$ into $\mathbb{R}$ is a convex function on $I$ if, whenever $\alpha \in [0, 1]$ and $x_1, x_2 \in I$, we have

$$c(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha c(x_1) + (1 - \alpha)c(x_2).$$

A convex function $c$ on an open interval $I$ is necessarily continuous on $I$ and thus, in particular, $c(X)$ is a random variable on $(\Omega, \mathcal{F}, P)$ when $X$ is a random variable on $(\Omega, \mathcal{F}, P)$ with $X(\omega) \in I$, $\forall \omega \in \Omega$.

**Theorem 1.4.2 (Jensen inequality)** Suppose that $I$ is an open interval in $\mathbb{R}$, being either finite or infinite in length, and $X$ is a random variable on the probability space $(\Omega, \mathcal{F}, P)$ such that

$$E|X| < \infty, \quad \text{and} \quad X(\omega) \in I, \forall \omega \in \Omega.$$

Then $EX \in I$. Moreover, if $c$ is a convex function on $I$ such that $E|c(X)| < \infty$, then

$$E[c(X)] \geq c(EX).$$

**Remark 1.4.3** Convex functions which often arise in applications are $c(x) \triangleq |x|^p$ for some $p \in [1, \infty)$, $c(x) \triangleq (x - a)^+$ and $c(x) \triangleq \exp(ax)$ for some constant $a \in \mathbb{R}$, and open intervals $I$ occurring in the use of Jensen’s inequality are typically $I = \mathbb{R}$ or $I = (0, \infty)$.

Abbreviating $[E|X|]^{1/p}$ by $E^{1/p}|X|$, $p \in (0, \infty)$, we can easily infer the following from Jensen’s inequality:

**Corollary 1.4.4 (Lyapunov inequality)** Suppose that $p, q \in (0, \infty)$ are constants with $q < p$, $(\Omega, \mathcal{F}, P)$ is a probability space, and $X \in L^p(\Omega, \mathcal{F}, P)$. Then $X \in L^q(\Omega, \mathcal{F}, P)$ and $E^{1/q}[|X|^q] \leq E^{1/p}[|X|^p]$. 

32
1.4.2 Independence

Suppose that \((\Omega, \mathcal{F}, P)\) is a probability space. We first recall the notion of independence of a pair of events from elementary probability: Two events \(A_1, A_2 \in \mathcal{F}\) are said to be independent when \(P(A_1 \cap A_2) = P(A_1)P(A_2)\). We next extend the notion of independence from a pair of events to a collection or family which is made up of finitely many events: A family of events \(\{A_1, A_2, \ldots, A_n\}\), with \(A_i \in \mathcal{F}\), \(i = 1, 2, \ldots n\), is called independent when, for each subset \(\{A_{i_1}, A_{i_2}, \ldots A_{i_m}\} \subset \{A_1, A_2, \ldots A_n\}\), we have

\[
P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_m}).
\]

There are many applications, particularly involving stochastic processes, where the preceding notion of independence is not quite general enough and we need to formulate a concept of independence not just for a family of individual events but rather for a family of collections of events. We begin with the case where we have finitely many such collections: Suppose \(\mathcal{G}_i \subset \mathcal{F}\) is a given collection of events indexed by \(i = 1, 2, \ldots n\). Then the family of collections \(\{\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n\}\) is called independent when, for each choice of events \(A_1 \in \mathcal{G}_1\), \(A_2 \in \mathcal{G}_2\), \(\ldots, A_n \in \mathcal{G}_n\), the resulting collection of events \(\{A_1, A_2, \ldots, A_n\}\) is independent (in the case where a pair of collections \(\mathcal{G}_1, \mathcal{G}_2\) is independent we usually say that \(\mathcal{G}_1\) and \(\mathcal{G}_2\) are independent). We are now ready to formulate a concept of independence which applies to an arbitrary family of collections of events and which is general enough for all of our requirements: Suppose that \(\mathcal{G}_\lambda \subset \mathcal{F}\) is a given collection of events in \(\mathcal{F}\) for each \(\lambda\) in some indexing set \(\Lambda\). Then the family of collections of events \(\{\mathcal{G}_\lambda, \lambda \in \Lambda\}\) is decreed to be independent when \(\{\mathcal{G}_{\lambda_1}, \mathcal{G}_{\lambda_2}, \ldots, \mathcal{G}_{\lambda_n}\}\) is independent for each finite subset \(\{\lambda_1, \lambda_2, \ldots, \lambda_n\}\) of \(\Lambda\).

Having formulated independence of a given family of collections, we are now able to introduce a concept of independence for arbitrarily many mappings: Suppose that \(\{X_\lambda, \lambda \in \Lambda\}\) is a given family of mappings defined on a common probability space \(\Omega, \mathcal{F}, P\), where each \(X_\lambda\) takes values in some measurable space \((F_\lambda, B_\lambda)\) and is \(\mathcal{F}/B_\lambda\)-measurable. This family of mappings is called independent when the family of \(\sigma\)-algebras \(\{\sigma\{X_\lambda\}, \lambda \in \Lambda\}\) is independent.

**Remark 1.4.5** The preceding formulation of independence for an arbitrary set of random variables is well suited to the needs of analysis, but we still need a criterion which allows us to verify when a given set of random vectors \(\{X_\lambda, \lambda \in \Lambda\}\) is independent in the sense just formulated. In view of the preceding definitions it is clearly enough to have a criterion which tests independence for only a finite collection of random vectors. Suppose that \(X_1, X_2, \ldots, X_n\) are random vectors on a common probability space \((\Omega, \mathcal{F}, P)\) where each \(X_k\) takes values in a finite-dimensional Euclidean space \(\mathbb{R}^{d_k}\), \(k = 1, 2, \ldots n\). Associated with \(X_k\) is the cumulative distribution function \(F_{X_k}(x_k) \triangleq P(\{X_k \leq x_k\})\), for \(x_k \triangleq (x_{k1}, x_{k2}, \ldots x_{kd_k}) \in \mathbb{R}^{d_k}\). Put \(d \triangleq \sum_{k=1}^n d_k\) and let \(F_{X_1, X_2, \ldots, X_n}\) denote the cumulative distribution function of the \(\mathbb{R}^d\)-valued random vector \((X_1, X_2, \ldots X_n)\) which is just a concatenation of the vectors.
\(X_1, X_2, \ldots X_n:\)

\[F_{X_1, X_2, \ldots X_n}(x_1, x_2, \ldots x_n) \triangleq P\left(\bigcap_{k=1}^{n} \{X_k \leq x_k\}\right), \quad \forall x_k \in \mathbb{R}^{d_k}, \ k = 1, 2, \ldots n.\]

Our criterion for checking independence is then as follows:

**Theorem 1.4.6** Suppose that \(\{X_1, X_2, \ldots X_n\}\) is a family of random vectors on a probability space \((\Omega, \mathcal{F}, P)\), where each \(X_k\) takes values in a finite-dimensional Euclidean space \(\mathbb{R}^{d_k}\), \(k = 1, 2, \ldots n\). Then \(\{X_1, X_2, \ldots X_n\}\) is independent if and only if we have

\[F_{X_1, X_2, \ldots X_n}(x_1, x_2, \ldots x_n) = \prod_{k=1}^{n} F_{X_k}(x_k) \quad \forall x_k \in \mathbb{R}^{d_k}, \ k = 1, 2, \ldots n.\]

There is some additional terminology which is useful in regard to independence. If \(X \triangleq (X^1, X^2, \ldots X^d)\) is an \(\mathbb{R}^d\)-valued random vector on a probability space \((\Omega, \mathcal{F}, P)\) and \(\mathcal{G} \subset \mathcal{F}\) is a given collection of events, we shall say that \(X\) is independent of \(\mathcal{G}\) when the two collections of events \(\sigma\{X\} = \sigma\{X^1, X^2, \ldots X^d\}\) and \(\mathcal{G}\) are independent.

### 1.4.3 Characteristic Functions

Suppose that \(X\) is an \(\mathbb{R}^d\)-valued random vector on some probability space \((\Omega, \mathcal{F}, P)\). For each \(\theta \in \mathbb{R}^d\) put

\[(1.49) \quad \phi_X(\theta) \triangleq E[\cos\{(\theta, X)\}] + iE[\sin\{(\theta, X)\}],\]

(here \((x, y)\) denotes the usual Euclidean inner product of \(x, y \in \mathbb{R}^d\)). Since the sine and cosine functions are uniformly bounded in magnitude (by 1) it follows that the expectations on the right hand side of (1.49) are defined and the complex number \(\phi_X(\theta)\) always exists, whether \(X\) is integrable or not. Using the Euler formula, one usually writes (1.49) in the more succinct form

\[(1.50) \quad \phi_X(\theta) \triangleq E[e^{i(\theta, X)}].\]

Thus, an \(\mathbb{R}^d\)-valued random vector \(X\) on some probability space \((\Omega, \mathcal{F}, P)\) determines a mapping \(\phi_X : \mathbb{R}^d \to \mathbb{C}\) which is called the **characteristic function** of \(X\). We see from (1.47) that the characteristic function of \(X\) can also be written in terms of \(\mu_X\), the distribution measure generated by \(X\), as follows:

\[(1.51) \quad \phi_X(\theta) = \int_{\mathbb{R}^d} e^{i(\theta, x)} \, d\mu_X(x).\]

Now suppose that \(\tilde{X}\) is an \(\mathbb{R}^d\)-valued random vector on some probability space \((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})\), with characteristic function

\[(1.52) \quad \phi_{\tilde{X}}(\theta) \triangleq E[e^{i(\theta, X)}] = \int_{\mathbb{R}^d} e^{i(\theta, x)} \, d\mu_{\tilde{X}}(x)\]
(here $\tilde{E}$ indicates integration with respect to the probability measure $\tilde{P}$). If $X$ and $\tilde{X}$ are identically distributed, then Proposition 1.4.1 ensures that $\mu_X(A) = \mu_{\tilde{X}}(A), \forall A \in \mathcal{B}(\mathbb{R}^d)$, thus comparison of (1.51) with (1.52) establishes

\begin{equation}
\phi_X(\theta) = \phi_{\tilde{X}}(\theta), \quad \forall \theta \in \mathbb{R}^d.
\end{equation}

In short, identically distributed $\mathbb{R}^d$-valued random vectors give rise to identical characteristic functions. With this in mind, the question of a converse assertion suggests itself, namely: If $X$ and $\tilde{X}$ are $\mathbb{R}^d$-valued random vectors such that (1.53) holds, does it follow that $X$ and $\tilde{X}$ are identically distributed? This question is answered in the affirmative by the following uniqueness result from harmonic analysis:

**Theorem 1.4.7 (Lévy)** Suppose that $\mu$ and $\nu$ are probability measures on the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that

$$
\int_{\mathbb{R}^d} \exp\{i(\theta, x)\} \, d\mu(x) = \int_{\mathbb{R}^d} \exp\{i(\theta, x)\} \, d\nu(x)
$$

for all $\theta \in \mathbb{R}^d$. Then $\mu(A) = \nu(A), \forall A \in \mathcal{B}(\mathbb{R}^d)$.

As a consequence of Theorem 1.4.7 we see that when $X$ and $\tilde{X}$ are $\mathbb{R}^d$-valued random vectors defined on possibly distinct probability spaces $(\Omega, \mathcal{F}, P)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, with $\phi_X(\theta) = \phi_{\tilde{X}}(\theta), \forall \theta \in \mathbb{R}^d$, then $\mu_X(A) = \mu_{\tilde{X}}(A), \forall A \in \mathcal{B}(\mathbb{R}^d)$, thus in particular $X$ and $\tilde{X}$ are identically distributed.

Characteristic functions provide a useful complement to Theorem 1.4.6 for verifying independence of a family of random vectors. With reference to Remark 1.4.5, let $\phi_{X_1, X_2, \ldots, X_n}(\theta), \theta \in \mathbb{R}^d$, be the characteristic function of the $\mathbb{R}^d$-valued random vector $(X_1, X_2, \ldots, X_n)$, and let $\phi_{X_k}(\theta_k), \theta_k \in \mathbb{R}^{dk}$, be the characteristic function of the $\mathbb{R}^{dk}$-valued random vector $X_k$, for all $k = 1, 2, \ldots, n$. Then we have the following:

**Theorem 1.4.8 (Kac)** Suppose that $\{X_1, X_2, \ldots, X_n\}$ is a family of random vectors on a probability space $(\Omega, \mathcal{F}, P)$, where each $X_k$ takes values in a finite-dimensional Euclidean space $\mathbb{R}^{dk}, k = 1, 2, \ldots, n$. Then $\{X_1, X_2, \ldots, X_n\}$ is independent if and only if we have

\begin{equation}
\phi_{X_1, X_2, \ldots, X_n}(\theta_1, \theta_2, \ldots, \theta_n) = \prod_{k=1}^{n} \phi_{X_k}(\theta_k), \quad \forall \theta_k \in \mathbb{R}^{dk}, \ k = 1, 2, \ldots, n.
\end{equation}

Characteristic functions also allow one to compute higher moments of random vectors:

**Theorem 1.4.9** Suppose $X = (X^1, X^2, \ldots, X^d)$ is an $\mathbb{R}^d$-valued random vector on some probability space $(\Omega, \mathcal{F}, P)$, with a characteristic function $\phi_X$. If $X$ is square integrable then $\phi_X(\theta)$ is twice continuously differentiable at all $\theta \triangleq (\theta^1, \theta^2, \ldots, \theta^d)$ with

$$
\frac{\partial \phi_X(\theta)}{\partial \theta^j} = iE[\exp\{i(\theta, X)\} X^j] \quad \text{and} \quad \frac{\partial^2 \phi_X(\theta)}{\partial \theta^j \partial \theta^k} = -E[\exp\{i(\theta, X)\} X^j X^k],
$$

for each $j, k = 1, 2, \ldots, d$. 

35
1.4.4 The Gaussian Distribution

Definition 1.4.10 Suppose that \( m \) is a vector in \( \mathbb{R}^d \) and \( Q \) is a \( d \) by \( d \) symmetric positive semidefinite matrix. We say that an \( \mathbb{R}^d \)-valued random vector \( X \) on a probability space \( (\Omega, \mathcal{F}, P) \) is Gaussian distributed with expectation \( m \) and covariance \( Q \) if its characteristic function has the specific form

\[
\phi_X(\theta) = \exp \left\{ i(\theta, m) - \frac{1}{2}(\theta, Q\theta) \right\}, \quad \forall \, \theta \in \mathbb{R}^d.
\] (1.55)

For this definition to make sense we must settle the following question: given a vector \( m \) in \( \mathbb{R}^d \) and a \( d \) by \( d \) symmetric positive semidefinite matrix \( Q \), does there exist some \( \mathbb{R}^d \)-valued random vector \( X \) on some probability space \( (\Omega, \mathcal{F}, P) \) whose characteristic function is given by the right-hand-side of (1.55)? Indeed, if this question cannot be answered in the affirmative, then we have defined a non-existent and rather useless concept. The next theorem shows that the preceding definition makes sense:

Theorem 1.4.11 Suppose that \( m \) is a vector in \( \mathbb{R}^d \) and \( Q \) is a \( d \) by \( d \) positive semidefinite matrix. Then there exists some \( \mathbb{R}^d \)-valued random vector \( X \) on some probability space \( (\Omega, \mathcal{F}, P) \) such that

\[
E[\exp\{i(\theta, X)\}] = \exp \left\{ i(\theta, m) - \frac{1}{2}(\theta, Q\theta) \right\}, \quad \forall \, \theta \in \mathbb{R}^d.
\] (1.56)

We still have to justify calling \( m \) the expectation and \( Q \) the covariance in Definition 1.4.10. From Theorem 1.4.9 along with (1.55) we easily get

\[
E[X^j] = m^j \quad \text{and} \quad E[(X^j - m^j)(X^k - m^k)] = Q_{j,k},
\]

so that \( E[X] = m \) and \( \text{Cov}(X) = Q \), as required to justify our terminology. Henceforth, we shall write \( X \sim N(m, Q) \) to indicate that \( X \) is Gaussian distributed with expectation \( m \) and covariance \( Q \).

Notice that Theorem 1.4.11 does not guarantee uniqueness of the random vector \( X \). Indeed, there may exist many \( \mathbb{R}^d \)-valued random vectors \( X \) on different probability spaces whose characteristic functions are all given by the right hand side of (1.56). However, Theorem 1.4.7 ensures that all such random vectors have the same cumulative distribution function, which is called the Gaussian or normal distribution. In the special case where the matrix \( Q \) is nonsingular we can associate a specific density function with this distribution:

Theorem 1.4.12 Suppose \( m \) is a vector in \( \mathbb{R}^d \) and \( Q \) is a \( d \) by \( d \) symmetric positive definite matrix. If an \( \mathbb{R}^d \)-valued random vector \( X \) is such that \( X \sim N(m, Q) \) then \( X \) has a density function given by

\[
f_X(x) \triangleq \frac{1}{(2\pi)^{d/2}(\det(Q))^{1/2}} \exp \left\{ -\frac{1}{2}(x-m)^T Q^{-1} (x-m) \right\}, \quad \forall \, x \in \mathbb{R}^d.
\]
1.4.5 Conditional Expectation

We start by formulating a theorem which justifies the definition of conditional expectation. In view of the importance of this result we give its proof:

**Theorem 1.4.13 (Kolmogorov)** Suppose that \( X \) is a random variable on a probability space \((\Omega, \mathcal{F}, P)\) such that \( E|X| < \infty \), and \( \mathcal{G} \) is a \( \sigma \)-algebra over \( \Omega \) with \( \mathcal{G} \subset \mathcal{F} \). Then there exists some \( \mathcal{G} \)-measurable random variable \( Z \) on \( \Omega \) such that \( E|Z| < \infty \) and

\[
\int_A Z dP = \int_A X dP, \quad \forall \ A \in \mathcal{G}.
\]

(1.57)

Moreover, if \( \tilde{Z} \) is any \( \mathcal{G} \)-measurable random variable on \( \Omega \) such that \( E|\tilde{Z}| < \infty \) and

\[
\int_A \tilde{Z} dP = \int_A X dP, \quad \forall \ A \in \mathcal{G},
\]

then \( P[\tilde{Z} = Z] = 1 \).

Proof: Define

\[
Q_+(A) \doteq \int_A X_+ dP, \quad \forall \ A \in \mathcal{G}.
\]

(1.59)

Then Corollary 1.2.13 and the integrability of \( X \) show that \( Q_+ \) is a finite measure on the measurable space \((\Omega, \mathcal{G})\) and it is clear that

\[
Q_+ \ll P[\mathcal{G}].
\]

(1.60)

From (1.60) and Theorem 1.2.30 (with \( (E, \mathcal{S}) \doteq (\Omega, \mathcal{G}), \nu \doteq Q_+, \) and \( \mu \doteq P \)) there exists some \( \mathcal{G} \)-measurable mapping \( Z_+ : \Omega \to [0, \infty] \) such that

\[
Q_+(A) = \int_A Z_+ dP, \quad \forall \ A \in \mathcal{G}.
\]

or, equivalently,

\[
E[Z_+; A] = E[X_+; A], \quad \forall \ A \in \mathcal{G}.
\]

(1.61)

Thus, \( Z_+ \) is a Radon-Nikodym derivative of \( Q_+ \) with respect to \( P \) relative to the \( \sigma \)-algebra \( \mathcal{G} \) (see Remark 1.2.32). Since \( E|X| < \infty \), we have \( E[X_+] < \infty \), hence \( E[Z_+] < \infty \). Thus, without loss of generality, we can take \( Z_+ \) to be \([0, \infty)\)-valued (see Remark 1.2.32). In exactly the same way, there exists a \( \mathcal{G} \)-measurable mapping \( Z_- : \Omega \to [0, \infty) \) such that \( E|Z_-| < \infty \) and

\[
E[Z_-; A] = E[X_-; A], \quad \forall \ A \in \mathcal{G}.
\]

(1.62)

Now put

\[
Z \doteq Z_+ - Z_-.
\]
In view of (1.61) and (1.62) we have constructed a $G$-measurable mapping $Z : \Omega \to \mathbb{R}$ such that $E|Z| < \infty$ and (1.57) holds. Now suppose that $\tilde{Z}$ is some $G$-measurable random variable such that $E|\tilde{Z}| < \infty$ and (1.58) holds. Then clearly

$$\int_A (Z - \tilde{Z}) \, dP = 0, \quad \forall A \in G.$$ 

But $Z - \tilde{Z}$ is $G$-measurable, hence Theorem 1.2.24(b) ensures $P[Z - \tilde{Z} = 0] = 1$, as required.

There is no possibility of showing that the random variable $Z$, whose existence is asserted by Theorem 1.4.13, is unique in a strict pointwise sense, since we see from Theorem 1.2.24(a) that, if $Y$ is any integrable $G$-measurable random variable such that $Z = Y$ a.s., then (1.57) holds with $Y$ in place of $Z$. Thus, the strongest form of uniqueness we can expect is a.s. uniqueness, and this is exactly what Theorem 1.4.13 provides. We shall use $E[X \mid G]$ to denote a fixed but arbitrarily chosen $G$-measurable $P$-integrable random variable $Z$ for which (1.57) holds, and we call this random variable a conditional expectation of $X$ given $G$, the latter entity being called the conditioning $\sigma$-algebra. Notice our use of the indefinite article: we speak of “a conditional expectation ...” rather than “the conditional expectation ...” precisely to remind ourselves of the arbitrary choice involved whenever we write down $E[X \mid G]$. We could have viewed $E[X \mid G]$ as an equivalence class of $G$-measurable integrable random variables subject to (1.57), but we deliberately refrain from this way of thinking about conditional expectation because it turns out to be completely inappropriate for the later study of continuous-parameter stochastic processes (in this regard the reader can do no better than read the comments on page xiii of Williams [31]). The need to make arbitrary choices from within a set of functions with some property, any two members of which agree a.s., seems to be intrinsic to modern probability, and we shall see similar situations arising in later chapters, for example when we formulate the so-called quadratic variation process of a martingale, and later still when we define stochastic integrals.

Conditional expectations turn out to have a particularly nice geometric interpretation when $X$ is square integrable. Recall from §1.2.5 that $L^2(\Omega, \mathcal{F}, P)$ is a Hilbert space with inner product defined by

$$(X, Y) \triangleq E[XY], \quad \forall X, Y \in L^2(\Omega, \mathcal{F}, P),$$

and norm $\|X\|_2 \triangleq (X, X)^{1/2}$. It is clear that $L^2(\Omega, \mathcal{G}, P)$, the normed vector space of all $\mathcal{G}$-measurable square integrable random variables on $\Omega$, is vector subspace of $L^2(\Omega, \mathcal{F}, P)$ (in fact, by Theorem 1.2.27, it is a closed vector subspace).

**Theorem 1.4.14** Suppose that $X$ is a random variable on probability space $(\Omega, \mathcal{F}, P)$ such that $E|X|^2 < \infty$, and $\mathcal{G}$ is a $\sigma$-algebra over $\Omega$ with $\mathcal{G} \subset \mathcal{F}$. Then $E[X \mid \mathcal{G}] \in L^2(\Omega, \mathcal{G}, P)$,

$$\|X - E[X \mid \mathcal{G}]\|_2 \leq \|X - Y\|_2, \quad \forall Y \in L^2(\Omega, \mathcal{G}, P),$$

and

$$(X - E[X \mid \mathcal{G}]) \perp Z, \quad \forall Z \in L^2(\Omega, \mathcal{G}, P).$$
In the parlance of Hilbert spaces, Theorem 1.4.14 displays $E[X | G]$ as a projection of $X$ onto the closed vector subspace $L^2(\Omega, \mathcal{G}, P)$ (see Theorem 1.1.1).

The following theorem identifies the most basic properties of conditional expectations:

**Theorem 1.4.15** Suppose $X$ and $Y$ are random variables on the probability space $(\Omega, \mathcal{F}, P)$, with $E|X| < \infty$, and let $\mathcal{G}$ be a $\sigma$-algebra over $\Omega$ with $\mathcal{G} \subset \mathcal{F}$. Then

(a) $E[X | G] \geq 0$ a.s. whenever $X \geq 0$ a.s.;

(b) $E[\alpha X + \beta Y | G] = \alpha E[X | G] + \beta E[Y | G]$ a.s. for $\alpha, \beta \in \mathbb{R}$ and $Y$ being $P$-integrable;

(c) If $XY$ is $P$-integrable and $Y$ is $\mathcal{G}$-measurable then $E[XY | G] = YE[X | G]$ a.s.;

(d) If $\mathcal{H}$ is a $\sigma$-algebra over $\Omega$ with $\mathcal{H} \subset \mathcal{G}$ then $E[E[X | G] | \mathcal{H}] = E[X | \mathcal{H}]$ a.s.;

(e) $E[E[X | G]] = EX$;

(f) If $\mathcal{H}$ is a $\sigma$-algebra over $\Omega$ which is independent of the $\sigma$-algebra $\sigma\{\sigma\{X\}, \mathcal{G}\}$ then $E[X | \sigma\{\mathcal{G}, \mathcal{H}\}] = E[X | G]$ a.s. In particular, $E[X | \mathcal{H}] = EX$ a.s. whenever $\sigma\{X\}$ and $\mathcal{H}$ are independent.

**Remark 1.4.16** Part (d) of the preceding theorem is usually called the composition rule for conditional expectations. Observe that the qualifications “a.s.” are essential in the preceding statements because, in each of parts (a) - (f), we are fixing some arbitrary choice of the conditional expectations involved. In fact, a completely accurate rendition of, say, Theorem 1.4.15(b) should really go as follows: Let $E[\alpha X + \beta Y | G]$, $E[X | G]$, and $E[Y | G]$ be arbitrarily chosen conditional expectations. Then $E[\alpha X + \beta Y | G] = \alpha E[X | G] + \beta E[Y | G]$ a.s. Needless to say, we avoid lengthy and tedious statements of this kind, but every assertion about conditional expectations should really be understood in this sense.

**Remark 1.4.17** The fact that conditional expectations are uniquely defined a.s. sometimes requires us to deal with situations in which one event is “almost” a subset of another, or two events are “almost” equal. To express this slight blurring of the standard set-theoretic ideas of set-inclusion and set-equality we use the following notation: For events $A$ and $B$ in the probability space $(\Omega, \mathcal{F}, P)$ write

$$A \subset B \quad \text{a.s. and} \quad A = B \quad \text{a.s.}$$

to denote $P(A - B) = 0$ and $P(A \Delta B) = 0$ respectively.

**Remark 1.4.18** Suppose $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathcal{G}$ and $\mathcal{H}$ are $\sigma$-algebras over $\Omega$ with $\mathcal{G} \subset \mathcal{H} \subset \mathcal{F}$, and $\tilde{P}$ is a probability measure on $(\Omega, \mathcal{F})$ such that $\tilde{P} \ll P[\mathcal{F}]$. Define the Radon-Nikodym derivatives

$$X_\mathcal{G} \triangleq \frac{d\tilde{P}}{dP} \bigg|_{\mathcal{G}}, \quad \text{and} \quad X_\mathcal{H} \triangleq \frac{d\tilde{P}}{dP} \bigg|_{\mathcal{H}}. \quad (1.63)$$

From Remark 1.2.34 it follows that $X_\mathcal{G}$ is a $[0, \infty]$-valued and $\mathcal{G}$-measurable mapping such that

$$\tilde{P}(A) = E[X_\mathcal{G}; A], \quad \forall \ A \in \mathcal{G}, \quad (1.64)$$

39
where we use \( \mathbb{E}[\cdot] \) to denote expectation with respect to the probability measure \( P \). In particular, with \( A \uparrow \Omega \), we get \( \mathbb{E}[X_G] = 1 \), so that \( X_G \) is integrable with respect to the probability measure \( P \). Identical observations hold for \( X_H \), which is therefore a \([0, \infty]\)-valued and \( \mathcal{H} \)-measurable mapping, integrable with respect to measure \( P \), such that

\[
\tilde{P}(A) = \mathbb{E}[X_H; A], \quad \forall A \in \mathcal{H}.
\]

The Radon-Nikodym derivatives in (1.63) are related by

\[
X_G = \mathbb{E}[X_H \mid G] \quad P - \text{a.s.} \quad (1.66)
\]

To see this, fix some \( A \in \mathcal{G} \). Then \( A \in \mathcal{H} \) (since \( \mathcal{G} \subset \mathcal{H} \)) and therefore, by (1.64) and (1.65), we get

\[
\mathbb{E}[X_G; A] = \mathbb{E}[X_H; A].
\]

From the arbitrary choice of \( A \in \mathcal{G} \) we find

\[
\mathbb{E}[X_G 1_A] = \mathbb{E}[X_H 1_A], \quad \forall A \in \mathcal{G},
\]

and therefore, from the \( \mathcal{G} \)-measurability of \( X_G \) and a.s. uniqueness asserted by Theorem 1.4.13, we get (1.66). This relation will be essential when we look at martingales in later chapters.

Next we formulate versions of Levi’s monotone convergence theorem, Fatou’s theorem, and the Lebesgue dominated convergence theorem in the context of conditional expectations:

**Theorem 1.4.19** Suppose \( \{X_n, n = 1, 2, \ldots\} \) is a sequence of random variables on a probability space \((\Omega, \mathcal{F}, P)\) with \( E|X_n| < \infty, \forall n = 1, 2, \ldots \), and \( \mathcal{G} \) is a \( \sigma \)-algebra over \( \Omega \) with \( \mathcal{G} \subset \mathcal{F} \).

(a) If \( 0 \leq X_n(\omega) \leq X_{n+1}(\omega), \forall n = 1, 2, \ldots, \forall \omega \in \Omega \), and \( E[\lim_{n \to \infty} X_n] < \infty \), then

\[
\lim_{n \to \infty} E[X_n \mid \mathcal{G}] = E\left[\lim_{n \to \infty} X_n \mid \mathcal{G}\right] \quad \text{a.s.}
\]

(b) If \( X_n(\omega) \geq 0, \forall n = 1, 2, \ldots, \forall \omega \in \Omega \), and \( E[\liminf_{n \to \infty} X_n] < \infty \), then

\[
\liminf_{n \to \infty} E[X_n \mid \mathcal{G}] \geq E\left[\liminf_{n \to \infty} X_n \mid \mathcal{G}\right] \quad \text{a.s.}
\]

(c) If \( \{X_n, n = 1, 2, \ldots\} \) converges a.s. to a random variable \( X \), and \( |X_n(\omega)| \leq Z(\omega), \forall n = 1, 2, \ldots, \forall \omega \in \Omega \), with \( EZ < \infty \), then \( X \) is \( P \)-integrable and

\[
\lim_{n \to \infty} E[X_n \mid \mathcal{G}] = E[X \mid \mathcal{G}] \quad \text{a.s.}
\]

There is also a version of Jensen’s inequality for conditional expectations:

**Theorem 1.4.20** Suppose that \( I \) is an open interval in \( \mathbb{R} \), being either either finite or infinite in length, and \( X \) is a random variable on \((\Omega, \mathcal{F}, P)\) such that

\[
E|X| < \infty, \quad \text{and} \quad X(\omega) \in I, \forall \omega \in \Omega.
\]

If \( c \) is a convex function on \( I \) such that \( E|c(X)| < \infty \), and \( \mathcal{G} \) is a \( \sigma \)-algebra over \( \Omega \) with \( \mathcal{G} \subset \mathcal{F} \), then

\[
E[X \mid \mathcal{G}] \in I \quad \text{a.s.} \quad \text{and} \quad E[c(X) \mid \mathcal{G}] \geq c(E[X \mid \mathcal{G}]) \quad \text{a.s.}
\]
The next proposition is a direct consequence of Theorem 1.4.20, Theorem 1.4.15(b) and the fact that \(c(x) \triangleq |x|^p, \ p \in [1, \infty)\), is a convex function:

**Proposition 1.4.21** Suppose \(X\) and \(\{X_n, \ n = 1, 2, \ldots\}\) are random variables in \(L^p(\Omega, \mathcal{F}, P)\), with \(p \in [1, \infty)\), and \(\mathcal{G}\) is a \(\sigma\)-algebra over \(\Omega\) with \(\mathcal{G} \subset \mathcal{F}\). Then:

(a) \(\|E[X | \mathcal{G}]\|_p \leq \|X\|_p\):

(b) If \(\lim_{n \to \infty} \|X_n - X\|_p = 0\) then

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

Conditional expectations assume a more concrete form in the special case where the conditioning \(\sigma\)-algebra \(\mathcal{G}\) is generated by a set of mappings on \((\Omega, \mathcal{F}, P)\). We shall use Theorem 1.3.1 to look at conditional expectations when the conditioning \(\sigma\)-algebra is generated by a finite number of random variables. Thus, suppose that \(X\) and \(Y_1, Y_2, \ldots Y_n\) are random variables on a probability space \((\Omega, \mathcal{F}, P)\), with \(E|X| < \infty\), and put \(\mathcal{G} \triangleq \sigma\{Y_1, Y_2, \ldots Y_n\}\). Also, let \(E \triangleq \mathbb{R}^n\), \(\mathcal{S} \triangleq \mathcal{B}(\mathbb{R}^n)\), and define \(Y : \Omega \to E\) by \(Y \triangleq (Y_1, Y_2, \ldots Y_n)^T\). It is easily checked that \(Y^{-1}[\mathcal{S}] = \mathcal{G}\). Thus, if \(E[X | \mathcal{G}]\) is a conditional expectation of \(X\) given \(\mathcal{G}\), then \(E[X | \mathcal{G}]\) is \(\mathcal{G}\)-measurable, and so it follows from Theorem 1.3.1 that there is some \(\mathcal{B}(\mathbb{R}^n)\)-measurable function \(\Psi\) from \(\mathbb{R}^n\) to \(\mathbb{R}\) such that

\[
(1.67) \quad E[X | \mathcal{G}](\omega) = \Psi(Y(\omega)), \quad \forall \omega \in \Omega.
\]

Hence, \(E[X | \mathcal{G}](\omega)\) is given by a fixed mapping \(\Psi\) of the “observation record” \(\{Y_1(\omega), Y_2(\omega), \ldots Y_n(\omega)\}\) for each \(\omega \in \Omega\). The point to understand is that the same mapping \(\Psi\) works for all \(\omega\) once we have fixed our (arbitrary) choice of the conditional expectation \(E[X | \mathcal{G}]\). One can regard the function \(\Psi\) as defining a deterministic method for “processing” the observed data values \(Y_1(\omega), Y_2(\omega), \ldots Y_n(\omega)\) to yield the value \(E[X | \mathcal{G}](\omega)\). One customarily uses the notation \(E[X | Y_1 = y_1, Y_2 = y_2, \ldots Y_n = y_n]\) to indicate the number \(\Psi(y_1, y_2, \ldots y_n)\) for arbitrary real arguments \(y_1, y_2, \ldots y_n\) and \(E[X | Y_1, Y_2, \ldots Y_n]\) to denote some arbitrarily chosen conditional expectation \(E[X | \mathcal{G}]\) when \(\mathcal{G} \triangleq \sigma\{Y_1, Y_2, \ldots Y_n\}\).

Explicit computation of the mapping \(\Psi\) in (1.67) usually presents significant difficulties. There is however, a considerable body of work within the general realm of filtering theory which provides algorithms for computing \(\Psi(Y_1(\omega), Y_2(\omega), \ldots Y_n(\omega))\), given an “observation record” \(\{Y_1(\omega), Y_2(\omega), \ldots Y_n(\omega)\}\). We choose not to take up these issues in this introductory chapter, and confine ourselves to mentioning just one case where an explicit form for \(\Psi\) is possible:

**Theorem 1.4.22** Suppose that \(X, Y_1, Y_2, \ldots Y_n\) are random variables on a probability space \((\Omega, \mathcal{F}, P)\) such that the \(\mathbb{R}^{n+1}\)-valued random vector \((X, Y_1, Y_2, \ldots Y_n)\) has a density function \(f_{XY} : \mathbb{R}^{n+1} \to [0, \infty]\), and such that \(E|X| < \infty\). For each \(x \in \mathbb{R}\) and \(y \in \mathbb{R}^n\) define

\[
f_X(x) \triangleq \int_{\mathbb{R}^n} f_{XY}(x, \xi_2) d\xi_2, \quad \text{and} \quad f_Y(y) \triangleq \int_{\mathbb{R}} f_{XY}(\xi_1, y) d\xi_1,
\]

41
and for each \( x \in \mathbb{R} \), put

\[
f_{X|Y}(x|y) \triangleq \frac{f_{XY}(x,y)}{f_Y(y)}
\]

for all \( y \in \mathbb{R}^n \) such that \( f_Y(y) \neq 0 \), and \( f_{X|Y}(x|y) \triangleq 0 \) for all \( y \in \mathbb{R}^n \) such that \( f_Y(y) = 0 \). Then:

(a) \( f_X \), \( f_Y \) and \( f_{X|Y} \) are \([0, \infty]\)-valued measurable functions on the measurable spaces \((\mathbb{R}, \mathcal{B}(\mathbb{R}))\), \((\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))\), and \((\mathbb{R}^{n+1}, \mathcal{B}(\mathbb{R}^{n+1}))\) respectively.

(b) \( f_X \) and \( f_Y \) are density functions of \( X \) and \( Y \) respectively.

(c) If \( E [X \mid Y_1, Y_2, \ldots Y_n] \) is some conditional expectation of \( X \) given \( \sigma \{Y_1, Y_2, \ldots Y_n\} \) then

\[
E [X \mid Y_1, Y_2, \ldots Y_n] = \Psi(Y_1, Y_2, \ldots Y_n) \quad \text{a.s.}
\]

for \( \Psi : \mathbb{R}^n \to \mathbb{R} \) defined by

\[
\Psi(y) \triangleq \int_{\mathbb{R}} \xi_1 f_{X|Y}(|y\xi_1|)d\xi_1.
\]

### 1.5 Monotone Class Theorems

Properly speaking, monotone class theorems are part of measure theory and should have been summarized in §1.2. Indeed, monotone class ideas are implicit in several of the topics to which we have alluded in preceding sections; for example, the Tonelli Theorem 1.2.40, and Theorem 1.4.6 on independence, are most easily established using monotone class arguments. We choose to place the topic of monotone class theorems in a section of its own because the application of monotone class notions involves a particular methodology which we wish to emphasize in view of its later importance, and also because monotone class theorems, at least in their modern reincarnation due to Dynkin, are sometimes omitted from the more traditional introductions to measure theory.

In this section we shall state monotone class theorems for both sets and functions. We do not assume previous familiarity with monotone class ideas and our account of this topic will be rather more complete than the brief summaries in the preceding sections. We shall then illustrate the particular style of argument which goes with the use of monotone class notions by establishing special cases Theorems 1.2.40 and 1.4.6. First recall the notion of a \( \pi \)-class, which has already been seen in regard to Theorem 1.2.4:

**Definition 1.5.1** A designated collection \( \mathcal{C} \) of subsets of a given set \( E \) is called a \textbf{\( \pi \)-class} over \( E \) if \( A \cap B \in \mathcal{C} \) whenever \( A \in \mathcal{C} \) and \( B \in \mathcal{C} \).

Thus, a \( \pi \)-class of sets is closed under pairwise intersections and hence, by induction, under finite intersections as well. The nice thing about \( \pi \)-classes is their simple structure; it is usually easy to check when a given collection of subsets is a \( \pi \)-class. Closely associated with \( \pi \)-classes of sets is the concept of a \( \lambda \)-class of sets:
Definition 1.5.2 A designated collection $C$ of subsets of a given set $E$ is called a $\lambda$-class over $E$ when the following conditions hold:

(i) $E \in C$;
(ii) If $A \in C$ and $B \in C$ with $A \subset B$ then $B - A \in C$;
(iii) If $\{A_n, \ n = 1, 2, \ldots\}$ is a monotonically increasing sequence of members of $C$ then its limit $\bigcup_{n=1}^{\infty} A_n$ is also a member of $C$.

The next proposition establishes some basic properties of $\lambda$-classes:

Proposition 1.5.3 Suppose that $C$ is a $\lambda$-class over a set $E$. Then the following hold:

(a) $A^c \in C$ whenever $A \in C$;
(b) If $A \in C$, $B \in C$ and $A \cap B = \emptyset$ then $A \cup B \in C$;
(c) If $C$ is a $\pi$-class as well as a $\lambda$-class then $C$ is a $\sigma$-algebra.

Proof: (a) We have $A \in C$ (by hypothesis), $E \in C$ (by Definition 1.5.2(i)), and $A^c = E - A$. Thus, by Definition 1.5.2(ii) we see that $A^c \in C$.

(b) Since $A \cap B = \emptyset$ clearly $A \subset B^c$. Now $A$ and $B$ are members of $C$ (by hypothesis), hence $B^c \in C$ (by part (a)), and thus $B^c - A \in C$ (by Definition 1.5.2(ii)). Moreover $[A \cup B]^c = B^c - A$ (since $A \cap B = \emptyset$), thus $[A \cup B]^c \in C$, and therefore $A \cup B \in C$ by part (a).

(c) We first establish that $C$ is an algebra. Fix members $A$ and $B$ in $C$. Then $A^c$ is a member of $C$ (by part (a)) hence $B \cap A^c \in C$, since $C$ is a $\pi$-class. Since $A$ and $B \cap A^c$ are disjoint, we see from part (b) and the set identity $A \cup B = A \cup [B \cap A^c]$, that $A \cup B$ is a member of $C$, hence $C$ is an algebra. To show that $C$ is a $\sigma$-algebra fix some sequence $\{A_n, \ n = 1, 2, \ldots\}$ of members of $C$, and let $B_N \triangleq \bigcup_{n=1}^{N} A_n$. Since $C$ is an algebra we see that $B_N$ is a member of $C$ for each $N = 1, 2, \ldots$, and of course $\{B_N, \ N = 1, 2, \ldots\}$ is an increasing sequence of sets whose limit $\bigcup_{n=1}^{\infty} A_n$ must belong to $C$ (by Definition 1.5.2(iii)).

It follows trivially from Definition 1.5.2 that the intersection of an arbitrary collection of $\lambda$-classes over a common set $E$ is likewise a $\lambda$-class, and the power set $\mathcal{P}(E)$ is of course itself a $\lambda$-class. Thus, if one forms the intersection of all $\lambda$-classes over $E$ which include the members of some given arbitrary collection of sets $\mathcal{D} \subset \mathcal{P}(E)$, then one obtains the smallest or minimal $\lambda$-class over $E$ which includes all members of $\mathcal{D}$. With this in mind, we are now able to prove a powerful theorem which relates $\pi$ and $\lambda$-classes, called Dynkin’s monotone class theorem:

Theorem 1.5.4 (Dynkin) Suppose $\mathcal{D}$ is a $\pi$-class and $\mathcal{C}$ is a $\lambda$-class, both over a common set $E$. If $\mathcal{D} \subset \mathcal{C}$ then $\sigma\{\mathcal{D}\} \subset \mathcal{C}$.

Proof: Let $\mathcal{G}$ denote the minimal $\lambda$-class which includes all members of $\mathcal{D}$. Clearly $\mathcal{D} \subset \mathcal{G} \subset \mathcal{C}$, and if we can show that $\mathcal{G}$ is a $\pi$-class then Proposition 1.5.3(c) says that $\mathcal{G}$ is a $\sigma$-algebra which includes $\mathcal{D}$,
hence \( \sigma\{D\} \subset G \subset C \), as required to establish the theorem. To show that \( G \) is indeed a \( \pi \)-class requires the introduction of two cleverly chosen collections of sets. Put:

\[
\mathcal{H} \triangleq \{ A \subset E : A \cap D \in G, \ \forall D \in \mathcal{D} \}.
\]

Since \( \mathcal{D} \subset G \) and \( \mathcal{D} \) is a \( \pi \)-class we see that \( \mathcal{D} \subset \mathcal{H} \). Moreover, one easily checks by straightforward set manipulation that \( \mathcal{H} \) is a \( \lambda \)-class over \( E \), whence \( G \subset \mathcal{H} \) (by the aforementioned minimality of \( G \)). In particular, we have established that \( A \cap D \in G \) whenever \( A \in G \) and \( D \in \mathcal{D} \), and hence, if we put

\[
\mathcal{F} \triangleq \{ B \subset E : B \cap A \in G, \ \forall A \in \mathcal{G} \},
\]

then we must have \( \mathcal{D} \subset \mathcal{F} \). By easy set manipulations one checks that \( \mathcal{F} \) is a \( \lambda \)-class over \( E \), hence by the minimality of \( G \) we must have \( G \subset \mathcal{F} \). This proves that \( G \) is a \( \pi \)-class, as required.

To illustrate the usefulness of this theorem we shall use it to establish a special case of Theorem 1.4.6. First we must prove the following two lemmata:

**Lemma 1.5.5** Suppose that \( \mathcal{D} \) and \( \mathcal{C} \) are independent collections of events in a probability space \((\Omega, \mathcal{F}, P)\). If \( \mathcal{D} \) is a \( \pi \)-class then \( \sigma\{\mathcal{D}\} \) and \( \mathcal{C} \) are independent.

Proof: Fix some arbitrary \( B \in \mathcal{C} \) and put

\[
\mathcal{D}_B^* \triangleq \{ A \in \sigma\{\mathcal{D}\} : P(A \cap B) = P(A)P(B) \}.
\]

Since \( \mathcal{D} \) and \( \mathcal{C} \) are independent it follows at once that \( \mathcal{D} \subset \mathcal{D}_B^* \). Moreover, \( \mathcal{D}_B^* \) is easily seen to be a \( \lambda \)-class of events. Indeed, one trivially observes that \( \Omega \in \mathcal{D}_B^* \), while if \( A_1 \) and \( A_2 \) are members of \( \mathcal{D}_B^* \) with \( A_1 \subset A_2 \) then clearly

\[
P[(A_2 - A_1)B] = P(A_2B) - P(A_1B) = P(A_2 - A_1)P(B),
\]

whence \( A_2 - A_1 \in \mathcal{D}_B^* \). Next, fix some monotonically increasing sequence \( \{A_n, \ n = 1, 2, \ldots\} \) of members of \( \mathcal{D}_B^* \). By Theorem 1.2.5(c) it follows that \( \lim_{n \to \infty} A_n \) belongs to \( \mathcal{D}_B^* \), hence we conclude that \( \mathcal{D}_B^* \) is a \( \lambda \)-class. By Theorem 1.5.4 one sees that \( \sigma\{\mathcal{D}\} \subset \mathcal{D}_B^* \subset \sigma\{\mathcal{D}\} \), and hence, by the arbitrary choice of \( B \in \mathcal{C} \), we get

\[
\sigma\{\mathcal{D}\} = \mathcal{D}_B^*, \quad \forall B \in \mathcal{C}.
\]

In particular, we have \( P(AB) = P(A)P(B) \) for all \( A \in \sigma\{\mathcal{D}\} \) and \( B \in \mathcal{C} \), whence it follows that \( \sigma\{\mathcal{D}\} \) and \( \mathcal{C} \) are independent.

**Lemma 1.5.6** Suppose \( \mathcal{D}_1, \mathcal{D}_2, \) and \( \mathcal{D}_3 \) are \( \pi \)-classes of events in a probability space \((\Omega, \mathcal{F}, P)\), such that \( \Omega \in \mathcal{D}_i, \ i = 1, 2, 3 \), and the family of collections \( \{ \mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3 \} \) is independent. Then the family of collections \( \{ \sigma\{\mathcal{D}_1\}, \sigma\{\mathcal{D}_2\}, \sigma\{\mathcal{D}_3\} \} \) is independent.
Proof: Since \( \{D_1, D_2, D_3\} \) is independent, we have
\[
(1.68) \quad P(D_1D_2D_3) = P(D_1)P(D_2)P(D_3), \quad \forall D_i \in D_i, \; i = 1, 2, 3,
\]
and, taking \( D_3 \triangleq \Omega \) in (1.68) it then follows that
\[
(1.69) \quad P(D_1D_2) = P(D_1)P(D_2), \quad \forall D_i \in D_i, \; i = 1, 2.
\]
Combining (1.68) and (1.69) we then get
\[
(1.70) \quad P[(D_1D_2)D_3] = P(D_1D_2)P(D_3), \quad \forall D_i \in D_i, \; i = 1, 2, 3.
\]
Now put
\[
\mathcal{E} \triangleq \{D_1D_2 : \; D_i \in D_i, \; i = 1, 2\}.
\]
In view of (1.70) one sees that \( D_3 \) and \( \mathcal{E} \) are independent and hence, by Lemma 1.5.5, it follows that \( \sigma\{D_3\} \) and \( \mathcal{E} \) are independent. Now \( \mathcal{E} \) is clearly a \( \pi \)-class, and therefore a second application of Lemma 1.5.5 shows that \( \sigma\{D_3\} \) and \( \sigma\{\mathcal{E}\} \) are likewise independent. We next fix arbitrary events \( A_i \in \sigma\{D_i\} \), \( i = 1, 2, 3 \). Since \( \Omega \in D_2 \) we have \( D_1 \subset \mathcal{E} \), whence of course \( \sigma\{D_1\} \subset \sigma\{\mathcal{E}\} \), and likewise we see that \( \sigma\{D_2\} \subset \sigma\{\mathcal{E}\} \). Thus \( A_1A_2 \in \sigma\{\mathcal{E}\} \), and so, by the independence of \( \sigma\{D_3\} \) and \( \sigma\{\mathcal{E}\} \) established earlier, we get
\[
(1.71) \quad P(A_1A_2A_3) = P(A_1A_2)P(A_3).
\]
Now, again by a two-fold application of Lemma 1.5.5, we see that \( \sigma\{D_1\} \) and \( \sigma\{D_2\} \) are independent, whence
\[
(1.72) \quad P(A_1A_2) = P(A_1)P(A_2).
\]
Combining (1.71) and (1.72) one has
\[
(1.73) \quad P(A_1A_2A_3) = P(A_1)P(A_2)P(A_3).
\]
Since we can repeat the preceding argument leading to (1.73) with \( A_1 \) chosen to be \( \Omega \), \( A_2 \) and \( A_3 \) as before, we get \( P(A_2A_3) = P(A_2)P(A_3) \), and, in an obviously similar manner, we also get \( P(A_1A_3) = P(A_1)P(A_3) \) and \( P(A_1A_2) = P(A_1)P(A_2) \). It follows that \( \{A_1, A_2, A_3\} \) is an independent collection of events, and in view of the arbitrary choice of the \( A_i \) we conclude that \( \{\sigma\{D_1\}, \sigma\{D_2\}, \sigma\{D_3\}\} \) is independent.

We can now prove a special case of Theorem 1.4.6:

**Theorem 1.5.7** Suppose that \( X, Y \) and \( Z \) are random variables on a probability space \( (\Omega, \mathcal{F}, P) \) with cumulative distribution functions \( F_X, F_Y \) and \( F_Z \) respectively, and joint cumulative distribution function \( F_{XYZ} \). If
\[
(1.74) \quad F_{XYZ}(x,y,z) = F_X(x)F_Y(y)F_Z(z), \quad \forall x, y, z \in \mathbb{R},
\]
then \( \{\sigma\{X\}, \sigma\{Y\}, \sigma\{Z\}\} \) is independent.
Proof: Define the collections of events \( D_1 \triangleq \{ \{ X \leq x \} : -\infty < x \leq \infty \} \), \( D_2 \triangleq \{ \{ Y \leq y \} : -\infty < y \leq \infty \} \), and \( D_3 \triangleq \{ \{ Z \leq z \} : -\infty < z \leq \infty \} \). It follows at once that the \( D_i \) are \( \pi \)-classes of events with \( \Omega \in D_i \) for \( i = 1, 2, 3 \) (this is why we used \( \leq \infty \) rather than \( < \infty \) in defining \( D_i \)). Moreover, (1.74) tells us that the family of collections \( \{ D_1, D_2, D_3 \} \) is independent. Since \( \sigma \{ X \} = \sigma \{ D_1 \} \), \( \sigma \{ Y \} = \sigma \{ D_2 \} \), and \( \sigma \{ Z \} = \sigma \{ D_3 \} \), the theorem follows at once from Lemma 1.5.6.

A simple inductive extension of the argument just used enables one to prove Theorem 1.4.6 in full generality. We leave this as an exercise for the reader.

Theorem 1.5.4 is a variant due to Dynkin on a much older result from measure theory called the **monotone class theorem**. To state the monotone class theorem we need the notion of a monotone class of sets:

**Definition 1.5.8** A designated collection \( \mathcal{M} \) of subsets of a given set \( E \) is called a **monotone class** over \( E \) if, whenever \( \{ A_n, n = 1, 2, \ldots \} \) is a monotone sequence of members of \( \mathcal{M} \), the limit \( \lim_{n \to \infty} A_n \) of the sequence is also a member of \( \mathcal{M} \).

It follows from Proposition 1.5.3(a) and the De Morgan identities of set theory that a \( \lambda \)-class is a monotone class, although the converse is not necessarily true.

**Theorem 1.5.9** Suppose \( \mathcal{A} \) is an algebra and \( \mathcal{M} \) is a monotone class, both over a common set \( E \). If \( \mathcal{A} \subset \mathcal{M} \) then \( \sigma \{ \mathcal{A} \} \subset \mathcal{M} \).

Theorem 1.5.9 is established by an argument which is an easy modification of that used for Theorem 1.5.4, and so we shall refrain from giving the proof. Comparing Theorems 1.5.4 and 1.5.9, we note a good deal of similarity. There are many applications, particularly in probability theory, where the “inner” collection of sets is only a \( \pi \)-class and not an algebra, and for such applications Theorem 1.5.4 is the appropriate tool. On the other hand, if the “inner” collection of sets is in fact an algebra, then one should use Theorem 1.5.9, since it is easier to verify that the “outer” collection of sets is a monotone class than to show that it is a \( \lambda \)-class.

So far we have concerned ourselves with monotone class theorems for sets. We now look at monotone class theorems for functions. Henceforth the notation \( f_n \uparrow f \) will indicate that a sequence of \( \mathbb{R} \)-valued functions \( f_n \) is increasing pointwise to the \( \mathbb{R} \)-valued function \( f \), we shall write \( f \geq 0 \) to mean that \( f \) is a \([0, \infty]\)-valued function, and we shall use 1 to denote a constant function with the value of unity.

**Definition 1.5.10** A set \( \mathcal{H} \) of functions mapping a set \( E \) into \([0, \infty]\) is called an **\( \mathcal{L} \)-class** when the following conditions hold:

(i) \( 1 \in \mathcal{H} \);

(ii) If \( f \) and \( g \) are members of \( \mathcal{H} \) such that \( 0 \leq f(x) \leq g(x) \leq B, \forall x \in E \) and some constant \( B \in (0, \infty) \), then \( g - f \in \mathcal{H} \);
(iii) If \( f \) and \( g \) are members of \( \mathcal{H} \) then \( \alpha f + \beta g \in \mathcal{H}, \forall \alpha, \beta \in [0, \infty) \);
(iv) If \( \{f_n, n = 1, 2, \ldots\} \) is a sequence of functions in \( \mathcal{H} \) such that \( f_n \uparrow f \), then \( f \in \mathcal{H} \).

The next result is a monotone class theorem for functions:

**Theorem 1.5.11** Suppose \( \mathcal{D} \) is a \( \pi \)-class of sets over some set \( E \) and \( \mathcal{H} \) is an \( \mathcal{L} \)-class of functions on \( E \) such that \( I_D \in \mathcal{H} \) for each \( D \in \mathcal{D} \). Then \( \mathcal{H} \) includes all \([0, \infty]\)-valued \( \sigma\{\mathcal{D}\}\)-measurable functions on \( E \).

Proof: Put \( \mathcal{G} \overset{\triangle}{=} \{A \subset E : I_A \in \mathcal{H}\} \). Then we see that \( \mathcal{D} \subset \mathcal{G} \). Moreover condition (i) of Definition 1.5.10 ensures \( E \in \mathcal{G} \), while condition (ii) of Definition 1.5.10 implies that for \( B, A \in \mathcal{G} \) with \( A \subset B \), we have \( B - A \in \mathcal{G} \). Moreover, (iv) of Definition 1.5.10 ensures that the limit of any monotone increasing sequence of members of \( \mathcal{G} \) itself belongs to \( \mathcal{G} \). We see that \( \mathcal{G} \) is a \( \lambda \)-class of sets over \( E \) which includes the \( \pi \)-class \( \mathcal{D} \), hence Theorem 1.5.4 ensures that \( \sigma\{\mathcal{D}\} \subset \mathcal{G} \). In view of the form of \( \mathcal{G} \) and (iii) of Definition 1.5.10, this implies that any non-negative \( \sigma\{\mathcal{D}\}\)-measurable simple function on \( E \) is a member of \( \mathcal{H} \), since we can always write such a function in the form \( \sum_{i=1}^{N} \alpha_i I_{A_i} \) for some \( \alpha_i \in [0, \infty) \) and \( A_i \in \sigma\{\mathcal{D}\} \).

Now fix an arbitrary \( \sigma\{\mathcal{D}\}\)-measurable mapping \( f : E \to [0, \infty] \). By Proposition 1.2.9 there is some sequence \( \{f_n, n = 1, 2, \ldots\} \) of \([0, \infty]\)-valued \( \sigma\{\mathcal{D}\}\)-measurable simple functions on \( E \) such that \( \lim_{n \to \infty} f_n \uparrow f \). As already noted, we have \( f_n \in \mathcal{H}, \forall n = 1, 2, \ldots \), thus Definition 1.5.10(iv) ensures that \( f \in \mathcal{H} \).

As a final illustration of the methodology that goes with monotone class ideas, we establish the Tonelli Theorem 1.2.40:

**Proof of Theorem 1.2.40** : Suppose, to begin with, that \( \mu_1 \) and \( \mu_2 \) are finite measures, namely \( \mu_1(E_1) < \infty \) and \( \mu_2(E_2) < \infty \). Define

\[
\mathcal{D} \overset{\triangle}{=} \{A_1 \otimes A_2 : A_i \in \mathcal{S}_i, i = 1, 2\},
\]

and let \( \mathcal{C} \) be the collection of all sets \( C \in \mathcal{S}_1 \otimes \mathcal{S}_2 \) such that, with \( f \overset{\triangle}{=} I_C \), the mapping

\[
x_2 \to \int_{E_1} f(x_1, x_2) d\mu_1(x_1) : E_2 \to [0, \infty],
\]

is \( \mathcal{S}_2 \)-measurable, and

\[
\int_{E_1 \otimes E_2} f(x_1, x_2) d\mu_1 \otimes \mu_2(x_1, x_2) = \int_{E_2} \left( \int_{E_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2).
\]

Now \( \mathcal{D} \) is clearly a \( \pi \)-class over \( E_1 \otimes E_2 \), and one easily checks that \( \mathcal{C} \) is a \( \lambda \)-class over \( E_1 \otimes E_2 \). Indeed, it is clear that \( E_1 \otimes E_2 \in \mathcal{C} \), and, by linearity of integration, it follows that \( C_1, C_2 \in \mathcal{C} \) with \( C_1 \subset C_2 \) implies \( C_2 - C_1 \in \mathcal{C} \) (note that, since \( \mu_1 \) and \( \mu_2 \) are finite measures, the undefined combination \( \infty - \infty \) can never occur in the difference of two integrals). Lastly, one immediately verifies (iii) of Definition 1.5.2
by use of the Monotone Convergence Theorem 1.2.12. By definition of the product measure $\mu_1 \otimes \mu_2$, it is clear that $D \subset \mathcal{C}$, thus, by Theorem 1.5.4, we have $\sigma\{D\} \subset \mathcal{C} \subset S_1 \otimes S_2$, and since $\sigma\{D\} = S_1 \otimes S_2$ (recall Section 1.2.7), it follows that $\mathcal{C} = S_1 \otimes S_2$. We have therefore shown that the mapping in (1.75) is $S_2$-measurable and (1.76) holds when $f \triangleq I_C$, for each $C \in S_1 \otimes S_2$, and hence, by the linearity of integration, the same statements hold true when $f$ is a non-negative $S_1 \otimes S_2$-measurable simple function on $E_1 \otimes E_2$. Finally, this fact, together with Proposition 1.2.9 and Theorem 1.2.12, shows that (1.75) is $S_2$-measurable and (1.76) holds for each $S_1 \otimes S_2$-measurable mapping $f : E_1 \otimes E_2 \to [0, \infty]$. Generalization to the case where $\mu_1$ and $\mu_2$ are $\sigma$-finite is straightforward and left to the reader. ■
1.6 Problems

Problem 1.6.1 Suppose that $(E, S, \mu)$ is a $\sigma$-finite measure space and $f : E \to [0, \infty]$ is an $S$-measurable function. Establish that $\{(x, y) \in E \otimes [0, \infty] : f(x) \geq y\}$ is a member of the product $\sigma$-algebra $S \otimes B([0, \infty])$, and then show
\[
\int_E f \ d\mu = \int_0^\infty \mu\{x \in E : f(x) \geq y\} \ dy.
\]

Problem 1.6.2 Suppose that $X$ is a random variable taking values in $(0, \infty)$ and $\alpha \in [1, \infty)$ is a constant. Show that
\[
\frac{1}{EX^\alpha} \leq E\left[\frac{1}{X^\alpha}\right].
\]

Problem 1.6.3 Suppose $X : E \to F$ is a mapping from a set $E$ into a set $F$. If $\mathcal{C}$ is a given collection of subsets of $F$ (i.e. $\mathcal{C} \subset \mathcal{P}(F)$) we shall write $X^{-1}[\mathcal{C}] \triangleq \{X^{-1}(C) : C \in \mathcal{C}\}$. Show that
\[
\sigma\{X^{-1}[\mathcal{C}]\} = X^{-1}[\sigma(\mathcal{C})].
\]

Problem 1.6.4 $X$ and $Y$ are random variables on a probability space $(\Omega, \mathcal{F}, P)$ with $E|X| < \infty$, and, likewise, $\tilde{X}$ and $\tilde{Y}$ are random variables on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ with $E|\tilde{X}| < \infty$. According to Theorem 1.3.1, there are Borel-measurable mappings $\Phi : \mathbb{R} \to \mathbb{R}$ and $\tilde{\Phi} : \mathbb{R} \to \mathbb{R}$ such that
\[
E[X \mid Y] = \Phi(Y) \quad P \text{ a.s.} \quad \tilde{E}[\tilde{X} \mid \tilde{Y}] = \tilde{\Phi}(\tilde{Y}) \quad \tilde{P} \text{ a.s.}
\]
Suppose that the bivariate random vectors $(X, Y)$ and $(\tilde{X}, \tilde{Y})$ are identically distributed, namely
\[
P[(X, Y) \in B] = \tilde{P}[(\tilde{X}, \tilde{Y}) \in B], \quad B \in \mathcal{B}(\mathbb{R}^2),
\]
and put
\[
\mu(A) \triangleq P[Y \in A] = \tilde{P}[\tilde{Y} \in A], \quad A \in \mathcal{B}(\mathbb{R}).
\]
Show that
\[
\mu[y \in \mathbb{R} \mid \Phi(y) \neq \tilde{\Phi}(y)] = 0.
\]

Problem 1.6.5 Suppose that the random variable $X : (\Omega, \mathcal{F}, P) \to \mathbb{R}$ is such that $E|X|^2 < \infty$ and $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ are sub-$\sigma$-algebras. Show that
\[
E\left[|X - E[X \mid \mathcal{G}]|^2\right] \leq E\left[|X - E[X \mid \mathcal{H}]|^2\right].
\]

Problem 1.6.6 Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathcal{G} \subset \mathcal{F}$ is a sub-$\sigma$-algebra, and $\mathcal{D} \subset \mathcal{G}$ is a $\pi$-class with $\Omega \in \mathcal{D}$ and $\sigma(\mathcal{D}) = \mathcal{G}$. If $X$ and $Y$ are integrable random variables on $(\Omega, \mathcal{F}, P)$, and $E[X \mid D] = E[Y \mid D]$, $\forall \ D \in \mathcal{D}$, prove that
\[
E[X \mid \mathcal{G}] = E[Y \mid \mathcal{G}], \quad \text{a.s.}
\]
Hint: use Theorem 1.5.4
Problem 1.6.7 Suppose that \( \mathcal{G}_1, \mathcal{G}_2 \) and \( \mathcal{G}_3 \) are sub-\( \sigma \)-algebras in a probability space \((\Omega, \mathcal{F}, P)\). Then \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) are said to be conditionally independent given \( \mathcal{G}_3 \) when \( E[I_{A_1}I_{A_2} | \mathcal{G}_3] = E[I_{A_1} | \mathcal{G}_3] E[I_{A_2} | \mathcal{G}_3] \) a.s. for each \( A_1 \in \mathcal{G}_1 \) and \( A_2 \in \mathcal{G}_2 \). Establish the following: (a) If \( \sigma(\mathcal{G}_1, \mathcal{G}_3) \) and \( \mathcal{G}_2 \) are independent then \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) are conditionally independent given \( \mathcal{G}_3 \). (b) \( \mathcal{G}_1 \) and \( \mathcal{G}_2 \) are conditionally independent given \( \mathcal{G}_3 \) if and only if

\[
E[I_{A_1} | \sigma(\mathcal{G}_2, \mathcal{G}_3)] = E[I_{A_1} | \mathcal{G}_3], \quad \text{a.s.}
\]

for each \( A_1 \in \mathcal{G}_1 \). Hint: Use Problem 1.6.6.

Problem 1.6.8 Suppose that \( X, Y : (\Omega, \mathcal{F}, P) \to \mathbb{R} \) are random variables, \( \mathcal{G} \subset \mathcal{F} \) is a sub-\( \sigma \)-algebra, and \( X \) is \( \mathcal{G} \)-measurable. If

\[
E[f(Y) | \mathcal{G}] = f(X), \quad \text{a.s.}
\]

for each uniformly bounded \( \mathcal{B}(\mathbb{R}) \)-measurable function \( f : \mathbb{R} \to \mathbb{R} \), then show that \( X = Y \) a.s.

Problem 1.6.9 Suppose that \( X : (\Omega, \mathcal{F}, P) \to \mathbb{R}^{d_1} \) and \( Y : (\Omega, \mathcal{F}, P) \to \mathbb{R}^{d_2} \) are random vectors, and \( \mathcal{G} \subset \mathcal{F} \) is a sub-\( \sigma \)-algebra. If \( Y \) and \( \mathcal{G} \) are \( P \)-independent, \( X \) is \( \mathcal{G} \)-measurable, and \( f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R} \) is \( \mathcal{B}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \)-measurable with \( E[|f(X, Y)|] < \infty \), show that

\[
E[f(X, Y) | \mathcal{G}] = E[f(X, Y) | X] \quad \text{a.s.}
\]

Hint: use Theorem 1.5.4 to show that the preceding holds when \( f = I_C \) for each \( C \in \mathcal{B}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \).

Problem 1.6.10 Suppose that the random vectors \( X : (\Omega, \mathcal{F}, P) \to \mathbb{R}^{d_1} \) and \( Y : (\Omega, \mathcal{F}, P) \to \mathbb{R}^{d_2} \) are independent, and the mapping \( f : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R} \) is \( \mathcal{B}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \)-measurable with \( E[|f(X, Y)|] < \infty \). Define \( \Phi : \mathbb{R}^{d_1} \to \mathbb{R} \) by

\[
\Phi(x) \triangleq \begin{cases} 
E[f(x, Y)] & \forall x \in \mathbb{R}^{d_1} \text{ such that } E[|f(x, Y)|] < \infty \\
0 & \text{otherwise}.
\end{cases}
\]

Prove that \( \Phi(\cdot) \) is \( \mathcal{B}(\mathbb{R}^{d_1}) \)-measurable, and \( \Phi(X) = E[f(X, Y) | X] \) a.s.

Problem 1.6.11 (a) Suppose that \( \mathcal{G} \) is a sub-\( \sigma \)-algebra in the probability space \((\Omega, \mathcal{F}, P)\), \( X \) and \( Y \) are random variables on \((\Omega, \mathcal{F}, P)\) with \( E|X| < \infty \) and \( E|Y| < \infty \), and \( A \in \mathcal{G} \). Show that if

\[
P\{\omega \in A : X(\omega) \neq Y(\omega)\} = 0
\]

then

\[
P\{\omega \in A : E[X | \mathcal{G}](\omega) \neq E[Y | \mathcal{G}](\omega)\} = 0.
\]

(b) If \( \mathcal{G} \) is a sub-\( \sigma \)-algebra in the probability space \((\Omega, \mathcal{F}, P)\) and \( A \in \mathcal{G} \), then we put

\[
A \cap \mathcal{G} \triangleq \{A \cap C : C \in \mathcal{G}\}
\]

50
and call \( A \cap G \) the trace of the \( \sigma \)-algebra \( G \) on the event \( A \). Now suppose that \( H \) is another sub-\( \sigma \)-algebra in \((\Omega, \mathcal{F}, P)\), and event \( A \) is such that \( A \in G \cap H \) and
\[
A \cap G = A \cap H.
\]
i.e. the trace of \( G \) on \( A \) is identical to the trace of \( H \) on \( A \). If \( X \) is a random variable on \((\Omega, \mathcal{F}, P)\) with \( E|X| < \infty \), show that
\[
P\{\omega \in A : E[X | G] (\omega) \neq E[X | H] (\omega)\} = 0.
\]
Hint: First show that \( I_A E[X | G] \) and \( I_A E[X | H] \) are \( G \cap H \)-measurable.

(c) Suppose that event \( A \) and the sub-\( \sigma \)-algebras \( G \) and \( H \) are as in (b). If \( X \) and \( Y \) are random variables on \((\Omega, \mathcal{F}, P)\) with \( E|X| < \infty \) and \( E|Y| < \infty \) and
\[
P\{\omega \in A : X(\omega) \neq Y(\omega)\} = 0
\]
then show that
\[
P\{\omega \in A : E[X | G] (\omega) \neq E[Y | H] (\omega)\} = 0.
\]
The significance of this result is that the restriction of \( E[X | G] \) to an event \( A \) depends only on the “localization” of \( X \) and \( G \) to the event \( A \).

**Problem 1.6.12** Suppose that \( \Phi, W : (\Omega, \mathcal{F}, P) \to \mathbb{R} \) are independent random variables and \( W \) is Gaussian distributed with zero-mean and unit variance. Put
\[
\tilde{W} \overset{\Delta}{=} W - \Phi, \quad \Lambda \overset{\Delta}{=} \exp \left[ \Phi W - \frac{1}{2} \Phi^2 \right],
\]
and define measure \( \tilde{P} \) on \((\Omega, \mathcal{F})\) by
\[
\tilde{P}(A) \overset{\Delta}{=} E[\Lambda; A], \quad \forall \ A \in \mathcal{F}.
\]
Establish the following:
(i) \( \tilde{P} \) is a probability measure on \((\Omega, \mathcal{F})\);
(ii) with respect to the measure \( \tilde{P} \) the random variable \( \tilde{W} \) is Gaussian with zero mean and unit variance (this result is a precursor of the Girsanov theorem which will be studied in Chapter 5);
(iii) \( \tilde{P}[\Phi \in \Gamma] = P[\Phi \in \Gamma], \forall \ \Gamma \in \mathcal{B}(\mathbb{R}) \), that is \( \Phi \) has the same distribution with respect to the measures \( P \) and \( \tilde{P} \);
(iv) \( \tilde{W} \) and \( \Phi \) are \( \tilde{P} \)-independent;
(vi) for each bounded Borel-measurable \( f : \mathbb{R} \to \mathbb{R} \) we have
\[
E \left[ f(\Phi) | \tilde{W} \right] = H(\tilde{W}),
\]
where
\[
H(w) \overset{\Delta}{=} \frac{E[f(\Phi) \exp\{-\Phi w - (1/2)\Phi^2\}]}{E[\exp\{-\Phi w - (1/2)\Phi^2\}]} , \quad \forall \ w \in \mathbb{R}.
\]
This result is a very special case of the Kallianpur-Striebel formula in nonlinear filtering.
Problem 1.6.13 Suppose that $X_n : (\Omega, \mathcal{F}, P) \to \mathbb{R}, n = 1, 2, \ldots$, is a sequence of random variables such that
\[
\sup_n |X_n| < \infty, \quad \text{a.s.}
\]
Establish the following: for each $\epsilon \in (0, \infty)$ there is some number $K \in (0, \infty)$ and event $A \in \mathcal{F}$ (depending on $\epsilon$) such that $P(A) \leq \epsilon$ and $|X_n(\omega)| \leq K$ for each $\omega \not\in A$ and $n = 1, 2, \ldots$.

That is, any sequence of random variables which is a.s. pointwise bounded is uniformly bounded outside sets of arbitrarily small probability.

Hint: Use Theorem 1.2.5(d).

Problem 1.6.14 Suppose that $X$ and $Y$ are random variables on the probability space $(\Omega, \mathcal{F}, P)$.

Establish the following:
(a) If $P[Y \leq y < X] = 0$ for all $y \in \mathbb{R}$, then $P[Y < X] = 0$.
(b) If $E|X| < \infty$, $E|Y| < \infty$, and $E[X|Y] = Y$ a.s. and $E[Y|X] = X$ a.s.

then $P[X = Y] = 1$. Hint for (b): Use $E[X|Y] = Y$ to show that $E[X-Y; Y > y] = 0$ for all $y \in \mathbb{R}$.

Now observe that $\{Y > y\} = \{X \leq y < Y\} \cup \{X > y\} \cap \{Y > y\}$.

Problem 1.6.15 Random variables $\Lambda$ and $X$ are defined on the probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$ is a sub-$\sigma$-algebra. The random variable $\Lambda$ is $P$-strictly positive (recall Remark 1.2.36) and $E\Lambda = 1$.

Define $\tilde{P}(A) \triangleq E[\Lambda; A], A \in \mathcal{F}$. Establish the following:
(i) $\tilde{P}$ is a probability measure on $(\Omega, \mathcal{F})$ equivalent to $P$ (i.e. $\tilde{P} \equiv P[\mathcal{F}]$);
(ii) $E[\Lambda | \mathcal{G}]$ is $P$-strictly positive as well as $\tilde{P}$-strictly positive, and
\[
\frac{d\tilde{P}}{dP} \bigg|_\mathcal{G} = E[\Lambda | \mathcal{G}] \quad \text{a.s.} \quad \frac{dP}{d\tilde{P}} \bigg|_\mathcal{G} = \frac{1}{E[\Lambda | \mathcal{G}]} \quad \text{a.s.}
\]
(iii) if $E|X| < \infty$ then $E[|\Lambda X|] < \infty$ and
\[
\tilde{E}[X | \mathcal{G}] = \frac{E[\Lambda X | \mathcal{G}]}{E[\Lambda | \mathcal{G}]} \quad \text{a.s.}
\]

Problem 1.6.16 Suppose that $X$ is a random variable on the probability space $(\Omega, \mathcal{F}, P)$ with $E|X| < \infty$, that $\mathcal{G} \subset \mathcal{F}$ is a sub-$\sigma$-algebra, and that $X$ and $E[X | \mathcal{G}]$ have identical cumulative distribution functions. Prove that $X = E[X | \mathcal{G}]$ a.s.

Problem 1.6.17 Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathcal{G} \subset \mathcal{F}$ is a sub-$\sigma$-algebra, and $(\Xi, S, \mu)$ is a finite measure space. Let $\xi \mapsto X(\xi, \omega) : \Xi \otimes \Omega \to \mathbb{R}$ be a $S \otimes \mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable mapping such that:
\[
(1.77) \quad E\left[\int_\Xi |X(\xi, \omega)| \, d\mu(\xi)\right] < \infty
\]
Show that there exists a $S⊗G/B(IR)$-measurable mapping $Z : Ξ⊗Ω → IR$ with the following properties:

(a) For $P$-almost all $ω$ we have
\[ \int_Ξ |Z(ξ, ω)| dμ(ξ) < ∞ \quad P - a.s; \]

(b) for each $ξ ∈ Ξ$ such that $E|X(ξ)| < ∞$ we have
\[ Z(ξ, ·) = E [X(ξ) | G] (·) \quad P - a.s; \]

and
\[ \int_Ξ Z(ξ)dμ(ξ) = E \left[ \int_Ξ X(ξ)dμ(ξ) \left| G \right. \right] \quad P - a.s. \]

(c) For each $G/S$-measurable mapping $φ : Ω → Ξ$ such that $E|Y| < ∞$, where $Y : Ω → IR$ is defined by $Y(ω) \overset{Δ}{=} X(φ(ω), ω), \forall ω ∈ Ω$, we have
\[ E [Y | G] (·) = Z(φ(·), ·) \quad P - a.s. \]

The result in (b) justifies the following *Fubini theorem for conditional expectations*: Subject to the preceding conditions we have
\[ \int_Ξ E [X(ξ) | G] dμ(ξ) = E \left[ \int_Ξ X(ξ)dμ(ξ) \left| G \right. \right] \quad P - a.s. \]

Hint: First establish the result when $X$ has the special product form $X(ξ, ω) = I_A(ξ)I_B(ω)$ for some $A ∈ S$ and $B ∈ F$ (this is just a semi-trivial observation). Then use Theorem 1.5.11 to show the result when $X$ is non-negative and (1.77) holds. Finally, generalize to the case where $X$ takes positive and negative values.
Chapter 2

Discrete-Parameter Martingales

The stochastic integral and its associated calculus was pioneered by Itô([15], [16], [17]) some half-century ago as a tool for constructing a special class of Markov processes called diffusions, and for some considerable time thereafter stochastic integration theory was seen as belonging properly to that general part of probability theory concerned specifically with diffusions. A significant change in this point of view occurred with the appearance of Kunita and Watanabe's celebrated paper [21], when it became clear that stochastic calculus was both significantly simplified and greatly extended in scope and generality when developed in the context of martingales. Indeed, modern stochastic calculus is essentially a “calculus of martingales” for which the basic tools of martingale theory are indispensable, and thus it is with martingales that we must begin. Stochastic calculus is concerned with continuous-parameter (or continuous-time) stochastic processes, and therefore we shall mostly be interested in martingales in a continuous-parameter setting. One cannot, however, hope to understand continuous-parameter martingales without first acquiring a thorough familiarity with some of the basic elements of discrete-parameter martingales. In this chapter our goal is to develop some of these basic elements, deferring a study of the continuous-parameter case to Chapter 4.

2.1 Filtrations and Stopping Times

Remark 2.1.1 In the following definitions we shall use $I$ to denote some fixed subset of the set of all integers, namely $I \subset \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. We emphasize that neither $-\infty$ nor $+\infty$ are integers, and hence are not members of $I$.

Definition 2.1.2 Suppose that $I$ is as in Remark 2.1.1 and $(\Omega, \mathcal{F}, P)$ is a probability space. An indexed collection $\{\mathcal{F}_n, n \in I\}$ of $\sigma$-algebras $\mathcal{F}_n$ over $\Omega$ is called a filtration in $(\Omega, \mathcal{F}, P)$ when $\mathcal{F}_m \subset \mathcal{F}_n \subset \mathcal{F}$ for all $m, n \in I$ with $m < n$. A collection of $\mathbb{R}^d$-valued random variables $\{X_n; n \in I\}$ on $(\Omega, \mathcal{F}, P)$ is said to be adapted to the filtration $\{\mathcal{F}_n, n \in I\}$ when $X_n$ is $\mathcal{F}_n/\mathcal{B}(\mathbb{R}^d)$-measurable for each $n \in I$. Under these conditions we call the indexed collection of pairs $\{(X_n, \mathcal{F}_n); n \in I\}$ an $\mathbb{R}^d$-valued adapted
sequence or an adapted process.

Remark 2.1.3 We shall mostly be dealing with filtrations and adapted processes indexed by the set of all non-negative integers, namely $I \triangleq \{0,1,2,3,\ldots\}$, and here we shall use the more expressive notations $\{\mathcal{F}_n, n = 0,1,2,\ldots\}$ and $\{(X_n, \mathcal{F}_n); n = 0,1,2,\ldots\}$. Although $+\infty$ is not a member of this indexing set we can nevertheless define the $\sigma$-algebra

$$\begin{equation}
\mathcal{F}_\infty \triangleq \sigma\{\mathcal{F}_n, n = 0,1,2,\ldots\},
\end{equation}$$

in order to get an “extended” filtration $\{\mathcal{F}_n, n = 0,1,2,\ldots, +\infty\}$ which is indexed by the extended non-negative integers. The $\sigma$-algebra in (2.1) will play an important role when we look at limiting properties of martingales later in this chapter.

Example 2.1.4 Filtrations typically arise in the following way: Suppose that $\{Y_n; n = 0,1,2\ldots\}$ is a given sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$ and put

$$\begin{equation}
\mathcal{F}_n \triangleq \sigma\{Y_0,Y_1,\ldots Y_n\}, \ \forall n = 0,1,2,\ldots
\end{equation}$$

Clearly $\{\mathcal{F}_n, n = 0,1,2,\ldots\}$ is a filtration in $(\Omega, \mathcal{F}, P)$ and $\{(Y_n, \mathcal{F}_n); n = 0,1,2,\ldots\}$ is an adapted sequence. The filtration defined by (2.2) is called the raw or natural filtration of the sequence of random variables $\{Y_n; n = 0,1,2\ldots\}$, and we usually write $\mathcal{F}^Y_n$ for the $\sigma$-algebra in (2.2). Raw filtrations have the following property: Suppose that $\{X_n; n = 0,1,2\ldots\}$ is another sequence of random variables on $(\Omega, \mathcal{F}, P)$ which is adapted to $\{\mathcal{F}^Y_n, n = 0,1,2,\ldots\}$. Then $X_n$ is measurable with respect to $\mathcal{F}^Y_n \triangleq \sigma\{Y_0,Y_1,\ldots Y_n\}$ and consequently Theorem 1.3.21 shows that there is some $B(\mathbb{R}^{n+1})$-measurable mapping $\Psi_n: \mathbb{R}^{n+1} \to \mathbb{R}$ such that

$$X_n(\omega) = \Psi_n(Y_0(\omega), Y_1(\omega),\ldots Y_n(\omega)), \ \forall \omega \in \Omega.$$ 

That is, when the sequence of random variables $\{X_n; n = 0,1,2\ldots\}$ is adapted to the raw filtration $\{\mathcal{F}^Y_n, n = 0,1,2,\ldots\}$ of $\{Y_n; n = 0,1,2,\ldots\}$, then each random variable $X_n$ is completely determined by the random variables $Y_0,Y_1,\ldots Y_n$.

Having formulated the notion of a filtration we are now ready to introduce the idea of a stopping time. Suppose that $(\Omega, \mathcal{F}, P)$ is a given probability space. Then a mapping $T: \Omega \to \{0,1,2,\ldots, +\infty\}$ is called a discrete-parameter random time when $\{T \leq n\} \in \mathcal{F}, \ \forall n = 0,1,2,\ldots$. Observe that $T$ is permitted by its very definition to assume the value $+\infty$, and the condition $\{T \leq n\} \in \mathcal{F}$ holds trivially at $n = +\infty$ since $\{T \leq +\infty\} = \Omega$. Of particular interest is the following refinement of a random time:

Definition 2.1.5 Suppose that $\{\mathcal{F}_n, n = 0,1,2,\ldots\}$ is a given filtration on a probability space $(\Omega, \mathcal{F}, P)$. Then a random time $T: \Omega \to \{0,1,2,\ldots, +\infty\}$ is called a discrete-parameter $\{\mathcal{F}_n\}$-stopping time when $\{T \leq n\} \in \mathcal{F}_n$ for all $n = 0,1,2,\ldots$
We shall usually omit the appellation “discrete-parameter” for random times and stopping times when this is clear from the context.

**Remark 2.1.6** Suppose that \( T : \Omega \rightarrow \{0,1,2,\ldots,+\infty\} \) is a discrete-parameter random time and \( \{\mathcal{F}_n, n = 0,1,2,\ldots\} \) is a filtration such that \( \{T = n\} \in \mathcal{F}_n \) for each \( n = 0,1,2,\ldots \). Since \( \mathcal{F}_k \subset \mathcal{F}_n \) for all \( k = 0,1,2,\ldots,n \), and
\[
\{T \leq n\} = \bigcup_{k=0}^{n}\{T = k\},
\]
we see that \( T \) is a \( \{\mathcal{F}_n\} \)-stopping time.

**Remark 2.1.7** If \( T \) is an \( \{\mathcal{F}_n\} \)-stopping time, then we see that
\[
\{T < \infty\} = \bigcup_{0 \leq n < \infty}\{T \leq n\} \in \mathcal{F}_\infty
\]
where \( \mathcal{F}_\infty \) is the \( \sigma \)-algebra defined in (2.1), and thus \( \{T = \infty\} \in \mathcal{F}_\infty \).

The next example gives some commonly occurring stopping times:

**Example 2.1.8** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0,1,2,\ldots\} \) is a given \( \mathbb{R} \)-valued adapted sequence on some probability space \((\Omega, \mathcal{F}, P)\), and let \( a \in \mathbb{R} \). For each \( \omega \) define
\[
S(\omega) \overset{\Delta}{=} \min\{k \geq 0 ; X_k(\omega) \leq a\}
\]
In this definition we use the convention that \( \min\{\emptyset\} = +\infty \), that is, if \( X_k(\omega) > a \) for all \( k = 0,1,2,\ldots \) then we put \( S(\omega) \overset{\Delta}{=} +\infty \). Then \( S \) is a \( \{\mathcal{F}_n\} \)-stopping time. Indeed, since \( X_n \) is \( \mathcal{F}_n \)-measurable for each \( n \), we have \( \{S = 0\} = \{X_0 \leq a\} \in \mathcal{F}_0 \) and
\[
\{S = n\} = \left[ \bigcap_{k=0}^{n-1}\{X_k > a\} \right] \cap \{X_n \leq a\} \in \mathcal{F}_n
\]
for all \( n = 1,2,\ldots \), whence \( S \) is an \( \{\mathcal{F}_n\} \)-stopping time. Next, suppose that \( b > a \) and define
\[
T(\omega) \overset{\Delta}{=} \min\{k > S(\omega) ; X_k(\omega) \geq b\}
\]
for all \( \omega \) (again we use the convention \( \min\{\emptyset\} = +\infty \)). Then \( \{T = 0\} = \emptyset \in \mathcal{F}_0 \). Now \( S < T \) and thus, for each \( n = 1,2,\ldots \) we have
\[
\{T = n\} = \bigcup_{k=0}^{n-1}\left[ \{T = n\} \cap \{S = k\} \right]
\]
But for all \( n = 1,2,\ldots \) and \( k = 0,1,2,\ldots,n-1 \) it is clear that
\[
\{T = n\} \cap \{S = k\} = \{S = k\} \cap \left[ \bigcap_{i=k+1}^{n-1}\{X_i < b\} \right] \cap \{X_n \geq b\} \in \mathcal{F}_n
\]
It follows from (2.6) and (2.7) that \( \{T = n\} \in \mathcal{F}_n \) for all \( n = 0,1,2,\ldots \), so that \( T \) is a \( \{\mathcal{F}_n\} \)-stopping time.
Remark 2.1.9 Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is an \( \mathbb{R}^d \)-valued adapted sequence on a probability space \( (\Omega, \mathcal{F}, P) \). For any given \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \) define
\[
D_\Gamma(\omega) \triangleq \min\{k \geq 0 : X_k(\omega) \in \Gamma\}, \quad \forall \omega \in \Omega,
\]
where we again take \( \min\{\emptyset\} = +\infty \). An obvious generalization of the argument used in Example 2.1.8 establishes that \( D_\Gamma \) is an \( \{\mathcal{F}_n\} \)-stopping time. As will be indicated in Chapter 3, generalizing this latter assertion to a continuous-parameter setting is very non-trivial and involves some profound issues in the modern theory of stochastic processes (see Remark 3.3.10).

Definition 2.1.10 Suppose that \( T : \Omega \to \{0, 1, 2, \ldots, +\infty\} \) is a stopping time with respect to some filtration \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) on a probability space \( (\Omega, \mathcal{F}, P) \). We denote by \( \mathcal{F}_T \) the collection of all sets \( A \subset \Omega \) which have the property that
\[
A \cap \{T \leq n\} \in \mathcal{F}_n
\]
for each \( n = 0, 1, 2, \ldots, +\infty \) (recalling the definition of \( \mathcal{F}_\infty \) in (2.1)). It is readily checked that \( \mathcal{F}_T \) is a \( \sigma \)-algebra over \( \Omega \). We call \( \mathcal{F}_T \) the pre-\( \sigma \)-algebra generated by \( T \).

The next proposition is almost trivial but is stated in full to emphasize the simplest properties of pre-\( \sigma \)-algebras. The elementary proofs are left to the reader:

**Proposition 2.1.11** Suppose \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) is a given filtration in the probability space \( (\Omega, \mathcal{F}, P) \) and \( T \) is a \( \{\mathcal{F}_n\} \)-stopping time. Then:
(a) \( \mathcal{F}_T \) is a \( \sigma \)-algebra over \( \Omega \) and \( \mathcal{F}_T \subset \mathcal{F}_\infty \);
(b) \( T \) is \( \mathcal{F}_T \)-measurable, that is \( \sigma\{T\} \subset \mathcal{F}_T \);
(c) \( \mathcal{F}_T \) is the collection of all sets \( A \subset \Omega \) with the property that \( A \cap \{T = n\} \in \mathcal{F}_n \) for each \( n = 0, 1, 2, \ldots, +\infty \).
(d) If, for some constant \( n \in \{0, 1, 2, \ldots, +\infty\} \), we have \( T(\omega) = n \), \( \forall \omega \in \Omega \), then the \( \sigma \)-algebras \( \mathcal{F}_T \) and \( \mathcal{F}_n \) are identical.

The next proposition collects more basic facts about stopping times and pre-\( \sigma \)-algebras:

**Proposition 2.1.12** Suppose that \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) is a filtration in the probability space \( (\Omega, \mathcal{F}, P) \), and \( S \) and \( T \) are \( \{\mathcal{F}_n\} \)-stopping times. Then
(a) \( S \wedge T, S \vee T \) and \( S + T \) are \( \{\mathcal{F}_n\} \)-stopping times;
(b) If \( S(\omega) \leq T(\omega) \) for all \( \omega \) then \( \mathcal{F}_S \subset \mathcal{F}_T \);
(c) \( \mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T \);
(d) The events \( \{S < T\}, \{S = T\} \) and \( \{S > T\} \) are members of \( \mathcal{F}_{S\wedge T} \);
(e) If \( A \in \mathcal{F}_S \) then the events \( A \cap \{S \leq T\}, A \cap \{S < T\} \) and \( A \cap \{S = T\} \) are members of \( \mathcal{F}_{S\wedge T} \).
Proof: (a) For each \( n = 0, 1, 2, \ldots \) we see that \( \{ S \land T \leq n \} = \{ S \leq n \} \cup \{ T \leq n \} \in \mathcal{F}_n \) and \( \{ S \lor T \leq n \} = \{ S \leq n \} \cap \{ T \leq n \} \in \mathcal{F}_n \). As for \( S + T \), observe that
\[
\{ S + T = n \} = \bigcup_{k=0}^{n} [\{ S = n - k \} \cap \{ T = k \}] \in \mathcal{F}_n
\]
for each \( n = 0, 1, 2, \ldots \).

(b) Fix some arbitrary \( A \in \mathcal{F}_S \). Now, for each \( n = 0, 1, 2, \ldots, \infty \), it is clear that \( \{ T \leq n \} \subset \{ S \leq n \} \) and hence
\[
A \cap \{ T \leq n \} = (A \cap \{ S \leq n \}) \cap \{ T \leq n \} \in \mathcal{F}_n,
\]
since \( \{ T \leq n \} \) and \( A \cap \{ S \leq n \} \) are members of \( \mathcal{F}_n \) (the assertion that \( A \cap \{ S \leq n \} \in \mathcal{F}_n \) follows from \( A \in \mathcal{F}_S \) and the definition of \( \mathcal{F}_S \) - see (2.8)).

(c) Observe from (b) that \( \mathcal{F}_{S \land T} \subset \mathcal{F}_S \cap \mathcal{F}_T \). To establish the opposite set inclusion, fix any \( A \in \mathcal{F}_S \cap \mathcal{F}_T \) and any \( n = 0, 1, 2, \ldots, \infty \). Then
\[
A \cap \{ S \land T \leq n \} = A \cap [\{ S \leq n \} \cup \{ T \leq n \}] \in \mathcal{F}_n
\]
since \( A \cap \{ S \leq n \} \) and \( A \cap \{ T \leq n \} \) are members of \( \mathcal{F}_n \). By the arbitrary choice of \( n \), we see that \( A \in \mathcal{F}_{S \land T} \).

(d) First we shall show that
\[
(2.9) \quad A \cap \{ S \leq T \} \in \mathcal{F}_T
\]
for all \( A \in \mathcal{F}_S \). Thus fix some event \( A \in \mathcal{F}_S \) as well as some \( n \in \{ 0, 1, 2, \ldots, +\infty \} \). Clearly \( \{ S \leq T \} \cap \{ T \leq n \} \subset \{ S \leq n \} \) and
\[
\{ S \leq n \} \cap \{ T \leq n \} \cap \{ S \leq T \} = \{ S \leq n \} \cap \{ T \leq n \} \cap \{ S \land n \leq T \land n \},
\]
thus
\[
(2.10) \quad (A \cap \{ S \leq T \}) \cap \{ T \leq n \} = A \cap \{ S \leq n \} \cap \{ S \leq T \} \cap \{ T \leq n \} = A \cap \{ S \leq n \} \cap \{ T \leq n \} \cap \{ S \land n \leq T \land n \}.
\]

Now \( A \cap \{ S \leq n \} \) and \( \{ T \leq n \} \) are members of \( \mathcal{F}_n \). In view of Proposition 2.1.11(b) we see that \( S \land n \) is \( \mathcal{F}_{S \land n} \)-measurable and hence, by (c), is also \( \mathcal{F}_n \)-measurable. Likewise \( T \land n \) is \( \mathcal{F}_n \)-measurable, so that \( \{ S \land n \leq T \land n \} \in \mathcal{F}_n \). It follows from (2.10) that \( (A \cap \{ S \leq T \}) \cap \{ T \leq n \} \in \mathcal{F}_n \), \( \forall n = 0, 1, 2, \ldots, \infty \), and (2.9) is now a consequence of the definition of \( \mathcal{F}_T \) (see (2.8)). Taking \( A \upharpoonright \Omega \) in (2.9) shows that \( \{ S \leq T \} \in \mathcal{F}_T \) and hence \( \{ S \leq T \} \in \mathcal{F}_T \). Next, put \( R \upharpoonright S \land T \) and observe from (a) that \( R \) is a stopping time and therefore must be \( \mathcal{F}_R \)-measurable (see Proposition 2.1.11(b)). Now, from (b), we see that \( \mathcal{F}_R \subset \mathcal{F}_S \) hence \( R \) is \( \mathcal{F}_S \)-measurable and thus \( \{ R < S \} \in \mathcal{F}_S \). But clearly \( \{ R < S \} = \{ T < S \} \) so we get \( \{ S > T \} \in \mathcal{F}_S \). Thus \( \{ S > T \} \in \mathcal{F}_S \cap \mathcal{F}_T \) and (from symmetry) \( \{ T > S \} \in \mathcal{F}_S \cap \mathcal{F}_T \). Finally \( \{ S = T \} = \{ S > T \}^c \cap \{ T > S \}^c \in \mathcal{F}_S \cap \mathcal{F}_T \).
Related to this observation is the following simple result: and it is called a martingale. If
\( \{X_n; n \in I\} \) is a supermartingale, then one sees from Definition 2.2.1 that
\( \mathbb{E}[X_n | F_m] = X_m \) a.s. \( \forall m, n \in I \) with \( m \leq n \).

On the other hand, an \( \mathbb{R} \)-valued adapted sequence \( \{X_n, F_n; n \in I\} \) is called a submartingale when
\( \{(-X_n, F_n); n \in I\} \) is a supermartingale, that is, when (a) holds and (b) is replaced by
\( (c) \mathbb{E}[X_n | F_m] \geq X_m \) a.s. \( \forall m, n \in I \) with \( m \leq n \);
and it is called a martingale when it is both a supermartingale and a submartingale. Clearly then, an
\( \mathbb{R} \)-valued adapted sequence \( \{X_n, F_n; n \in I\} \) is a martingale when (a) holds and \( \mathbb{E}[X_n | F_m] = X_m \)
a.s. \( \forall m, n \in I \) with \( m \leq n \).

Remark 2.2.2 Condition (b) in Definition 2.2.1 of a supermartingale suggests a \emph{downward} trend in \( X_n \)
as the indexing element \( n \) increases, hence we may regard a supermartingale as a probabilistic analogue
of a non-increasing sequence of real numbers indexed by \( I \). Similarly, a submartingale [martingale] may
be viewed as the probabilistic analogue of a non-decreasing [constant] sequence of real numbers. These
analogues will be quite useful for motivating the martingale convergence theorems in later sections (see
Remark 2.5.8).

Remark 2.2.3 Using the basic properties of conditional expectations one easily sees that an \( \mathbb{R} \)-valued
adapted sequence \( \{X_n, F_n; n \in I\} \) is a martingale if and only if
\( (a) \mathbb{E}|X_n| < \infty, \forall n \in I, \) and
\( (b) \mathbb{E}[X_n I_A] = \mathbb{E}[X_m I_A] \) whenever \( m, n \in I \) with \( m \leq n \) and \( A \in F_m \).
Likewise, the \( \mathbb{R} \)-valued adapted sequence \( \{X_n, F_n; n \in I\} \) is a [sub/super]-martingale if and only if
(a) and (b) hold with equality in (b) replaced by \( \geq \) and \( \leq \) respectively.

If \( \{X_n, F_n; n \in I\} \) is supermartingale, then one sees from Definition 2.2.1 that \( \mathbb{E}[X_n] \leq \mathbb{E}[X_m] \) when
\( m, n \in I \) and \( m \leq n \), with equality of the expectations holding when \( \{X_n, F_n; n \in I\} \) is a martingale.
Related to this observation is the following simple result:

2.2 Martingales; Basic Definitions and Properties

The next definition formulates the notions of martingale, submartingale and supermartingale as \( \mathbb{R} \)-valued adapted processes with particular additional properties:

Definition 2.2.1 Suppose that \( I \) is as in Remark 2.1.1. A \emph{supermartingale} is an \( \mathbb{R} \)-valued adapted sequence
\( \{(X_n, F_n); n \in I\} \) on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\) such that
\( (a) \mathbb{E}[X_n] < \infty, \forall n \in I, \) and
\( (b) \mathbb{E}[X_n | F_m] \leq X_m \) a.s. \( \forall m, n \in I \) with \( m \leq n \).

On the other hand, an \( \mathbb{R} \)-valued adapted sequence \( \{(X_n, F_n); n \in I\} \) is called a \emph{submartingale} when
\( \{(-X_n, F_n); n \in I\} \) is a supermartingale, that is, when (a) holds and (b) is replaced by
\( (c) \mathbb{E}[X_n | F_m] \geq X_m \) a.s. \( \forall m, n \in I \) with \( m \leq n \);
and it is called a \emph{martingale} when it is both a supermartingale and a submartingale. Clearly then, an
\( \mathbb{R} \)-valued adapted sequence \( \{(X_n, F_n); n \in I\} \) is a martingale when (a) holds and \( \mathbb{E}[X_n | F_m] = X_m \)
a.s. \( \forall m, n \in I \) with \( m \leq n \).

Remark 2.2.2 Condition (b) in Definition 2.2.1 of a supermartingale suggests a \emph{downward} trend in \( X_n \)
as the indexing element \( n \) increases, hence we may regard a supermartingale as a probabilistic analogue
of a non-increasing sequence of real numbers indexed by \( I \). Similarly, a submartingale [martingale] may
be viewed as the probabilistic analogue of a non-decreasing [constant] sequence of real numbers. These
analogues will be quite useful for motivating the martingale convergence theorems in later sections (see
Remark 2.5.8).

Remark 2.2.3 Using the basic properties of conditional expectations one easily sees that an \( \mathbb{R} \)-valued
adapted sequence \( \{(X_n, F_n); n \in I\} \) is a martingale if and only if
\( (a) \mathbb{E}|X_n| < \infty, \forall n \in I, \) and
\( (b) \mathbb{E}[X_n I_A] = \mathbb{E}[X_m I_A] \) whenever \( m, n \in I \) with \( m \leq n \) and \( A \in F_m \).
Likewise, the \( \mathbb{R} \)-valued adapted sequence \( \{(X_n, F_n); n \in I\} \) is a [sub/super]-martingale if and only if
(a) and (b) hold with equality in (b) replaced by \( \geq \) and \( \leq \) respectively.

If \( \{(X_n, F_n); n \in I\} \) is supermartingale, then one sees from Definition 2.2.1 that \( \mathbb{E}[X_n] \leq \mathbb{E}[X_m] \) when
\( m, n \in I \) and \( m \leq n \), with equality of the expectations holding when \( \{(X_n, F_n); n \in I\} \) is a martingale.
Related to this observation is the following simple result:
Proposition 2.2.4 If \( \{(X_n, \mathcal{F}_n); n \in I\} \) is a supermartingale with constant expectation, namely \( E[X_m] = E[X_n], \forall m, n \in I, \) then it is a martingale.

Proof: Fix non-negative integers \( m, n \in I \) with \( m \leq n \), and let \( A \in \mathcal{F}_m \). Then the supermartingale property implies \( E[X_m; A^c] \geq E[X_n; A^c] \), so that \( E[X_m] - E[X_m; A] \geq E[X_n] - E[X_n; A] \), and therefore \( E[X_m; A] \leq E[X_n; A] \). But, again from the supermartingale property, \( E[X_m; A] \geq E[X_n; A] \), so that in fact \( E[X_m; A] = E[X_n; A] \). Since this holds for all \( A \in \mathcal{F}_m \) the result follows.

Next, we look at two particularly important examples of martingales:

Example 2.2.5 Suppose \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) is a filtration in a probability space \( (\Omega, \mathcal{F}, P) \) and \( X \) is a random variable on \( (\Omega, \mathcal{F}, P) \) such that \( E|X| < \infty \). Put

\[
X_n \triangleq E[X \mid \mathcal{F}_n], \quad \forall n = 0, 1, 2, \ldots
\]

Then, it follows at once from the composition rule for conditional expectations (Theorem 1.4.15(d)) that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a martingale.

Example 2.2.6 Suppose \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) is a filtration in a probability space \( (\Omega, \mathcal{F}, P) \) and \( \tilde{P} \) is a probability measure on \( (\Omega, \mathcal{F}) \) such that \( \tilde{P} \ll P[\mathcal{F}] \). Define the non-negative random variables

\[
X_n \triangleq \frac{d\tilde{P}}{dP} \bigg|_{\mathcal{F}_n}, \quad \forall n = 0, 1, 2, \ldots
\]

In view of Remark 1.4.18 we see that \( X_m = E[X_n \mid \mathcal{F}_m] \) \( P \)-a.s. \( \forall m, n \in \{0, 1, 2, \ldots\} \) with \( m \leq n \), hence the sequence of Radon-Nikodym derivatives \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a martingale on the probability space \( (\Omega, \mathcal{F}, P) \).

Remark 2.2.7 If \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a [sub/super]-martingale and \( p \in [1, \infty) \) a constant such that \( E|X_n|^p < \infty \) for each \( n = 0, 1, 2, \ldots \), then it is called an \( L^p \) [sub/super]-martingale, and if it is such that \( \sup_n E|X_n|^p < \infty \) then it is called an \( L^p \)-bounded [sub/super]-martingale.

For \( L^2 \)-martingales we have the following useful result:

Proposition 2.2.8 Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is an \( L^2 \)-martingale on the probability space \( (\Omega, \mathcal{F}, P) \). Then the following hold:

(a) For non-negative integers \( m, n, p, q \) with \( m < n \leq p < q \), we have \( E[(X_n - X_m)(X_q - X_p)] = 0 \);
(b) When \( m \leq n \) then \( E[(X_n - X_m)^2] = E[X_n^2 - X_m^2] \).

Proof: (a) Since \( X_n - X_m \) is \( \mathcal{F}_n \)-measurable and the martingale property of \( \{(X_n, \mathcal{F}_n); n \in I\} \) ensures that \( E[X_q - X_p \mid \mathcal{F}_p] = 0 \) a.s., we see from Theorem 1.4.15(c,e) that

\[
E[(X_n - X_m)(X_q - X_p)] = E[E[(X_n - X_m)(X_q - X_p) \mid \mathcal{F}_p]] = E[(X_n - X_m)E[(X_q - X_p) \mid \mathcal{F}_p]] = 0.
\]
(b) Clearly
\[
E[(X_n - X_m)^2] = E[X_n^2 - 2X_nX_m + X_m^2]
\]
and Theorem 1.4.15(c,e) together with \(E[X_n | F_m] = X_m\) a.s. says
\[
E[X_mX_n] = E[E[X_mX_n | F_m]] = E[X_mE[X_n | F_m]] = E[X_m^2].
\]
(2.13)

Combining (2.12) and (2.13) establishes part (b).

Remark 2.2.9 Proposition 2.2.8(a) says that an \(L^2\)-martingale always has orthogonal increments.

The next proposition shows that a submartingale results from a suitable convex mapping of a martingale or submartingale. This proposition is just a consequence of Jensen’s inequality for conditional expectations (Theorem 1.4.20):

Proposition 2.2.10 Suppose that \(\{(X_n, F_n); n = 0, 1, 2, \ldots\}\) is a martingale [resp. submartingale], and \(c : \mathbb{R} \to \mathbb{R}\) is a convex [resp. convex and non-decreasing] function such that \(E|c(X_n)| < \infty\) for all \(n = 0, 1, \ldots\) Then \(\{(c(X_n), F_n); n = 0, 1, 2, \ldots\}\) is a submartingale.

2.3 The Optional Sampling Theorem

Our goal in this section is to establish a special case of Doob’s celebrated optional sampling theorem. The idea is to show that the supermartingale property
\[
E[X_n | F_m] \leq X_m \quad \text{a.s. for all } n \geq m,
\]
persists when the fixed integers \(n\) and \(m\) are replaced by stopping times \(T\) and \(S\) respectively, with \(T \geq S\). To this end, we introduce the notion of a “sampling” random variable:

Remark 2.3.1 Suppose that \(\{X_n; n = 0, 1, 2, \ldots, +\infty\}\) is a collection of \(\mathbb{R}^d\)-valued random variables on the probability space \((\Omega, F, P)\) and \(T : \Omega \to \{0, 1, 2, \ldots, +\infty\}\) is a random time. The mapping \(X_T : \Omega \to \mathbb{R}^d\) defined by
\[
X_T(\omega) \overset{\Delta}{=} X_n(\omega) \quad \text{when } \omega \in \{T = n\}, \quad \forall n = 0, 1, 2, \ldots, +\infty.
\]
is said to sample the indexed collection \(\{X_n; n = 0, 1, 2, \ldots, +\infty\}\) at the random time \(T\).

Next, we show that \(X_T\) has nice measurability properties when \(T\) is a stopping time:

Proposition 2.3.2 Suppose that \(\{(X_n, F_n); n = 0, 1, 2, \ldots\}\) is an \(\mathbb{R}^d\)-valued adapted sequence on a probability space \((\Omega, F, P)\), \(F_\infty\) is defined by (2.1), and \(X_\infty : \Omega \to \mathbb{R}^d\) is a \(F_\infty\)-measurable function. If \(T : \Omega \to \{0, 1, 2, \ldots, +\infty\}\) is an \(\{F_n\}\)-stopping time then \(X_T\) is \(F_T\)-measurable.
Proof: Fix some \( n \in \{0, 1, 2, \ldots, +\infty\} \). For an arbitrary \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \) we see from Remark 2.3.1 that

\[
\{X_T \in \Gamma \} \cap \{T = n\} = \{X_n \in \Gamma \} \cap \{T = n\} \in \mathcal{F}_n.
\]

Since this holds for each \( n = 0, 1, 2, \ldots, +\infty \), we conclude from Proposition 2.1.11(c) that \( \{X_T \in \Gamma\} \in \mathcal{F}_T \).

We can now establish the optional sampling theorem for \emph{uniformly bounded} stopping times:

**Theorem 2.3.3** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a supermartingale and \( S \) and \( T \) are \( \{\mathcal{F}_n\}\)-stopping times on \( (\Omega, \mathcal{F}, P) \) such that \( 0 \leq S(\omega) \leq T(\omega) \leq N, \forall \omega \in \Omega \), for some constant \( N < \infty \). Then:

(a) \( E|X_T| < \infty \);

(b) \( E[X_T | \mathcal{F}_S] \leq X_S \quad \text{a.s.} \)

Proof: (a) Since \( T \leq N \) we see that \( |X_T| \leq |X_0| + |X_1| + \ldots + |X_N| \) hence \( E|X_T| < \infty \). To establish (b) notice that it is enough to show that

\[
E[X_T; A] \leq E[X_S; A], \quad \forall \ A \in \mathcal{F}_S.
\]

Indeed, having proved (2.15), we then see (from the definition of conditional expectations) that

\[
E[(E[X_T | \mathcal{F}_S] - X_S); A] \leq 0, \quad \forall \ A \in \mathcal{F}_S,
\]

so that Theorem 1.2.24(b), together with the \( \mathcal{F}_S \)-measurability of \( (E[X_T | \mathcal{F}_S] - X_S) \), ensures that (b) follows. To establish (2.15) fix some \( A \in \mathcal{F}_S \), and observe that

\[
\sum_{n=1}^{N} I\{S < n \leq T\}(\omega) \ (X_n(\omega) - X_{n-1}(\omega)) = \sum_{n=S(\omega)+1}^{T(\omega)} (X_n(\omega) - X_{n-1}(\omega))
\]

\[
\quad \quad \quad = X_T(\omega) - X_S(\omega), \quad \forall \omega \in \Omega.
\]

Taking expectations over \( A \) on each side of (2.16) then gives

\[
E[X_T - X_S; A] = E \left[ \sum_{n=1}^{N} I\{S < n \leq T\} \ (X_n - X_{n-1}); \ A \right]
\]

\[
= \sum_{n=1}^{N} E[X_n - X_{n-1}; \ A \cap \{S < n \leq T\}].
\]

Next, one clearly has \( A \cap \{S < n \leq T\} = (A \cap \{S \leq n - 1\}) \cap \{n - 1 < T\}, \forall n = 1, 2, \ldots, N \). Since \( S \) and \( T \) are stopping times we have \( \{n - 1 < T\} = \{T \leq n - 1\}^c \in \mathcal{F}_{n-1} \), and, since \( A \in \mathcal{F}_S \), we see from Proposition 2.1.12(b,e) that \( A \cap \{S \leq n - 1\} \in \mathcal{F}_{n-1} \), \( \forall n = 1, 2, \ldots, N \). It then follows that \( A \cap \{S < n \leq T\} \in \mathcal{F}_{n-1}, \forall n = 1, 2, \ldots, N \), and so, since \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a supermartingale, we get \( E[X_n - X_{n-1}; \ A \cap \{S < n \leq T\}] \leq 0 \). Using this inequality in (2.17) gives (2.15).
Theorem 2.3.3 holds for uniformly bounded stopping times. In applications one often comes across stopping times which can take the value $+\infty$, and the remainder of this section will therefore be devoted to establishing a variant of Theorem 2.3.3 for \textit{unbounded} stopping times. The first issue to be dealt with is the following: if $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a given supermartingale and $T$ is a $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$-stopping time which can take the value $+\infty$ then, in order for the “sampling” random variable $X_T$ in Remark 2.3.1 to be defined, we must also have available some $\mathcal{F}_\infty$-measurable random variable $X_{\infty}$. This motivates the following definition:

\textbf{Definition 2.3.4} The collection of pairs $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}$, indexed by the extended non-negative integers, is a \textbf{closed supermartingale} when:

\begin{enumerate}[(a)]
    \item $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a supermartingale on a probability space $(\Omega, \mathcal{F}, P)$;
    \item $\mathcal{F}_\infty$ is given by (2.1) and $X_\infty : \Omega \rightarrow \mathbb{R}$ is a $\mathcal{F}_\infty$-measurable random variable such that
\end{enumerate}

\begin{equation}
E[X_\infty] < \infty \quad \text{and} \quad X_n \geq E[X_\infty | \mathcal{F}_n] \quad \text{a.s.} \quad \forall \ n = 0, 1, 2, \ldots
\end{equation}

Likewise, the collection of pairs $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}$ is a \textbf{closed submartingale} when $\{(-X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}$ is a closed supermartingale; and it is a \textbf{closed martingale} when it is both a closed supermartingale and a closed submartingale.

\textbf{Remark 2.3.5} If $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a supermartingale and there exists some $\mathcal{F}_\infty$-measurable random variable $X_\infty$ such that $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}$ is a closed supermartingale, then the supermartingale $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is called \textbf{closable} and is said to be \textbf{closed} by the random variable $X_\infty$, while $X_\infty$ is called a \textbf{closing} random variable. Similar terminology will be used for submartingales and martingales.

\textbf{Remark 2.3.6} The following remarks are an immediate consequence of Definition 2.3.4:

\begin{enumerate}[(a)]
    \item If $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a supermartingale with $X_n \geq 0$ a.s. then it is closable, and $X_\infty(\omega) \overset{\Delta}{=} 0$, $\forall \ \omega \in \Omega$, defines a closing random variable.
    \item If a supermartingale $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is closed by a random variable $X_\infty$ then it is also closed by any $\mathcal{F}_\infty$-measurable random variable $Y_\infty$ such that $E[Y_\infty] < \infty$ and $X_\infty \geq Y_\infty$ a.s.
    \item If $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a closable \textit{martingale} then its closing random variable is a.s. unique. Moreover, this martingale has a very specific structure: namely, if $X_\infty$ is a closing random variable then
\end{enumerate}

$$X_n = E[X_\infty | \mathcal{F}_n] \quad \text{a.s.} \quad \forall \ n = 0, 1, 2, \ldots$$

\begin{enumerate}[(a)]
    \item Not every supermartingale $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is closable. For a trivial example, take $X_n \overset{\Delta}{=} -n$ and $\mathcal{F}_n \overset{\Delta}{=} \{\emptyset, \Omega\}$, $\forall \ n = 0, 1, 2, \ldots$. Then $\mathcal{F}_\infty = \{\emptyset, \Omega\}$, and there clearly fails to exist a $\mathcal{F}_\infty$-measurable random variable $X_\infty$ such that (2.18) holds.
\end{enumerate}
We are going to extend Theorem 2.3.3 to allow for unbounded stopping times. A first step in this direction is the next result, which deals with the special closed supermartingales of Remark 2.3.6(a):

**Proposition 2.3.7** Suppose that \( \{(X_n, F_n); n = 0, 1, 2, \ldots, +\infty\} \) is a closed supermartingale on the probability space \((\Omega, F, P)\) with \( X_n \geq 0 \) a.s. \( \forall n = 0, 1, 2, \ldots \), and \( X_\infty = 0 \) a.s. If \( S, T : \Omega \to \{0, 1, 2, \ldots, +\infty\} \) are \( \{F_n\} \)-stopping times with \( S(\omega) \leq T(\omega) \), \( \forall \omega \in \Omega \). then:

(a) \( E[X_T] < \infty \);

(b) \( E[X_T|F_S] \leq X_S \) a.s.

**Proof:** (a) Since the \( X_n \) are non-negative, clearly

\[
X_T \leq \liminf_{n \to \infty} X_{T \wedge n}, \quad \forall \omega \in \Omega.
\]

By Theorem 2.3.3(b) one has \( E[X_{T \wedge n}] \leq E[X_0], \forall n = 0, 1, 2, \ldots \), thus, from (2.19) and Fatou’s Theorem 1.2.15,

\[
E[X_T] \leq \liminf_{n \to \infty} E[X_{T \wedge n}] \leq E[X_0] < \infty.
\]

(b) Fix some \( A \in F_S \). For each (finite) integer \( n = 0, 1, 2 \ldots \) we easily see that

\[
E[X_T; A \cap \{T \leq n\}] = E[X_{T \wedge n}; A \cap \{T \leq n\}] \leq E[X_{T \wedge n}; A \cap \{S \leq n\}]
\]

\[
\leq E[X_{S \wedge n}; A \cap \{S \leq n\}] = E[X_S; A \cap \{S \leq n\}].
\]

The two equalities in (2.20) are obvious, the first inequality in (2.20) follows since \( \{T \leq n\} \subset \{S \leq n\} \) and \( X_{T \wedge n} \geq 0 \), while the second inequality is a consequence of Theorem 2.3.3(b) along with the facts that \( S \wedge n \) and \( T \wedge n \) are bounded stopping times with \( S \wedge n \leq T \wedge n \), and \( A \cap \{S \leq n\} \in F_{S \wedge n} \) (see Proposition 2.1.12(e)). Now \( I\{T \leq n\}(\omega) \) increases monotonically to \( I\{T < \infty\}(\omega) \) for each \( \omega \) as \( n \to \infty \), thus by the Monotone Convergence Theorem 1.2.12,

\[
\lim_{n \to \infty} E[X_T; A \cap \{T \leq n\}] = E[X_T; A \cap \{T < \infty\}].
\]

An obviously analogous expression, but where \( T \) is everywhere replaced with \( S \), also holds, and combining this with (2.21) and (2.20) we get

\[
E[X_T; A \cap \{T < \infty\}] \leq E[X_S; A \cap \{S < \infty\}].
\]

Since we suppose that \( X_\infty = 0 \) a.s. clearly \( E[X_T; A \cap \{T = \infty\}] = E[X_S; A \cap \{S = \infty\}] = 0 \), and adding this to (2.22) gives \( E[X_T; A] \leq E[X_S; A] \). In view of the arbitrary choice of \( A \in F_S \) we get part (b). \( \square \)

Next, we prove another “optional sampling” result, this time for the special martingale of Example 2.2.5:
Proposition 2.3.8 Suppose that \( \{ \mathcal{F}_n, n = 0, 1, 2, \ldots \} \) is some filtration in the probability space \((\Omega, \mathcal{F}, P)\), let \( \mathcal{F}_\infty \) be defined by (2.1), and suppose that \( X \) is a random variable on \((\Omega, \mathcal{F}, P)\) with \( E|X| < \infty \). Put
\[
X_n \triangleq E[X | \mathcal{F}_n], \quad \forall \ n = 0, 1, 2, \ldots, +\infty.
\]
If \( S, T : (\Omega, \mathcal{F}, P) \to \{0, 1, 2, \ldots, +\infty\} \) are \( \{ \mathcal{F}_n \} \)-stopping times with \( S(\omega) \leq T(\omega), \forall \ \omega \in \Omega \), then we have:
\[
\begin{align*}
(\text{a}) \ E|X_T| &< \infty; \\
(\text{b}) \ X_T &= E[X | \mathcal{F}_T] \ a.s.; \\
(\text{c}) \ X_S &= E[X_T | \mathcal{F}_S] \ a.s.
\end{align*}
\]
Proof : (a) Clearly the events \( \{T = n\}, \forall \ n = 0, 1, 2, \ldots, +\infty\), are a partition of \( \Omega \), namely \( \Omega = \bigcup_{0 \leq n \leq \infty} \{T = n\} \) and \( \{T = m\} \cap \{T = n\} = \emptyset \ \forall \ m, n \in \{0, 1, 2, \ldots, +\infty\} \) with \( m \neq n \), and thus
\[
X_T = \sum_{0 \leq n \leq \infty} X_n I \{T = n\}. \tag{2.24}
\]
From (2.23), along with Jensen’s inequality for conditional expectations (Theorem 1.4.20), we see that
\[
|X_n| \leq E[|X| | \mathcal{F}_n], \forall \ n = 0, 1, 2, \ldots, +\infty,
\]
thus, since \( \{T = n\} \in \mathcal{F}_n \), we have
\[
E[|X_n|; \{T = n\}] \leq E[|X|; \{T = n\}], \tag{2.25}
\]
for each \( n = 0, 1, 2, \ldots, +\infty \). In view of (2.24) and (2.25) we get
\[
E|X_T| = \sum_{0 \leq n \leq \infty} E[|X_n|; \{T = n\}] \leq \sum_{0 \leq n \leq \infty} E[|X|; \{T = n\}] = E|X| < \infty, \tag{2.26}
\]
where the two equalities in (2.26) involve interchanging the sum over \( 0 \leq n \leq \infty \) with expectation and are justified by the Monotone Convergence Theorem 1.2.12.
(b) Suppose initially that \( X(\omega) \geq 0, \forall \ \omega \in \Omega \). Then \( X_n \geq 0 \ a.s. \) for each \( n = 0, 1, 2, \ldots, +\infty \), thus \( X_T \geq 0 \ a.s. \). Now fix any \( A \in \mathcal{F}_T \). Since \( A \cap \{T = n\} \in \mathcal{F}_n, \forall \ n = 0, 1, 2, \ldots, +\infty \), we see from (2.23) that
\[
E[X_n; A \cap \{T = n\}] = E[X; A \cap \{T = n\}], \quad \forall \ n = 0, 1, 2, \ldots, +\infty.
\]
Thus, from (2.24) and the Monotone Convergence Theorem 1.2.12, we have
\[
E[X_T; A] = \sum_{0 \leq n \leq \infty} E[X_n; A \cap \{T = n\}] = \sum_{0 \leq n \leq \infty} E[X; A \cap \{T = n\}] = E[X; A]. \tag{2.27}
\]
Since (2.27) holds for each \( A \in \mathcal{F}_T \) and \( X_T \) is \( \mathcal{F}_T \)-measurable, we see that \( E[X | \mathcal{F}_T] = X_T \ a.s. \). Next, consider the general case where \( X \) takes positive and negative values, and let \( X_+ \triangleq \max\{X(\omega), 0\} \), \( X_- \triangleq \max\{-X(\omega), 0\}, \forall \ \omega \in \Omega \), denote the positive and negative parts of \( X \). Put \( X_n^+ \triangleq E[X_+ | \mathcal{F}_n] \),
\( X_n^2 \overset{\Delta}{=} E[X_\cdot|\mathcal{F}_n], \forall \ n = 0, 1, 2, \ldots, +\infty. \) Then \( X_n = X_n^1 - X_n^2 \) a.s., \( \forall \ n = 0, 1, 2, \ldots, +\infty, \) and thus \( X_T = X_T^1 - X_T^2 \) a.s. Now we have already shown that \( X_T^1 = E[X_+|\mathcal{F}_T] \) a.s. and \( X_T^2 = E[X_-|\mathcal{F}_T] \) a.s. Combining these, we obtain \( X_T = E[X|\mathcal{F}_T] \) a.s.

(c) Since \( \mathcal{F}_S \subset \mathcal{F}_T \) when \( S \leq T, \) this part of the proposition is an immediate consequence of (b) and the composition rule for conditional expectations.

Finally, we can put Proposition 2.3.7 and Proposition 2.3.8 together to get Doob’s optional sampling theorem for unbounded stopping times and closed supermartingales:

**Theorem 2.3.9** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\} \) is a closed supermartingale on probability space \((\Omega, \mathcal{F}, P),\) and \( S, T : \Omega \to \{0, 1, 2, \ldots, +\infty\} \) are \( \{\mathcal{F}_n\}\)-stopping times with \( S(\omega) \leq T(\omega), \forall \omega \in \Omega. \) Then:

(a) \( E[X_T|\mathcal{F}_T] < \infty; \)

(b) \( E[X_T|\mathcal{F}_S] \leq X_S \) a.s.

**Proof:** (a) Put \( Z_n \overset{\Delta}{=} E[X_\infty|\mathcal{F}_n] \) and \( Y_n \overset{\Delta}{=} X_n - Z_n, \forall \ n = 0, 1, 2, \ldots, +\infty. \) In view of (2.18) we easily see that \( \{(Y_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\} \) is a closed supermartingale with \( Y_n \geq 0 \) a.s. \( \forall \ n = 0, 1, 2, \ldots, \) and \( Y_\infty = 0 \) a.s. Thus Proposition 2.3.7(a) ensures that \( E[Y_T|\mathcal{F}_T] < \infty, \) while, from Proposition 2.3.8(a) and the definition of \( Z_n, \) we see that \( E[Z_T|\mathcal{F}_T] < \infty. \) Thus \( E[X_T|\mathcal{F}_T] < \infty. \)

(b) Fix some \( A \in \mathcal{F}_S. \) Then, from Proposition 2.3.7(b),

\[
E[Y_T|\mathcal{F}_S] \leq Y_S \ a.s.
\]

Also, from Proposition 2.3.8(c), one has

\[
E[Z_T|\mathcal{F}_S] = Z_S \ a.s.
\]

Addition of (2.28) and (2.29) then gives the assertion in (b).

**Remark 2.3.10** Theorem 2.3.9 has an obvious analogue when \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\} \) is a closed submartingale, namely we have \( E[X_T|\mathcal{F}_T] < \infty \) and \( E[X_T|\mathcal{F}_S] \geq X_S \) a.s.; likewise, if \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\} \) is a closed martingale, then we have \( E[X_T|\mathcal{F}_T] < \infty \) and \( E[X_T|\mathcal{F}_S] = X_S \) a.s. Throughout these notes we shall be establishing results such as the above in either the setting of supermartingales or submartingales, depending on convenience. It will always be apparent when a result established for supermartingales can be obviously modified to cover submartingales (or vice-versa) by means of a sign change.

**Remark 2.3.11** In applications one must often use Theorem 2.3.9 for unbounded stopping times \( S \) and \( T \) when a supermartingale \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) (rather than a closed supermartingale) is given. It is therefore important to determine conditions which imply that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is closable. As seen in Remark 2.3.6(b), a non-negative supermartingale is closable. In § 2.7 we shall look at alternative conditions which ensure closability of a supermartingale (see Theorem 2.7.1).
**Remark 2.3.12** We next use stopping times to formulate a useful criterion for an adapted sequence to be a martingale. We recall that a mapping \( T : \Omega \rightarrow \{0, 1, 2, \ldots \} \) is said to be **bounded** when there is some non-negative integer \( N < \infty \) such that \( T(\omega) \leq N \), \( \forall \omega \in \Omega \).

**Lemma 2.3.13** Suppose that \( \{(X_n, F_n); n = 0, 1, 2, \ldots \} \) is a \( \mathbb{R} \)-valued adapted process with \( E|X_T| < \infty \) and \( E[X_T] = 0 \) for each bounded \( \{F_n\}\)-stopping time \( T \). Then \( \{(X_n, F_n); n = 0, 1, 2, \ldots \} \) is a martingale.

Proof: Fix integers \( s, t \in \{0, 1, \ldots \} \) with \( s < t < \infty \), and let \( A \in \mathcal{F}_s \). Define the mapping \( T : \Omega \rightarrow \{0, 1, 2, \ldots, t\} \) as follows:

\[
T(\omega) \triangleq \begin{cases} 
  s & \text{for all } \omega \in A \\
  t & \text{for all } \omega \notin A.
\end{cases}
\]

Since \( A \in \mathcal{F}_s \), it is easily checked that \( T \) is a bounded \( \{\mathcal{F}_n\}\)-stopping time. Thus, by hypothesis, we have

\[
0 = E[X_T] = E[X_s; A] + E[X_t; A'].
\]

Since the constant \( t \) is (trivially) a bounded \( \{\mathcal{F}_n\}\)-stopping time, by the hypothesis of the lemma we also have

\[
0 = E[X_t] = E[X_t; A] + E[X_t; A'].
\]

Combining (2.30) and (2.31), we get

\[
E[X_s; A] = E[X_t; A]
\]

Since this holds for all \( A \in \mathcal{F}_s \) and \( X_s \) is \( \mathcal{F}_s \)-measurable it follows that \( X_s = E[X_t | \mathcal{F}_s] \) a.s.

Using Lemma 2.3.13 we can establish the following:

**Theorem 2.3.14** Suppose that \( \{(X_n, F_n); n = 0, 1, 2, \ldots \} \) is a martingale and \( T \) is a \( \{\mathcal{F}_n\}\)-stopping time on \( (\Omega, \mathcal{F}, P) \). Then \( \{(X_{n\wedge T}, F_n); n = 0, 1, 2, \ldots \} \) is a martingale.

Proof: Proposition 2.3.2 ensures that \( X_{n\wedge T} \) is \( \mathcal{F}_{n\wedge T} \)-measurable, hence \( \mathcal{F}_n \)-measurable, for each \( n = 0, 1, 2, \ldots \), thus \( \{(X_{n\wedge T}, F_n); n = 0, 1, 2, \ldots \} \) is an adapted sequence. Moreover, since \( \{|X_n|; F_n); n = 0, 1, 2, \ldots \} \) is a submartingale (by Proposition 2.2.10), we see from Theorem 2.3.3 that \( E|X_{n\wedge T}| < \infty \), \( \forall n = 0, 1, 2, \ldots \). It remains to verify the martingale property for \( \{(X_{n\wedge T}, F_n); n = 0, 1, 2, \ldots \} \). To this end, let \( S \) be some arbitrary but bounded \( \{\mathcal{F}_n\}\)-stopping time. Then \( S \wedge T \) is also a bounded \( \{\mathcal{F}_n\}\)-stopping time, thus Theorem 2.3.3 ensures that \( E|X_{S\wedge T}| < \infty \) and \( E[X_{S\wedge T}] = E[X_0] \). Now put \( Y_n \triangleq X_{n\wedge T} - X_0 \), \( \forall n = 0, 1, 2, \ldots \). In view of the preceding observations it follows that \( \{Y_n, F_n); n = 0, 1, 2, \ldots \) is an adapted process with \( E|Y_S| \leq E|X_{S\wedge T}| + E[X_0] < \infty \) and \( E[Y_S] = E[X_{S\wedge T}] - E[X_0] = 0 \). The arbitrary choice of the bounded stopping time \( S \), along with Lemma 2.3.13, ensures that \( \{(Y_n, F_n); n = 0, 1, 2, \ldots \} \) is a martingale, as required to establish the result.
2.4 The Basic Martingale Inequalities

Theorem 2.4.1 Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a submartingale on \( (\Omega, \mathcal{F}, P) \). Then, for each \( x \in (0, \infty) \) and \( n = 0, 1, 2, \ldots \) we have

\[
x \cdot P \left[ \max_{0 \leq k \leq n} X_k \geq x \right] \leq E \left[ X_n; \max_{0 \leq k \leq n} X_k \geq x \right] \leq E[X_n^+]
\]

and

\[
x \cdot P \left[ \min_{0 \leq k \leq n} X_k \leq -x \right] \leq E \left[ X_n; \min_{0 \leq k \leq n} X_k \geq x \right] - E[X_0] \leq E[X_n^+] - E[X_0].
\]

Proof: We first establish (2.32). Fix some arbitrary \( x \in (0, \infty) \) and \( n = 1, 2, \ldots \) For each \( \omega \) define

\[
S(\omega) \triangleq \min \{k \geq 0 : X_k(\omega) \geq x\} \wedge n.
\]

It follows from Example 2.1.8 and Proposition 2.1.12 (a) that \( S \) is a \( \{\mathcal{F}_n\}\)-stopping time, and clearly \( S \leq n \). From Theorem 2.3.3(b) one sees that

\[
E[X_n] \geq E[X_S]; \max_{0 \leq k \leq n} X_k \geq x
\]

and

\[
E[X_S]; \max_{0 \leq k \leq n} X_k \geq x \geq xP \left[ \max_{0 \leq k \leq n} X_k \geq x \right].
\]

Now, by the definition of \( S \), we see that

\[
\left\{ \max_{0 \leq k \leq n} X_k \geq x \right\} \subset \{X_S \geq x\},
\]

hence

\[
E \left[ X_S; \max_{0 \leq k \leq n} X_k \geq x \right] \geq xP \left[ \max_{0 \leq k \leq n} X_k \geq x \right].
\]

Again, by the definition of \( S \), we have

\[
\left\{ \max_{0 \leq k \leq n} X_k < x \right\} \subset \{S = n\},
\]

and so

\[
E \left[ X_S; \max_{0 \leq k \leq n} X_k < x \right] = E \left[ X_n; \max_{0 \leq k \leq n} X_k < x \right].
\]

Combining (2.34), (2.35), and (2.36) we get the first inequality of (2.32), while the second inequality of (2.32) is just a consequence of the obvious inequalities

\[
E \left[ X_n; \max_{0 \leq k \leq n} X_k \geq x \right] \leq E \left[ X_n^+; \max_{0 \leq k \leq n} X_k \geq x \right] \leq E[X_n^+].
\]

As for (2.33), for each \( \omega \) define

\[
T(\omega) \triangleq \min \{k \geq 0 : X_k(\omega) \leq -x\} \wedge n.
\]
Again, we see that $T$ is an $\{F_n\}$-stopping time, hence, from Theorem 2.3.3(b), we have $E[X_0] \leq E[X_T]$, or

$$E[X_0] \leq E\left[ X_T; \min_{0 \leq k \leq n} X_k \leq -x \right] + E\left[ X_T; \min_{0 \leq k \leq n} X_k > -x \right].$$

$$\leq -x \cdot P\left[ \min_{0 \leq k \leq n} X_k \leq -x \right] + E\left[ X_T; \min_{0 \leq k \leq n} X_k > -x \right].$$

Again, the first inequality in (2.33) follows upon re-arranging (2.37), while the second inequality is just an obvious upper bound. \[\blacksquare\]

**Lemma 2.4.2** Suppose that $X$ and $Y$ are non-negative random variables on the probability space $(\Omega, \mathcal{F}, P)$ such that

$$x \cdot P(X \geq x) \leq E[Y; X \geq x]$$

for all $x \in [0, \infty)$. Then, for each $p \in (1, \infty)$, we have

$$\|X\|_p \leq \left( \frac{p}{p-1} \right) \|Y\|_p.$$ 

In particular, if $\|Y\|_p < \infty$ then $\|X\|_p < \infty$.

**Proof:** Fix an arbitrary $p \in (1, \infty)$. One easily checks that $\{(x, \omega) \in [0, \infty) \otimes \Omega : X(\omega) \geq x\} \in \mathcal{B}([0, \infty)) \otimes \mathcal{F}$, hence the mapping $(x, \omega) \rightarrow I\{X \geq x\}$ on $[0, \infty) \otimes \Omega$ is $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$-measurable. We then see that

$$\int_0^\infty px^{p-1}P(X \geq x)dx = \int_0^\infty px^{p-1}E[I\{X \geq x\}]dx$$

$$= E[\int_0^\infty px^{p-1}I\{X \geq x\}dx] = E[\int_0^X px^{p-1}dx] = E[X^p]$$

where we have used the Fubini-Tonelli Theorem 1.2.40 to interchange the order of integration at the second equality of (2.40). In the same way, writing $q \triangleq p/(p-1)$, we see that

$$\int_0^\infty px^{q-2}E[Y; I\{X \geq x\}]dx = E[Y \int_0^\infty px^{q-2}I\{X \geq x\}dx]$$

$$= E[Y \int_0^X px^{q-2}dx] = E[qY X^{p-1}] \leq q\|Y\|_p \cdot \|X^{p-1}\|_{q}.$$ 

Here, we have used the Fubini-Tonelli Theorem 1.2.40 at the first equality and Hölder’s inequality (Theorem 1.2.25) to get the final inequality in (2.41). Now, by (2.40), (2.38), and (2.41), it follows that

$$E[X^p] = \int_0^\infty px^{p-1}P(X \geq x)dx$$

$$\leq \int_0^\infty px^{q-2}E[Y; X \geq x]dx \leq q\|Y\|_p \cdot \|X^{p-1}\|_{q}.$$
Since \( p = q(p - 1) \), we see that \( \|X^{p-1}\|_q = E^{1/q}[X^p] \) hence, from (2.42):

\[
E[X^p] \leq q\|Y\|_p \cdot E^{1/q}[X^p].
\]

Thus (2.39) follows at once from (2.43) when \( E[X^p] < \infty \). Now suppose that \( E[X^p] = +\infty \), and fix some arbitrary integer \( n \geq 1 \). For each \( x \geq 0 \) we have two possibilities namely (i) \( n \geq x \) in which case \( \{X \geq x\} = \{X \land n \geq x\} \) or (ii) \( n < x \) in which case \( \{X \land n \geq x\} = \emptyset \). In both of these two cases we see from (2.38) that

\[
x \cdot P(X \land n \geq x) \leq E[Y; X \land n \geq x]
\]

for all \( x \geq 0 \). Since (2.39) has been shown to follow from (2.38) whenever \( X \) is \( L_p \)-bounded, and since \( X \land n \) is obviously \( L_p \) bounded, we see from (2.44) that

\[
\|X \land n\|_p \leq q\|Y\|_p
\]

for all integers \( n \geq 1 \). Taking \( n \to \infty \) and using the Monotone Convergence Theorem 1.2.12 in (2.45) gives (2.39).

\[\blacksquare\]

**Remark 2.4.3** The argument which established (2.40) of course continues to hold when \( p = 1 \), and hence, for any non-negative random variable \( X \), we see that

\[
E[X^p] = \int_0^\infty px^{p-1}P(X \geq x) \, dx
\]

for each \( p \in [1, \infty) \). This result is called Robin’s identity and is extremely useful for doing concrete calculations in probability.

Theorem 2.4.1 and Lemma 2.4.2 clearly give the following result known as Doob’s \( L^p \)-inequality:

**Theorem 2.4.4** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a non-negative submartingale. Then, for each \( p \in (1, \infty) \), we have that

\[
\|\max_{0 \leq k \leq n} X_k\|_p \leq \left( \frac{p}{p-1} \right) \cdot \|X_n\|_p,
\]

for all \( n = 0, 1, 2, \ldots \).
2.5 The Martingale Convergence Theorem

Suppose that \( \{x_k, k = 0, 1, 2, \ldots\} \) is a sequence of real numbers, and \( a \) and \( b \) are real numbers such that \( a < b \). Define the non-negative integers

\[
\begin{align*}
\tau_0 &\triangleq 0 \\
\sigma_1 &\triangleq \min\{k \geq \tau_0; x_k \leq a\} \\
\tau_1 &\triangleq \min\{k > \sigma_1; x_k \geq b\} \\
\sigma_2 &\triangleq \min\{k > \tau_1; x_k \leq a\} \\
\tau_2 &\triangleq \min\{k > \sigma_2; x_k \geq b\} \\
&\vdots \\
\sigma_n &\triangleq \min\{k > \tau_{n-1}; x_k \leq a\} \\
\tau_n &\triangleq \min\{k > \sigma_n; x_k \geq b\} \\
&\vdots
\end{align*}
\]

(2.47)

In (2.47) we use the convention that \( \min\{\emptyset\} \triangleq +\infty \). For each positive integer \( N \) define

\[ U_N^x(a, b) \triangleq \max\{n \geq 0; \tau_n \leq N\}, \quad (2.48) \]

the so-called number of upcrossings of the interval \([a, b]\) by the sequence \( \{x_0, x_1, x_2, \ldots x_N\} \). The significance of \( U_N^x(a, b) \) is that it counts the total number of times that the sequence \( \{x_0, x_1, \ldots x_N\} \) passes completely from being less than or equal to \( a \) to greater than or equal to \( b \). In the next two remarks we collect some simple observations concerning upcrossings and the numbers \( \sigma_n \) and \( \tau_n \) in (2.47). These observations will be useful for later proofs:

**Remark 2.5.1** Clearly \( U_N^x(a, b) \) is non-decreasing as \( N \) increases, hence the limit

\[ U_{\infty}^x(a, b) \triangleq \lim_{N \to \infty} U_N^x(a, b), \quad (2.49) \]

always exists in \([0, \infty]\). Moreover, from (2.47) we see that

\[ 0 = \tau_0 \leq \sigma_1 < \tau_1 < \sigma_2 < \tau_2 < \sigma_3 < \ldots \]

Since \( \sigma_n \) and \( \tau_n \) are integer-valued it follows that \( \tau_n \geq 2n - 1 \) for all \( n = 1, 2, 3, \ldots \) (equality holds in the extreme case where the maximum possible number of upcrossings occurs, namely \( \tau_0 = \sigma_1 = 0 \) and \( \tau_n = 1 + \sigma_n, \sigma_{n+1} = 1 + \tau_n \) for all \( n = 1, 2, \ldots \)). Now fix positive integers \( n \) and \( N \) such that \( 2n > N \). From (2.48) we see that

\[ \tau_n \geq 2n - 1 \geq N \geq U_N^x(a, b) \]

(2.50)

and hence, since \( \tau_n \) increases with increasing \( n \), it follows from (2.50) that \( n \geq U_N^x(a, b) \) whenever \( 2n > N \), and hence

\[ U_N^x(a, b) \leq \lfloor 1 + N/2 \rfloor, \quad \forall N = 1, 2, 3, \ldots \]

(2.51)
For each \( N = 1, 2, 3, \ldots \) put \( n(N) \triangleq [1 + N/2] \), and observe that \( 2n(N) > N \). Thus, from (2.50), we see that \( \tau_{n(N)} \geq N \), hence \( x_N = x_{N \wedge \tau_{n(N)}} \), \( \forall \ N = 1, 2, 3, \ldots \), and therefore

\[
(2.52) \quad x_N - x_0 = \sum_{k=1}^{n(N)} [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}] + \sum_{k=U+1}^{n(N)} [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}], \quad \forall \ N = 1, 2, 3, \ldots
\]

Considering the first term on the right hand side of (2.52) and using (2.51) we get (writing \( U \) for \( U^x_N(a, b) \)):

\[
(2.53) \quad \sum_{k=1}^{n(N)} [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}] = \sum_{k=1}^{U} [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}] + \sum_{k=U+1}^{n(N)} [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}] 
\geq (b - a) U^x_N(a, b) + [x_N - x_{N \wedge \sigma_{(1+U)}}], \quad \forall \ N = 1, 2, 3, \ldots
\]

Here the last inequality follows since (2.48) ensures that \( \tau_k > \sigma_k > \sigma_{1+U} \), \( \forall \ k \geq U + 2 \), thus \( x_{N \wedge \tau_{(1+U)}} = x_N \) and \( [x_{N \wedge \tau_k} - x_{N \wedge \sigma_k}] = [x_N - x_N] = 0 \) for each \( k \geq U + 2 \). The decomposition of \( x_N - x_0 \) represented by (2.52) and (2.53) will be needed later when we prove the upcrossing inequality for submartingales (see Theorem 2.5.4).

**Remark 2.5.2** Upcrossings can be used to characterize when a given sequence of real numbers fails to converge to a limit. Indeed, suppose that the sequence of real numbers \( \{x_n, \ n = 0, 1, 2, \ldots \} \) fails to converge to a limit in \([-\infty, \infty]\) as \( n \to \infty \). Then it follows that

\[
-\infty \leq \liminf_{n \to \infty} x_n < \limsup_{n \to \infty} x_n \leq +\infty
\]

and hence there are rational numbers \( a \) and \( b \) such that

\[
(2.54) \quad \liminf_{n \to \infty} x_n < a < b < \limsup_{n \to \infty} x_n
\]

Thus, from the definitions of limit superior and limit inferior, there must exist infinite increasing sequences of integers \( 0 \leq n_1 < n_2 < n_3 \ldots \) and \( 0 \leq m_1 < m_2 < m_3 \ldots \) such that

\[
x_{m_r} \leq a < b \leq x_{n_r}
\]

for all \( r = 1, 2, \ldots \). From the definition of upcrossings it then follows that

\[
(2.55) \quad U^x_\infty(a, b) = +\infty.
\]

This fact will be used when we establish the martingale convergence theorem (see Theorem 2.5.7).

**Remark 2.5.3** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is an \( \mathcal{F} \)-valued adapted sequence on the probability space \((\Omega, \mathcal{F}, P)\). Then we have a sequence of real numbers \( \{X_n(\omega), \ n = 0, 1, 2, \ldots \} \) for each \( \omega \).
Now fix real numbers $a$ and $b$ such that $a < b$. By analogy with the integers $\sigma_n$ and $\tau_n$ defined in (2.47) we define non-negative integers $\sigma_n(\omega)$ and $\tau_n(\omega)$ for each $\omega$ as follows:

\[
\begin{align*}
\tau_0(\omega) &\triangleq 0 \\
\sigma_1(\omega) &\triangleq \min\{k \geq \tau_0(\omega); X_k(\omega) \leq a\} \\
\tau_1(\omega) &\triangleq \min\{k > \sigma_1(\omega); X_k(\omega) \geq b\} \\
\sigma_2(\omega) &\triangleq \min\{k > \tau_1(\omega); X_k(\omega) \leq a\} \\
\tau_2(\omega) &\triangleq \min\{k > \sigma_2(\omega); X_k(\omega) \geq b\} \\
\vdots \\
\sigma_n(\omega) &\triangleq \min\{k > \tau_{n-1}(\omega); X_k(\omega) \leq a\} \\
\tau_n(\omega) &\triangleq \min\{k > \sigma_n(\omega); X_k(\omega) \geq b\} \\
\vdots 
\end{align*}
\]

(2.56)

and let $U_N^X(a, b; \omega)$ denote the number of upcrossings of the interval $[a, b]$ by the sequence of real numbers $\{X_0(\omega), X_1(\omega), \ldots X_N(\omega)\}$, namely

\[
U_N^X(a, b; \omega) \triangleq \max\{n; \tau_n(\omega) \leq N\}
\]

(2.57)

It is clear from Example 2.1.8 that each $\tau_n$ and $\sigma_n$ in (2.56) are $\{\mathcal{F}_n\}$-stopping times, and it follows from (2.57) that

\[
\{U_N^X(a, b) = n\} = \{\tau_n \leq N\} \cap \{\tau_{n+1} > N\}
\]

for all $n = 0, 1, 2, \ldots$. Since $\tau_n$ and $\tau_{n+1}$ are stopping times we then get $\{U_N^X(a, b) = n\} \in \mathcal{F}_N$ for all $n = 0, 1, 2, \ldots$. In particular, $U_N^X(a, b)$ is a $\mathcal{F}_N$-measurable random variable. We next show that there is a simple upper bound for the expectation of $U_N^X(a, b)$ when $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a submartingale.

**Theorem 2.5.4** Suppose that $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a submartingale on $(\Omega, \mathcal{F}, P)$. Then, for each $N = 1, 2, 3, \ldots$, and real numbers $a$ and $b$ with $a < b$, we have

\[
(b - a) \cdot E[U_N^X(a, b)] \leq E[(X_N - a)^+]
\]

(2.58)

Proof: Fix a positive integer $N$ and real numbers $a$ and $b$ with $a < b$. Put $Y_n \triangleq (X_n - a)^+$ for all $n = 0, 1, 2, \ldots$. From Proposition 2.2.10 we see that $\{(Y_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a non-negative submartingale. Define stopping times $\sigma_n$ and $\tau_n$ as in (2.56) and let $U_N^X(0, b - a; \omega)$ denote the number of upcrossings of the interval $[0, b - a]$ by the sequence of real numbers $\{Y_0(\omega), Y_1(\omega), \ldots Y_N(\omega)\}$. Since $Y_{\sigma_k}(\omega) = 0$ when $\sigma_k(\omega) < \infty$, it is clear that

\[
U_N^X(a, b; \omega) = U_N^X(0, b - a; \omega), \quad \forall \omega \in \Omega.
\]

(2.59)
Then, applying (2.52) and (2.53) to the sequence of numbers \( \{Y_0(\omega), Y_1(\omega), \ldots, Y_N(\omega)\} \), we get (writing \( U \) for \( U_N^X(0, b - a) \), \( n(N) \triangleq [1 + N/2] \)):

\[
Y_N - Y_0 \geq (b - a) \cdot U + [Y_N - Y_{N \wedge \sigma_{1+U}}] + \sum_{k=1}^{n(N)} [Y_{N \wedge \sigma_k} - Y_{N \wedge \tau_{k-1}}]. 
\]

(2.60)

Now, when \( \sigma_{1+U} < N \), then \( Y_{N \wedge \sigma_{1+U}} = Y_{\sigma_{1+U}} = 0 \), so that the second term on the right side of (2.60) is non-negative, and when \( \sigma_{1+U} \geq N \) then the second term on the right of (2.60) is clearly zero. Moreover, since \( N \wedge \sigma_k \) and \( N \wedge \tau_{k-1} \) are stopping times such that \( N \wedge \sigma_k \geq N \wedge \tau_{k-1} \) and \( \{(Y_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a submartingale, we see from Theorem 2.3.3(b) that \( E[Y_{N \wedge \sigma_k}] \geq E[Y_{N \wedge \tau_{k-1}}] \).

Thus, from (2.60) and (2.59), it follows that

\[
E[Y_N] - E[Y_0] \geq (b - a) \cdot E[U_N^X(a, b)]
\]

(2.61)

Now (2.58) follows from (2.61), along with the definition of \( Y_n \) and the fact that \( E[Y_0] \geq 0 \). \( \square \)

Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is an \( \mathcal{B} \)-valued adapted sequence on the probability space \((\Omega, \mathcal{F}, P)\). We have already observed that \( U_N^X(a, b; \omega) \) is \( \mathcal{F}_N \)-measurable when viewed as a function of \( \omega \), and that the limit

\[
U_\infty^X(a, b; \omega) = \lim_{N \to \infty} U_N^X(a, b; \omega)
\]

(2.62)

exists in \([0, \infty]\) for each \( \omega \). It follows that \( U_\infty^X(a, b; \omega) \) is \( \mathcal{F}_\infty \)-measurable with respect to \( \omega \) and hence the probability of events such as \( \{U_\infty^X(a, b) = +\infty\} \) is defined. We next show that this event actually has a probability equal to zero when \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is an appropriately bounded submartingale. To this end we need the following:

**Remark 2.5.5** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a submartingale on \((\Omega, \mathcal{F}, P)\) such that

\[
\sup_n E[X_n^+] < \infty.
\]

(2.63)

For an arbitrary \( n = 0, 1, 2, \ldots \) one has \( X_n = X_n^+ - X_n^- \) and \( |X_n| = X_n^+ + X_n^- \), thus \( |X_n| = 2X_n^+ - X_n \).

By the submartingale property we have \( E[X_n] \geq E[X_0] \), whence \( \sup_n E[X_n^+] \leq 2 \sup_n E[X_n^+] - E[X_0] \).

It follows that (2.63) implies the seemingly stronger condition \( \sup_n E[X_n^+] < \infty \).

**Lemma 2.5.6** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a submartingale on \((\Omega, \mathcal{F}, P)\) such that (2.63) holds. Then, for any pair of real numbers \( a \) and \( b \) such that \( a < b \), we have

\[
P[U_\infty^X(a, b) = +\infty] = 0
\]

(2.64)

Proof : Fix real numbers \( a \) and \( b \) such that \( a < b \). By Theorem 2.5.4 we see that

\[
(b - a) \cdot E[U_N^X(a, b)] \leq |a| + E[X_N]
\]

(2.65)
for all \( N = 1, 2, \ldots \). Now (2.63), together with Remark 2.5.5, ensures that

\[
(2.66) \quad \sup_N E|X_N| < \infty.
\]

Since the convergence in (2.62) is monotonic for each \( \omega \), we see from (2.65), (2.66) and the Monotone Convergence Theorem 1.2.12 that

\[
(b - a) \cdot E[U_\infty^X(a, b)] < \infty
\]

which in turn implies (2.64) since \( U_\infty^X(a, b) \) is non-negative.

With the preceding results established we are now able to prove the main result of this section, namely the following martingale convergence theorem:

**Theorem 2.5.7** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a submartingale on \((\Omega, \mathcal{F}, P)\) such that

\[
(2.67) \quad \sup_n E[X_n^+] < \infty.
\]

Then there exists some a.s. unique random variable \( X_\infty : \Omega \to \mathbb{R} \), measurable with respect to \( \mathcal{F}_\infty \), such that \( E|X_\infty| < \infty \) and \( \lim_{n \to \infty} X_n = X_\infty \) a.s.

**Proof:** Let \( \Lambda \) be the set of all \( \omega \) such that the sequence of real numbers \( \{X_n(\omega) ; n = 0, 1, 2, \ldots\} \) fails to converge to a limit in \([-\infty, +\infty]\). Thus, if the functions \( \overline{X}_\infty \) and \( \underline{X}_\infty \) are defined by

\[
(2.68) \quad \overline{X}_\infty(\omega) \overset{\triangle}{=} \limsup_{n \to \infty} X_n(\omega) \quad \text{and} \quad \underline{X}_\infty(\omega) \overset{\triangle}{=} \liminf_{n \to \infty} X_n(\omega), \quad \forall \omega \in \Omega,
\]

then

\[
\Lambda \overset{\triangle}{=} \{\omega ; -\infty \leq \underline{X}_\infty(\omega) < \overline{X}_\infty(\omega) \leq +\infty\}.
\]

Since \( \overline{X}_\infty \) and \( \underline{X}_\infty \) are clearly \( \mathcal{F}_\infty \)-measurable we see that \( \Lambda \in \mathcal{F}_\infty \subset \mathcal{F} \). We are going to show that \( P(\Lambda) = 0 \). In view of Remark 2.5.2 we see that, corresponding to each \( \omega \in \Lambda \), there are rational numbers \( a(\omega) \) and \( b(\omega) \) such that

\[
U_\infty^X(a(\omega), b(\omega); \omega) = +\infty
\]

It follows that

\[
(2.69) \quad \Lambda \subset \bigcup_{a,b} \{\omega; U_\infty^X(a, b; \omega) = +\infty\}
\]

where the union is over all pairs of rational numbers \( a \) and \( b \) such that \( a < b \). Since there are only countably many such pairs it follows that \( \Lambda \) is contained within a countable union of sets each of which, by (2.67) and Lemma 2.5.6, is of probability zero. Thus \( P(\Lambda) = 0 \) and so, in view of (2.68), we obtain

\[
\underline{X}_\infty = \lim_{n \to \infty} X_n \quad \text{a.s.}
\]
To see that $X_\infty$ is integrable, observe from (2.67) and Remark 2.5.5 that $\sup_n E|X_n| < \infty$, and hence, by Fatou’s Theorem 1.2.15,

$$\int_\Omega |X_\infty| dP \leq \lim inf_{n \to \infty} \int_\Omega |X_n| dP \leq \sup_n E|X_n| < \infty.$$

Finally, since we want the limit to be $\mathbb{R}$-valued (rather than $\overline{\mathbb{R}}$-valued), define $X_\infty$ such that $X_\infty(\omega) \triangleq X_\infty^-(\omega)$ whenever $|X_\infty(\omega)| < +\infty$, and $X_\infty(\omega) \triangleq 0$ whenever $|X_\infty(\omega)| = +\infty$. Since $P(|X_\infty| = +\infty) = 0$ we have established the result.

**Remark 2.5.8** (a) In Remark 2.2.2 we noted that a submartingale can be viewed as the probabilistic analogue of a non-decreasing sequence of real numbers. Theorem 2.5.7 is then an analogue of the elementary result that a non-decreasing sequence of real numbers with finite upper-bound converges to a finite limit; in Theorem 2.5.7 the condition (2.67) plays the role of a “finite upper-bound” on the submartingale.

(b) Recalling the relation between submartingales and supermartingales in Definition 2.2.1 we see that if $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a supermartingale with $\sup_n E[X_n^-] < \infty$ then the conclusions of Theorem 2.5.7 continue to hold.

(c) If $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a submartingale with $X_n \leq 0$ a.s. $\forall n = 0, 1, 2, \ldots$, then clearly (2.67) holds, and so there exists a random variable $X_\infty$ with the properties asserted in Theorem 2.5.7. Likewise, if $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a supermartingale such that $X_n \geq 0$ a.s. $\forall n = 0, 1, 2, \ldots$ then, by a sign-change, one sees that there is a random variable $X_\infty$ again having the properties asserted in Theorem 2.5.7.

**Example 2.5.9** It is natural to ask if one gets $L^1$ as well as a.s. convergence in Theorem 2.5.7. Here we shall give an example showing that $L^1$-convergence does not necessarily hold. Suppose that $\{Y_n; n = 0, 1, 2, \ldots\}$ is a sequence of independent and identically distributed random variables on $(\Omega, \mathcal{F}, P)$, with $Y_n \geq 0$ and $E[Y_n] = 1$, $\forall n = 0, 1, 2, \ldots$ and put $a \triangleq E[Y_n^{1/2}]$. Clearly $a > 0$ and $a^2 = E[Y_n^{1/2}]^2 \leq E[Y_n] = 1$ where we have used Jensen’s inequality (Theorem 1.4.2). We shall henceforth suppose that $0 < a < 1$. Next, define

$$X_n \triangleq \prod_{0 \leq i \leq n} Y_i, \quad \mathcal{F}_n \triangleq \sigma\{Y_0, Y_1, \ldots, Y_n\}, \quad \forall n = 0, 1, 2, \ldots$$

Clearly $X_n \geq 0$, $\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is an $\mathbb{R}$-valued adapted sequence, and

$$E|X_n| = \prod_{0 \leq i \leq n} E[Y_i] = 1, \quad \forall n = 0, 1, 2, \ldots$$

Moreover, for each $n = 0, 1, 2, \ldots$ we have

$$E[X_{n+1} \mid \mathcal{F}_n] = E[X_n Y_{n+1} \mid \mathcal{F}_n] = X_n E[Y_{n+1} \mid \mathcal{F}_n] = X_n \text{ a.s.}$$
The second equality in (2.73) follows from Theorem 1.4.15(c), since \( X_n \) is \( \mathcal{F}_n \)-measurable, while the third equality is a consequence of Theorem 1.4.15(f), since \( \sigma \{ Y_{n+1} \} \) and \( \mathcal{F}_n \) are independent, and hence \( E [ Y_{n+1} \mid \mathcal{F}_n] = E[Y_{n+1}]=1 \) a.s. Consequently, \( \{ (X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a martingale with \( X_n \geq 0 \), therefore it is obviously a non-negative supermartingale. In the light of Remark 2.5.8 the sequence \( \{ X_n; n = 0, 1, 2 \ldots \} \) converges a.s. to some integrable random variable \( X_\infty \). To see that this convergence fails to hold in the \( L^1 \)-sense write

\[
Z_n \triangleq \frac{X_n^{1/2}}{a^{n+1}}, \quad \forall n = 0, 1, 2, \ldots
\]

By the same reasoning used at (2.73) we see that \( \{(Z_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a non-negative martingale (hence a supermartingale), and therefore Remark 2.5.8 ensures that \( \{Z_n; n = 0, 1, 2 \ldots \} \) converges a.s. to some limiting integrable random variable \( Z_\infty \). Now \( |Z_\infty(\omega)| < \infty \), and since \( 0 < a < 1 \) we have \( \lim_{n \to \infty} a^n = 0 \), thus it is clear from (2.74) that \( \lim_{n \to \infty} X_n = 0 \) a.s. whence \( X_\infty = 0 \) a.s. But \( E|X_n| = 1, \forall n = 0, 1, 2, \ldots \) thus \( \{ X_n; n = 0, 1, 2 \ldots \} \) cannot converge in \( L^1 \) to \( X_\infty \).

**Remark 2.5.10** By the Dominated Convergence Theorem 1.2.21 we see that \( L^1 \)-convergence follows from a.s. convergence in Theorem 2.5.7 when there is some fixed integrable random variable which uniformly dominates the \( X_n \) in magnitude. The Dominated Convergence Theorem is not always the most appropriate tool for determining when \( L^1 \)-convergence takes place, because the existence of a dominating function is sometimes not clearly apparent. In Section 2.6 we shall introduce the notion of uniform integrability which, at least in the context of integration over probability spaces, provides more general conditions than dominated convergence for determining when a.s. convergence yields \( L^1 \)-convergence. As will be seen in §2.6 this notion is ideally suited for use in conjunction with Theorem 2.5.7.

So far we have confined attention to submartingales which are defined on the set of non-negative integers, and indeed the great majority of discrete-parameter submartingales that arise in applications are of this form. However, in Chapter 4, we shall also need the limiting properties as \( n \to -\infty \) of a submartingale \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) defined over the non-positive integers i.e. take \( I \triangleq \{ \ldots, -2, -1, 0 \} \) in Definition 2.2.1. Submartingales of this kind will be called **negatively indexed** for emphasis. The limiting properties of negatively indexed submartingales are easily established by making some simple changes in the argument which led to Theorem 2.5.7. Indeed, suppose that \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) is a negatively indexed submartingale. Fix an arbitrary negative integer \( N \) and arbitrary real numbers \( a \) and \( b \) such that \( a < b \). For an arbitrary \( \omega \in \Omega \) put \( x_0 \triangleq X_N(\omega), x_1 \triangleq X_{N+1}(\omega), \ldots x_{-N} \triangleq X_0(\omega) \). We use \( V_N^X(a, b; \omega) \) to denote the number of upcrossings of the interval \([a, b]\) by the real numbers \( \{x_0, x_1, \ldots x_{-N}\} \). Clearly \( V_N^X(a, b; \omega) \) is non-decreasing as \( N \) decreases to \(-\infty\), hence the limit

\[
V_{-\infty}^X(a, b; \omega) \triangleq \lim_{N \to -\infty} V_N^X(a, b; \omega)
\]
exists in \([0, \infty]\) for each \(\omega\). In fact, if we define \(Y_k \triangleq X_{N+k}\) and \(G_k \triangleq \mathcal{F}_{N+k}\) for all \(k = 0, 1, 2, \ldots - N\), and \(Y_k \triangleq X_0, G_k \triangleq \mathcal{F}_0\) for all \(k = 1 – N, 2 – N, \ldots\), then \(\{(Y_k, G_k), k = 0, 1, 2, \ldots\}\) is a submartingale and \(V_N^X(a, b; \omega)\) is equal to \(U_N^Y(a, b; \omega)\), the number of upcrossings of \([a, b]\) by \(\{Y_0(\omega), Y_1(\omega) \ldots Y_N(\omega)\}\). Now \(U_N^Y(a, b; \omega)\) is \(\mathcal{F}_N\)-measurable with respect to \(\omega\) hence so also is \(V_N^X(a, b; \omega)\), and from Theorem 2.5.4 we see that

\[
E[V_{-\infty}^X(a, b)] < \infty
\]

which further implies

\[
P[V_{-\infty}^X(a, b) = +\infty] = 0
\]

Observe that, since each \(V_N^X(a, b; \omega)\) is \(\mathcal{F}_0\)-measurable with respect to \(\omega\), it follows from (2.75) that \(V_{-\infty}^X(a, b; \omega)\) is also \(\mathcal{F}_0\)-measurable.

**Remark 2.5.11** Notice that we need an upper bound of the form (2.63) in order to prove (2.64) for a submartingale defined over the non-negative integers, whereas the seemingly similar result (2.77) has been established for the case of negatively indexed submartingales without having to postulate any similar upper bound.

Now suppose that \(\Lambda\) is the set of all \(\omega\) such that the sequence \(\{X_0(\omega), X_{-1}(\omega), X_{-2}(\omega), \ldots\}\) fails to converge to a limit in \([-\infty, +\infty]\), namely

\[
\Lambda \triangleq \{\omega ; \liminf_{n \to -\infty} X_n(\omega) < \limsup_{n \to -\infty} X_n(\omega)\}.
\]

Then, if \(\omega \in \Lambda\), there must be rational numbers \(a(\omega)\) and \(b(\omega)\), along with infinite decreasing sequences of integers \(0 \geq n_1 > n_2 > \ldots\) and \(0 \geq m_1 > m_2 > \ldots\) such that \(a(\omega) < b(\omega)\) and

\[
X_{m_r}(\omega) \leq a(\omega) < b(\omega) \leq X_{n_r}(\omega)
\]

for all \(r = 0, 1, 2, \ldots\) (for typographical convenience we suppress the obvious dependence of \(\{m_r\}\) and \(\{n_r\}\) upon \(\omega\)). From (2.78) we see that

\[
V_{-\infty}^X(a(\omega), b(\omega); \omega) = +\infty
\]

for all \(\omega \in \Lambda\). Now, using (2.77) and an argument identical to that in the proof of Theorem 2.5.7, we conclude that \(P(\Lambda) = 0\) and hence

\[
\lim_{n \to -\infty} X_n = X_{-\infty} \quad a.s.
\]
where
(2.80) \[ X_{-\infty}(\omega) \triangleq \limsup_{n \to -\infty} X_n(\omega), \quad \forall \omega \in \Omega. \]

Next, observe from the definition of limit superior that for any negative integer \( N \) we have
\[ X_{-\infty}(\omega) = \limsup_{n \to -\infty} X_{n+N}(\omega) \]
for each \( \omega \). Since the random variables \( X_{n+N} \) are \( \mathcal{F}_N \)-measurable for all \( n = 0, -1, -2, \ldots \) we see that \( X_{-\infty} \) is \( \mathcal{F}_N \)-measurable. But the choice of \( N \) is arbitrary and negative, hence \( X_{-\infty} \) in (2.80) is actually measurable with respect to
(2.81) \[ \mathcal{F}_{-\infty} \triangleq \bigcap_{-\infty < n \leq 0} \mathcal{F}_n \]

We summarize the preceding considerations as follows:

**Theorem 2.5.12** Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots\} \) is a negatively indexed submartingale on \((\Omega, \mathcal{F}, P)\). Then there exists some a.s. unique mapping \( X_{-\infty} : \Omega \rightarrow \mathbb{R} \), measurable with respect to \( \mathcal{F}_{-\infty} \) defined in (2.81), such that \( \lim_{n \to -\infty} X_n = X_{-\infty} \) a.s.

Observe that we have not called \( X_{-\infty} \) in this theorem a random variable since there is no guarantee that \( X_{-\infty} \) is a \( \mathbb{R} \)-valued function; indeed we could well have \( P[|X_{-\infty}| = +\infty] > 0 \) (e.g. take \( X_n \triangleq n \) and \( \mathcal{F}_n \triangleq \{\emptyset, \Omega\} \) for all \( n = 0, -1, -2, \ldots \), in which case \( X_{-\infty} \equiv -\infty \)). In §2.7 we shall develop conditions which are sufficient to ensure that \( X_{-\infty} \) is an integrable random variable (see Theorem 2.7.8).

### 2.6 Uniform Integrability

A classical problem in measure and integration theory concerns conditions which ensure that a.s. convergence of a sequence of functions to some limit implies \( L^1 \)-convergence of the sequence to the same limit. Such conditions are given by the Lebesgue Dominated Convergence Theorem 1.2.21 which postulates existence of an integrable “dominating” function for the given sequence of functions. In probability theory and stochastic processes one must often show that a.s. convergence implies \( L^1 \)-convergence when no such dominating function is clearly apparent. Our goal in this section is to determine alternative conditions for this to happen which do not require the presence of a dominating function. The essential condition is expressed by the concept of uniform integrability, which we now introduce.

**Proposition 2.6.1** Suppose that \( X \) is a random variable defined on the probability space \((\Omega, \mathcal{F}, P)\) such that \( E|X| < \infty \). Then, corresponding to each \( \epsilon \in (0, \infty) \), there is some \( c(\epsilon) \in [0, \infty) \) such that
(2.82) \[ E[|X|; |X| \geq c] < \epsilon \]
for all \( c \in [c(\epsilon), \infty) \).
Proof: Fix arbitrary \( \epsilon \in (0, \infty) \); in view of Theorem 1.2.20 there is some \( \delta(\epsilon) \in (0, \infty) \) such that
\[
(2.83) \quad E[|X|; A] < \epsilon
\]
for all \( A \in F \) with \( P(A) < \delta(\epsilon) \). By the Markov inequality (see Theorem 1.2.23) we have \( P[|X| \geq c] \leq c^{-1}E[|X|] \) for all \( c \in (0, \infty) \), and since \( E[|X|] < \infty \) there exists some \( c(\epsilon) \in (0, \infty) \) such that \( P[|X| \geq c] < \delta(\epsilon) \) for all \( c \in (c(\epsilon), \infty) \). Fixing \( c \in (c(\epsilon), \infty) \) and writing \( A \triangleq [|X| \geq c] \) we obtain (2.82) from (2.83).

We now formulate the notion of uniform integrability for a collection of integrable random variables. Essentially this requires that the condition expressed by Proposition 2.6.1 hold uniformly for all random variables in the collection:

**Definition 2.6.2** A collection of random variables \( \mathcal{C} \) on a probability space \((\Omega, F, P)\) is uniformly integrable when the following holds: corresponding to each \( \epsilon \in (0, \infty) \) there is some \( c(\epsilon) \in [0, \infty) \) such that
\[
\sup_{X \in \mathcal{C}} E[|X|; |X| \geq c] < \epsilon
\]
for all \( c \in [c(\epsilon), \infty) \).

**Remark 2.6.3** Suppose that \( \mathcal{C} \) is a uniformly integrable collection of random variables on \((\Omega, F, P)\). Fix some \( \epsilon \in (0, \infty) \), let \( c(\epsilon) \) be as in Definition 2.6.2, and fix some \( c \in (c(\epsilon), \infty) \). Then
\[
E|X| = E[|X|; |X| < c] + E[|X|; |X| \geq c] < (c + \epsilon), \quad \forall X \in \mathcal{C}.
\]

It follows that a uniformly integrable collection of random variables is uniformly \( L^1 \)-bounded. The converse is, however, false. For an example of a collection of random variables which is \( L^1 \)-uniformly bounded but not uniformly integrable, take \( \Omega \triangleq [0, 1], \quad F \triangleq \mathcal{B}([0, 1]), \quad P \triangleq \text{Lebesgue measure on } \mathcal{B}([0, 1]) \). For all \( n = 1, 2, \ldots \) define \( X_n : \Omega \to \mathbb{R} \) by \( X_n(\omega) \triangleq n \) for all \( \omega \in [0, 1/n], \quad X_n(\omega) \triangleq 0 \) for all \( \omega \in ((1/n), 1], \) and put \( \mathcal{C} \triangleq \{X_n, \ n = 1, 2, \ldots\} \). Since \( E|X_n| = 1 \) for all \( n \) we see that \( \mathcal{C} \) is uniformly \( L^1 \)-bounded. However, for any \( c \in (0, \infty) \), we have \( E[|X_n|; |X_n| \geq c] = 1 \) for all integers \( n > c \). Thus \( \mathcal{C} \) fails to be uniformly integrable.

**Remark 2.6.4** Suppose that \( \{X_n, \ n = 0, 1, 2, \ldots\} \) is a sequence of random variables on a probability space \((\Omega, F, P)\), with \( 0 \leq X_n \leq X_{n+1} \) a.s. for each \( n = 0, 1, 2, \ldots \) (this is called an increasing sequence). If this sequence is \( L^1 \)-bounded, namely \( \sup_n E|X_n| < \infty \), then it follows at once from the Monotone Convergence Theorem 1.2.12 that the collection \( \mathcal{C} \) is uniformly integrable. Notice that, although the sequence constructed in Remark 2.6.3 is \( L^1 \)-bounded, it fails to be increasing, and this is what causes uniform integrability to fail.

The next theorem gives two conditions which are sufficient to ensure that a given collection of random variables is uniformly integrable:
Theorem 2.6.5 Suppose that \( \mathcal{C} \) is a given collection of random variables defined on the probability space \((\Omega, \mathcal{F}, P)\).

(a) If there is some non-negative random variable \( Y \) on \( \Omega \) such that \( E[Y] < \infty \) and \( |X| \leq Y \) a.s. for all \( X \in \mathcal{C} \), then \( \mathcal{C} \) is uniformly integrable.

(b) If \( \sup_{X \in \mathcal{C}} E[|X|^p] < \infty \) for some \( p \in (1, \infty) \) then \( \mathcal{C} \) is uniformly integrable.

Proof: (a) For each \( X \in \mathcal{C} \) and \( c \in (0, \infty) \) it is clear that \( E[|X|; |X| \geq c] \leq E[Y; Y \geq c] \). That \( \mathcal{C} \) is uniformly integrable now follows from Proposition 2.6.1.

(b) Suppose that \( E[|X|^p] \leq B < \infty \), \( \forall X \in \mathcal{C} \). For each \( c \in (0, \infty) \) we have

\[
(2.84) \quad E[|X|; |X| \geq c] = E[|X|^p; |X|^{1-p}] \leq c^{1-p} \cdot E[|X|^p; |X| \geq c] \leq B \cdot c^{1-p}, \quad \forall X \in \mathcal{C}.
\]

Since \( p \in (1, \infty) \) we have \( c^{1-p} \to 0 \) as \( c \to \infty \) and hence \( (2.84) \) implies that \( \mathcal{C} \) is uniformly integrable.

Remark 2.6.6 Suppose that \( \{X_\lambda, \lambda \in \Lambda\} \) and \( \{Z_\lambda, \lambda \in \Lambda\} \) are collections of random variables defined on \((\Omega, \mathcal{F}, P)\) and indexed by a common set \( \Lambda \). The following observations follow trivially from Definition 2.6.2:

(a) \( \{X_\lambda, \lambda \in \Lambda\} \) is uniformly integrable if and only if \( \{|X_\lambda|, \lambda \in \Lambda\} \) is uniformly integrable.

(b) If \( \mathcal{C} \equiv \{X_\lambda, \lambda \in \Lambda\} \) is uniformly integrable then so is any subset of \( \mathcal{C} \).

(c) If \( \{X_\lambda, \lambda \in \Lambda\} \) and \( \{Z_\lambda, \lambda \in \Lambda\} \) are both uniformly integrable then so also is \( \{X_\lambda + Z_\lambda, \lambda \in \Lambda\} \).

(d) If, for each \( \lambda \in \Lambda \), \( Y_\lambda \) is a random variable defined on \( \Omega \) such that \( X_\lambda = Y_\lambda \) a.s. then uniform integrability of \( \{X_\lambda, \lambda \in \Lambda\} \) implies uniform integrability of \( \{Y_\lambda, \lambda \in \Lambda\} \).

The next theorem gives another condition for uniform integrability. This condition will be especially useful when we look at uniform integrability of martingales:

Theorem 2.6.7 Suppose that \( X \) is a random variable on the probability space \((\Omega, \mathcal{F}, P)\) such that \( E|X| < \infty \), and \( \{\mathcal{G}_\lambda, \lambda \in \Lambda\} \) is a collection of sub \( \sigma \)-algebras of \( \mathcal{F} \) (that is, \( \mathcal{G}_\lambda \subset \mathcal{F} \) for each \( \lambda \in \Lambda \)).

For each \( \lambda \in \Lambda \) put \( Y_\lambda \triangleq E[X|\mathcal{G}_\lambda] \). Then the collection of random variables \( \{Y_\lambda, \lambda \in \Lambda\} \) is uniformly integrable.

Proof: Fix an arbitrary \( \epsilon \in (0, \infty) \). By Theorem 1.2.20 there exists some \( \delta(\epsilon) \in (0, \infty) \) such that

\[
(2.86) \quad E[|X|; A] < \epsilon
\]

for all \( A \in \mathcal{F} \) subject to \( P(A) < \delta(\epsilon) \). Now, for each \( \lambda \in \Lambda \), one has \( Y_\lambda = E[X|\mathcal{G}_\lambda] \) a.s. thus Jensen’s inequality (Theorem 1.4.20) gives

\[
(2.87) \quad |Y_\lambda| \leq E[|X| | \mathcal{G}_\lambda] \quad \text{a.s.}
\]

In view of \( (2.87) \) and the Markov inequality (Theorem 1.2.23) we get:

\[
(2.88) \quad P[|Y_\lambda| \geq c] \leq c^{-1} E|Y_\lambda| \leq c^{-1} E|X|
\]
for all \( c \in (0, \infty) \) and all \( \lambda \in \Lambda \). Since \( E|X| < \infty \) we see from (2.88) that there is some \( c(\epsilon) \in (0, \infty) \) such that \( P[|Y_\lambda| \geq c] < \delta(\epsilon) \) for all \( c \in (c(\epsilon), \infty) \) and all \( \lambda \in \Lambda \), whence, from (2.86) we get

\[
E[|X|; |Y_\lambda| \geq c] < \epsilon
\]

for all \( c \in (c(\epsilon), \infty) \) and \( \lambda \in \Lambda \). Now \( \{||Y_\lambda| \geq c\} \in \mathcal{G}_\lambda \) and thus by (2.87) and the definition of conditional expectation:

\[
E[|Y_\lambda|; |Y_\lambda| \geq c] \leq E[E[|X| \mid \mathcal{G}_\lambda]; |Y_\lambda| \geq c] = E[|X|; |Y_\lambda| \geq c]
\]

for all \( c \in (0, \infty) \) and \( \lambda \in \Lambda \). From (2.89) and (2.90) we get \( E[|Y_\lambda|; |Y_\lambda| \geq c] < \epsilon \) for all \( c \in (c(\epsilon), \infty) \) and \( \lambda \in \Lambda \), as required.

\[\Box\]

**Remark 2.6.8** In view of Remark 2.6.6 (d) it is clear that the specific choice of conditional expectation is immaterial as far as uniform integrability of \( \{Y_\lambda, \lambda \in \Lambda\} \) is concerned.

For the next proposition we shall need mappings \( \theta_c : \mathbb{R} \to \mathbb{R} \) defined for each \( c \in (0, \infty) \) as follows:

\[
\begin{align*}
\theta_c(x) & \triangleq x \quad \forall x \in [-c, c] \\
\theta_c(x) & \triangleq c \quad \forall x \in (c, \infty) \\
\theta_c(x) & \triangleq -c \quad \forall x \in (-\infty, -c)
\end{align*}
\]

\[(2.91)\]

**Proposition 2.6.9** Suppose that \( \mathcal{C} \) is a uniformly integrable collection of random variables defined on the probability space \( (\Omega, \mathcal{F}, P) \). Then, for each \( \epsilon \in (0, \infty) \), there exists some \( c(\epsilon) \in [0, \infty) \) such that \( E|X - \theta_c(X)| < \epsilon \) for all \( X \in \mathcal{C} \) and all \( c \in (c(\epsilon), \infty) \).

Proof: Fix some arbitrary \( c \in (0, \infty) \), and observe that

\[
|x - \theta_c(x)| = (x - c)^+ + (x + c)^-
\]

for all \( x \in \mathbb{R} \). Now

\[
E[(X - c)^+] = E[(X - c); X \geq c] \leq E[|X|; |X| \geq c]
\]

and

\[
E[(X + c)^-] = E[ -(X + c); X \leq -c] \leq E[|X|; |X| \geq c]
\]

The proposition follows from (2.92), (2.93) and (2.94).

\[\Box\]

Using Proposition 2.6.9 we can establish the main result of this section:

**Theorem 2.6.10 (Vitali)** Suppose that \( X \) and \( \{X_n, n = 0, 1, 2, \ldots\} \) are random variables defined on the probability space \( (\Omega, \mathcal{F}, P) \), and that \( \lim_{n \to \infty} X_n = X \) a.s. Then the following are equivalent:

\( (a) \) the collection \( \{X_n, n = 0, 1, 2, \ldots\} \) is uniformly integrable;

\( (b) \) \( E|X| < \infty \) and \( \lim_{n \to \infty} E|X_n - X| = 0 \);

\( (c) \) \( E|X| < \infty \) and \( \lim_{n \to \infty} E|X_n| = E|X| \).
Proof (a) implies (b): In view of (a) and Remark 2.6.3 there is some $B < \infty$ such that $E|X_n| < B$ for all $n = 0, 1, 2, \ldots$. Thus, from (a) and Fatou’s Theorem 1.2.15:

$$E|X| = E\liminf_{n \to \infty} |X_n| \leq \liminf_{n \to \infty} E|X_n| < B$$

Now fix some arbitrary $\epsilon \in (0, \infty)$, and for each $c \in (0, \infty)$ let $\theta_c(\cdot)$ be defined as in (2.91). Then, by Proposition 2.6.9 and (a), there is some $c(\epsilon) \in (0, \infty)$ such that

$$E|X_n - \theta_c(X_n)| < (\epsilon/3) \quad \text{and} \quad E|X - \theta_c(X)| < (\epsilon/3)$$

for all $n = 0, 1, 2, \ldots$ and $c \in (c(\epsilon), \infty)$. Now fix any $c \in (c(\epsilon), \infty)$. Clearly $\theta_c(\cdot)$ is a continuous mapping and thus $\{\theta_c(X_n), n = 0, 1, 2, \ldots\}$ is a sequence of random variables bounded in magnitude by $c$, and $\lim_{n \to \infty} \theta_c(X_n) = \theta_c(X)$ a.s. (from (a)). Thus, by the Dominated Convergence Theorem 1.2.21, there is some integer $N(\epsilon)$ such that

$$E|\theta_c(X_n) - \theta_c(X)| < (\epsilon/3)$$

for all $n \geq N(\epsilon)$. We see from (2.95) and (2.96) that

$$E|X - X_n| \leq E|X_n - \theta_c(X_n)| + E|\theta_c(X_n) - \theta_c(X)| + E|\theta_c(X) - X| < \epsilon$$

for all $n \geq N(\epsilon)$.

(b) implies (c): Since $|a| - |b| \leq |a - b|$, $\forall \ a, b \in \mathbb{R}$, one has

$$|E|X_n| - E|X|| \leq E||X_n| - |X|| \leq E|X_n - X|.$$

(c) implies (a): For each $c \in (1, \infty)$ define the continuous function $\psi_c : [0, \infty) \to [0, \infty)$ as follows:

$$\psi_c(x) \triangleq \begin{cases} x & \forall x \in [0, c] \\ 0 & \forall x \in [c + 1, \infty) \\ \text{linear} & \forall x \in [c, c + 1] \end{cases}$$

and observe that

$$xI_{[c+1,\infty)}(x) \leq x - \psi_c(x) \leq xI_{[c,\infty)}(x), \quad \forall \ x \in [0, \infty).$$

Now fix some $\epsilon \in (0, \infty)$. In view of Proposition 2.6.1, the fact that $E|X| < \infty$, and the second inequality of (2.97), there exists some $c_1 \in (0, \infty)$ such that

$$0 \leq E|X| - E\psi_{c_1}(|X|) < \epsilon/2.$$

Since $X_n \to X$ a.s. and $\psi_{c_1}(\cdot)$ is continuous and uniformly bounded, the Dominated Convergence Theorem ensures $\lim_{n \to \infty} E\psi_{c_1}(|X_n|) = E\psi_{c_1}(|X|)$; this, together with the convergence in (c), shows that there is some positive integer $n(\epsilon)$ such that

$$E|X_n| - E\psi_{c_1}(|X_n|) - (E|X| - E\psi_{c_1}(|X|)) < \epsilon/2, \quad \forall \ n \geq n(\epsilon).$$
Upon combining (2.99), (2.98) and the first inequality of (2.97) there results:

\begin{equation}
E[|X_n|; |X_n| \geq c_1 + 1] \leq E|X_n| - E\psi_{c_1}(|X_n|) < \epsilon, \quad \forall n \geq n(\epsilon).
\end{equation}

Additionally, one sees from Proposition 2.6.1 that there is some \(c_2 \in (0, \infty)\) such that

\begin{equation}
E[|X_n|; |X_n| \geq c_2 + 1] < \epsilon, \quad \forall n = 0, 1, 2, \ldots, n(\epsilon).
\end{equation}

By (2.100) and (2.101), for each \(c > 1 + \max\{c_1, c_2\}\) one has

\begin{equation}
E[|X_n|; |X_n| \geq c] < \epsilon, \quad \forall n = 0, 1, 2, \ldots
\end{equation}

showing that \(\{X_n; n = 0, 1, 2 \ldots\}\) is uniformly integrable.

\textbf{Remark 2.6.11} Theorem 2.6.10 shows that, when \(X_n \to X\) a.s. on a probability space \((\Omega, \mathcal{F}, P)\), then uniform integrability is both necessary and sufficient to get \(E|X_n - X| \to 0\); consequently, uniform integrability is the most \textit{general possible} condition which secures \(L^1\)-convergence. To this extent Theorem 2.6.10 is a more general result than the Dominated Convergence Theorem 1.2.21 (which it actually implies on the basis of Theorem 2.6.5(a)). Notice however, that the results of this section to \textit{not} apply to integration over \textit{unbounded} measure spaces, for which the Dominated Convergence Theorem is still the basic tool for getting \(L^1\)-convergence from a.e. convergence.

\section{2.7 Uniformly Integrable Supermartingales}

A supermartingale \(\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}\) will be called \textit{uniformly integrable} when the collection of random variables \(\{X_n; n = 0, 1, 2, \ldots\}\) is uniformly integrable. The terms uniformly integrable submartingale and uniformly integrable martingale will have an obviously analogous interpretation.

\textbf{Theorem 2.7.1} (a) Suppose \(\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}\) is a uniformly integrable supermartingale defined on \((\Omega, \mathcal{F}, P)\). Then there exists an a.s. unique \(\mathcal{F}_\infty\)-measurable random variable \(X_\infty: \Omega \to \mathbb{R}\) such that \(E|X_\infty| < \infty\) and the following hold:

\begin{equation}
\lim_{n \to \infty} X_n = X_\infty \quad \text{a.s.}
\end{equation}

\begin{equation}
\lim_{n \to \infty} E|X_n - X_\infty| = 0;
\end{equation}

\begin{equation}
X_n \geq E[X_\infty | \mathcal{F}_n] \quad \text{a.s.} \quad \forall n = 0, 1, 2, \ldots
\end{equation}

In particular, \(\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}\) is a closed supermartingale.

(b) If \(\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}\) is a uniformly integrable martingale on \((\Omega, \mathcal{F}, P)\), then the preceding statements continue to hold, except that (2.104) is replaced by

\begin{equation}
X_n = E[X_\infty | \mathcal{F}_n] \quad \text{a.s.} \quad \forall n = 0, 1, 2, \ldots
\end{equation}

and \(\{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots, +\infty\}\) is a closed martingale.
Proof: (a) Since the collection of random variables \( \{X_n, n = 0, 1, 2, \ldots \} \) is uniformly integrable, we see from Remark 2.6.3 that \( \sup_n E|X_n| < \infty \) and hence, from Theorem 2.5.7, there exists some \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty \) such that the a.s. convergence in (2.102) holds, and thus, in view of Theorem 2.6.10 and a second application of the postulated uniform integrability of the sequence \( \{X_n\} \), we get the \( L^1 \)-convergence in (2.103) as well. It remains to establish that (2.104) holds. Fix an arbitrary non-negative integer \( n \) and some arbitrary event \( A \in \mathcal{F}_n \). For all integers \( k \geq n \) we have that

\[
E[X_n; A] \geq E[X_k; A],
\]

and clearly

\[
|E[X_k; A] - E[X_\infty; A]| \leq E|X_k - X_\infty|.
\]

Since we have already established that \( \lim_{k \to \infty} E|X_k - X_\infty| = 0 \), it follows from (2.106) and (2.107) that

\[
E[X_n; A] \geq E[X_\infty; A].
\]

Now (2.108) holds for all \( A \in \mathcal{F}_n \), hence we obtain (2.104) from the \( \mathcal{F}_n \)-measurability of \( X_n \). The asserted a.s. uniqueness of \( X_\infty \) is an immediate consequence of (2.102).

Remark 2.7.2 The preceding theorem establishes closability of a uniformly integrable supermartingale \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) (recall Remark 2.3.6), and shows that a closing random variable \( X_\infty \) is the a.s. and \( L^1 \)-limits of the sequence \( \{X_n; n = 0, 1, 2, \ldots \} \).

Remark 2.7.3 From Theorem 2.6.7, Remark 2.3.6(c), and Theorem 2.7.1(b) we see that a martingale is closable if and only if it is uniformly integrable. Notice that there is not such a tight equivalence in the case of supermartingales: Theorem 2.7.1(a) shows that uniformly integrable supermartingales are closable, but a closable supermartingale need not be uniformly integrable. In fact Example 2.5.9 gives a nonnegative martingale (hence supermartingale) \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) which is closable (by Remark 2.3.6(a)), and converges a.s. to an integrable limiting random variable \( X_\infty \); since \( L^1 \)-convergence to \( X_\infty \) has been shown not to hold, the supermartingale cannot be uniformly integrable.

When \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is an \( L^p \)-bounded martingale for some \( p \in (1, \infty) \) then we can strengthen the conclusions of Theorem 2.7.1 as follows:

Theorem 2.7.4 Suppose \( p \in (1, \infty) \) and \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is an \( L^p \)-bounded martingale on a probability space \((\Omega, \mathcal{F}, P)\). Then there exists an a.s. unique \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty : \Omega \to \mathbb{R} \) such that \( E|X_\infty|^p < \infty \) and the following hold:

\[
\lim_{n \to \infty} X_n = X_\infty \quad \text{a.s.}
\]

\[
\lim_{n \to \infty} E|X_n - X_\infty|^p = 0;
\]

85
Remark 2.7.5 A useful observation based on (2.112) is that, if \( \{ (X_n, \mathcal{F}_n) ; n = 0, 1, 2, \ldots \} \) is a submartingale (by Proposition 2.2.10 and convexity of \( c(x) \triangleq |x|^p \)), whence, by (2.113) and Theorem 2.5.7, there is some a.s. unique \( \mathcal{F}_\infty \)-measurable random variable \( Z_\infty \) such that \( E|Z_\infty| < \infty \) and \( \lim_{n \to \infty} |X_n|^p = Z_\infty \) a.s. In view of (2.109) and a.s. uniqueness we see that \( Z_\infty = |X_\infty|^p \) a.s. thus \( E|X_\infty|^p < \infty \). It remains to establish (2.110) and (2.112). Since \( \{ (|X_n|, \mathcal{F}_n) ; n = 0, 1, 2, \ldots \} \) is clearly a non-negative submartingale, we have

\[
E \left[ \max_{0 \leq k \leq n} |X_k|^p \right] \leq \left( \frac{p}{p-1} \right)^p E|X_n|^p \leq \left( \frac{p}{p-1} \right)^p E|X_\infty|^p, \quad \forall \ n = 1, 2, \ldots
\]

where the first inequality follows from Theorem 2.4.4 and the second from (2.111) along with Jensen’s inequality for conditional expectations (Theorem 1.4.20). Taking \( n \to \infty \) and using the monotone convergence theorem then gives (2.112). As for (2.110), observe that

\[
|X_n - X_\infty|^p \leq 2^p \sup_{0 \leq k < \infty} |X_k|^p, \quad \forall \ n = 1, 2, \ldots
\]

Now (2.112) ensures that the right side of (2.114) is integrable, whence (2.110) follows from (2.109) and the Lebesgue Dominated Convergence Theorem 1.2.21.

Theorem 2.7.6 (Lévy) Suppose that \( \{ \mathcal{F}_n, n = 0, 1, 2, \ldots \} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\) and \( X \) is a random variable on \((\Omega, \mathcal{F}, P)\) with \( E|X| < \infty \). Then

\[
\lim_{n \to \infty} E[X \mid \mathcal{F}_n] = E[X \mid \mathcal{F}_\infty]
\]

in the sense of both a.s. and \( L^1 \)-convergence.

Proof: Put \( X_n \triangleq E[X \mid \mathcal{F}_n], \forall n = 0, 1, 2, \ldots \) That \( \{ (X_n, \mathcal{F}_n) ; n = 0, 1, 2, \ldots \} \) is a martingale has been seen in Example 2.2.5, while uniform integrability of the collection \( \{ X_n, n = 0, 1, 2, \ldots \} \) follows from
Theorem 2.6.7. Then, from Theorem 2.7.1(b), there exists some $F_\infty$-measurable random variable $X_\infty$ such that $E|X_\infty| < \infty$, with

$$\lim_{n \to \infty} X_n = X_\infty,$$

the convergence taking place in both a.s. and $L^1$. It remains to show that

$$X_\infty = E[X | F_\infty] \text{ a.s.}$$

in order to conclude (2.115) from (2.116). Put $D \triangleq \bigcup_{0 \leq n < \infty} F_n$ and fix some $A \in D$. Then there is some positive integer $n_1$ such that $A \in F_n, \forall n \geq n_1$, thus

$$E[X; A] = E[E[X | F_n]; A] = E[X_n; A], \quad \forall n \geq n_1.$$ (2.118)

By $L^1$-convergence in (2.116) one also has

$$\lim_{n \to \infty} E[X_n; A] = E[X_\infty; A],$$ (2.119)

thus (2.118) and (2.119) show that

$$E[X; A] = E[X_\infty; A], \quad \forall A \in D.$$ (2.120)

Now put

$$C \triangleq \{ A \in F_\infty : E[X; A] = E[X_\infty; A] \}.$$ (2.121)

One trivially checks that $C$ is a $\lambda$-class over $\Omega$ and (2.120) shows that $D \subset C$. Since $D$ is clearly a $\pi$-class and $\sigma\{D\} = F_\infty$ (see (2.1)), Theorem 1.5.4 shows that $C = F_\infty$. In view of (2.121), we see that $E[X; A] = E[X_\infty; A], \forall A \in F_\infty$; since $X_\infty$ is $F_\infty$-measurable, we get (2.117) as required. 

The next theorem is an important result which ensures that a negatively indexed submartingale is necessarily uniformly integrable provided that it is lower-bounded. This theorem will be used several times in Chapter 4, when we develop the basic properties of martingales in a continuous-parameter context.

**Theorem 2.7.7** Suppose that \{(X_n, F_n); n = 0, -1, -2, \ldots\} is a negatively indexed submartingale on $(\Omega, F, P)$ such that

$$\inf_{-\infty < n \leq 0} E[X_n] > -\infty.$$ (2.122)

Then \{X_n, n = 0, -1, -2, \ldots\} is a uniformly integrable collection of random variables.

Proof: Put $\alpha \triangleq \inf_{-\infty < n \leq 0} E[X_n]$ and fix some arbitrary $\epsilon \in (0, \infty)$. Since $-\infty < \alpha \leq E[X_{n-1}] \leq E[X_n]$ there exists some negative integer $N(\epsilon)$ such that

$$0 \leq E[X_{N(\epsilon)}] - E[X_n] < \epsilon$$ (2.123)
for all integers \( n \leq N(\epsilon) \). Next, fix some \( c \in (0, \infty) \) and observe that
\[
|x| \cdot 1\{|x| > c\} = x \cdot 1\{x > c\} + x \cdot 1\{x \geq -c\} - x
\]
for all \( x \in \mathbb{R} \). Thus
\[
E[|X_n|; |X_n| > c] = E[X_n; X_n > c] + E[X_n; X_n \geq -c] - E[X_n]
\]
for all \( n = 0, -1, -2 \ldots \) Since \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) is a submartingale,
\[
E[X_n; X_n > c] \leq E[X_{N(\epsilon)}; X_n > c] \quad \text{and} \quad E[X_n; X_n \geq -c] \leq E[X_{N(\epsilon)}; X_n \geq -c]
\]
for all integers \( n \leq N(\epsilon) \). Putting together (2.123), (2.124) and (2.125), we obtain:
\[
E[|X_n|; |X_n| > c] \leq E[|X_{N(\epsilon)}|; |X_n| > c] + \epsilon
\]
for all integers \( n \leq N(\epsilon) \). Next, observe that \( |X_n| = 2X_n^+ - X_n \) and \( E[X_n] \geq \alpha > -\infty \) for all \( n \). Moreover, by Theorem 1.4.20 with \( c(x) \triangleq \max\{x, 0\} \), we see that \( \{(X_n^+, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) is a submartingale, hence \( E[X_n^+] \leq E[X_0^+] \). We have thus obtained
\[
E[|X_n|; |X_n| > c] \leq 2E[X_0^+] - \alpha < \infty;
\]
In view of Theorem 1.2.20, corresponding to \( \epsilon \in (0, \infty) \) there is some \( \delta(\epsilon) \in (0, \infty) \) such that
\( E[|X_{N(\epsilon)}|; A] < \epsilon \) for all \( A \in \mathcal{F} \) subject to \( P(A) < \delta(\epsilon) \). Fixing \( c_1(\epsilon) \in (0, \infty) \) such that \( c^{-1}(2E[X_0^+] - \alpha) < \delta(\epsilon) \) for all \( c \in (c_1(\epsilon), \infty) \), and using (2.127) along with the Markov inequality (Theorem 1.2.23), we see that \( P[|X_n| > c] < \delta(\epsilon) \), hence in particular \( E[|X_{N(\epsilon)}|; |X_n| > c] < \epsilon \), for all \( c \in (c_1(\epsilon), \infty) \) and \( n = 0, -1, -2 \ldots \). Thus from (2.126) it follows that
\[
E[|X_n|; |X_n| > c] < 2\epsilon
\]
for all \( c \in (c_1(\epsilon), \infty) \) and integers \( n \leq N(\epsilon) \). Finally, observe from Proposition 2.6.1 that there exists some \( c_2(\epsilon) \in (0, \infty) \) such that
\[
E[|X_n|; |X_n| > c] < 2\epsilon
\]
for all \( c \in (c_2(\epsilon), \infty) \) and all \( n = 0, -1, -2 \ldots N(\epsilon) \). Uniform integrability of \( \{X_n, n = 0, -1, -2 \ldots \} \) follows from (2.128) and (2.129).

Theorem 2.7.7, together with Theorem 2.6.10, immediately yields the following refinement of Theorem 2.5.12:

**Theorem 2.7.8** Suppose that (2.122) holds for some negatively indexed submartingale \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) on \( (\Omega, \mathcal{F}, P) \). Then there exists some a.s. unique random variable \( X_{-\infty} : \Omega \rightarrow \mathbb{R} \), measurable with respect to \( \mathcal{F}_{-\infty} \), such that \( E|X_{-\infty}| < \infty \) and \( \lim_{n \rightarrow -\infty} X_n = X_{-\infty} \) holds in both the a.s. and \( L^1 \) sense.
We next state a theorem which complements Theorem 2.7.6 but applies to negatively indexed martingales. We shall use this result in Chapter 4 on continuous-parameter martingales.

**Theorem 2.7.9** Suppose that \( \{F_n, n = 0, -1, -2, \ldots \} \) is a sequence of \( \sigma \)-algebras in the probability space \((\Omega, \mathcal{F}, P)\), with \( F_n \subset F_{n+1} \subset \mathcal{F} \) for all \( n = -1, -2, \ldots \), and \( X \) is a random variable on \((\Omega, \mathcal{F}, P)\) such that \( E|X| < \infty \). Define \( X_n \overset{\Delta}{=} E[X | F_n] \) for all \( n = 0, -1, -2, \ldots \) Then \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) is a uniformly integrable negatively indexed martingale, and

\[
\lim_{n \to -\infty} X_n = E[X | \mathcal{F}_{-\infty}] \quad (2.130)
\]

in the sense of both a.s. and \( L^1 \)-convergence, \( \mathcal{F}_{-\infty} \) being defined in (2.81).

**Proof:** Clearly \( \{(X_n, \mathcal{F}_n); n = 0, -1, -2, \ldots \} \) is a negatively indexed martingale, as follows from the composition rule for conditional expectations (Theorem 1.4.15(d)), and since \( E[X_n] = E[X_0], \forall n = 0, -1, -2, \ldots \), we see that (2.122) holds trivially, and therefore, by Theorem 2.7.8, there is some \( \mathcal{F}_{-\infty} \)-measurable random variable \( X_{-\infty} \) such that \( E|X_{-\infty}| < \infty \) and

\[
\lim_{n \to -\infty} X_n = X_{-\infty} \quad (2.131)
\]

holds in both the a.s. and \( L^1 \)-sense. It remains to show that \( X_{-\infty} = E[X | \mathcal{F}_{-\infty}] \) a.s. To this end let \( A \) be an arbitrary event in \( \mathcal{F}_{-\infty} \). Then \( A \in \mathcal{F}_n, \forall n = 0, 1, 2, \ldots \), hence, by the definition of \( X_n \), we have

\[
E[X; A] = E[X_n; A] \quad (2.132)
\]

for all \( n = 0, -1, -2, \ldots \) It follows that

\[
|E[X; A] - E[X_{-\infty}; A]| = |E[X_n; A] - E[X_{-\infty}; A]| \leq E|X_n - X_{-\infty}| \quad (2.133)
\]

for all \( n = 0, -1, -2, \ldots \) Now, since (2.131) has been seen to hold in the \( L^1 \)-sense, we see from (2.133) and the arbitrary choice of \( A \) in \( \mathcal{F}_{-\infty} \) that

\[
E[X; A] = E[X_{-\infty}; A]
\]

for all \( A \in \mathcal{F}_{-\infty} \). Since \( X_{-\infty} \) is \( \mathcal{F}_{-\infty} \)-measurable it follows that \( X_{-\infty} = E[X | \mathcal{F}_{-\infty}] \) a.s., which establishes (2.130).
2.8 Problems

Problem 2.8.1 (a) Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is an adapted sequence on a probability space \((\Omega, \mathcal{F}, P)\) with \( E|X_n| < \infty, \forall n = 0, 1, 2, \ldots \) Show that \( \{X_n; n = 0, 1, 2 \ldots \} \) can be written as

\[
(2.134) \quad X_n = X_0 + A_n + M_n, \quad \forall \ n = 0, 1, 2, \ldots
\]

where \( \{(M_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a martingale with \( M_0 \equiv 0, A_0 \equiv 0, \) and \( A_n \) is \( \mathcal{F}_{n-1} \)-measurable for all \( n = 1, 2, \ldots \) Also, show that this decomposition is unique in the sense that, if \( \{(\tilde{M}_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a martingale with \( \tilde{M}_0 \equiv 0, \{\tilde{A}_n; n = 0, 1, 2 \ldots \} \) is a process such that \( \tilde{A}_0 \equiv 0 \) and \( \tilde{A}_n \) is \( \mathcal{F}_{n-1} \)-measurable for each \( n = 1, 2, \ldots, \) and

\[
X_n = X_0 + \tilde{A}_n + \tilde{M}_n, \quad \forall \ n = 0, 1, 2, \ldots
\]

then

\[
P[M_n = \tilde{M}_n] = 1 \quad \text{and} \quad P[A_n = \tilde{A}_n] = 1, \quad \forall \ n = 0, 1, 2, \ldots
\]

Hint: Observe that

\[
X_n - X_0 = \sum_{k=1}^{n} \{X_k - E[X_k | F_{k-1}]\} + \sum_{k=1}^{n} \{E[X_k | F_{k-1}] - X_k\},
\]

for all \( n = 1, 2, \ldots \)

(b) In the decomposition given by (2.134) show that

\[
0 \leq A_n \leq A_{n+1} \quad \text{a.s.} \quad \forall \ n = 0, 1, 2, \ldots
\]

if and only if \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a submartingale.

(c) Now suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots \} \) is a uniformly integrable submartingale. Show that \( \{M_n; n = 0, 1, 2 \ldots \} \) and \( \{A_n; n = 0, 1, 2 \ldots \} \) in (2.134) are uniformly integrable and \( E[A_\infty] < \infty, \) where \( A_\infty \triangleq \liminf_{n \to \infty} A_n. \)

Problem 2.8.2 Suppose that \( \{X_n; n = 0, 1, 2 \ldots \} \) is a uniformly integrable sequence of nonnegative random variables on \((\Omega, \mathcal{F}, P).\) Establish the following:

\[
\limsup_{n \to \infty} E[X_n] \leq E[\limsup_{n \to \infty} X_n].
\]

Hint: Establish the result supposing that the \( X_n \) are uniformly bounded: \( 0 \leq X_n(\omega) \leq c < \infty, \) \( \forall n = 0, 1, 2, \ldots, \forall \omega \in \Omega. \) Now generalize to the case where \( \{X_n; n = 0, 1, 2 \ldots \} \) is u.i.

Problem 2.8.3 (a) Suppose that \( \{X_n; n = 0, 1, 2 \ldots \} \) is a sequence of independent random variables on \((\Omega, \mathcal{F}, P),\) such that \( P[X_n = c] < 1, \forall n = 0, 1, 2, \ldots, \forall c \in \mathbb{R} \) i.e. each \( X_n \) is not a.s. equal to a constant value. Put

\[
\mathcal{F}_n \triangleq \sigma\{X_0, X_1, \ldots X_n\}, \quad \forall n = 0, 1, 2, \ldots
\]
Show that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ with $\mathcal{F}_n \neq \mathcal{F}_{n+1}$ i.e. $\mathcal{F}_{n+1}$ is strictly larger than $\mathcal{F}_n$.

(b) In Proposition 2.1.12 we saw the following: suppose $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$ is a given filtration in a probability space $(\Omega, \mathcal{F}, P)$, and $S, T$ are stopping times with respect to $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$. If $A \in \mathcal{F}_S$ then

$$A \cap \{S \leq T\} \in \mathcal{F}_{S\wedge T}. \tag{2.135}$$

Show by a counterexample that (2.135) is generally false when $A \in \mathcal{F}_T$.

**Problem 2.8.4** Suppose that $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$ is a filtration in the probability space $(\Omega, \mathcal{F}, P)$, and $T$ is a $\{\mathcal{F}_n\}$-stopping time such that

$$P[T \leq n + 1 \mid \mathcal{F}_n] > a, \quad \forall n = 0, 1, 2, \ldots$$

for some constant $a \in (0, 1)$.

(a) Show that

$$P[T > k] \leq (1 - a)^k, \quad \forall k = 1, 2, \ldots$$

Hint: Observe that $P[T > k] = P[T > k, T > (k - 1)], \forall k = 1, 2, \ldots$, and use induction.

(b) Use the result from (a) to conclude that $E[T] < \infty$.

**Problem 2.8.5** (a) Suppose that $Y$ is a zero-mean random variable on $(\Omega, \mathcal{F}, P)$, with $|Y(\omega)| \leq a$ for all $\omega \in \Omega$ and some constant $a \in [0, \infty)$. Show that

$$E \exp\{\lambda Y\} \leq e^{a\lambda} + e^{-a\lambda}, \quad \forall \lambda \in \mathbb{R}.$$ 

Hint: Put $g(x) \triangleq \exp\{\lambda x\}$, $x \in [-a, a]$, and observe from convexity of $g(\cdot)$ that

$$g(x) \leq \frac{a - x}{2a} g(-a) + \frac{a + x}{2a} g(a), \quad \forall x \in [-a, a].$$

(b) Suppose that $\{Y_k, k = 1, 2, \ldots\}$ are independent integrable random variables on $(\Omega, \mathcal{F}, P)$ such that $EY_k = 0$ and $|Y_k(\omega)| \leq a$ for all $\omega \in \Omega$ and $k = 1, 2, \ldots$. Put

$$X_n \triangleq \sum_{k=1}^{n} Y_k, \quad \mathcal{F}_n \triangleq \sigma\{Y_1, Y_2, \ldots, Y_n\}, \quad \forall n = 1, 2, \ldots$$

Use the result in (a) to establish that

$$P\left[\max_{1 \leq k \leq n} X_k \geq x\right] \leq \exp\left\{\frac{-x^2}{2a^2(n)}\right\}, \quad \forall x \in [0, \infty), \quad \forall n = 1, 2, \ldots.$$ 

Hint: Put $Z_n \triangleq e^{\lambda X_n}$ for arbitrary $\lambda \in (0, \infty)$, show that $\{(Z_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}$ is a non-negative submartingale, and apply Theorem 2.4.1. Then find the value of parameter $\lambda$ which minimizes the resulting upper-bound. Note also the useful inequality $(e^x + e^{-x})/2 \leq \exp[x^2/2]$, $x \in \mathbb{R}.$
Problem 2.8.6 Let \( \{X_n, n = 1, 2, \ldots\} \) be an i.i.d. sequence on \((\Omega, \mathcal{F}, P)\) with \(E|X_n| < \infty \) and \( \mu \triangleq E[X_n] \). Put
\[
S_n \triangleq \sum_{i=1}^{n} X_i, \quad \mathcal{F}_n \triangleq \sigma\{X_1, \ldots, X_n\}, \forall n = 1, 2, \ldots
\]
and let \( T : \Omega \to \{1, 2, \ldots, \infty\} \) be an \( \mathcal{F}_n \)-stopping time (i.e. \( \{T \leq n\} \in \mathcal{F}_n \), for all \( n = 1, 2, \ldots \)) such that \( E[T] < \infty \). Establish the following:
(a) \( E[|X_n| I\{n \leq T\}] = E[|X_1|] P\{T \geq n\} \) and \( E[X_n I\{n \leq T\}] = \mu P\{T \geq n\} \), for all \( n = 1, 2, \ldots \)
(b) \( E[S_T] = \mu E[T] \) (hint: use the result from (a) together with the monotone and dominated convergence theorems).

Problem 2.8.7 Suppose that \( \{X_n; n = 0, 1, 2 \ldots\} \) is a sequence of independent and identically distributed r.v.’s on \((\Omega, \mathcal{F}, P)\). Put
\[
\mu(\Gamma) \triangleq P(X_n \in \Gamma), \forall \Gamma \in \mathcal{B}(\mathbb{R}),
\]
and
\[
\mathcal{F}_n \triangleq \sigma\{X_0, X_1, X_2, \ldots X_n\}
\]
for all \( n = 0, 1, 2, \ldots \). If \( T : \Omega \to \{0, 1, 2, \ldots, +\infty\} \) is a stopping time with respect to \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) such that \( \{T < \infty\} = \Omega \), establish the following: For any \( A \in \mathcal{F}_T \), integer \( m \geq 1 \), and sets \( \Gamma_n \in \mathcal{B}(\mathbb{R}) \), \( \forall n = 1, 2, \ldots, m \), we have
\[
P(\omega \in A : X_{T(\omega)+n}\in \Gamma_n, \forall n = 1, 2, \ldots m) = P(A) \prod_{n=1}^{m} \mu(\Gamma_n).
\]

Problem 2.8.8 Suppose that \( \{(X_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\} \) is a supermartingale on \((\Omega, \mathcal{F}, P)\) with \( X_n(\omega) \in [0, \infty), \forall n = 0, 1, 2, \ldots, \forall \omega \in \Omega \), and put
\[
T(\omega) \triangleq \min\{n : X_n(\omega) = 0\}, \forall \omega \in \Omega.
\]
Prove that
\[
P[\omega : T(\omega) < \infty \text{ and } X_n(\omega) = 0, n \geq T(\omega)] = P[T < \infty].
\]

Problem 2.8.9 Suppose that \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots\} \) is a filtration in \((\Omega, \mathcal{F}, P)\), and \( \{X_n, n = 0, 1, 2, \ldots\} \) is a sequence of random variables on \((\Omega, \mathcal{F}, P)\) converging a.s. to a random variable \( X \). If \( |X_n(\omega)| \leq Z(\omega) \), for all \( n = 0, 1, 2 \ldots \) and \( \omega \in \Omega \), for some random variable \( Z \) such that \( E|Z| < \infty \), establish that
\[
\lim_{n \to \infty} E[X_n | \mathcal{F}_n] = E[X | \mathcal{F}_\infty] \quad \text{a.s.}
\]
Hint: Put \( Y_m \triangleq \sup_{n \geq m} |X_n - X| \), for all \( m = 0, 1, 2, \ldots \) and
\[
\Delta_n \triangleq |E[X_n | \mathcal{F}_n] - E[X | \mathcal{F}_\infty]|, \quad n = 0, 1, 2, \ldots
\]
and show that
\[
\limsup_{n \to \infty} \Delta_n \leq E[Y_m | \mathcal{F}_\infty], \quad \text{for each } m.
\]
Problem 2.8.10 Suppose that \( \{X_n; n = 0, 1, 2, \ldots \} \) is a discrete-parameter process on \((\Omega, \mathcal{F}, P)\), let \( \{\mathcal{F}_n, n = 0, 1, 2, \ldots \} \) be the raw filtration generated by \( X_n \), namely
\[
\mathcal{F}_n \triangleq \sigma\{X_0, X_1, X_2, \ldots, X_n\}, \quad \forall n = 0, 1, 2, \ldots
\]
and let \( T : \Omega \to \{0, 1, 2, \ldots, +\infty\} \) be a \( \{\mathcal{F}_n\}\)-stopping time. Establish the following:

(a) If \( \omega_1, \omega_2 \in \Omega \) are such that \( T(\omega_1) < +\infty \), and \( X_k(\omega_1) = X_k(\omega_2), \forall k = 0, 1, 2, \ldots, T(\omega_1) \), then \( T(\omega_1) = T(\omega_2) \).

(b) If \( \omega_1, \omega_2 \in \Omega \) are such that \( T(\omega_1) = +\infty \), and \( X_k(\omega_1) = X_k(\omega_2), \forall k = 0, 1, 2, \ldots \), then \( T(\omega_2) = +\infty \).

(c) For each \( n = 0, 1, 2, \ldots \), there exists a set \( \Gamma_n \in \mathcal{B}(\mathbb{R}^{n+1}) \) such that
\[
\{\omega \in \Omega : T(\omega) = n\} = \{\omega \in \Omega : (X_0(\omega), X_1(\omega), X_2(\omega), \ldots, X_n(\omega)) \in \Gamma_n\}, \quad \forall n = 0, 1, 2, \ldots,
\]
and, if \( C_n \) is defined by
\[
C_n \triangleq \{(x_0, x_1, x_2, \ldots) \in \mathbb{R}^\infty : (x_0, x_1, x_2, \ldots, x_n) \in \Gamma_n\}, \quad \forall n = 0, 1, 2, \ldots,
\]
then \( C_m \cap C_n = \emptyset \) when \( m \neq n \).

(d) For each \( n = 0, 1, 2, \ldots \) one has
\[
\{T = n\} = \{(X_{0\wedge T}, X_{1\wedge T}, \ldots, X_{n\wedge T}) \in \Gamma_n\},
\]
where \( \Gamma_n \) is the set constructed in (c).

Now use (d) to establish the following characterization of the pre \( \sigma \)-algebra generated by the stopping time \( T \): \[(2.136) \quad \mathcal{F}_T = \sigma\{X_{k\wedge T}, k = 0, 1, 2, \ldots\}.
\]
Briefly explain the intuitive significance of the results in 7(a), 7(b) and (2.136).

Hint : use Proposition 1.3.19 for parts (a), (b), (c), and Theorem 1.3.22 for part (d).
Chapter 3

Elements of Continuous-Parameter Stochastic Processes

For the purposes of stochastic integration we must adapt and extend the theory of discrete-parameter martingales established in Chapter 2 to a continuous-parameter setting. In order to attain this goal it is necessary to first establish continuous-parameter “analogues” of the notions of filtration, adapted process, and stopping times, which were presented in Section 2.1 in discrete-parameter form. It turns out that the study of these ideas in a continuous-parameter setting involves non-trivial issues which simply do not arise in the discrete-parameter case, and fully merits a chapter of its own. Accordingly, a major goal of this chapter is to formulate definitions and basic properties of filtrations, stopping times, and the various notions of “adaptedness” which are natural in the continuous-parameter case. A second major goal of this chapter is to introduce the crucial notion of a Wiener process (or Brownian motion), both for the purpose of illustrating the concepts we are about to introduce and also because of the privileged role which it plays in the stochastic calculus. Once these ideas are firmly in place we shall then be able to devote Chapter 4 to establishing continuous-parameter analogues of the discrete-parameter martingale results of Chapter 2.

3.1 Basic Definitions

We start with several basic definitions which pertain to continuous-parameter stochastic processes taking values in a finite dimensional Euclidean space $\mathbb{R}^d$.

**Definition 3.1.1** Suppose that $(\Omega, \mathcal{F}, P)$ is a given probability space, and $I \subset [0, \infty)$ is an interval being either open, semi-open or closed, and either finite or infinite in length. A collection $\{X_t, t \in I\}$ of mappings $X_t : \Omega \to \mathbb{R}^d$ is called an $\mathbb{R}^d$-valued continuous-parameter stochastic process or, simply, a process when $X_t$ is $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$-measurable for each $t \in I$. If $I$ is closed on the left with left end-point $t = 0$ then a process $\{X_t, t \in I\}$ is called null at the origin when $X_0(\omega) = 0$ for each $\omega \in \Omega$. 

94
Remark 3.1.2 When the continuous-parameter context is clear we shall usually omit the appellation “continuous-parameter” when speaking of such processes. We shall also reserve the term “stochastic process” or “process” for the special case where the $X_t$ are real-valued. When we want to indicate that $X_t$ may take values in some finite-dimensional Euclidean space $\mathbb{R}^d$, $d = 1, 2, \ldots$, then we shall always include the adjective “$\mathbb{R}^d$-valued” and speak of an $\mathbb{R}^d$-valued process.

Remark 3.1.3 Definition 3.1.1 is formulated with reference to some interval $I \subset [0, \infty)$. We shall incur no major loss of generality by taking the special case $I = [0, \infty)$ and henceforth we shall assume that this is so unless otherwise indicated.

Remark 3.1.4 In the context of Definition 3.1.1 we see that, corresponding to each $\omega \in \Omega$, we have a mapping $t \mapsto X_t(\omega) : [0, \infty) \to \mathbb{R}^d$ which is usually called the sample path of the process at the sample point $\omega$. There are questions concerning, for example, the Borel-measurability or continuity of sample paths which have no natural counterpart in the discrete-parameter setting. Such issues make continuous-parameter stochastic processes intrinsically more subtle and interesting than discrete-parameter stochastic processes, and we shall look at some of these in later sections of this chapter.

Remark 3.1.5 For typographical reasons we shall often denote the $\mathbb{R}^d$-valued random vector $X_t$ in Definition 3.1.1 by $X(t)$ and its value $X_t(\omega)$ at any $\omega \in \Omega$ by $X(t, \omega)$.

There are several senses in which we can regard two given stochastic processes on the same probability space as being “equal”:

Definition 3.1.6 Two $\mathbb{R}^d$-valued stochastic processes $\{X_t, t \in [0, \infty)\}$ and $\{Y_t, t \in [0, \infty)\}$, defined on a common probability space $(\Omega, \mathcal{F}, P)$, are indistinguishable when the set $A \Delta \{\omega : X_t(\omega) = Y_t(\omega), \forall t \in [0, \infty)\}$ is a member of $\mathcal{F}$ and $P(A) = 1$.

Remark 3.1.7 Since $\omega \to X_t(\omega)$ and $\omega \to Y_t(\omega)$ are $\mathcal{F}$-measurable mappings, the set $\{X_t = Y_t\}$ is in $\mathcal{F}$ for each $t \in [0, \infty)$. Now $[0, \infty)$ includes uncountably many points, and thus we see that $A = \cap _{t \in [0, \infty)} \{X_t = Y_t\}$ is given by uncountably many intersections of sets in $\mathcal{F}$. Since a $\sigma$-algebra is closed only with respect to countable intersections, there is no guarantee that $A$ is a member of $\mathcal{F}$, and so we must include the requirement that $A \in \mathcal{F}$ as part of the definition of indistinguishability. To express the fact that two processes $\{X_t; t \in [0, \infty)\}$ and $\{Y_t; t \in [0, \infty)\}$ are indistinguishable we shall usually say “we have a.s.: $X_t = Y_t$, $\forall t \in [0, \infty)$”.

Definition 3.1.8 Two $\mathbb{R}^d$-valued stochastic processes $\{X_t, t \in [0, \infty)\}$ and $\{Y_t, t \in [0, \infty)\}$, defined on a common probability space $(\Omega, \mathcal{F}, P)$, are modifications of each other when $P(X_t = Y_t) = 1$ for each $t \in [0, \infty)$.
We see at once that, if \( \{X_t, t \in [0, \infty)\} \) and \( \{Y_t, t \in [0, \infty)\} \) are indistinguishable, then they must be modifications of each other. However, the converse does not necessarily hold true, as the next example illustrates:

**Example 3.1.9** Suppose that \((\Omega, \mathcal{F}, P)\) is a probability space for which \(\Omega = [0, 1]\), \(\mathcal{F} = \mathcal{B}([0, 1])\) and \(P\) denotes Lebesgue measure. Define processes \(\{X_t, t \in [0, 1]\}\) and \(\{Y_t, t \in [0, 1]\}\) on this probability space as follows: for all \(\omega \in \Omega\), \(t \in [0, 1]\), put (i) \(Y_t(\omega) = 0\), and (ii) \(X_t(\omega) = 1\) when \(t = \omega\) and \(X_t(\omega) = 0\) otherwise. Clearly \(\{X_t, t \in [0, 1]\}\) and \(\{Y_t, t \in [0, 1]\}\) are modifications of each other while \(\{\omega : X_t(\omega) = Y_t(\omega) \ \forall t \in [0, 1]\}\) is the empty set, hence the processes cannot be indistinguishable.

**Remark 3.1.10** Although two processes which are modifications of each other need not be indistinguishable, it turns out that if the sample paths of the two processes are regular enough then indistinguishability is indeed a consequence of the fact that one process is a modification of the other. To specify this regularity we use the following terminology: An \(\mathbb{R}^d\)-valued stochastic process \(\{X_t; t \in [0, \infty)\}\) on the probability space \((\Omega, \mathcal{F}, P)\) is called right-continuous when the mapping \(t \to X_t(\omega) : [0, \infty) \to \mathbb{R}^d\) is right-continuous on \([0, \infty)\) for each \(\omega \in \Omega\), left-continuous when \(t \to X_t(\omega) : [0, \infty) \to \mathbb{R}^d\) is left-continuous on \([0, \infty)\) for each \(\omega \in \Omega\), and continuous when \(t \to X_t(\omega) : [0, \infty) \to \mathbb{R}^d\) is continuous on \([0, \infty)\) for each \(\omega \in \Omega\).

**Proposition 3.1.11** Suppose that \(\{X_t, t \in [0, \infty)\}\) and \(\{Y_t, t \in [0, \infty)\}\) are either right-continuous or left-continuous \(\mathbb{R}^d\)-valued stochastic processes defined on a common probability space \((\Omega, \mathcal{F}, P)\). If the processes are modifications of each other then they are indistinguishable.

Proof: We give the proof in the case of right-continuous processes. Let \(\{t_n, n = 1, 2, \ldots\}\) be the set of all rational numbers in \([0, \infty)\), and put

\[
\Omega^* \triangleq \bigcap_{1 \leq n < \infty} \{\omega : X_{t_n}(\omega) = Y_{t_n}(\omega)\}.
\]

Clearly \(\Omega^* \in \mathcal{F}\). Since the rational numbers are dense in \([0, \infty)\) and \(\{X_t; t \in [0, \infty)\}\) and \(\{Y_t; t \in [0, \infty)\}\) are right-continuous, it is easily seen that

\[
\Omega^* = \{\omega : X_t(\omega) = Y_t(\omega), \ \forall t \in [0, \infty)\}.
\]

Now \(\{Y_t; t \in [0, \infty)\}\) is a modification of \(\{X_t; t \in [0, \infty)\}\), and thus \(P[X_{t_n} = Y_{t_n}] = 1\), \(\forall n = 1, 2, \ldots\), whence \(P(\Omega^*) = 1\).

There is a very natural way of thinking about a given \(\mathbb{R}^d\)-valued stochastic process \(\{X_t; t \in [0, \infty)\}\) on the probability space \((\Omega, \mathcal{F}, P)\), namely as a mapping of the two variables \(t\) and \(\omega\) into \(\mathbb{R}^d\):

\[
(t, \omega) \to X(t, \omega) : [0, \infty) \otimes \Omega \to \mathbb{R}^d
\]

96
(here we are using the notation in Remark 3.1.5). There is a Borel $\sigma$-algebra $\mathcal{B}([0, \infty))$ over $[0, \infty)$ and hence the product of sets $[0, \infty) \otimes \Omega$ carries the product $\sigma$-algebra $\mathcal{B}([0, \infty)) \otimes \mathcal{F}$. With this product $\sigma$-algebra in mind we can formulate another desirable property of stochastic processes:

**Definition 3.1.12** An $\mathbb{R}^d$-valued stochastic process $\{X_t; t \in [0, \infty)\}$ on a probability space $(\Omega, \mathcal{F}, P)$ is **jointly measurable** when

$$\{(t, \omega) \in [0, \infty) \otimes \Omega : X(t, \omega) \in \Gamma\} \in (\mathcal{B}([0, \infty)) \otimes \mathcal{F})$$

for each $\Gamma \in \mathcal{B}(\mathbb{R}^d)$.

**Remark 3.1.13** This definition specifies that an $\mathbb{R}^d$-valued stochastic process $\{X_t; t \in [0, \infty)\}$ is jointly measurable when the mapping in (3.1) is $\mathcal{B}([0, \infty)) / \mathcal{B}(\mathbb{R}^d)$-measurable. Joint measurability of a stochastic process is a basic property essential to most of the developments which follow. Among its most useful consequences are the following:

(a) the mapping $t \rightarrow X_t : [0, \infty) \rightarrow \mathbb{R}^d$ is $\mathcal{B}([0, \infty)) / \mathcal{B}(\mathbb{R}^d)$-measurable for each $\omega \in \Omega$ (see Lemma 1.2.38);

(b) when $d = 1$ and $X(t, \omega) \geq 0$ for all $(t, \omega) \in [0, \infty) \otimes \Omega$ then, by the Tonelli Theorem 1.2.40, the mapping $\omega \rightarrow \int_{[0, \infty)} X(t, \omega)dt$ is $\mathcal{F}$-measurable and

$$E \left[ \int_{[0, \infty)} X(t, \omega)dt \right] = \int_{[0, \infty)} E[X_t]dt.$$  

(3.2)

When $X(t, \omega)$ assumes both positive and negative values and

$$E \left[ \int_{[0, \infty)} |X(t, \omega)|dt \right] < \infty$$

then, by the Fubini Theorem 1.2.41, (3.2) continues to hold.

In view of the usefulness of joint measurability, there is an obvious interest in seeking verifiable conditions which imply the joint measurability of a given process. A classical theorem due to Doob essentially asserts that if a process is continuous in probability then it has some modification which is jointly measurable. Despite the importance of this theorem for the basic theory of stochastic processes we shall say no more about it here because, at least as far as stochastic calculus is concerned, joint measurability is usually not quite good enough, and we shall need stronger measurability hypotheses on our processes. We formulate the first of these hypotheses in the next definition, which is an obvious continuous-parameter analogue of Definition 2.1.2:

**Definition 3.1.14** Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space. A collection $\{\mathcal{F}_t, t \in [0, \infty)\}$ of $\sigma$-algebras over $\Omega$ is called a **filtration** in $(\Omega, \mathcal{F}, P)$ when $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $s, t \in [0, \infty)$ such that $s < t$. An $\mathbb{R}^d$-valued process $\{X_t; t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ is said to be **adapted** to the filtration $\{\mathcal{F}_t, t \in [0, \infty)\}$ when the mapping $\omega \rightarrow X_t(\omega) : \Omega \rightarrow \mathbb{R}^d$ is $\mathcal{F}_t / \mathcal{B}(\mathbb{R}^d)$-measurable for each $t \in [0, \infty)$; under these conditions we call the collection of pairs $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ an **adapted process**.
Remark 3.1.15 Analogously with (2.1) in the discrete-parameter case we extend the filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) to \( t = +\infty \) by
\[
\mathcal{F}_\infty \triangleq \sigma \{ \mathcal{F}_t, t \in [0, \infty) \}.
\]
Thus, \( \mathcal{F}_t \subset \mathcal{F}_\infty \subset \mathcal{F}, \forall t \in [0, \infty) \), and \( \mathcal{F}_\infty \) is the smallest \( \sigma \)-algebra which includes \( \mathcal{F}_t \) for each \( t \in [0, \infty) \).

Example 3.1.16 Filtrations naturally arise in the continuous-parameter context as follows: Suppose that \( \{ Y_t; t \in [0, \infty) \} \) is a given \( \mathbb{R}^d \)-valued process on a probability space \((\Omega, \mathcal{F}, P)\), and put
\[
\mathcal{F}_t^Y \triangleq \sigma \{ Y_s, 0 \leq s \leq t \}, \forall t \in [0, \infty).
\]
Then \( \{ \mathcal{F}_t^Y, t \in [0, \infty) \} \) is a filtration in \((\Omega, \mathcal{F}, P)\) and \( \{ (Y_t, \mathcal{F}_t^Y); t \in [0, \infty) \} \) is an adapted process. The filtration defined by (3.4) is called the raw or natural filtration of the process \( \{ Y_t; t \in [0, \infty) \} \). As in Example 2.1.4, the raw filtration has the following property: Suppose that \( \{ X_t; t \in [0, \infty) \} \) is another stochastic process on \((\Omega, \mathcal{F}, P)\) which is adapted to \( \{ \mathcal{F}_t^Y, t \in [0, \infty) \} \). Then, for each \( t \in [0, \infty) \), \( X_t \) is measurable with respect to \( \mathcal{F}_t^Y \triangleq \sigma \{ Y_s, 0 \leq s \leq t \} \), and in view of Theorem 1.3.22 there is some sequence \( \{ t_0, t_0, t_2, \ldots \} \subset [0, t] \), and some measurable mapping \( \Psi_t : (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R})^\infty) \to \mathbb{R} \), such that
\[
X_t(\omega) = \Psi_t(Y_{t_0}(\omega), Y_{t_1}(\omega), Y_{t_2}(\omega), \ldots), \forall \omega \in \Omega.
\]
Consequently, for each \( \omega \in \Omega \), \( X_t(\omega) \) is completely determined by the restriction of the sample-path \( s \to Y_s(\omega) \) to the interval \([0, t]\). In fact, \( X_t(\omega) \) is determined by \( \{ Y_{t_0}(\omega), Y_{t_1}(\omega), Y_{t_2}(\omega), \ldots \} \) for some sequence \( \{ t_0, t_1, t_2, \ldots \} \subset [0, t] \).

Remark 3.1.17 There are often situations in which we have an adapted process together with some modification of the adapted process, and we want to assert that the modification is also adapted with respect to the same filtration. In order to do so we must impose a certain condition on the filtration in question. To formulate this condition suppose that \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\), and put
\[
\mathcal{N} \triangleq \{ N \in \mathcal{F} : P(N) = 0 \}
\]
We say that \( \mathcal{F}_0 \) includes all \textbf{P-null events in} \( \mathcal{F} \) when \( \mathcal{N} \subset \mathcal{F}_0 \). Under this condition it follows that \( \mathcal{N} \subset \mathcal{F}_t \) for each \( t \in [0, \infty) \).

Proposition 3.1.18 Suppose that \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is an \( \mathbb{R}^d \)-valued adapted process on the probability space \((\Omega, \mathcal{F}, P)\), and \( \{ Y_t; t \in [0, \infty) \} \) is another \( \mathbb{R}^d \)-valued process, likewise defined on \((\Omega, \mathcal{F}, P)\), which is a modification of \( \{ X_t; t \in [0, \infty) \} \). If \( \mathcal{F}_0 \) includes all \textbf{P-null events in} \( \mathcal{F} \) then \( \{ (Y_t, \mathcal{F}_t); t \in [0, \infty) \} \) is an adapted process.
Proof: We must show that $\omega \to Y_t(\omega) : \Omega \to \mathbb{R}^d$ is $\mathcal{F}_t/\mathcal{B}(\mathbb{R}^d)$-measurable for each $t \in [0, \infty)$. Fix some arbitrary $t \in [0, \infty)$ and observe that $\{X_t \neq Y_t\} \in \mathcal{F}$, $P(X_t \neq Y_t) = 0$. Thus, $\{X_t \neq Y_t\} \in \mathcal{N} \subset \mathcal{F}_t$, hence also $\{X_t = Y_t\} = \{X_t \neq Y_t\}^c \in \mathcal{F}_t$. Now fix some arbitrary $\Gamma \in \mathcal{B}(\mathbb{R}^d)$. Then

(3.6) $\{Y_t \in \Gamma\} = \{\{Y_t \in \Gamma\} \cap \{X_t = Y_t\}\} \cup \{\{Y_t \in \Gamma\} \cap \{X_t \neq Y_t\}\}$

Since $\{X_t = Y_t\} \in \mathcal{F}_t$ we see that

(3.7) $\{Y_t \in \Gamma\} \cap \{X_t = Y_t\} = \{X_t \in \Gamma\} \cap \{X_t = Y_t\} \in \mathcal{F}_t$,

and, since $P(X_t \neq Y_t) = 0$ we also have

(3.8) $\{Y_t \in \Gamma\} \cap \{X_t \neq Y_t\} \in \mathcal{N} \subset \mathcal{F}_0 \subset \mathcal{F}_t$.

From (3.6), (3.7) and (3.8) we obtain $\{Y_t \in \Gamma\} \in \mathcal{F}_t$. $\blacksquare$

**Proposition 3.1.19** Suppose that the following conditions hold:

(a) $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a filtration in the probability space $(\Omega, \mathcal{F}, P)$, and $\mathcal{F}_0$ includes all $P$-null events in $\mathcal{F}$ (recall Remark 3.1.17);

(b) $\{(X_n(t), \mathcal{F}_t); t \in [0, \infty)\}$ is an adapted $\mathbb{R}$-valued process on $(\Omega, \mathcal{F}, P)$ for each $n = 0, 1, 2, \ldots$;

(c) for each $t \in [0, \infty)$ the sequence of random variables $\{X_n(t), n = 0, 1, 2, \ldots\}$ converges in probability to a random variable $X(t)$ as $n \to \infty$.

Then $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is an $\mathbb{R}$-valued adapted processs.

Proof: Fix some arbitrary $t \in [0, \infty)$. By Theorem 1.2.11(b), there is some sub-sequence $\{X_{nk}(t), k = 0, 1, 2, \ldots\}$ of the sequence $\{X_n(t), n = 0, 1, 2, \ldots\}$ and some $N_t^1 \in \mathcal{F}$ such that $P(N_t^1) = 0$ and

(3.9) $\lim_{k \to \infty} X_{nk}(t, \omega) = X(t, \omega), \forall \omega \notin N_t^1$.

Now put

(3.10) $X^*(t, \omega) \overset{\Delta}{=} \limsup_{k \to \infty} X_{nk}(t, \omega), \forall \omega \in \Omega$.

Since $\omega \to X_{nk}(t, \omega)$ is $\mathcal{F}_t$-measurable for each $k$, we see that $\omega \to X^*(t, \omega)$ is $\mathcal{F}_t$-measurable (recall Proposition 1.2.6), hence, by the arbitrary choice of $t \in [0, \infty)$, it follows that $\{(X^*(t), \mathcal{F}_t), t \in [0, \infty)\}$ is an adapted process. Now let

(3.11) $N_t^2 \overset{\Delta}{=} \{\omega : X(t, \omega) \neq X^*(t, \omega)\}$.

Clearly $N_t^2 \in \mathcal{F}$, and by (3.9) we see that $N_t^2 \subset N_t^1$ hence $P(N_t^2) = 0$. It follows, again from the arbitrary choice of $t \in [0, \infty)$, that $\{X(t), t \in [0, \infty)\}$ is a modification of $\{X^*(t), t \in [0, \infty)\}$. Now Proposition 3.1.18 together with (a) shows that $\{(X(t), \mathcal{F}_t), t \in [0, \infty)\}$ is an adapted process. $\blacksquare$
Remark 3.1.20 Frequently one is given some adapted process \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\), and one wants to be sure that \( \{(Y_t, \mathcal{F}_t); t \in [0, \infty)\} \) is also an adapted process when \( Y_t \) is defined on \( \Omega \) by
\[
Y_t(\omega) \triangleq \int_0^t X_s(\omega)ds, \quad \forall \ t \in [0, \infty), \quad \forall \ \omega \in \Omega.
\]
In order to use the Fubini-Tonelli Theorems 1.2.40 and 1.2.41 to conclude that \( Y_t \) is \( \mathcal{F}_t \)-measurable we must clearly insist, in addition to the obvious non-negativity or integrability conditions required by Theorems 1.2.40 and 1.2.41, that the mapping
\[
(s, \omega) \to X(s, \omega) : ([0, t] \otimes \Omega, \mathcal{B}([0, t]) \otimes \mathcal{F}_t) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))
\]
be measurable for each \( t \in [0, \infty) \). We are thus led to formulate another measurability concept for stochastic processes:

Definition 3.1.21 An \( \mathbb{R}^d \)-valued stochastic process \( \{X_t; t \in [0, \infty)\} \), on the probability space \((\Omega, \mathcal{F}, P)\), is \textbf{progressively measurable} with respect to a given filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \) in \((\Omega, \mathcal{F}, P)\) when, for each \( t \in [0, \infty) \) and \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \), we have
\[
\{(s, \omega) \in [0, t] \otimes \Omega : X(s, \omega) \in \Gamma\} \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t
\]
We shall usually say that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is progressively measurable to indicate that \( \{X_t; t \in [0, \infty)\} \) is progressively measurable with respect to \( \{\mathcal{F}_t, t \in [0, \infty)\} \). Progressive measurability will soon be seen to be an essential notion in the stochastic calculus, for essentially the reason outlined in Remark 3.1.20.

Progressive measurability can also be characterized in terms of measurability with respect to a particular \( \sigma \)-algebra over the product space \([0, \infty) \otimes \Omega\):

Proposition 3.1.22 Suppose that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\). Define
\[
PM\{\mathcal{F}_t\} \triangleq \{A \subset [0, \infty) \otimes \Omega : \{(I_A(t), \mathcal{F}_t); t \in [0, \infty)\} \text{ is progressively measurable}\}.
\]
Then \( PM\{\mathcal{F}_t\} \) is a \( \sigma \)-algebra over \([0, \infty) \otimes \Omega\) with \( PM\{\mathcal{F}_t\} \subset \mathcal{B}([0, \infty)) \otimes \mathcal{F} \). Moreover, if \( \{X_t; t \in [0, \infty)\} \) is a given \( \mathbb{R}^d \)-valued process on \((\Omega, \mathcal{F}, P)\) then \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is progressively measurable if and only if the mapping
\[
(t, \omega) \to X(t, \omega) : [0, \infty) \otimes \Omega \to \mathbb{R}^d
\]
is \( PM\{\mathcal{F}_t\} \)-measurable.

Proof: The proof is a simple application of Theorem 1.5.4 and is left to the reader.

Remark 3.1.23 The \( \sigma \)-algebra \( PM\{\mathcal{F}_t\} \) in (3.13) is called the \textbf{progressively measurable} \( \sigma \)-algebra of the filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \).
Remark 3.1.24 If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is progressively measurable then Proposition 3.1.22 shows that \( \{X_t, t \in [0, \infty]\} \) is jointly measurable, and Definition 3.1.21 along with Lemma 1.2.38 obviously ensures that \( \{X_t, t \in [0, \infty)\} \) is adapted to the filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \). Thus, a progressively measurable process is both jointly measurable and adapted, and one is naturally led to ask if the converse assertion is true, namely: is a jointly measurable and adapted process necessarily progressively measurable? There are in fact subtle examples of jointly measurable adapted processes which fail to be progressively measurable. Such counterexamples typically involve ideas which are well beyond the level of this basic introduction, and hence we refrain from considering any of them here (an example may be found on page 62 of Chung and Williams [5]). Moreover, a profound theorem of Chung and Doob [4] asserts a partial converse, namely, a jointly measurable adapted process always has a progressively measurable modification. In this sense the set of progressively measurable processes may be regarded as “nearly” as large as the set of jointly measurable adapted processes. We shall never have occasion to use the theorem of Chung and Doob, and hence we do not give a precise statement of it here. For our purposes the following very easy proposition will always be sufficient to ensure progressive measurability:

Proposition 3.1.25 Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted \( \mathbb{R}^d \)-valued process defined on the probability space \((\Omega, \mathcal{F}, P)\). If \( \{X_t; t \in [0, \infty)\} \) is either right-continuous or left-continuous then it is progressively measurable with respect to \( \{\mathcal{F}_t, t \in [0, \infty)\} \).

Proof: Suppose that \( \{X_t; t \in [0, \infty)\} \) is a right-continuous process, and fix some arbitrary \( t \in [0, \infty) \). We must show that the map

\[
(s, \omega) \rightarrow X_s(\omega) : [0, t] \otimes \Omega \rightarrow \mathbb{R}^d
\]

is \( \mathcal{B}([0, t]) \otimes \mathcal{F}_t/\mathcal{B}(\mathbb{R}^d) \)-measurable. For each \( n = 1, 2, \ldots \) define \( X_0^{(n)}(\omega) \overset{\Delta}{=} X_0(\omega) \) and

\[
X_s^{(n)}(\omega) \overset{\Delta}{=} X_{t(k+1)2^{-n}t}(\omega)
\]

\( \forall \omega \in \Omega, \forall s \in (k2^{-n}t, (k+1)2^{-n}t] \), with \( k = 0, 1, 2, \ldots 2^n - 1 \). It is easily seen from the definition of product \( \sigma \)-algebra (see § 1.2.7) that the maps

\[
(s, \omega) \rightarrow X_s^{(n)}(\omega) : [0, t] \otimes \Omega \rightarrow \mathbb{R}^d
\]

are \( \mathcal{B}([0, t]) \otimes \mathcal{F}_t/\mathcal{B}(\mathbb{R}^d) \)-measurable for all \( n = 1, 2, \ldots \). But, since \( \{X_t; t \in [0, \infty)\} \) is right-continuous, we have that

\[
\lim_{n \to \infty} X_s^{(n)}(\omega) = X_s(\omega), \quad \forall \omega \in \Omega, \forall s \in [0, t].
\]

The proposition is now a consequence of Proposition 1.2.6. The proof for left-continuous processes is similar.

We end this section on basic definitions by formulating notions of left- and right-continuity for a
given filtration. For motivation we first recall some elementary facts from real analysis. Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a given non decreasing function, namely \( \varphi(s) \leq \varphi(t) \), \( \forall \, s, t \in [0, \infty) \) with \( s \leq t \), and define the number
\[
\varphi(t+) \triangleq \inf_{s<t} \varphi(s)
\]
for each \( t \in [0, \infty) \). Likewise, for each \( t \in (0, \infty) \) define the number
\[
\varphi(t-) \triangleq \sup_{0<s<t} \varphi(s). \]
Since the function \( \varphi(\cdot) \) is non-decreasing it is clear that, if \( \varphi(t+) = \varphi(t) \), \( \forall \, t \in [0, \infty) \), then \( \varphi(\cdot) \) is a right-continuous mapping. Likewise, if \( \varphi(t-) = \varphi(t) \), \( \forall \, t \in (0, \infty) \), then \( \varphi(\cdot) \) is a left-continuous mapping. A filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is analogous to a non-decreasing function \( \varphi(\cdot) \) on \([0, \infty)\) in the sense that the \( \mathcal{F}_t \) are non-decreasing with increasing \( t \). Motivated by the preceding considerations, we now define left- and right-continuity in the context of filtrations:

**Definition 3.1.26** Suppose that \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\).

For each \( t \in [0, \infty) \) put
\[
\mathcal{F}_{t+} \triangleq \bigcap_{s>t} \mathcal{F}_s,
\]
and for each \( t \in (0, \infty) \) put
\[
\mathcal{F}_{t-} \triangleq \sigma\{\mathcal{F}_s, \ 0 \leq s < t\}.
\]
The filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is called **right-continuous** when \( \mathcal{F}_{t+} = \mathcal{F}_t, \ \forall \, t \in [0, \infty) \), and is called **left-continuous** when \( \mathcal{F}_{t-} = \mathcal{F}_t, \ \forall \, t \in (0, \infty) \).

**Remark 3.1.27** The statement that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued right-continuous [left-continuous, continuous] adapted process, will always mean that the sample paths \( t \to X_t(\omega) : [0, \infty) \to \mathbb{R}^d \) are right-continuous [left-continuous, continuous] for each \( \omega \in \Omega \), and the collection of pairs \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted process (see Definition 3.1.14). Notice that this terminology implies nothing about left or right-continuity of \( \{ \mathcal{F}_t, t \in [0, \infty) \} \). If the filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is indeed left or right-continuous then this will be specifically and separately stated in each case.

**Remark 3.1.28** We see from (3.14) that \( \varphi(t+) \) is defined to be the greatest number which is less than or equal to \( \psi(s) \) for each \( s \in (t, \infty) \). In the same way, \( \mathcal{F}_{t+} \) is defined in (3.16) to be the largest \( \sigma \)-algebra over \( \Omega \) which is included within \( \mathcal{F}_s \) for each \( s \in (t, \infty) \). Similarly, whereas \( \varphi(t-) \) is defined in (3.15) to be the least number which is greater than or equal to \( \varphi(s) \) for each \( s \in [0, t) \), so \( \mathcal{F}_{t-} \) is defined in (3.17) as the smallest \( \sigma \)-algebra over \( \Omega \) which includes \( \mathcal{F}_s \) for each \( s \in [0, t) \).

**Remark 3.1.29** Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a given non-decreasing function. Then the following statements follow easily from elementary real analysis: (i) \( \varphi(t-) \leq \varphi(t) \leq \varphi(t+) \) for all \( t \in (0, \infty) \), and the second inequality holds at \( t = 0 \) as well; (ii) if \( t \in [0, \infty) \), and \( \{t_n, \ n = 1,2,\ldots\} \) is a sequence
in \((t, \infty)\) such that \(t < t_{n+1} \leq t_n\) and \(\lim_{n \to \infty} t_n = t\), then \(\varphi(t+) = \inf_{1 \leq n < \infty} \varphi_{t_n}\); (iii) if \(t \in (0, \infty)\), and \(\{t_n, n = 1, 2\ldots\}\) is a sequence in \([0, t)\) such that \(0 \leq t_n \leq t_{n+1} < t\) and \(\lim_{n \to \infty} t_n = t\), then \(\varphi(t-) = \sup_{1 \leq n < \infty} \varphi_{t_n}\); (iv) if \(\psi(t) \overset{\Delta}{=} \varphi(t+)\) for all \(t \in [0, \infty)\) then \(\psi(\cdot)\) is right-continuous. It is just as easy to establish the following analogous statements for filtrations, the elementary proofs being left to the reader:

**Proposition 3.1.30** Suppose that \(\{F_t, t \in [0, \infty)\}\) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\). Then: (i) \(F_{t-} \subset F_t \subset F_{t+} \subset F\) for all \(t \in (0, \infty)\), and the second and third of the set inclusions holds at \(t = 0\) as well;

(ii) if \(t \in [0, \infty)\), and \(\{t_n, n = 1, 2\ldots\}\) is a sequence in \((t, \infty)\) such that \(t < t_{n+1} \leq t_n\) and \(\lim_{n \to \infty} t_n = t\), then

\[
F_{t+} = \bigcap_{1 \leq n < \infty} F_{t_n}.
\]

(iii) if \(t \in (0, \infty)\), and \(\{t_n, n = 1, 2\ldots\}\) is a sequence in \([0, t)\) such that \(0 \leq t_n \leq t_{n+1} < t\) and \(\lim_{n \to \infty} t_n = t\), then

\[
F_{t-} = \sigma\{F_{t_n}, n = 1, 2, \ldots\}.
\]

(iv) if \(G_t \overset{\Delta}{=} F_{t+}\) for all \(t \in [0, \infty)\) then \(\{G_t, t \in [0, \infty)\}\) is a right-continuous filtration (i.e. \(G_t = G_{t+}\) for each \(t \in [0, \infty)\)).

**Remark 3.1.31** The filtration \(\{G_t, t \in [0, \infty)\}\) in Proposition 3.1.30(iv) is called the **right-continuous enlargement** of \(\{F_t, t \in [0, \infty)\}\) and is denoted by \(\{F_{t+}, t \in [0, \infty)\}\).

### 3.2 The Wiener Process

In this section we introduce what is undoubtedly the most interesting and important example of a continuous-parameter stochastic process, namely the celebrated Wiener process. The Wiener process is ubiquitous in pure and applied probability, on account of its subtle mathematical properties and the fact that it is an essential building-block in numerous models for random phenomena in science, engineering and economics. Our goal is to provide an introduction to only the most basic aspects of the Wiener process which are pertinent to the special role that it plays in stochastic calculus.

Before giving a formal definition of the Wiener process we formulate a general property possessed by many stochastic processes:

**Definition 3.2.1** An \(\mathbb{R}^d\)-valued stochastic process \(\{X(t); t \in [0, \infty)\}\) on a probability space \((\Omega, \mathcal{F}, P)\) is said to have **independent increments** when, for each finite set of indices \(0 = t_0 < t_1 < \ldots < t_n < \infty\), the set of random variables \(\{X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})\}\) is independent.

The independent increments property of a stochastic process \(\{X(t); t \in [0, \infty)\}\) basically asserts that the increment \(X(t) - X(s)\) is not influenced in any way by the increments of the process outside the
interval \([s, t]\). The next definition sets forth the Wiener process as an independent increments process with certain other distinguishing properties as well:

**Definition 3.2.2** An \(\mathbb{R}\)-valued stochastic process \(\{W(t); t \in [0, \infty)\}\) on a probability space \((\Omega, \mathcal{F}, P)\) is a **scalar standard Wiener process** when

(i) \(\{W(t); t \in [0, \infty)\}\) is null at the origin;

(ii) the sample-paths \(t \to W(t, \omega) : [0, \infty) \to \mathbb{R}\) are continuous for each \(\omega \in \Omega\);

(iii) for each \(s, t \in [0, \infty)\) with \(s < t\), we have \(W(t) - W(s) \sim N(0, t - s)\);

(iv) \(\{W(t); t \in [0, \infty)\}\) has independent increments.

**Remark 3.2.3** Definition 3.2.2 is motivated by our desire to precisely formulate a mathematical model for some of the basic attributes of a wide variety of physical phenomena. We make no attempt to look at modeling issues here, and only observe that the Wiener process is basic to the construction of mathematical models of random phenomena in an amazing range of applications. To mention but a few, these include quantum mechanics, stochastic control, communication systems, financial economics, asymptotic analysis of queues and inventory systems, large-sample statistics, and hydrology and water resource management. The reason for this broad applicability seems to be the ability of the Wiener process to model with reasonable accuracy the main features inherent in natural random phenomena within a rather simple mathematical framework. The reader who wishes to pursue these issues is referred to Nelson [23] and the collection of papers in Wax [30].

**Remark 3.2.4** There is as yet no guarantee that a stochastic process with the properties postulated in Definition 3.2.2 even exists. For example, these properties could be incompatible in the sense that it is impossible for a process to have all of them at once. To clear away doubts of this kind and assure ourselves that we are not dealing with a meaningless concept, we must explicitly construct *some* process \(\{W(t); t \in [0, \infty)\}\) on *some* probability space \((\Omega, \mathcal{F}, P)\) which demonstrably has all of the properties set forth in Definition 3.2.2. Despite the clear importance of the existence issue, we shall not construct a Wiener process here. Indeed, we are never really disadvantaged in studying the role of the Wiener process in stochastic calculus if we merely take the properties in Definition 3.2.2 at face value and uncritically accept that a Wiener process exists. That said, we nevertheless caution the reader that anyone with a serious interest in probability and stochastic processes ought to have a thorough understanding of how to construct a Wiener process, if only for “cultural” reasons. Detailed constructions can be found, for example, in Bhattacharya and Waymire ([2], Chap. 1), Durrett ([10], Chap. 7), and Karatzas and Shreve ([19], Chap. 2).

The notion of a Wiener process formulated in Definition 3.2.2 is amply motivated on physical grounds and completely satisfactory for the requirements of elementary and classical probability theory. From the point of view of stochastic calculus however, it has one significant disadvantage, namely it includes no reference to a *filtration* in \((\Omega, \mathcal{F}, P)\). We are now going to formulate an equivalent definition of
a Wiener process but stated in terms of an underlying filtration. For this purpose we need the next
proposition, which will be useful elsewhere in this chapter as well:

**Proposition 3.2.5** Suppose that \( \{X(t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued process on a probability space
\((\Omega, \mathcal{F}, P)\), and define

\[
\mathcal{F}_t^X \triangleq \sigma\{X(u), u \in [0, t]\}, \quad \mathcal{H}_t^X \triangleq \sigma\{X(v) - X(t), v \in [t, \infty)\}, \forall t \in [0, \infty).
\]

Then \( \{X(t); t \in [0, \infty)\} \) has independent increments if and only if \( \mathcal{F}_t^X \) and \( \mathcal{H}_t^X \) are independent for
each \( t \in [0, \infty) \).

**Proof:** Suppose \( \{X(t); t \in [0, \infty)\} \) has independent increments, and fix some \( t \in [0, \infty) \). Then, for
arbitrary points \( u_i \in [0, t], i = 0, 1, \ldots, n, \) and \( v_j \in (t, \infty), j = 1, 2, \ldots, k, \) with \( 0 = u_0 < u_1 < u_2 < \ldots < u_n \) and \( t \leq v_1 < v_2 < \ldots < v_k \), we see that the \( \sigma \)-algebras \( \sigma\{X(u_0), X(u_1) - X(u_0), \ldots, X(u_n) - X(u_{n-1})\} \) and \( \sigma\{X(v_1) - X(t), X(v_2) - X(v_1), \ldots, X(v_k) - X(v_{k-1})\} \) are \( P \)-independent. Since, trivially,
one has

\[
\sigma\{X(u_0), X(u_1), X(u_2), \ldots, X(u_n)\} = \sigma\{X(u_0), X(u_1) - X(u_0), \ldots, X(u_n) - X(u_{n-1})\},
\]

and

\[
\sigma\{X(v_1) - X(t), X(v_2) - X(t), \ldots, X(v_k) - X(t)\} = \sigma\{X(v_1) - X(t), X(v_2) - X(v_1), \ldots, X(v_k) - X(v_{k-1})\},
\]

we see that \( \sigma \)-algebras \( \sigma\{X(u_0), X(u_1), X(u_2), \ldots, X(u_n)\} \) and \( \sigma\{X(v_1) - X(t), X(v_2) - X(t), \ldots, X(v_k) - X(t)\} \) are \( P \)-independent. Now let \( \mathcal{D} \) be the collection of events \( A \in \mathcal{F} \), each of which can be written in
the special form

\[
A = \bigcap_{r=0}^{n} \{X(u_r) \in \Gamma_r\},
\]

for some nonnegative integer \( n \), some \( \Gamma_r \in \mathcal{B}(\mathbb{R}^d) \), and some \( u_r \in [0, t], r = 0, 1, \ldots, n, \) with \( 0 = u_0 < u_1 < \ldots u_n \leq t \) (events having this particular form are usually called **cylinder sets**). It is clear that
\( \mathcal{D} \) is a \( \pi \)-class over \( \Omega \), with

\[
\sigma\{\mathcal{D}\} = \mathcal{F}_t^X.
\]

Similarly, let \( \mathcal{E} \) be the collection of events \( B \in \mathcal{F} \), each of which can be written as

\[
B = \bigcap_{i=1}^{k} \{X(v_i) - X(t) \in \Delta_i\}
\]

for some positive integer \( k \), some \( \Delta_i \in \mathcal{B}(\mathbb{R}^d) \), and some \( v_i \in (t, \infty), i = 1, 2, \ldots k, \) with \( t < v_1 < v_2 < \ldots v_k \). Again, it is clear that \( \mathcal{E} \) is a \( \pi \)-class over \( \Omega \), with

\[
\sigma\{\mathcal{E}\} = \mathcal{H}_t^X.
\]
Now fix an arbitrary \( A \in \mathcal{D} \) with the typical form given by (3.19), and fix arbitrary \( B \in \mathcal{E} \) with the typical form given by (3.21). Clearly, \( A \) is a member of the \( \sigma \)-algebra \( \sigma\{X(u_0), X(u_1), X(u_2), \ldots X(u_n)\} \), and \( B \) is a member of the \( \sigma \)-algebra \( \sigma\{X(v_1) - X(t), X(v_2) - X(t), \ldots X(v_k) - X(t)\} \). Since we have observed that these \( \sigma \)-algebras are \( P \)-independent, it follows that \( A \) and \( B \) are \( P \)-independent, and hence, by the arbitrary choices of \( A \in \mathcal{D} \) and \( B \in \mathcal{E} \), it follows that the \( \pi \)-classes \( \mathcal{D} \) and \( \mathcal{E} \) are \( P \)-independent. Then, by Lemma 1.5.5, we see that \( \mathcal{D} \) and \( \sigma\{\mathcal{E}\} \) are \( P \)-independent (using the fact that \( \mathcal{E} \) is a \( \pi \)-class), and then, by a second application of Lemma 1.5.5, we see that \( \sigma\{\mathcal{D}\} \) and \( \sigma\{\mathcal{E}\} \) are \( P \)-independent (using the fact that \( \mathcal{D} \) is a \( \pi \)-class). In view of this fact and (3.20), (3.22), it follows that \( \mathcal{F}_t^X \) and \( \mathcal{H}_t^X \) are \( P \)-independent, as required.

The converse assertion, namely that independence of \( \mathcal{F}_t^X \) and \( \mathcal{H}_t^X \) for every \( t \in [0, \infty) \) implies that \( \{X(t); t \in [0, \infty)\} \) has independent increments, is a trivial observation.

Using this proposition we can re-write Definition 3.2.2 in the following equivalent form:

**Definition 3.2.6** An \( \mathbb{R} \)-valued stochastic process \( \{W(t); t \in [0, \infty)\} \) on a probability space \((\Omega, \mathcal{F}, P)\) is a **scalar standard Wiener process** when (i), (ii) and (iii) of Definition 3.2.2 hold, and (iv) for each \( s, t \in [0, \infty) \) with \( s < t \), the increment \( W(t) - W(s) \) is independent of \( \mathcal{F}_s^W \triangleq \sigma\{W(u), u \in [0, s]\} \).

**Remark 3.2.7** We see at once that, if \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process, then \( EW(t) = 0, \forall t \in [0, \infty) \), and hence

\[
\text{Cov}(W(s), W(t)) = E[W(s)W(t)] = s \wedge t, \quad \forall s, t \in [0, \infty).
\]

It is easy to extend the concept of a Wiener process to a multidimensional setting. If \( \{W(t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued process on some probability space \((\Omega, \mathcal{F}, P)\), we shall expand the random vector \( W(t) \) into \( d \)-scalar random variables, namely

\[
(3.23) \quad W(t) = (W^1(t), W^2(t), \ldots W^d(t)) \quad \forall t \in [0, \infty),
\]

so that we have stochastic processes \( \{W^k(t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\), and put

\[
(3.24) \quad \mathcal{F}_\infty^W \triangleq \sigma\{W^k(t), t \in [0, \infty)\}, \quad \forall k = 1, 2, \ldots d.
\]

**Definition 3.2.8** An \( \mathbb{R}^d \)-valued stochastic process \( \{W(t); t \in [0, \infty)\} \) on a probability space \((\Omega, \mathcal{F}, P)\) is a **\( d \)-dimensional standard Wiener process** when

(i) \( \{W^k(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on \((\Omega, \mathcal{F}, P)\), for each \( k = 1, 2, \ldots d \);

(ii) the collection of \( \sigma \)-algebras \( \{\mathcal{F}_\infty^W, \mathcal{F}_\infty^{W^2}, \ldots \mathcal{F}_\infty^{W^d}\} \) is independent.

**Remark 3.2.9** Suppose that \( \{W(t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process. In view of Remark 3.2.7 we see that

\[
\text{Cov}(W(s), W(t)) = (s \wedge t)I_d \quad \text{and} \quad W(t) - W(s) \sim N(0, |t - s|I_d)
\]
for each \( s, t \in [0, \infty) \), where \( I_d \) denotes the \( d \) by \( d \) unit matrix. It is then an easy application of basic probability theory to verify that an alternative, but entirely equivalent, definition of a multidimensional Wiener process is as follows:

**Definition 3.2.10** An \( \mathbb{R}^d \)-valued stochastic process \( \{ W(t); t \in [0, \infty) \} \) on a probability space \((\Omega, \mathcal{F}, P)\) is a \( d \)-dimensional standard Wiener process when

1. \( \{ W(t); t \in [0, \infty) \} \) is null at the origin;
2. the sample-paths \( t \to W(t, \omega) : [0, \infty) \to \mathbb{R}^d \) are continuous for each \( \omega \in \Omega \);
3. for each \( s, t \in [0, \infty) \), with \( s < t \), the \( \mathbb{R}^d \)-valued increment \( W(t) - W(s) \) is distributed according to \( W(t) - W(s) \sim N(0, |t - s|I_d) \), and independent of \( \mathcal{F}_s^W \triangleq \sigma\{ W(u), u \in [0,s] \} \).

There are problems in stochastic calculus, particularly in the realm of stochastic differential equations, where significant advantages result if we have available a slightly generalized concept of the Wiener process in which there is some flexibility in the choice of the underlying filtration and we are not limited only to the raw filtration, as in Definition 3.2.10(iii). Motivated by these considerations we present a variant of the preceding definition of a Wiener process, which involves not just a stochastic process but an underlying filtration as well:

**Definition 3.2.11** Given an \( \mathbb{R}^d \)-valued stochastic process \( \{ W(t); t \in [0, \infty) \} \) and a filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) on a probability space \((\Omega, \mathcal{F}, P)\), the indexed family \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process when

1. \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted process and null at the origin;
2. the sample-paths \( t \to W(t, \omega) : [0, \infty) \to \mathbb{R}^d \) are continuous for each \( \omega \in \Omega \);
3. for each \( s, t \in [0, \infty) \), with \( s < t \), the \( \mathbb{R}^d \)-valued increment \( W(t) - W(s) \) is distributed according to \( W(t) - W(s) \sim N(0, |t - s|I_d) \), and independent of \( \mathcal{F}_s \).

**Remark 3.2.12** It is clear that there is the possibility for some confusion, since we are now using the term “Wiener process” for the two distinct, although closely related, concepts set forth in Definitions 3.2.10 and 3.2.11. Henceforth, whenever we say that \( \{ W(t); t \in [0, \infty) \} \) is a \( d \)-dimensional standard Wiener process we shall mean that \( \{ W(t); t \in [0, \infty) \} \) satisfies the conditions of Definition 3.2.10 (or, equivalently, of Definition 3.2.8), and when we say that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process we shall mean that the filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) and process \( \{ W(t); t \in [0, \infty) \} \) together satisfy the conditions of Definition 3.2.11.

**Remark 3.2.13** Clearly any \( \mathbb{R}^d \)-valued stochastic process \( \{ W(t); t \in [0, \infty) \} \) is adapted to a filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) if and only if \( \mathcal{F}_t^W \subset \mathcal{F}_t \) for each \( t \in [0, \infty) \). It follows that, if \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process in the sense of Definition 3.2.11, then \( W(t) - W(s) \) is independent of \( \mathcal{F}_s \), hence independent of \( \mathcal{F}_s^W \subset \mathcal{F}_s \), for each \( s, t \in [0, \infty) \) with \( s < t \), so that the process \( \{ W(t); t \in [0, \infty) \} \) has all the properties set forth in the equivalent Definitions 3.2.8 and 3.2.10.
Remark 3.2.14 If \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process on \( (\Omega, \mathcal{F}, P) \) then, clearly, each \( \{(W^k(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process on \( (\Omega, \mathcal{F}, P) \).

Remark 3.2.15 Suppose \( T \in [0, \infty) \) is a constant. We shall say that \( \{(W(t), \mathcal{F}_t); t \in [0, T]\} \) is a \( d \)-dimensional standard Wiener process over the interval \( [0, T] \) when \( \{W(t); t \in [0, T]\} \) is an \( \mathbb{R}^d \)-valued process on some probability space \( (\Omega, \mathcal{F}, P) \), \( \{\mathcal{F}_t, t \in [0, T]\} \) is a filtration in \( (\Omega, \mathcal{F}, P) \) (i.e. \( \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}, \forall s, t \in [0, T] \) with \( s < t \)), and all conditions in Definition 3.2.11 hold with the obvious restriction of \( s, t \) to interval \( [0, T] \).

3.3 Continuous-Parameter Stopping Times

In this section our goal is to formulate the notion of a stopping time in a continuous-parameter context, with a view to establishing in Chapter 4 a continuous-parameter analogue of the optional sampling Theorem 2.3.9. Continuous-parameter stopping times and the continuous-parameter version of Theorem 2.3.9 are crucial tools of analysis in the advanced theory of stochastic processes. For our purposes they are central to the idea of localization, which will be needed when we develop the concept of stochastic integration in Chapter 5.

Definition 3.3.1 Suppose that \( (\Omega, \mathcal{F}, P) \) is some probability space. A \( \mathcal{F} \)-measurable mapping \( T : \Omega \rightarrow [0, \infty] \) is called a continuous-parameter random time. If \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a given filtration in \( (\Omega, \mathcal{F}, P) \), then a continuous-parameter random time \( T \) is called a continuous-parameter stopping time with respect to \( \{\mathcal{F}_t, t \in [0, \infty)\} \) when \( \{T \leq t\} \in \mathcal{F}_t \) for all \( t \in [0, \infty) \). Finally, a continuous-parameter random time \( T \) is called a continuous parameter optional time with respect to \( \{\mathcal{F}_t, t \in [0, \infty)\} \) when \( \{T < t\} \in \mathcal{F}_t \) for all \( t \in (0, \infty) \).

To abbreviate the terminology we shall often say that a random time \( T \) is a \( \{\mathcal{F}_t\} \)-stopping time to mean that \( T \) is a stopping time with respect to the filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \). A similar abbreviation of terminology is understood when we say that \( T \) is a \( \{\mathcal{F}_t\} \)-optional time. Also, we shall omit the apellation “continuous-parameter” when this is clear from the context.

Remark 3.3.2 If \( T \) is a \( \{\mathcal{F}_t\} \)-stopping time then obviously \( \{T \leq \infty\} = \Omega \in \mathcal{F}_\infty \), hence the defining property of a stopping time continues to hold at \( t = +\infty \). Likewise, if \( T \) is a \( \{\mathcal{F}_t\} \)-optional time, then
\[
\{T < \infty\} = \bigcup_{0 \leq n < \infty} \{T < n\} \in \mathcal{F}_\infty,
\]
thus the defining property of an optional time also holds at \( t = +\infty \).

The next result relates the notions of optional times and stopping times:
Proposition 3.3.3 Suppose that \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in \((\Omega, \mathcal{F}, P)\), and let \( \alpha \in [0, \infty] \) be a constant.

(a) If \( T \) is a \( \{ \mathcal{F}_t \} \)-stopping time then it is also a \( \{ \mathcal{F}_t \} \)-optional time;

(b) \( T \) is an optional time with respect to \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) if and only if it is a stopping time with respect to the filtration \( \{ \mathcal{F}_{t+}, t \in [0, \infty) \} \);

(c) Suppose that \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a right-continuous filtration. Then \( T \) is a \( \{ \mathcal{F}_t \} \)-stopping time if and only if \( T \) is a \( \{ \mathcal{F}_t \} \)-optional time;

(d) If \( T(\omega) = \alpha \) for all \( \omega \in \Omega \) then \( T \) is both an optional time and a stopping time with respect to \( \{ \mathcal{F}_t, t \in [0, \infty) \} \).

Proof: (a) Fix some \( t \in (0, \infty) \). Then \( \{ T \leq t - 1/n \} \in \mathcal{F}_{t-1/n} \subset \mathcal{F}_t \) for all \( n = 0, 1, 2 \ldots \) (since \( T \) is a stopping time). Now clearly

\[
\{ T < t \} = \bigcup_{1 \leq n < \infty} \{ T \leq t - 1/n \} \in \mathcal{F}_t.
\]

(b) Suppose that \( T \) is an optional time with respect to \( \{ \mathcal{F}_t, t \in [0, \infty) \} \), and fix some arbitrary \( t \in [0, \infty) \). Then \( \{ T < t + 1/n \} \in \mathcal{F}_{t+1/n} \) for all \( n = 1, 2 \ldots \) Now clearly

\[
\{ T \leq t \} = \bigcap_{N \leq n < \infty} \{ T < t + 1/n \}, \quad \forall N = 1, 2, \ldots
\]

thus \( \{ T \leq t \} \in \mathcal{F}_{t+1/N} \) for all \( N = 1, 2 \ldots \). In view of Proposition 3.1.30(ii) we see that \( \{ T \leq t \} \in \mathcal{F}_{t+} \).

As for the converse implication, suppose that \( T \) is a stopping time with respect to \( \{ \mathcal{F}_{t+}, t \in [0, \infty) \} \), and fix some \( t \in (0, \infty) \). Then \( \{ T \leq t - 1/n \} \in \mathcal{F}_{(t-1/n)+} \subset \mathcal{F}_t \) for all \( n = 1, 2 \ldots \) From (3.25) we see that \( \{ T < t \} \in \mathcal{F}_t \). This establishes (b). Finally (c) is an immediate consequence of (a) and (b), and (d) is just a trivial observation.

Remark 3.3.4 The significance of Proposition 3.3.3(a) is that the notion of a stopping time is more restrictive than that of an optional time. Indeed, in many applications one comes across random times \( T \) which are easily seen to be optional times with respect to some filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \), but where it is difficult to show directly that \( T \) is a \( \{ \mathcal{F}_t \} \)-stopping time. On the other hand, the greater specialization enjoyed by stopping times usually make them the preferred concept in terms of which to formulate general results dealing with the “stopped” properties of martingales and other stochastic processes (we shall see several such results in this and later chapters). If the underlying filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is right-continuous then Proposition 3.3.3(c) says that the distinction between stopping times and optional times disappears, and an \( \{ \mathcal{F}_t \} \)-optional time is also a \( \{ \mathcal{F}_t \} \)-stopping time. This is fortunate since we can now apply any result which has been established for stopping times to the optional times which typically arise in applications. This illustrates one reason for the usefulness of right-continuous filtrations, and we shall see other reasons in later chapters.
Remark 3.3.5 In Remark 2.1.6 we observed that if $T : \Omega \to \{0, 1, 2, \ldots, \infty\}$ is a discrete-parameter random time and $\{\mathcal{F}_n, n = 0, 1, 2, \ldots\}$ is a discrete-parameter filtration such that $\{T = n\} \in \mathcal{F}_n$ for each $n = 0, 1, 2, \ldots$, then $T$ is a $\{\mathcal{F}_n\}$-stopping time. Since the real line is uncountable, there is of course no comparable assertion in the continuous-parameter case: If $T : \Omega \to [0, \infty]$ is a continuous-parameter random time and $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a continuous-parameter filtration with $\{T = t\} \in \mathcal{F}_t$ for all $t \in [0, \infty)$, then it does not generally follow that $T$ is a $\{\mathcal{F}_t\}$-stopping time.

Remark 3.3.6 Suppose that $\Gamma$ is a non-empty subset of $\mathbb{R}^d$, and $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued adapted process defined on the probability space $(\Omega, \mathcal{F}, P)$. Define

\begin{equation}
D_T(\omega) \overset{\Delta}{=} \inf\{t \in [0, \infty) : X_t(\omega) \in \Gamma\}, \quad \forall \omega \in \Omega,
\end{equation}

where we adopt the usual convention that $\inf\{\emptyset\} = +\infty$. The mapping $D_T : \Omega \to [0, \infty]$ is called the \textbf{début} or \textbf{first contact time} of $\Gamma$. We are now going to look at conditions on the process $\{X_t; t \in [0, \infty)\}$ as well as the set $\Gamma \subset \mathbb{R}^d$ which ensure that $D_T$ is either an optional time or a stopping time:

**Proposition 3.3.7** Suppose that $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued adapted process and $\Gamma \subset \mathbb{R}^d$ is non-empty.

(a) If $\{X_t; t \in [0, \infty)\}$ is right-continuous and $\Gamma$ is an open set then $D_T$ is a $\{\mathcal{F}_t\}$-optional time.

(b) If $\{X_t; t \in [0, \infty)\}$ is continuous and $\Gamma$ is a closed set then $D_T$ is a $\{\mathcal{F}_t\}$-stopping time.

Proof: (a) Denote by $Q_+$ the set of non-negative \textit{rational} numbers. Fix an arbitrary $t \in [0, \infty)$, and observe that, if $X_s(\omega) \in \Gamma$ for some $s \in Q_+ \cap [0, t)$ then $D_T(\omega) < t$. Thus

\begin{equation}
\bigcup_{s \in Q_+ \cap [0, t)} \{X_s \in \Gamma\} \subset \{D_T < t\}
\end{equation}

To establish the opposite set inclusion, suppose that $D_T(\omega) < t$ for some $\omega$. It follows from (3.26) that $X_u(\omega) \in \Gamma$ for some $u \in [0, t)$. Since $\Gamma$ is open and $t \to X(t, \omega)$ is right-continuous we see that there is some $s \in Q_+ \cap [u, t)$ such that $X_s(\omega) \in \Gamma$. Thus $D_T(\omega) < t$ implies that $\omega \in \bigcup_{s \in Q_+ \cap [0, t)} \{X_s \in \Gamma\}$. Combining this with (3.27) yields

$$\{D_T < t\} = \bigcup_{s \in Q_+ \cap [0, t)} \{X_s \in \Gamma\}.$$  

Now the right hand side is made up of a \textit{countable} union of events $\{X_s \in \Gamma\} \in \mathcal{F}_s \subset \mathcal{F}_t$, hence $\{D_T < t\} \in \mathcal{F}_t$, $\forall t \in [0, \infty)$. Thus $D_T$ is a $\{\mathcal{F}_t\}$-optional time.

(b) For each $\epsilon \in (0, \infty)$ put

$$\Gamma^\epsilon \overset{\Delta}{=} \bigcup_{x \in \Gamma} B(x, \epsilon),$$

110
where \( B(x, \epsilon) \overset{\Delta}{=} \{ y \in \mathbb{R}^d : \ d(x, y) < \epsilon \} \) and \( d(\cdot, \cdot) \) denotes the usual Euclidean distance in \( \mathbb{R}^d \), and observe that \( \Gamma^c \) is an open subset of \( \mathbb{R}^d \). For each \( t \in [0, \infty) \) we shall prove that
\[
\{ D_\Gamma \leq t \} = \bigcap_{1 \leq n < \infty} \bigcup_{s \in [0, t] \cap Q_+} \{ X_s \in \Gamma^{1/n} \}.
\]

Since \( \Gamma \) is closed and \( u \to X_u(\omega) \) is continuous over \([0, \infty)\) for each \( \omega \in \Omega \), we see that \( \{ D_\Gamma = 0 \} = \{ X_0 \in \Gamma \} \), hence (3.28) holds when \( t = 0 \). Thus, fix some \( t \in (0, \infty) \), and suppose that \( \omega \) is outside the set on the right side of (3.28). Then, for some \( n \in \{1, 2, 3, \ldots\} \), we have \( X_s(\omega) \in [\Gamma^{1/n}]^c, \forall \ s \in [0, t] \cap Q_+ \). Since \( u \to X_u(\omega) \) is continuous and \( [\Gamma^{1/n}]^c \) is closed, it follows that \( X_u(\omega) \in [\Gamma^{1/n}]^c, \forall \ u \in [0, t] \), and therefore (again using continuity of \( u \to X_u(\omega) \)) we have \( X_u(\omega) \in [\Gamma^{1/n}]^c, \forall \ u \in [0, t + \delta] \), for some \( \delta \in (0, \infty) \), whence \( D_\Gamma(\omega) > t \). Thus, we have shown that the set on the left side of (3.28) is contained within the set on the right side of (3.28). For the opposite set inclusion let \( \omega \) be a member of the set on the right side of (3.28). Then, for each \( n = 1, 2, 3, \ldots \), there is some \( s_n \in [0, t] \cap Q_+ \) such that \( X_{s_n}(\omega) \in \Gamma^{1/n} \). Since \([0, t]\) is closed and bounded, the Bolzano-Weierstrass property ensures that there is some subsequence \( \{s_{n_k}\} \subset \{s_n\} \) and some \( s \in [0, t] \) such that \( \lim_{k \to \infty} s_{n_k} = s \). Then, since \( \Gamma \) is closed and \( u \to X_u(\omega) \) is continuous, we get \( X_s(\omega) \in \Gamma \), so that \( D_\Gamma(\omega) \leq t \), and (3.28) follows. Since \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty) \} \) is adapted, we have \( \{X_s \in \Gamma^{1/n}\} \subset \mathcal{F}_s \subset \mathcal{F}_t, \forall \ s \in [0, t] \cap Q_+ \), and since there are only countably many events on the right side of (3.28), we get \( \{D_\Gamma \leq t\} \in \mathcal{F}_t, \forall \ t \in [0, \infty) \), as required. \[\square\]

**Remark 3.3.8** Suppose that \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty) \} \) is a continuous, \( \mathbb{R}^d \)-valued, adapted process on \( (\Omega, \mathcal{F}, P) \). For each \( n = 1, 2, \ldots \) define
\[T_n(\omega) \overset{\Delta}{=} \inf\{ t \in [0, \infty) : |X_t(\omega)| \geq n \}, \forall \omega \in \Omega, \]
where \( |x| \) denotes the usual Euclidean norm of the vector \( x \in \mathbb{R}^d \). Then \( T_n \) is the début of the closed set \( \Gamma_n \overset{\Delta}{=} \{ x \in \mathbb{R}^d : |x| \geq n \} \), hence, in view of Proposition 3.3.7(b), is a \( \{\mathcal{F}_t\} \)-stopping time for each \( n = 1, 2, \ldots \). Moreover, for later applications we observe that
\[ (i) \ 0 \leq T_n(\omega) < T_{n+1}(\omega), \forall \omega \in \Omega, \forall n = 1, 2, \ldots ; \]
(ii) \( \lim_{n \to \infty} T_n(\omega) = +\infty, \forall \omega \in \Omega. \)
Clearly (i) just follows from the definition of \( T_n \). As for (ii), suppose by way of contradiction that there is some \( \omega \in \Omega \) and some constant \( A \in (0, \infty) \) such that \( T_n(\omega) \leq A, \forall n = 1, 2, \ldots \). Thus \( \lim_{n \to \infty} |X_{T_n(\omega)}(\omega)| = +\infty \) which contradicts the fact that \( X(\cdot, \omega) \) is continuous, hence uniformly bounded on the bounded interval \([0, A]\).

**Remark 3.3.9** In Proposition 3.3.7(a) we have seen that \( D_\Gamma \), defined by (3.26), is a \( \{\mathcal{F}_t\} \)-optional time when \( \Gamma \) is an open set in \( \mathbb{R}^d \) and the associated process \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty)\} \) has right-continuous sample-paths. In this case it is generally not possible to show that \( D_\Gamma \) is a \( \{\mathcal{F}_t\} \)-stopping time (a counter-example which illustrates this point is developed as part of Exercise 3.5.11). However, if
\( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a right-continuous filtration then Proposition 3.3.3(c) shows that \( D_\Gamma \) is also a \( \{ \mathcal{F}_t \} \)-stopping time.

**Remark 3.3.10** In Remark 2.1.9 it was seen that \( D_\Gamma \) is easily shown to be a stopping time in the discrete-parameter case for arbitrary Borel sets \( \Gamma \subset \mathbb{R}^d \), and one would naturally like to establish a result of similar generality in the continuous-parameter setting. This constitutes a profound issue in the general theory of stochastic processes, and illustrates a general phenomenon: questions about stopping times which are simple in a discrete-parameter context often become highly non-trivial when posed in the setting of continuous-parameter processes. The reason for this is is to be found in Remark 3.3.5: knowing that \( \{ T = t \} \in \mathcal{F}_t \) for all \( t \in [0, \infty) \) does not generally ensure that \( T \) is a \( \{ \mathcal{F}_t \} \)-stopping time. Using the methods of transfinite induction together with the theory of capacities and Suslin sets, it can in fact be shown that the début \( D_\Gamma \) in (3.26) is a \( \{ \mathcal{F}_t \} \)-stopping time for each \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \) provided that the probability space \((\Omega, \mathcal{F}, P)\) is complete, \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is progressively measurable, the \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is right-continuous, and \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \) (see Theorem IV-50 in Dellacherie and Meyer [6]). The proof of this powerful theorem is far beyond the scope of our simple introduction and we shall never have occasion to make use of it. Indeed, every optional time or stopping time that we shall come across in the sequel will be included within the two special cases studied in Proposition 3.3.7.

We next formulate an obvious continuous-parameter analogue of Definition 2.1.10:

**Definition 3.3.11** Suppose that \( T : \Omega \to [0, +\infty) \) is a stopping time with respect to a filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) in a probability space \((\Omega, \mathcal{F}, P)\). We denote by \( \mathcal{F}_T \) the collection of all sets \( A \subset \Omega \) having the property that
\begin{equation}
A \cap \{ T \leq t \} \in \mathcal{F}_t
\end{equation}
for each \( t \in [0, +\infty) \) (recall the definition of \( \mathcal{F}_\infty \) in (3.3)). It is easily seen that \( \mathcal{F}_T \) is a σ-algebra over \( \Omega \). Exactly as in the discrete-parameter case, we call \( \mathcal{F}_T \) the **pre-σ-algebra** generated by \( T \).

The next proposition extends Proposition 2.1.11 to the continuous-parameter context:

**Proposition 3.3.12** Suppose \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\) with respect to which \( T \) is a stopping time. Then:

(a) \( \mathcal{F}_T \) is a σ-algebra over \( \Omega \) and \( \mathcal{F}_T \subset \mathcal{F}_\infty \);

(b) \( T \) is \( \mathcal{F}_T \)-measurable, that is \( \sigma\{T\} \subset \mathcal{F}_T \);

(c) If, for some constant \( t \in [0, +\infty] \), we have \( T(\omega) = t, \forall \omega \in \Omega \), then the σ-algebras \( \mathcal{F}_T \) and \( \mathcal{F}_t \) are identical.

Basic properties of stopping times and their associated pre-σ-algebras are summarized in the next proposition, which is, again, just a continuous-parameter analogue of Proposition 2.1.12:
Proposition 3.3.13 Suppose that \( \{F_t, \ t \in [0, \infty)\} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\) with respect to which \(S\) and \(T\) are stopping times. Then

(a) \(S \wedge T, S \vee T\) and \(S + T\) are stopping times with respect to \(\{F_t, \ t \in [0, \infty)\}\);

(b) If \(S(\omega) \leq T(\omega)\) for all \(\omega\) then \(F_S \subset F_T\);

(c) \(F_{S\wedge T} = F_S \cap F_T\);

(d) The events \(\{S < T\}, \{S = T\}\) and \(\{S > T\}\) are members of \(F_{S\wedge T}\);

(e) If \(A \in F_S\) then the events \(A \cap \{S \leq T\}, A \cap \{S < T\}\) and \(A \cap \{S = T\}\) are members of \(F_{S\wedge T}\).

Proof: With the exception of the assertion that \(S + T\) is a stopping time, all other parts of this proposition are established by following arguments identical to those used for the corresponding parts of Proposition 2.1.12. To show that \(S + T\) is a stopping time we use a modification of the standard proof which shows that the sum of two measurable functions is a measurable function (see e.g. Royden [27]). For \(t = 0\), clearly \(\{S + T > t\} = (\{S = 0\} \cap \{T = 0\})^c \in F_0\), thus fix some arbitrary \(t \in (0, \infty)\). Then, clearly,

\[
(S + T > t) = \{S + T > t\} \cap \{T = 0\} \cup \{0 < T < t\} \cup \{T \geq t\}.
\]

Now

\[
(S + T > t) \cap \{T = 0\} = \{S > t\} \cap \{T = 0\} \in F_t.
\]

As for the second event on the right hand side of (3.30), suppose that \(\omega \in \{S + T > t\} \cap \{0 < T < t\}\). Since the set of rational numbers is dense in \(\mathbb{R}\), there is some rational \(r(\omega)\) such that \(S(\omega) > r(\omega) > t - T(\omega) > 0\) and \(t > r(\omega)\) (this is possible since \(T(\omega) > 0\)). Thus

\[
(S + T > t) \cap \{0 < T < t\} = \bigcup_{r \in Q \cap (0,t)} \{r < S\} \cap \{T > t - r\} \cap \{T < t\}
\]

whence \(\{S + T > t\} \cap \{0 < T < t\} \in F_t\), since all events in the countable union on the right hand side of (3.32) are clearly members of \(F_t\). Finally, consider the third event on the right hand side of (3.30) for which we have

\[
(S + T > t) \cap \{T \geq t\} = ([S = 0] \cap \{T > t\}) \cup ([S > 0] \cap \{T \geq t\}) \in F_t.
\]

From (3.33), (3.32), (3.31) and (3.30) it follows that \(\{S + T > t\} \in F_t\), which shows that \(S + T\) is a stopping time with respect to \(\{F_t, \ t \in [0, \infty)\}\). \(\blacksquare\)

Next, we settle some measure-theoretic technicalities somewhat along the lines of those in Proposition 2.3.2 in the discrete-parameter context. If \(\{X_t; \ t \in [0, \infty)\}\) is a given \(\mathbb{R}^d\)-valued process and \(T\) is a random time on \((\Omega, \mathcal{F}, P)\), then we put \(X_T(\omega) \overset{\Delta}{=} X_{T(\omega)}(\omega)\) for all \(\omega\) such that \(T(\omega) < \infty\). In this way, we define a mapping \(X_T: \{T < \infty\} \to \mathbb{R}^d\). The next proposition looks, among other things, at conditions which ensure the measurability of this mapping:
Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued progressively measurable process and \( T \) is an \( \{\mathcal{F}_t\} \)-stopping time defined on the probability space \((\Omega, \mathcal{F}, P)\). Then:

(a) \( \{(X_{t \wedge T}, \mathcal{F}_t), t \in [0, \infty)\} \) is a progressively measurable process;
(b) \( \{X_T \in A\} \cap \{T < \infty\} \in \mathcal{F}_T \) for all \( A \in \mathcal{B}(\mathbb{R}^d) \).

Proof: (a) Fix some \( t \in [0, \infty) \). Then \( t \wedge T \) is a stopping time upper-bounded by \( t \) hence \( \omega \rightarrow t \wedge T(\omega) \) is a measurable mapping from \((\Omega, \mathcal{F}_t)\) into \([0, t], \mathcal{B}([0, t])\)). Thus, from the definition of a product \( \sigma \)-algebra, we see that \( (s, \omega) \rightarrow (s, t \wedge T(\omega)) \) is a measurable mapping from \( ([0, t] \square [0, \infty], \mathcal{B}([0, t]) \square \mathcal{F}_t) \) into \( ([0, t] \square [0, t], \mathcal{B}([0, t] \square [0, t])) \). Now, since \( (x, y) \rightarrow x \wedge y \) is clearly a measurable mapping from \( ([0, t] \square [0, t], \mathcal{B}([0, t] \square [0, t])) \) into \( ([0, t], \mathcal{B}([0, t])) \), it follows from the preceding, together with the fact that the composition of two measurable mappings is measurable, that \( (s, \omega) \rightarrow s \wedge (t \wedge T(\omega)) \equiv s \wedge T(\omega) \) is a measurable mapping from \( ([0, t] \square [0, \infty], \mathcal{B}([0, t]) \square \mathcal{F}_t) \) into \( ([0, t], \mathcal{B}([0, t])) \), and therefore (again from the definition of a product \( \sigma \)-algebra) the mapping \( (s, \omega) \rightarrow (s \wedge T(\omega), \omega) \) is a measurable mapping from \( ([0, t] \square [0, \infty], \mathcal{B}([0, t]) \square \mathcal{F}_t) \) into itself. Now, by the assumed progressive measurability, we see that \( (s, \omega) \rightarrow X(s, \omega) \) is a measurable mapping from \( ([0, t] \square [0, \infty], \mathcal{B}([0, t]) \square \mathcal{F}_t) \) into \( \mathbb{R}^d \). Again using the fact that the composition of two measurable maps is measurable, we see that \( (s, \omega) \rightarrow (X(s \wedge T(\omega), \omega) \) is a measurable mapping from \( ([0, t] \square [0, \infty], \mathcal{B}([0, t]) \square \mathcal{F}_t) \) into \( \mathbb{R}^d \).

(b) Fix some arbitrary \( t \in [0, \infty) \) and \( A \in \mathcal{B}(\mathbb{R}^d) \). Then:

\[
(3.34) \quad \{X_T \in A\} \cap \{T < \infty\} \cap \{T \leq t\} = \{X_{t \wedge T} \in A\} \cap \{T \leq t\}.
\]

Now \( \{T \leq t\} \in \mathcal{F}_t \), and a consequence of (a) is that \( \omega \rightarrow X_{t \wedge T(\omega)}(\omega) \) is \( \mathcal{F}_t \)-measurable, whence \( \{X_{t \wedge T} \in A\} \in \mathcal{F}_t \). Therefore, from (3.34), we obtain

\[
\{X_T \in A\} \cap \{T < \infty\} \cap \{T \leq t\} \in \mathcal{F}_t
\]

for all \( t \in [0, \infty) \), as required to establish (b).

The next result, which follows at once from Proposition 3.3.14, is just a continuous-parameter analogue of Proposition 2.3.2:

**Proposition 3.3.15** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued progressively measurable process and \( X_\infty \) is a \( \mathcal{F}_\infty \)-measurable \( \mathbb{R}^d \)-valued random vector on the probability space \((\Omega, \mathcal{F}, P)\), with \( \mathcal{F}_\infty \) defined by (3.3). If \( T : \Omega \rightarrow [0, +\infty] \) is a \( \{\mathcal{F}_t\} \)-stopping time, then the mapping \( X_T : \Omega \rightarrow \mathbb{R}^d \) defined by

\[
(3.35) \quad X_T(\omega) \triangleq \begin{cases} 
X_T(\omega) & \text{when } T(\omega) < \infty \\
X_\infty(\omega) & \text{when } T(\omega) = \infty
\end{cases}
\]

is \( \mathcal{F}_T \)-measurable.
3.4 Fine-Tuning a Continuous-Parameter Raw Filtration

The notion of a filtration in a probability space was introduced in a discrete-parameter setting in Definition 2.1.2, and then extended to the continuous-parameter context in Definition 3.1.14. Our attitude towards filtrations, in both the discrete and continuous-parameter cases, has thus far been rather uncritical, and in particular we have not addressed the issue of how such filtrations originate from actual problems. Our goal in the present section is to look at the simplest aspects of this question. We are especially interested in the continuous-parameter case, both because it is the one most pertinent to our study of stochastic calculus, and also because important regularity properties such as right-continuity (recall Definition 3.1.26) really only make sense in a continuous-parameter context. Accordingly we shall, throughout this section, confine ourselves to the continuous-parameter case.

Definition 3.4.1 Suppose that \(\{ Y(t); \ t \in [0, \infty) \} \) is some \(\mathbb{R}^d\)-valued process on a probability space \((\Omega, \mathcal{F}, P)\), and put

\[
F_t^Y \triangleq \sigma \{ Y(s), \ s \in [0, t] \}, \ \forall \ t \in [0, \infty).
\]

Clearly \(\{ F_t^Y, \ t \in [0, \infty) \}\) is a filtration in \((\Omega, \mathcal{F}, P)\) and will be called the raw or natural filtration generated by the process \(\{ Y(t); \ t \in [0, \infty) \}\).

It will become clear with hindsight that the sort of problems to which one typically applies the stochastic calculus are usually characterized by some given (finite) family of stochastic processes on some common probability space \((\Omega, \mathcal{F}, P)\). Without loss of generality we can assemble this family of processes into a vector of processes \(\{ Y(t); \ t \in [0, \infty) \}\) taking values in some Euclidean space \(\mathbb{R}^d\) of appropriate dimension. It is natural to introduce into \((\Omega, \mathcal{F}, P)\) a filtration \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) which, at the very least, is such that \(\{(Y(t), \mathcal{F}_t); \ t \in [0, \infty)\}\) is adapted, and this suggests that we simply make \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) the raw filtration generated by \(\{ Y(t); \ t \in [0, \infty) \}\), namely \(\mathcal{F}_t \triangleq F_t^Y, \ \forall \ t \in [0, \infty)\).

Several of our previous results depend on \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) satisfying certain regularity hypotheses. Typically, \(\mathcal{F}_0\) must include all \(P\)-null events in \(\mathcal{F}\) (recall Remark 3.1.17 and Proposition 3.1.18), and elsewhere we saw that the property of right-continuity of \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) brought other significant advantages; in particular, an \(\{ \mathcal{F}_t\}\)-optional time is an \(\{ \mathcal{F}_t\}\)-stopping time (Remark 3.3.4). Later, in Chapter 4, we shall see further evidence for the usefulness of right-continuous filtrations (note, in particular, Theorem 4.2.18). These considerations lead us to formulate the notion of a “nice” filtration which has both of the preceding regularity properties:

Definition 3.4.2 A filtration \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) in a probability space \((\Omega, \mathcal{F}, P)\) is a standard filtration when

(i) \(\mathcal{F}_0\) includes all \(P\)-null events in \(\mathcal{F}\), and
(ii) \(\{ \mathcal{F}_t, \ t \in [0, \infty) \}\) is right-continuous.

In view of the advantages which arise from being able to work with standard filtrations, the question naturally arises as to whether or not the raw filtration \(F_t^Y\) generated by a given \(\mathbb{R}^d\)-valued process
\( \{Y(t); \ t \in [0, \infty)\} \) is a standard filtration. Unfortunately this is generally not the case, even when \( \{Y(t); \ t \in [0, \infty)\} \) is a process with continuous sample-paths. For example, in Exercise 3.5.11 it is seen that the raw filtration of a continuous process need not be right-continuous. The main goal in this section is to learn a standard constructive technique which will enable us to “fine-tune” the raw filtration \( \mathcal{F}_t^Y \) of a given process \( \{Y(t); \ t \in [0, \infty)\} \) in order to get a “minimally” enlarged filtration \( \{F_t, \ t \in [0, \infty)\} \) (i.e. \( \mathcal{F}_t^Y \subset F_t, \ \forall \ t \in [0, \infty) \)) which is a standard filtration.

The technique that we shall use to get standard filtrations is really a generalization of the construction one uses to complete a given measure space, a topic which properly belongs to elementary measure theory. Because of its importance for the goals of this section we shall now recall this construction in some detail. Recall that the symmetric difference \( A \triangle B \) of two sets \( A \) and \( B \) is defined by

\[
A \triangle B = (A - B) \cup (B - A).
\]

The next proposition lists the simplest properties of symmetric differences. The elementary proofs are left to the reader.

**Proposition 3.4.3** (a) For sets \( A, B, C \) and \( D \) we have:

(i) \( A \triangle B = C \) if and only if \( A = B \triangle C \);

(ii) \( A \triangle B = A^c \triangle B^c \);

(iii) \( (A \triangle B) \triangle (A \triangle C) \subset B \cup C \);

(iv) \( (A \triangle B) \cap (D \triangle C) \subset B \cup C \) when \( A \cap D = \emptyset \).

(b) Suppose \( \{A_n, \ n = 0, 1, 2, \ldots\} \) and \( \{B_n, \ n = 0, 1, 2, \ldots\} \) are sequences of sets. Then

\[
\left( \bigcup_{1 \leq n < \infty} A_n \right) \triangle \left( \bigcup_{1 \leq n < \infty} B_n \right) \subset \bigcup_{1 \leq n < \infty} (A_n \triangle B_n),
\]

and, for any set \( A \),

\[
A \triangle \left( \liminf_{n \to \infty} B_n \right) \subset \bigcup_{1 \leq n < \infty} (A \triangle B_n).
\]

(c) Suppose that \( X \) and \( Y \) are \( \mathbb{R} \)-valued mappings defined on a common set \( E \). Then

\[
\{X \in \Gamma\} \triangle \{Y \in \Gamma\} \subset \{X \neq Y\}
\]

for each \( \Gamma \subset \mathbb{R} \).

**Definition 3.4.4** A measure space \((E, \mathcal{S}, \mu)\) is complete when the conditions

\[
A \in \mathcal{S}, \ B \subset A, \ \text{and} \ \mu(A) = 0
\]

imply that \( B \) is a member of \( \mathcal{S} \) (and, of course, \( \mu(B) = 0 \)).
Thus, $Z(3.37)$ measure space, $H \subset S$ even though $H$ is a measurable set of $\mu$ measurable set of $\mu$ with measure which is complete. We next look at a construction for doing this. Suppose that $(E, S, \mu)$ is a measure space, $H \subset S$ is a sub-$\sigma$-algebra of $S$, and put

$$(3.37) \quad Z^\mu[H] \triangleq \{ N \subset E : \exists H \in H \text{ such that } N \subset H \text{ and } \mu(H) = 0 \}.$$ 

Thus, $Z^\mu[H]$ is the collection of subsets of $E$ each member of which is included within some $H$-measurable set of $\mu$-measure zero.

Remark 3.4.5 It is easily checked that completeness of $(E, S, \mu)$ entails the following: if $A, B \in S$ with $A \subset B$ and $\mu(A) = \mu(B)$, then any set $C$ for which $A \subset C \subset B$ is necessarily a member of $S$ with $\mu(A) = \mu(B) = \mu(C)$. Although completeness is a reasonable property for a measure space to have, one should be aware that several commonly occurring measure spaces fail to be complete. One such example is $(\mathbb{R}, B(\mathbb{R}), \lambda)$, where $\lambda$ denotes Lebesgue measure (the interested reader may consult Halmos ([14], §19, Ex.(4)) for a demonstration of this fact).

If a given measure space fails to be complete then we can always fine-tune it in order to get a measure space which is complete. We next look at a construction for doing this. Suppose that $(E, S, \mu)$ is a measure space, $H \subset S$ is a sub-$\sigma$-algebra of $S$, and put

$$(3.37) \quad Z^\mu[H] \triangleq \{ N \subset E : \exists H \in H \text{ such that } N \subset H \text{ and } \mu(H) = 0 \}.$$ 

Thus, $Z^\mu[H]$ is the collection of subsets of $E$ each member of which is included within some $H$-measurable set of $\mu$-measure zero.

Remark 3.4.6 If $\mathcal{A}$ and $\mathcal{B}$ are given collections of subsets of $E$ (not necessarily $\sigma$-algebras) we shall write $\sigma\{\mathcal{A}, \mathcal{B}\}$ for the $\sigma$-algebra generated by $\mathcal{C} \triangleq \mathcal{A} \cup \mathcal{B}$. i.e. $\sigma\{\mathcal{A}, \mathcal{B}\}$ is the minimal $\sigma$-algebra over $E$ which includes all members of $\mathcal{A}$ and all members of $\mathcal{B}$.

Proposition 3.4.7 Suppose $(E, S, \mu)$ is a measure space and $\mathcal{G}, \mathcal{H}$ are $\sigma$-algebras over $E$ with $\mathcal{G} \subset \mathcal{H} \subset S$. Then

$$(3.38) \quad \sigma\{\mathcal{G}, Z^\mu[\mathcal{H}]\} = \{ A \subset E : \exists G \in \mathcal{G} \text{ such that } A \Delta G \in Z^\mu[\mathcal{H}] \}.$$ 

Proof: Let $\mathcal{A}$ be the collection of sets on the right-hand side of (3.38) and observe that $\mathcal{A}$ is a $\sigma$-algebra over $E$. Indeed, one clearly has $E \in \mathcal{A}$. Moreover, if $A \in \mathcal{A}$ with $G \in \mathcal{G}$ such that $A \Delta G \in Z^\mu[\mathcal{H}]$, then by Proposition 3.4.3(a)(ii), we have $A^c \Delta G^c = A \Delta G \in Z^\mu[\mathcal{H}]$; since $G^c \in \mathcal{G}$ it follows that $A^c \in \mathcal{A}$. Next, let $\{A_n, n = 1, 2, \ldots\}$ be a sequence of sets in $\mathcal{A}$, and let $G_n \in \mathcal{G}$ be such that $A_n \Delta G_n \in Z^\mu[\mathcal{H}]$, $\forall n = 0, 1, 2 \ldots$ Then, in view of Proposition 3.4.3(b),

$$(3.39) \quad \left( \bigcup_{1 \leq n < \infty} A_n \right) \Delta \left( \bigcup_{1 \leq n < \infty} G_n \right) \subset \bigcup_{1 \leq n < \infty} (A_n \Delta G_n).$$ 

From (3.37) we see that $\bigcup_{1 \leq n < \infty} (A_n \Delta G_n) \in Z^\mu[\mathcal{H}]$, and since $\bigcup_{1 \leq n < \infty} G_n \in \mathcal{G}$, it follows from (3.37) and (3.39) that $\bigcup_{1 \leq n < \infty} A_n \in \mathcal{A}$. Thus $\mathcal{A}$ is a $\sigma$-algebra over $E$. By the definition of $\mathcal{A}$ it is at once clear that $\mathcal{G} \subset \mathcal{A}$ and $Z^\mu[\mathcal{H}] \subset \mathcal{A}$, hence we conclude that $\sigma\{\mathcal{G}, Z^\mu[\mathcal{H}]\} \subset \mathcal{A}$. For the opposite set inclusion, fix some $A \in \mathcal{A}$ and let $G \in \mathcal{G}$ be such that $A \Delta G \triangleq N \in Z^\mu[\mathcal{H}]$. By Proposition 3.4.3(a)(i) we find $A = G \Delta N \in \sigma\{\mathcal{G}, Z^\mu[\mathcal{H}]\}$, as required.

Remark 3.4.8 Observe from (3.37) that $Z^\mu[\mathcal{H}]$ generally includes sets which are not members of $S$, even though $\mathcal{H}$ is a sub-$\sigma$-algebra of $S$, and thus we cannot assert that $\sigma\{\mathcal{G}, Z^\mu[\mathcal{H}]\} \subset S$ (this set-inclusion would, of course, be true if $(E, S, \mu)$ were a complete measure space). Thus, although $\mu$ is
a measure on $S$, $\mu(A)$ is not necessarily defined for all $A \in \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$. The next proposition deals with the issue of using $\mu$ to define a measure on $\sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$.

**Proposition 3.4.9** Suppose $(E, \mathcal{S}, \mu)$ is a measure space and $\mathcal{G}$, $\mathcal{H}$ are $\sigma$-algebras over $E$ with $\mathcal{G} \subset \mathcal{H} \subset \mathcal{S}$. Then there is a unique measure $\overline{\mu} : \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\} \to [0, \infty]$ such that

\[
\overline{\mu}(A) \triangleq \begin{cases} 
\mu(A), & \forall A \in \mathcal{G} \\
0, & \forall A \in \mathcal{Z}^\mu[\mathcal{H}].
\end{cases}
\]

Moreover, $(E, \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}, \overline{\mu})$ is a complete measure space.

Proof: Fix some $A \in \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$. Then, by Proposition 3.4.7, there is some $G \in \mathcal{G}$ such that $A \triangle G \in \mathcal{Z}^\mu[\mathcal{H}]$. Define $\overline{\mu}(A) \triangleq \mu(G)$. For this definition to make sense we must check that, if $\tilde{G} \in \mathcal{G}$ is such that $A \triangle \tilde{G} \in \mathcal{Z}^\mu[\mathcal{H}]$, then $\mu(G) = \mu(\tilde{G})$. To this end, put $N \triangleq A \triangle G$, and $\tilde{N} \triangleq A \triangle \tilde{G}$, and observe from Proposition 3.4.3(a)(i),(iii), that $G = A \triangle N$, $\tilde{G} = A \triangle \tilde{N}$, and $G \triangle \tilde{G} \subset N \cup \tilde{N}$. Since $N, \tilde{N} \in \mathcal{Z}^\mu[\mathcal{H}]$, it follows that $N \subset H$ and $\tilde{N} \subset \tilde{H}$ for some $H$ and $\tilde{H} \in \mathcal{G}$ such that $\mu(H) = \mu(\tilde{H}) = 0$, and therefore $\mu(G \triangle \tilde{G}) \leq \mu(H \cup \tilde{H}) = 0$. Then, by the obvious inequality

\[
|\mu(S_1) - \mu(S_2)| \leq \mu(S_1 \triangle S_2), \quad \forall S_1, S_2 \in S,
\]

we obtain $\mu(G) = \mu(\tilde{G})$ as required. Hence, we have defined a mapping $\overline{\mu} : \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\} \to [0, \infty]$ for which (3.40) clearly holds. To see that $\overline{\mu}$ is a measure fix a sequence of sets $\{A_n, n = 1, 2, \ldots\}$ such that $A_n \in \sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$ and $A_m \cap A_n = \emptyset$ when $m \neq n$. For each $A_n$ there is some $G_n \in \mathcal{G}$ such that $A_n \triangle G_n \triangleq N_n \in \mathcal{Z}^\mu[\mathcal{H}]$. Thus, by Proposition 3.4.3(a)(i) and (iv), we have $G_n = A_n \triangle N_n$, and

\[
\mu(G_m \cap G_n) = 0, \quad \text{when } m \neq n.
\]

By Proposition 3.4.3(b) we see that

\[
\left( \bigcup_{1 \leq n < \infty} A_n \right) \triangle \left( \bigcup_{1 \leq n < \infty} G_n \right) \in \mathcal{Z}^\mu[\mathcal{H}],
\]

thus, the definition of $\overline{\mu}$ ensures that

\[
\overline{\mu}\left( \bigcup_{1 \leq n < \infty} A_n \right) = \mu\left( \bigcup_{1 \leq n < \infty} G_n \right) = \sum_{1 \leq n < \infty} \mu(G_n) = \sum_{1 \leq n < \infty} \overline{\mu}(A_n),
\]

where the second equality in (3.43) follows from (3.42). Thus $\overline{\mu}$ is a measure on $\sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$. To see uniqueness of $\overline{\mu}$ suppose that $\overline{\nu}$ is a measure on $\sigma\{\mathcal{G}, \mathcal{Z}^\mu[\mathcal{H}]\}$ such that

\[
\overline{\nu}(A) \triangleq \begin{cases} 
\mu(A), & \forall A \in \mathcal{G} \\
0, & \forall A \in \mathcal{Z}^\mu[\mathcal{H}].
\end{cases}
\]
For arbitrary \( A \in \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\} \) there is some \( G \in \mathcal{G} \) such that \( A \triangle G \in \mathcal{Z}^\mu[\mathcal{H}] \). By (3.44) we find that 
\[
\nu(A \triangle G) = 0,
\]
and thus \( |\nu(A) - \nu(G)| \leq \nu(A \triangle G) = 0 \) whence, by (3.44) and (3.40), we find that \( \nu(A) = \nu(G) = \mu(G) = \bar{\mu}(A) \), as required to establish the uniqueness of \( \bar{\mu} \). To see that \( (E, \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\}, \bar{\mu}) \) is a complete measure space fix some \( A \in \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\} \) with \( \bar{\mu}(A) = 0 \) and let \( B \subset A \). It is clearly sufficient to show that \( A \in \mathcal{Z}^\mu[\mathcal{H}] \), for then (3.37) ensures that \( B \in \mathcal{Z}^\mu[\mathcal{H}] \) and hence \( B \in \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\} \), as required. By Proposition 3.4.7 there is some \( G \in \mathcal{G} \) such that \( A \triangle G = N \in \mathcal{Z}^\mu[\mathcal{H}] \). By the definition of \( \bar{\mu} \), clearly \( \mu(G) = \bar{\mu}(A) = 0 \). Moreover \( A \subset N \cup G \), thus \( A \subset H \cup G \) for some \( H \in \mathcal{H} \) with \( N \subset H \) and \( \mu(H) = 0 \). Since \( H \cup G \in \mathcal{H} \) and \( \mu(H \cup G) = 0 \) it follows that \( A \in \mathcal{Z}^\mu[\mathcal{H}] \).

\[
\text{Remark 3.4.10} \quad \text{If the measure space } (E, \mathcal{S}, \mu) \text{ in Proposition 3.4.9 is complete then of course } \mathcal{Z}^\mu[\mathcal{H}] \subset \mathcal{S}, \text{ and so } \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\} \subset \mathcal{S}. \text{ It follows that } \bar{\mu}(A) = \mu(A), \forall A \in \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\}.
\]

\[
\text{Remark 3.4.11} \quad \text{If } \mathcal{G} \subset \mathcal{S} \text{ is a sub-\( \sigma \)-algebra of } \mathcal{S}, \text{ and we take } \mathcal{H} = \mathcal{G} \text{ in Proposition 3.4.9 then the complete measure space } (E, \sigma\{G, \mathcal{Z}^\mu[\mathcal{G}]\}, \bar{\mu}) \text{ is called the completion of the measure space } (E, \mathcal{G}, \mu). \text{ In particular, } (E, \sigma\{\mathcal{S}, \mathcal{Z}^\mu[\mathcal{S}]\}, \bar{\mu}) \text{ is the completion of } (E, \mathcal{S}, \mu). \text{ On the other hand, when } \mathcal{G} \text{ and } \mathcal{H} \text{ are } \sigma \text{-algebras over } E \text{ with } \mathcal{G} \subset \mathcal{H} \subset \mathcal{S}, \text{ then the complete measure space } (E, \sigma\{G, \mathcal{Z}^\mu[\mathcal{H}]\}, \bar{\mu}) \text{ is called the augmentation of } (E, \mathcal{G}, \mu) \text{ with the } \mu \text{-null sets of } \mathcal{H}. \text{ The concept of augmentation is thus more general than that of completion because it allows us to deal with a } \sigma \text{-algebra } \mathcal{H} \text{ which is strictly larger than } \mathcal{G}. \text{ As will soon become clear, this extra generality is needed when we fine-tune a given raw filtration to get a filtration with nice regularity properties.}
\]

\[
\text{Remark 3.4.12} \quad \text{If } (E, \mathcal{S}, \mu) \text{ is already a complete measure space then it is at once clear that its completion, namely } (E, \sigma\{\mathcal{S}, \mathcal{Z}^\mu[\mathcal{S}]\}, \bar{\mu}), \text{ is identical with } (E, \mathcal{S}, \mu).
\]

\[
\text{Remark 3.4.13} \quad \text{Recalling Remark 3.4.5 we see that } (\mathcal{R}, \sigma\{\mathcal{B}(\mathcal{R}), \mathcal{Z}^\lambda[\mathcal{B}(\mathcal{R})]\}, \bar{\lambda}) \text{ is the completion of } (\mathcal{R}, \mathcal{B}(\mathcal{R}), \lambda). \text{ The members of } \sigma\{\mathcal{B}(\mathcal{R}), \mathcal{Z}^\lambda[\mathcal{B}(\mathcal{R})]\} \text{ are called Lebesgue measurable sets.}
\]

Having established these simple measure-theoretic preliminaries, we next look at the issue of augmentation of a filtration. We begin with a simple technical result:

\[
\text{Proposition 3.4.14} \quad \text{Suppose that } \{\mathcal{G}_t, t \in [0, \infty)\} \text{ is a filtration in a complete probability space } (\Omega, \mathcal{F}, P). \text{ For each } t \in [0, \infty) \text{ put}
\]
\[
(3.45) \quad \mathcal{F}_t \triangleq \sigma\{\mathcal{G}_t, \mathcal{Z}^P[\mathcal{F}]\}.
\]

\text{Then} \( \{\mathcal{F}_t, t \in [0, \infty)\} \) \text{ is a filtration in } (\Omega, \mathcal{F}, P) \text{, and for each } t \in [0, \infty) \text{ we have}
\]
\[
(3.46) \quad \mathcal{F}_{t^+} = \sigma\{\mathcal{G}_{t^+}, \mathcal{Z}^P[\mathcal{F}]\}.
\]

\text{Proof : Completeness of } (\Omega, \mathcal{F}, P) \text{ implies } \mathcal{Z}^P[\mathcal{F}] \subset \mathcal{F}, \text{ and hence } \mathcal{F}_t \subset \mathcal{F}, \forall t \in [0, \infty). \text{ It is now clear that } \{\mathcal{F}_t, t \in [0, \infty)\} \text{ is a filtration in } (\Omega, \mathcal{F}, P). \text{ To verify (3.46) fix some } t \in [0, \infty). \text{ In view of (3.45) we see that } \mathcal{G}_u \subset \mathcal{F}_u, \forall u > t, \text{ and hence } \mathcal{G}_{t^+} \subset \mathcal{F}_{t^+}. \text{ By (3.45) again, we see that } \mathcal{Z}^P[\mathcal{F}] \subset \mathcal{F}_t \subset \mathcal{F}_{t^+}, \text{ as required.}
and hence $\sigma\{G_{t+}, Z^P[F]\} \subset F_{t+}$. To establish the opposite set inclusion fix any $A \in F_{t+}$. Then, $A \in F_{t+1/n}$, $\forall n = 1, 2, \ldots$, and in view of (3.45) and Proposition 3.4.7, there is some $B_n \in G_{t+1/n}$ such that $A \triangle B_n \in Z^P[F]$, $\forall n = 1, 2, \ldots$. Put $B \triangleq \liminf_{n \to \infty} B_n$. It follows that

$$B \in \bigcap_{1 \leq n < \infty} G_{t+1/n} = G_{t+},$$

(recall Proposition 3.1.30(ii)), and in view of Proposition 3.4.3 we also have $A \triangle B \in Z^P[F]$. Then Proposition 3.4.7 ensures that $A \in \sigma\{G_{t+}, Z^P[F]\}$, as required.

**Remark 3.4.15** Using Proposition 3.4.14 we can easily get a standard filtration from a raw filtration. Indeed, suppose that $\{Y(t); t \in [0, \infty)\}$ is some $IR^d$-valued process on a complete probability space $(\Omega, F, P)$, and let $\{F^Y_t; t \in [0, \infty)\}$ be the raw filtration generated by $\{Y(t); t \in [0, \infty)\}$ (recall Definition 3.4.1). Put

$$F_t \triangleq \sigma\{F^Y_t, Z^P[F]\}, \ \forall t \in [0, \infty).$$

From Proposition 3.1.30(iv) it follows that the filtration $\{G_t, t \in [0, \infty)\}$, defined by $G_t \triangleq F^Y_{t+}, \ \forall t \in [0, \infty)$, is right-continuous (i.e. $G_t = G_{t+}$), whence, for each $t \in [0, \infty)$, we have

$$F_t = \sigma\{G_t, Z^P[F]\} = \sigma\{G_{t+}, Z^P[F]\} = F_{t+},$$

the final equality following from Proposition 3.4.14. Thus the filtration $\{F_t, t \in [0, \infty)\}$ defined by (3.47), is right-continuous. In view of completeness of $(\Omega, F, P)$ and (3.47) we also see that $F_0$ includes all $P$-null events in $F$, hence $\{F_t, t \in [0, \infty)\}$ is a standard filtration (recall Definition 3.4.2).

The filtration $\{F_t, t \in [0, \infty)\}$ in Remark 3.4.15 really just inherits the right-continuity property from the right-continuous filtration $\{F^Y_{t+}, t \in [0, \infty)\}$. Of much greater interest and importance is the fact that $\{F_t, t \in [0, \infty)\}$ turns out to be a standard filtration even when $F_t$ is defined by

$$F_t \triangleq \sigma\{F^Y_t, Z^P[F]\}, \ \forall t \in [0, \infty),$$

provided $\{Y(t); t \in [0, \infty)\}$ has certain regularity properties. We establish this next, beginning with the following simple technical result:

**Proposition 3.4.16** Suppose that $\{Y(t); t \in [0, \infty)\}$ is an $IR^d$-valued right-continuous process on a probability space $(\Omega, F, P)$, and put

$$H^Y_t \triangleq \sigma\{Y(v) - Y(t), v \in [t, \infty)\}, \ \forall t \in [0, \infty).$$

Then, for each $t \in [0, \infty)$, we have

$$H^Y_{t+1/m} \subset H^Y_{t+1/n} \subset H^Y_t,$$

for all $n, m = 1, 2, \ldots$ with $n \geq m$, and

$$H^Y_t = \sigma\left( \bigcup_{1 \leq n < \infty} H^Y_{t+1/n} \right).$$
Proof: Clearly $\mathcal{H}^Y_t$ can be written as

$$\mathcal{H}^Y_t \triangleq \sigma \{Y(v_2) - Y(v_1), v_1, v_2 \in [t, \infty)\},$$

from which we get (3.50). To establish (3.51) fix some arbitrary $t \in [0, \infty)$, and, to simplify the notation, let $\mathcal{A}$ denote the $\sigma$-algebra on the right-hand side of (3.51). In view of (3.50) we see that $\mathcal{A} \subset \mathcal{H}^Y_t$. To prove the opposite set inclusion fix some arbitrary $v \in [t, \infty)$. Since $\{Y(t); t \in [0, \infty)\}$ is right-continuous, we have

$$Y(v, \omega) - Y(t, \omega) = \lim_{n \to \infty} [Y(v + 1/n, \omega) - Y(t + 1/n, \omega)], \quad \forall \omega \in \Omega. \quad (3.52)$$

Now $Y(v + 1/n) - Y(t + 1/n)$ is $\mathcal{H}^Y_{t+1/n}$-measurable, and thus $\mathcal{A}$-measurable, for each $n = 1, 2, \ldots$ By (3.52) we see that $Y(v) - Y(t)$ is $\mathcal{A}$-measurable, and in view of the arbitrary choice of $v \in [t, \infty)$ and (3.49), we get $\mathcal{H}^Y_t \subset \mathcal{A}$. \hfill \blacksquare

When $\{Y(t); t \in [0, \infty)\}$ has independent increments then Proposition 3.2.5 shows that $\mathcal{F}^Y_t$ and $\mathcal{H}^Y_t$ are independent. If $\{Y(t); t \in [0, \infty)\}$ is also right-continuous then we can non-trivially strengthen this statement:

**Proposition 3.4.17** Suppose that $\{Y(t); t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued right-continuous process with independent increments defined on probability space $(\Omega, \mathcal{F}, P)$. Then $\mathcal{F}^Y_t$ and $\mathcal{H}^Y_t$ are independent for each $t \in [0, \infty)$, for $\mathcal{F}^Y_t$ and $\mathcal{H}^Y_t$ defined by

$$\mathcal{F}^Y_t \triangleq \sigma \{Y(u), u \in [0, t]\}, \quad \mathcal{H}^Y_t \triangleq \sigma \{Y(v) - Y(t), v \in [t, \infty)\}, \quad \forall \ t \in [0, \infty). \quad (3.53)$$

Proof: Fix some $t \in [0, \infty)$ and some $A \in \mathcal{F}^Y_{t+}$, and put

$$X \triangleq I_A. \quad (3.54)$$

Since $\mathcal{F}^Y_t \subset \mathcal{F}^Y_{t+\epsilon}$, $\forall \epsilon \in (0, \infty)$, we see from the independent increments of $\{Y(t); t \in [0, \infty)\}$ together with Proposition 3.2.5 that $\mathcal{F}^Y_t$ and $\mathcal{H}^Y_{t+1/n}$ are independent for each $n = 1, 2, \ldots$ and thus

$$E [X \mid \mathcal{H}^Y_{t+1/n}] = E[X] \quad \text{a.s.} \quad (3.55)$$

for all $n = 1, 2, \ldots$ Moreover, by Proposition 3.4.16, we have

$$\mathcal{H}^Y_{t+1/m} \subset \mathcal{H}^Y_{t+1/n} \subset \mathcal{H}^Y_t \quad (3.56)$$

for all integers $n \geq m$, and

$$\mathcal{H}^Y_t = \sigma \left( \bigcup_{1 \leq n < \infty} \mathcal{H}^Y_{t+1/n} \right). \quad (3.57)$$

In view of (3.56), (3.57) and Theorem 2.7.6 we find

$$\lim_{n \to \infty} E [X \mid \mathcal{H}^Y_{t+1/n}] = E [X \mid \mathcal{H}^Y_t] \quad \text{a.s.} \quad (3.58)$$
Combining (3.55) and (3.58) we see that
\[ E[X; B] = E[E[X \mid \mathcal{H}_t^Y]; B] = E[X]P(B) \]
for all \( B \in \mathcal{H}_t^Y \). In view of (3.54) and the arbitrary choice of \( A \in \mathcal{F}_{t+}^Y \), this implies that \( P(A \cap B) = P(A)P(B) \), \( \forall A \in \mathcal{F}_{t+}^Y, \forall B \in \mathcal{H}_t^Y \), as required to establish the independence of \( \mathcal{F}_{t+}^Y \) and \( \mathcal{H}_t^Y \).

**Proposition 3.4.18** Suppose that \( \{Y(t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued right-continuous process with independent increments defined on probability space \( (\Omega, \mathcal{F}, P) \). Then, we have

\[ \mathcal{F}_{t+}^Y \subset \sigma(\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}_\infty^Y]), \]

for each \( t \in [0, \infty) \), for \( \mathcal{F}_t^Y \) given by (3.36) and

\[ \mathcal{F}_\infty^Y \triangleq \sigma\{Y(s), s \in [0, \infty)\} \].

Proof: Fix some \( t \in [0, \infty) \) and let \( A \in \mathcal{F}_{t+}^Y \). Put \( X \triangleq I_A, Y \triangleq E[X \mid \mathcal{F}_t^Y] \). Clearly \( \{X \neq Y\} \in \mathcal{F}_\infty^Y \subset \mathcal{F} \). In order to establish (3.59) it is enough to show

\[ P[X \neq Y] = 0. \]

Indeed, fixing an arbitrary \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \), we see from Proposition 3.4.3(c) that

\[ \{X \in \Gamma\} \triangle \{Y \in \Gamma\} \subset \{X \neq Y\} \]

and therefore, in view of (3.60),

\[ \{X \in \Gamma\} \triangle \{Y \in \Gamma\} \in \mathcal{Z}^P[\mathcal{F}_\infty^Y]. \]

Now clearly \( \{Y \in \Gamma\} \in \mathcal{F}_t^Y \), whence (3.61) and Proposition 3.4.7 imply that \( \{X \in \Gamma\} \in \sigma(\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}_\infty^Y]) \). Since this holds for all \( \Gamma \in \mathcal{B}(\mathbb{R}^d) \), we see that \( A \in \sigma(\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}_\infty^Y]) \), as required for (3.59) To establish (3.60), fix arbitrary events \( B \in \mathcal{F}_t^Y \) and \( C \in \mathcal{H}_t^Y \). Clearly \( (X - Y)I_B \) is \( \mathcal{F}_{t+}^Y \)-measurable, thus Proposition 3.4.17 says that \( (X - Y)I_B \) and \( I_C \) are independent. Thus

\[ E[(X - Y)I_BI_C] = E[(X - Y)I_B]P(C) = 0, \]

\( \forall B \in \mathcal{F}_t^Y, \forall C \in \mathcal{H}_t^Y \), where the second equality of (3.62) follows since \( (X - Y) \perp I_B \) (see Theorem 1.4.14). Now put

\[ D \triangleq \{B \cap C: B \in \mathcal{F}_t^Y, C \in \mathcal{H}_t^Y\}, \]

and

\[ C \triangleq \{D \in \mathcal{F}_\infty^Y: E[(X - Y)I_D] = 0\}. \]
Clearly $\mathcal{D}$ is a $\pi$-class over $\Omega$, one easily checks that $\mathcal{C}$ is a $\lambda$-class over $\Omega$, and (3.62) ensures that $\mathcal{D} \subset \mathcal{C}$, whence it follows from Theorem 1.5.4 that $\sigma\{\mathcal{D}\} \subset \mathcal{C}$. In view of (3.63) we have $\mathcal{F}_t^Y \subset \mathcal{D}$ and $\mathcal{H}_t^Y \subset \mathcal{D}$, thus

$$\sigma\{\mathcal{F}_t^Y, \mathcal{H}_t^Y\} \subset \sigma\{\mathcal{D}\} \subset \mathcal{C} \subset \mathcal{F}_\infty^Y.$$ 

Now, clearly, $\sigma\{\mathcal{F}_t^Y, \mathcal{H}_t^Y\} = \mathcal{F}_\infty^Y$, and thus we have $\mathcal{C} = \mathcal{F}_\infty^Y$. In view of (3.64) it follows that $E[(X - Y)I_D] = 0$, $\forall D \in \mathcal{F}_\infty^Y$, and therefore (see Theorem 1.2.24(b)) we get (3.60).

Using Proposition 3.4.18 we easily get the main result of this section:

**Theorem 3.4.19** Suppose that $\{Y(t); t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued right-continuous process with independent increments defined on a complete probability space $(\Omega, \mathcal{F}, P)$, and put

$$(3.65) \quad \mathcal{F}_t \overset{\Delta}{=} \sigma\{\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}]\}, \quad \forall t \in [0, \infty),$$

for $\mathcal{F}_t^Y$ defined by (3.36). Then $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a standard filtration in $(\Omega, \mathcal{F}, P)$.

**Proof:** Clearly $\mathcal{F}_0$ includes all $P$-null events in $\mathcal{F}$, so it remains to show that $\{\mathcal{F}_t, t \in [0, \infty)\}$ is right-continuous. Fix some $t \in [0, \infty)$. Since $\mathcal{F}_\infty^Y \subset \mathcal{F}$, we see that $\mathcal{Z}^P[\mathcal{F}_\infty^Y] \subset \mathcal{Z}^P[\mathcal{F}]$ and hence, using Proposition 3.4.18, we get

$$\mathcal{F}_t^Y \subset \sigma\{\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}_\infty^Y]\} \subset \sigma\{\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}]\};$$

whence of course

$$(3.66) \quad \sigma\{\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}]\} = \sigma\{\mathcal{F}_t^Y, \mathcal{Z}^P[\mathcal{F}_\infty^Y]\}.$$ 

In view of (3.65), (3.66) and Proposition 3.4.14 (which uses the postulated completeness of $(\Omega, \mathcal{F}, P)$) we get $\mathcal{F}_t^Y = \mathcal{F}_t$, as required.

**Corollary 3.4.20** Suppose that $\{W(t); t \in [0, \infty)\}$ is a $d$-dimensional standard Wiener process on a complete probability space $(\Omega, \mathcal{F}, P)$, and define

$$(3.67) \quad \mathcal{F}_t \overset{\Delta}{=} \sigma\{\mathcal{F}_t^W, \mathcal{Z}^P[\mathcal{F}]\}, \quad \forall t \in [0, \infty).$$

Then $\mathcal{F}_t$ is a standard filtration in $(\Omega, \mathcal{F}, P)$.

**Remark 3.4.21** The postulate of completeness of the probability space $(\Omega, \mathcal{F}, P)$ in Proposition 3.4.14, Theorem 3.4.19, and Corollary 3.4.20 has been made to ensure that $\mathcal{Z}^P[\mathcal{F}] \subset \mathcal{F}$ and hence that $\mathcal{F}_t \subset \mathcal{F}$, $\forall t \in [0, \infty)$, for the filtrations defined by (3.45), (3.65), and (3.67), i.e. $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a filtration in $(\Omega, \mathcal{F}, P)$ in each case.

**Proposition 3.4.22** Suppose $\{Y(t); t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued process on the complete probability space $(\Omega, \mathcal{F}, P)$, and $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a filtration in $(\Omega, \mathcal{F}, P)$ defined by (3.67). If $\mathcal{F}_\infty \overset{\Delta}{=} \sigma\{\mathcal{F}_t, t \in [0, \infty)\}$ then

$$\mathcal{F}_\infty = \sigma\{\mathcal{F}_\infty^Y, \mathcal{Z}^P[\mathcal{F}]\}.$$
Proof: By the definition of $\mathcal{F}_t$, we have $\mathcal{F}_t \subset \sigma\{\mathcal{F}^Y_\infty, Z^P[\mathcal{F}]\}, \forall t \in [0,\infty)$, whence $\mathcal{F}_\infty \subset \sigma\{\mathcal{F}^Y_\infty, Z^P[\mathcal{F}]\}$. For the opposite inclusion, fix arbitrary $A \in \sigma\{\mathcal{F}^Y_\infty, Z^P[\mathcal{F}]\}$. By Proposition 3.4.7, there is some $F \in \mathcal{F}^Y_\infty$ such that $A \triangle F \in Z^P[\mathcal{F}]$. Now $\mathcal{F}^Y_t \subset \mathcal{F}_t \subset \mathcal{F}_\infty$, $\forall t \in [0,\infty)$, thus $\mathcal{F}^Y_\infty \subset \mathcal{F}$. It follows that $F \in \mathcal{F}_\infty$. Moreover, since $Z^P[\mathcal{F}] \subset \mathcal{F}_t$, $\forall t \in [0,\infty)$, we see that $Z^P[\mathcal{F}] \subset \mathcal{F}_\infty$, hence $A \triangle F \in \mathcal{F}_\infty$. But, $A = F \triangle (A \triangle F)$ (see Proposition 3.4.3(a)) hence $A \in \mathcal{F}_\infty$. \hfill \Box

3.5 Problems

Problem 3.5.1 Establish Proposition 3.1.22.

Problem 3.5.2 Suppose that $\{(X(t); t \in [0,\infty))\}$ is an $\mathbb{R}^d$-valued process on $(\Omega, \mathcal{F}, P)$ with left-continuous sample-paths. Show that the raw filtration is left-continuous: $\mathcal{F}_t = \mathcal{F}_{t-}, \forall t \in (0,\infty)$, when

$$\mathcal{F}_t \triangleq \sigma\{X_s, \ s \in [0,t]\}. \quad (3.68)$$

Problem 3.5.3 Suppose that $\{(X(t, \mathcal{F}_t); t \in [0,\infty))\}$ is a progressively measurable $\mathbb{R}$-valued process on $(\Omega, \mathcal{F}, P)$, and that $f : [0,\infty) \otimes \mathbb{R} \rightarrow \mathbb{R}$ is $\mathcal{B}([0,\infty)) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable. Put $Y_t \triangleq f(t, X_t)$, and show that $\{(Y_t, \mathcal{F}_t); t \in [0,\infty)\}$ is progressively measurable.

Hint: use Theorem 1.5.11 from notes.

Problem 3.5.4 Suppose that $\{(X(t, \mathcal{F}_t); t \in [0,\infty))\}$ is an $\mathbb{R}^d$-valued adapted process on $(\Omega, \mathcal{F}, P)$. For each $\epsilon \in (0,\infty)$ define the filtration $\{\mathcal{G}_t^\epsilon, t \in [0,\infty)\}$ by $\mathcal{G}_t^\epsilon \triangleq \mathcal{F}_{t+\epsilon}$, for all $t \in [0,\infty)$. Prove the following: If $\{(X(t, \mathcal{G}_t^\epsilon), t \in [0,\infty))\}$ is progressively measurable for each $\epsilon \in (0,\infty)$ then $\{(X(t, \mathcal{F}_t); t \in [0,\infty))\}$ is progressively measurable.

Problem 3.5.5 (a) Suppose that $\{X_t; t \in [0,\infty)\}$ is an $\mathbb{R}$-valued independent increments process on $(\Omega, \mathcal{F}, P)$. Establish the following Markov property for $\{X_t; t \in [0,\infty)\}$: For each $\Gamma \in \mathcal{B}(\mathbb{R})$ and $0 \leq t < u < \infty$ we have

$$P[X_u \in \Gamma \mid \mathcal{F}_t] = P[X_u \in \Gamma \mid \sigma\{X_t\}], \quad \text{a.s.} \quad (3.69)$$

where

$$\mathcal{F}_t \triangleq \sigma\{X_s, \ s \in [0,t]\}.$$

Hint: Use Problem 1.6.9 and Proposition 3.2.5.

(b) Now suppose that $\{X_t; t \in [0,\infty)\}$ is a standard Wiener process on $(\Omega, \mathcal{F}, P)$. Show that, for each $a \in \mathbb{R}$ and $0 \leq t < u < \infty$, we have

$$P[X_u \leq a \mid \mathcal{F}_t] = \Psi(X_t), \quad \text{a.s.}$$
where
\[ \Psi(x) \triangleq \frac{1}{\sqrt{2\pi(u-t)}} \int_{-\infty}^{a-x} \exp \left\{ \frac{-z^2}{2(u-t)} \right\} \, dz, \quad \forall \ x \in \mathbb{R}. \]

Hint: Use Problem 1.6.10 and Theorem 1.4.12.

**Problem 3.5.6** Suppose that \( \{B(t)\} \) is a scalar standard Wiener process on the probability space \((\Omega, \mathcal{F}, P)\), and put
\[ \mathcal{F}_t \triangleq \sigma\{B(s), \ 0 \leq s \leq t\} \lor \mathcal{N}(P), \quad \mathcal{G}_t \triangleq \mathcal{F}_t \lor \sigma\{B(1)\}, \]
for all \( t \in [0, \infty) \). Show that
\[ E\left[ B(t) - B(s) \mid \mathcal{G}_s \right] = \frac{t - s}{1 - s} (B(1) - B(s)), \quad 0 \leq s \leq t \leq 1. \]

Hint: Use Theorem 1.4.15(f).

**Problem 3.5.7** Suppose that \( \{\mathcal{F}_t, \ t \in [0, \infty)\} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\), that \( S \) and \( T \) are \( \{\mathcal{F}_t\} \)-stopping times, and \( X \) is an integrable random variable defined on \((\Omega, \mathcal{F}, P)\).
(a) Show that
\[ E[X \mid \mathcal{F}_T] I\{T \leq S\} = E[X \mid \mathcal{F}_{S\wedge T}] I\{T \leq S\} \quad \text{a.s.} \]
(b) Use (a) to establish that the conditional expectation operators \( E[\cdot \mid \mathcal{F}_T] \) and \( E[\cdot \mid \mathcal{F}_S] \) commute in the sense that
\[ E\left[ E[X \mid \mathcal{F}_T] \mid \mathcal{F}_S \right] = E\left[ E[X \mid \mathcal{F}_S] \mid \mathcal{F}_T \right] = E[X \mid \mathcal{F}_{S\wedge T}] \quad \text{a.s.} \]

**Problem 3.5.8** Suppose that \( \{X_t; \ t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued process on \((\Omega, \mathcal{F}, P)\), and let \( \{\mathcal{F}_t, \ t \in [0, \infty)\} \) be the raw filtration given by (3.68). If \( T: \Omega \to [0, \infty] \) is a \( \{\mathcal{F}_t\} \)-stopping time, establish the following:

(a) Suppose that \( u_1 \in [0, \infty) \) is fixed, and \( A \in \mathcal{F}_{u_1} \). If \( \omega_1 \in A \) and \( X_s(\omega_1) = X_s(\omega_2), \forall \ s \in [0, u_1] \), then \( \omega_2 \in A \). Hint: use Theorem 1.3.22.

(b) If \( \omega_1, \omega_2 \in \Omega \) are such that \( T(\omega_1) < \infty \) and \( X_s(\omega_1) = X_s(\omega_2), \forall \ s \in [0, T(\omega_1)] \), then \( T(\omega_1) = T(\omega_2) \). Hint: use result in (a).

(c) If \( \omega_1, \omega_2 \in \Omega \) are such that \( T(\omega_1) = \infty \) and \( X_s(\omega_1) = X_s(\omega_2), \forall \ s \in [0, \infty) \), then \( T(\omega_2) = \infty \).

(d) Suppose that \( A \in \mathcal{F}_T \). If \( \omega_1 \in A \) and \( \omega_2 \in \Omega \) are such that \( T(\omega_1) < \infty \) and \( X_s(\omega_1) = X_s(\omega_2), \forall \ s \in [0, T(\omega_1)] \), then \( T(\omega_1) = T(\omega_2) \) and \( \omega_2 \in A \). Hint: use Theorem 1.3.22.

(e) Suppose that \( A \in \mathcal{F}_T \). If \( \omega_1 \in A \) and \( \omega_2 \in \Omega \) are such that \( T(\omega_1) = \infty \) and \( X_s(\omega_1) = X_s(\omega_2), \forall \ s \in [0, \infty) \), then \( T(\omega_2) = \infty \) and \( \omega_2 \in A \).
Problem 3.5.9 Suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty) \} \) and \( \{(A(t), \mathcal{F}_t); t \in [0, \infty) \} \) are continuous \( IR \)-valued adapted processes on \((\Omega, \mathcal{F}, P)\), null at the origin with \( A(t) \geq 0, X(t) \geq 0 \), and the sample-paths \( t \to A(t) : [0, \infty) \to [0, \infty) \) are non-decreasing, \( \forall \omega \). Put \( X^*(t) \overset{\Delta}{=} \max_{0 \leq s \leq t} X(s), \forall t \in [0, \infty), X^*(\infty) \overset{\Delta}{=} \sup_{s \in [0, \infty)} X(s) \), and \( A(\infty) \overset{\Delta}{=} \lim_{t \to \infty} A(t) \). If
\[
E[X(U)] \leq E[A(U)]
\]
for each bounded \( \{\mathcal{F}_t\} \)-stopping time \( U \), establish the following Lenglart-Rebolledo inequalities:

(a) \( P[X^*(T) \geq \epsilon] \leq \epsilon^{-1} E[A(T)] \)

Hint: consider the random variable \( X(T \land S_\epsilon)I\{X^*(T) \geq \epsilon\} \) where \( S_\epsilon \overset{\Delta}{=} \inf\{t \in [0, \infty) : X(s) \geq \epsilon\} \).

(b) \( P[X^*(T) \geq \epsilon, A(T) < \delta] \leq \epsilon^{-1} E[A(T) \land \delta] \)

(c) \( P[X^*(T) \geq \epsilon] \leq \epsilon^{-1} E[A(T) \land \delta] + P[A(T) \geq \delta] \).

Problem 3.5.10 Suppose that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\) and \( T \) is a \( \{\mathcal{F}_t\} \)-stopping time.

(a) For an event \( A \in \mathcal{F}_T \) define
\[
T_A(\omega) \overset{\Delta}{=} \begin{cases} 
T(\omega) & \text{for all } \omega \in A \\
+\infty & \text{for all } \omega \notin A.
\end{cases}
\]
Show that \( T_A \) is a \( \{\mathcal{F}_t\} \)-stopping time.

(b) Let \( \mathcal{C} \) be the collection of all events in \( \mathcal{F}_0 \) as well as all events of the form
\[
A \cap \{T > t\} \quad \text{for each } t \in [0, \infty) \text{ and each } A \in \mathcal{F}_t,
\]
and put
\[
\mathcal{F}_{T-} \overset{\Delta}{=} \sigma(\mathcal{C}),
\]
(\( \mathcal{F}_{T-} \) is called the pre-\( \sigma \)-algebra of events strictly prior to the stopping time \( T \)). Establish the following:

(i) \( T \) is \( \mathcal{F}_{T-} \)-measurable and \( \mathcal{F}_{T-} \subset \mathcal{F}_T \).

(ii) if \( S \) is a \( \{\mathcal{F}_t\} \)-stopping time with pre-\( \sigma \)-algebra \( \mathcal{F}_S \) (recall Definition 3.3.11) then \( A \cap \{S < T\} \in \mathcal{F}_{T-} \) for each \( A \in \mathcal{F}_S \).

Problem 3.5.11 Let \( \Omega \overset{\Delta}{=} C[0, \infty) \) (the set of all continuous functions \( \omega : [0, \infty) \to IR \)), for each \( t \in [0, \infty) \) define \( X_t : \Omega \to IR \) by
\[
X_t(\omega) \overset{\Delta}{=} \omega(t), \quad \forall \omega \in \Omega,
\]
and put
\[
\mathcal{F}_{T-} \overset{\Delta}{=} \sigma(\mathcal{C}),
\]
(\( \mathcal{F}_{T-} \) is called the pre-\( \sigma \)-algebra of events strictly prior to the stopping time \( T \)). Establish the following:

(i) \( T \) is \( \mathcal{F}_{T-} \)-measurable and \( \mathcal{F}_{T-} \subset \mathcal{F}_T \).

(ii) if \( S \) is a \( \{\mathcal{F}_t\} \)-stopping time with pre-\( \sigma \)-algebra \( \mathcal{F}_S \) (recall Definition 3.3.11) then \( A \cap \{S < T\} \in \mathcal{F}_{T-} \) for each \( A \in \mathcal{F}_S \).

Problem 3.5.11 Let \( \Omega \overset{\Delta}{=} C[0, \infty) \) (the set of all continuous functions \( \omega : [0, \infty) \to IR \)), for each \( t \in [0, \infty) \) define \( X_t : \Omega \to IR \) by
\[
X_t(\omega) \overset{\Delta}{=} \omega(t), \quad \forall \omega \in \Omega,
\]
and put
\[ \mathcal{F} \triangleq \sigma\{X_s, \ s \in [0, \infty)\}, \quad \mathcal{F}_t \triangleq \sigma\{X_s, \ s \in [0, t]\}. \]

In this way we get a continuous adapted process \( \{(X_t, \mathcal{F}_t); \ t \in [0, \infty)\} \) on \((\Omega, \mathcal{F})\).

(a) Show by example that, in general, \( \mathcal{F}_t \neq \mathcal{F}_{t+} \) i.e. \( \mathcal{F}_t \) is a strict subset of \( \mathcal{F}_{t+} \).

Hint: If \( \Gamma \triangleq (1, \infty) \) then Example 3.3.7(a) shows that the début \( D_\Gamma \) is a \( \{\mathcal{F}_t\}\)-optional time; now suppose that it is a \( \{\mathcal{F}_t\}\)-stopping time, and use Problem 3.5.8(b) to establish a contradiction.

(b) For each \( \omega \in \Omega \) define \( \hat{\omega}_{T(\omega)} \in \Omega \) as follows: when \( T(\omega) < \infty \) then \( \hat{\omega}_{T(\omega)}(t) \triangleq \omega(t), \ \forall \ t \in [0, T(\omega)] \), and \( \hat{\omega}_{T(\omega)}(t) \triangleq \omega(T(\omega)), \ \forall \ t \in (T(\omega), \infty) \). On the other hand, when \( T(\omega) = +\infty \), then \( \hat{\omega}_{T(\omega)}(t) \triangleq \omega(t), \ \forall \ t \in [0, \infty) \). Now establish the following: if \( A \in \mathcal{F}_T \) then
\[ I_A(\omega) = I_A(\hat{\omega}_{T(\omega)}), \quad \forall \ \omega \in \Omega. \]

Finally, use this identity to establish the following Galmarino representation of the pre-\( \sigma \)-algebra \( \mathcal{F}_T \):
\[ \mathcal{F}_T = \sigma\{X_{s\wedge T}, \ s \in [0, \infty)\}. \]
Chapter 4

Continuous Parameter-Martingales

Our goal in this chapter is to present that part of the theory of continuous-parameter martingales which is essential for an introduction to stochastic integration. Our primary concerns are twofold. First, we shall generalize the main results of Chapter 2, such as the optional sampling theorem and the martingale convergence theorems, to the continuous-parameter context, using as basic tools the elements of continuous-parameter stochastic processes set forth in Chapter 3. Then we shall define the so-called quadratic variation process of a given continuous-parameter martingale and establish the principal properties of this process.

4.1 Basic Definitions

The definition which follows is an obvious analogue of Definition 2.2.1:

**Definition 4.1.1** A real-valued adapted process \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) on some probability space \((\Omega, \mathcal{F}, P)\) is a continuous-parameter supermartingale when
(a) \( E|X_t| < \infty, \forall t \in [0, \infty) \), and
(b) \( E[X_t | \mathcal{F}_s] \leq X_s \) a.s. \( \forall s, t \in [0, \infty) \) such that \( s < t \);

and is a continuous-parameter submartingale when (a) holds and (b) is replaced with
(c) \( E[X_t | \mathcal{F}_s] \geq X_s \) a.s. \( \forall s, t \in [0, \infty) \) such that \( s < t \).

Finally, \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a continuous-parameter martingale when it is both a continuous-parameter supermartingale and submartingale, that is when (a) holds and \( E[X_t | \mathcal{F}_s] = X_s \) a.s. \( \forall s, t \in [0, \infty) \) such that \( s < t \).

**Remark 4.1.2** If \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a supermartingale then we see from Definition 4.1.1 that \( E[X_s] \geq E[X_t] \forall s, t \in [0, \infty), s \leq t \), the two expectations being equal when \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a martingale. Moreover, an obvious modification of the proof of Proposition 2.2.4 establishes that, if \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a supermartingale with constant expectation, namely \( E[X_0] = E[X_t], \forall t \in [0, \infty) \), then it is in fact a martingale.
Example 4.1.3 Suppose \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a filtration in a probability space \((\Omega, \mathcal{F}, P)\), and \( \tilde{P} \) is a probability measure on \((\Omega, \mathcal{F})\) with \( \tilde{P} \ll P[\mathcal{F}] \). Put
\[
X_t \triangleq \left. \frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_t}, \quad \forall t \in [0, \infty).
\]
Exactly as in Example 2.2.6 it follows that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale on \((\Omega, \mathcal{F}, P)\). Later we shall see that this martingale is the starting point for an essential result in stochastic calculus called the Girsanov theorem.

Remark 4.1.4 When \( p \in [1, \infty) \), then \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is called an \( L^p\)-supermartingale [submartingale, martingale] when it is both a supermartingale [submartingale, martingale] and \( E|X_t|^p < \infty \) for each \( t \in [0, \infty) \). If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^p\)-supermartingale [submartingale, martingale] such that \( \sup_{t \in [0, \infty)} E|X_t|^p < \infty \) then it is said to be an \( L^p\)-bounded supermartingale [submartingale, martingale].

An argument identical to that used for Proposition 2.2.8 establishes the following result for \( L^2\)-martingales:

Proposition 4.1.5 Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^2\)-martingale on the probability space \((\Omega, \mathcal{F}, P)\). Then the following hold:
(a) When \( s, t, u, v \in [0, \infty) \) with \( s < t \leq u < v \), then \( E[(X_t - X_s)(X_v - X_u)] = 0 \);
(b) When \( s, t \in [0, \infty) \) with \( s \leq t \) then \( E[(X_t - X_s)^2] = E[X_t^2 - X_s^2] \).

Remark 4.1.6 Suppose that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\). We shall use the following notation to indicate various classes of martingales:
(a) \( \mathcal{M}(\{\mathcal{F}_t\}, P) \) denotes the set of all \( \mathbb{R}\)-valued processes \( \{X_t; t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) such that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale. We include reference to the ambient probability measure \( P \) in this notation because \( P \) determines the conditional expectations which arise in the definition of a martingale, and because later we shall be exchanging the measure \( P \) for other probability measures on \((\Omega, \mathcal{F})\).
(b) For any \( p \in [1, \infty) \), \( \mathcal{M}_p(\{\mathcal{F}_t\}, P) [\mathcal{M}_{p,b}(\{\mathcal{F}_t\}, P)] \) will denote the set of all \( \mathbb{R}\)-valued processes \( \{X_t; t \in [0, \infty)\} \) such that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^p\)-martingale [\( L^p\)-bounded martingale] on \((\Omega, \mathcal{F}, P)\).
(c) Our study of stochastic calculus will focus particularly on martingales with continuous sample-paths. For later use we define the notation \( \mathcal{M}^c(\{\mathcal{F}_t\}, P) [\mathcal{M}_{p}^{c,0}(\{\mathcal{F}_t\}, P), \mathcal{M}_{p,b}^{c,0}(\{\mathcal{F}_t\}, P)] \) to indicate the set of all \( \mathbb{R}\)-valued processes \( \{X_t; t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) such that \( t \rightarrow X_t(\omega) \) is continuous on \([0, \infty)\), \( \forall \omega \in \Omega \), and \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale [\( L^p\)-martingale, \( L^p\)-bounded martingale].
(d) Extending the notation in (c), we write \( \mathcal{M}^{c,0}(\{\mathcal{F}_t\}, P) [\mathcal{M}_{p}^{c,0}(\{\mathcal{F}_t\}, P), \mathcal{M}_{p,b}^{c,0}(\{\mathcal{F}_t\}, P)] \) for the set of all members of \( \mathcal{M}^c(\{\mathcal{F}_t\}, P) [\mathcal{M}^c_p(\{\mathcal{F}_t\}, P), \mathcal{M}^c_{p,b}(\{\mathcal{F}_t\}, P)] \) which are null at the origin.
Proposition 4.1.8 has a simple generalization to the multidimensional setting. In the following proposition we write \( X \in M(\{\mathcal{F}_t\}, P) [X \in M_p(\{\mathcal{F}_t\}, P), X \in M_{p,b}(\{\mathcal{F}_t\}, P), \text{etc.}] \) to indicate that \( \{X_t; t \in [0, \infty)\} \) is a member of \( M(\{\mathcal{F}_t\}, P) [M_p(\{\mathcal{F}_t\}, P), M_{p,b}(\{\mathcal{F}_t\}, P), \text{etc.}] \).

In the the various classes of martingales formulated inRemark 4.1.6, the superscripts such as \( c, \) and \( c, 0, \) refer to the sample-path properties of the martingale (continuous, continuous and null at the origin), while the subscripts \( p, \) and \( p, b \) indicate its \( L^p \)-boundedness properties. Later, when we introduce the so-called local martingales, we shall extend this notation a bit further (see Remark 4.6.5).

**Remark 4.1.7** A routine application of the linearity of conditional expectations, together with the Minkowski inequality (see Theorem 1.2.26), readily shows that each of these classes of martingales is a real vector space with the usual pointwise notions of vector addition and scalar multiplication. In each case the zero element of the real vector space is the process \( \{X(t), \mathcal{F}_t\}; t \in [0, \infty) \) with \( X(t, \omega) = 0, \forall t \in [0, \infty), \forall \omega \in \Omega. \)

The next result illustrates how typical continuous-parameter martingales arise from the Wiener process:

**Proposition 4.1.8** Suppose \( \{W(t), \mathcal{F}_t\}; t \in [0, \infty) \) is a scalar standard Wiener process on a probability space \( (\Omega, \mathcal{F}, P) \). Then \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{W(t)^2 - t, \mathcal{F}_t); t \in [0, \infty)\} \) are continuous martingales.

**Proof:** Fix \( s, t \in [0, \infty) \) with \( s < t \). By Definition 3.2.11(iii) we see that \( W(t) - W(s) \) and \( \mathcal{F}_s \) are independent, and \( E[W(t) - W(s)] = 0 \). Therefore, by Theorem 1.4.15(f), we have

\[
E\left[ W(t) - W(s) \mid \mathcal{F}_s \right] = 0 \quad \text{a.s.}
\]

which shows that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale. It then follows that

\[
E\left[ W(s)W(t) \mid \mathcal{F}_s \right] = W(s)E\left[ W(t) \mid \mathcal{F}_s \right] = W(s)^2 \quad \text{a.s.}
\]

and hence

\[
(4.1) \quad E\left[ W(t)^2 - W(s)^2 \mid \mathcal{F}_s \right] = E\left[ \{W(t) - W(s)\}^2 \mid \mathcal{F}_s \right] \quad \text{a.s.}
\]

Now, by Definition 3.2.11(iii), we see that \( \{W(t) - W(s)\}^2 \) and \( \mathcal{F}_s \) are independent, hence

\[
(4.2) \quad E\left[ \{W(t) - W(s)\}^2 \mid \mathcal{F}_s \right] = E[\{W(t) - W(s)\}^2] = t - s.
\]

Combining (4.1) and (4.2) shows that \( \{(W(t)^2 - t, \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale. \( \square \)

**Remark 4.1.9** If \( W \overset{\Delta}{=} \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process on \( (\Omega, \mathcal{F}, P) \) then clearly \( W \) is an \( L^p \)-martingale for each \( p \in [1, \infty) \). In particular \( W \in M^{<0}_2(\{\mathcal{F}_t\}, P) \). However, we see from Remark 3.2.7 that \( E[W(t)^2] = t, \forall t \in [0, \infty) \), hence \( W \) is not a member of \( M^{<0}_{2,b}(\{\mathcal{F}_t\}, P) \).

Proposition 4.1.8 has a simple generalization to the multidimensional setting. In the following proposition we use \( \delta_{j,k}, j, k = 1, 2, \ldots, d \), to denote the familiar Kronecker delta, namely \( \delta_{j,k} \overset{\Delta}{=} 1 \) when \( j = k \) and \( \delta_{j,k} \overset{\Delta}{=} 0 \) when \( j \neq k \):
**Proposition 4.1.10** Suppose \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a d-dimensional standard Wiener process on a probability space \((\Omega, \mathcal{F}, P)\), with \( W(t) \overset{\Delta}{=} (W^1(t), W^2(t), \ldots, W^d(t)) \), \( \forall t \in [0, \infty) \). Then \( \{(W^k(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(W^j(t)W^k(t) - t \delta_{j,k}, \mathcal{F}_t); t \in [0, \infty)\} \) are continuous martingales for each \( j, k \in \{1, 2, \ldots d\} \).

**Proof:** Fix indices \( j, k \).

At present, Propositions 4.1.8 and 4.1.10 serve merely to show some of the ways in which continuous martingales arise from a Wiener process. Later, when we are familiar with the notion of the quadratic variation process introduced in Section 4.7, these propositions will take on a greater significance.

Suppose that the following conditions hold:

Proposition 4.1.13

Suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale [resp. submartingale], and \( c : \mathbb{R} \to \mathbb{R} \) is a convex [resp. convex and non-decreasing] function such that \( E|c(X_t)| < \infty \) for all \( t \in [0, \infty) \). Then \( \{(c(X_t), \mathcal{F}_t); t \in [0, \infty)\} \) is a submartingale.

Proposition 4.1.14

Suppose that the following conditions hold:

(a) \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\), and \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \);

(b) for some constant \( p \in [1, \infty) \) the process \( \{(X_n(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^p \)-martingale on \((\Omega, \mathcal{F}, P)\) for each \( n = 0, 1, 2, \ldots \);

(c) for each \( t \in [0, \infty) \) the sequence of random variables \( \{X_n(t), n = 0, 1, 2, \ldots\} \) converges in \( L^p(\Omega, \mathcal{F}, P) \) to a random variable \( X(t) \) as \( n \to \infty \).

Then \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^p \)-martingale.
Proof: By the Markov inequality (Theorem 1.2.23) and (c) one sees that \( X_n(t) \) converges in probability to \( X(t) \) for each \( t \in [0, \infty) \), hence Proposition 3.1.19 establishes that \( \{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is an adapted process. It remains to check the martingale property. Fix \( s, t \in [0, \infty) \) such that \( t > s \geq 0 \). Then
\[
E[X_n(t) \mid \mathcal{F}_s] = X_n(s) \quad \text{a.s.} \tag{4.6}
\]
for all \( n = 0, 1, 2, \ldots \) Now, by (c), the right side of (4.6) converges in \( L^p \) to \( X_s(s) \) a.s. (4.6) for all \( n \). Finally, observe from (c) that \( X(t) \in L^p(\Omega, \mathcal{F}, P) \), \( \forall t \in [0, \infty) \), showing that \( \{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is an \( L^p \)-martingale.

\[\square\]

### 4.2 Sample Path Properties

In this section our goal is to study the sample path properties of submartingales. We shall see that a continuous-parameter submartingale always has some modification with particularly well-behaved sample paths, provided that the associated filtration is a standard filtration (Definition 3.4.2). This in turn will turn out to be crucial when we establish continuous-parameter analogues of the main results on martingales which were developed in the discrete-parameter context in Chapter 2. In order to study the sample paths of continuous parameter martingales we shall need some preliminary ideas which have nothing to do with probability and stochastic processes, but which belong to the general realm of real analysis. We begin this section by summarizing these ideas.

**Remark 4.2.1** In the following definitions we shall let \( S \) denote a fixed dense subset of \([0, \infty)\). Typically, \( S \) will stand either for the set of all rational numbers in \([0, \infty)\), which we henceforth denote by \( Q_+ \), or for \([0, \infty)\) itself. We begin by formulating a weak form of boundedness:

**Definition 4.2.2** A mapping \( \varphi : S \to \mathbb{R} \) is **bounded on bounded intervals** when, for each positive integer \( n \), there exists some \( B_n \in [0, \infty) \) such that \( |\varphi(s)| < B_n \) for all \( s \in S \cap [0, n] \).

Thus, for example, the mapping \( \varphi(s) \triangleq s^2, s \in [0, \infty) \), is clearly bounded on bounded intervals, whereas \( \varphi(s) \triangleq (1-s)^{-1}, s \in [0, 1) \cup (1, \infty), \varphi(1) \triangleq 0 \), fails to be so bounded.

**Definition 4.2.3** A mapping \( \varphi : S \to \mathbb{R} \) has a **finite left limit** at \( t \in (0, \infty) \) if there exists some \( \xi \in \mathbb{R} \) such that the following holds: corresponding to each \( \epsilon > 0 \) there exists some \( \delta(\epsilon) > 0 \) such that \( |\varphi(s) - \xi| < \epsilon \) for all \( s \in (t - \delta, t) \cap S \).

**Remark 4.2.4** Consider the important special case where \( S \) is some strict subset of \([0, \infty)\) such as \( S \triangleq Q_+ \). The definition specifies the notion of a left limit at any \( t \in (0, \infty) \) and not just at values of \( t \) in \( S \), the domain of \( \varphi \). The definition is meaningful because \( S \) is dense in \([0, \infty)\). The number \( \xi \) in (Definition 4.2.3) is denoted by \( \varphi(t-) \) and called the **left limit** of \( \varphi \) at \( t \).
We next introduce the concepts of left inferior and left superior limit for a mapping $\varphi : S \to \mathbb{R}$. For each $t \in (0, \infty)$ and $\delta > 0$ put

$$
\psi(\delta) \overset{\Delta}{=} \inf_{s \in (t-\delta, t) \cap S} \varphi(s)
$$

Clearly $\psi(\delta)$ is non-decreasing as $\delta \downarrow 0$, hence $\lim_{\delta \downarrow 0} \psi(\delta)$ exists in $\overline{\mathbb{R}}$. We call this limit the **left inferior limit** of $\varphi$ at $t$ and denote it by $\liminf_{s \uparrow t} \varphi(s)$:

$$
\liminf_{s \uparrow t} \varphi(s) \overset{\Delta}{=} \lim_{\delta \downarrow 0} \left[ \inf_{s \in (t-\delta, t) \cap S} \varphi(s) \right]
$$

Similarly, we define the **left superior limit** at $t \in (0, \infty)$ by

$$
\limsup_{s \uparrow t} \varphi(s) \overset{\Delta}{=} \lim_{\delta \downarrow 0} \left[ \sup_{s \in (t-\delta, t) \cap S} \varphi(s) \right];
$$

in this case the function of $\delta > 0$ in square brackets on the right hand side is non-increasing as $\delta \downarrow 0$, and the limit again exists in $\overline{\mathbb{R}}$. The double upwards arrow in $s \uparrow t$ indicates that $s$ must be strictly less than $t$ in the above infimum and supremum.

**Example 4.2.5** In this example we shall identify $S$ with $Q_+$.  

(a) put $\varphi_1(s) \overset{\Delta}{=} (1 - s)^{-1}$, $\forall$ $0 \leq s < 1$ and $\varphi_1(s) \overset{\Delta}{=} 0 \forall$ $s > 1$, $s \in Q_+$. Then $\liminf_{s \uparrow 1} \varphi_1(s) = \limsup_{s \uparrow 1} \varphi_1(s) = +\infty$.

(b) put $\varphi_2(s) \overset{\Delta}{=} \sin[(1 - s)^{-1}]$, $\forall$ $0 \leq s < 1$ and $\varphi_2(s) \overset{\Delta}{=} 0 \forall s > 1$, $s \in Q_+$. Then $\liminf_{s \uparrow 1} \varphi_2(s) = -1$ and $\limsup_{s \uparrow 1} \varphi_2(s) = +1$.

(c) put $\varphi_3(s) \overset{\Delta}{=} \varphi_1(s)\varphi_2(s)$, $\forall s \in Q_+$. Then $\liminf_{s \uparrow t} \varphi_3(s) = -\infty$ and $\limsup_{s \uparrow t} \varphi_3(s) = +\infty$.

The next proposition lists some rather obvious properties of left inferior and superior limits, but is stated in full because it is useful to have such properties clearly articulated for later use. The proof, which involves just a manipulation and restatement of the basic definitions, is left as an exercise:

**Proposition 4.2.6** Suppose that $S$ is some fixed dense subset of $[0, \infty)$, and $\varphi : S \to \mathbb{R}$ is a given mapping. Then, for each $t \in (0, \infty)$, we have:

(a) $\liminf_{s \uparrow t} \varphi(s) \leq \limsup_{s \uparrow t} \varphi(s)$;

(b) if $\liminf_{s \uparrow t} \varphi(s) < a < b < \limsup_{s \uparrow t} \varphi(s)$ then there exist infinite sequences $\{s_k, k = 0, 1, 2, \ldots\}$ and $\{\overline{s}_k, k = 0, 1, 2, \ldots\}$ in $[0, t) \cap S$, both converging towards $t$ from the left, such that $\varphi(s_k) < a < b < \varphi(\overline{s}_k) \forall k = 0, 1, 2, \ldots$

(c) $\varphi$ has a finite left limit at $t$ if and only if

$$
\liminf_{s \uparrow t} \varphi(s) = \limsup_{s \uparrow t} \varphi(s) \overset{\Delta}{=} \xi
$$

for some $\xi \in \mathbb{R}$. In this case $\varphi(t-) = \xi$. 

133
Remark 4.2.7  So far we have considered finite left limits along with left inferior and superior limits. Of course, one can formulate precisely analogous concepts from the right rather than the left. Thus, suppose again that \( S \) denotes some fixed subset of \([0, \infty)\) which is dense in \([0, \infty)\). The mapping \( \varphi : S \to \mathbb{R} \) is said to have a **finite right limit** at \( t \in [0, \infty) \) when there exists some \( \zeta \in \mathbb{R} \) such that the following holds: for each \( \epsilon > 0 \) there exists some \( \delta > 0 \) such that \( |\varphi(s) - \zeta| < \epsilon \) for all \( s \in (t, t + \delta) \cap S \). The number \( \zeta \) is denoted by \( \varphi(t+) \) and is called the **right limit** of \( \varphi \) at \( t \). Likewise, given a mapping \( \varphi : S \to \mathbb{R} \), we can define right inferior and superior limits in the obvious way:

\[
\liminf_{s \downarrow t} \varphi(s) \triangleq \lim_{\delta \downarrow 0} \inf_{s \in (t, t + \delta) \cap S} \varphi(s)
\]

and

\[
\limsup_{s \downarrow t} \varphi(s) \triangleq \lim_{\delta \downarrow 0} \sup_{s \in (t, t + \delta) \cap S} \varphi(s).
\]

One can then formulate an exact analogue of Proposition 4.2.6, a simple task which is left to the reader.

Remark 4.2.8  Recalling Remark 4.2.1, let \( S \) be some fixed dense subset of \([0, \infty)\). A mapping \( \varphi : S \to \mathbb{R} \) which has a finite left limit at each \( t \in (0, \infty) \) and a finite right limit at each \( t \in [0, \infty) \) will henceforth be called a **regulated function**. A regulated function is well behaved, in the sense that it does not show any of the pathologies demonstrated in Example 4.2.5. We are going to show a relationship between the sample paths of a continuous-parameter submartingale and regulated functions. In order to do so we must extend the notion of upcrossings, developed in § 2.5 for sequences of real numbers, to real-valued mappings defined on \( Q_+ \), the set of non-negative rational numbers. To this end, fix some \( T \in (0, \infty) \) and let \( \pi \triangleq \{s_0, s_1 \ldots s_N\} \) be some finite strictly increasing sequence of real numbers \( s_j \in [0, T] \cap Q_+ \) i.e. \( 0 \leq s_j < s_k \leq T \) \( \forall j, k = 0, 1, 2, \ldots, N \) such that \( j < k \). We denote by \( \Pi_T \) the collection of all such finite-length strictly increasing sequences in \([0, T] \cap Q_+\). Suppose that \( \varphi : Q_+ \to \mathbb{R} \) is a given mapping, and for real numbers \( a, b \) with \( a < b \) and \( \pi \triangleq \{s_0, s_1 \ldots s_N\} \in \Pi_T \), let \( U_{\pi}(a, b; \varphi) \) denote the number of upcrossings of \([a, b]\) by the sequence of real numbers \( \{\varphi(s_0), \varphi(s_1) \ldots \varphi(s_N)\} \), and put

\[
(4.7) \quad U_{[0,T] \cap Q_+}(a, b; \varphi) \triangleq \sup_{\pi \in \Pi_T} U_{\pi}(a, b; \varphi).
\]

The significance of the number \( U_{[0,T] \cap Q_+}(a, b; \varphi) \) is shown by the next result:

**Proposition 4.2.9**  Suppose that \( \varphi : Q_+ \to \mathbb{R} \) is bounded on bounded intervals. If

\[
(4.8) \quad U_{[0,n] \cap Q_+}(a, b; \varphi) < \infty
\]

for all \( n = 1, 2, \ldots \) and all rationals \( a, b \) with \( a < b \), then \( \varphi \) is a regulated function.

Proof: We argue by contradiction and suppose that \( \varphi \) is not a regulated function. Then the existence of a finite left or right limit fails somewhere in \([0, \infty)\); for the sake of argument suppose that \( \varphi \) fails to
have a finite left limit at some $t \in (0, \infty)$. From Proposition 4.2.6 (c) we see that one of two possibilities must occur, namely (i) $\liminf_{s \uparrow t} \varphi(s)$ and $\limsup_{s \uparrow t} \varphi(s)$ are equal, with their common value being either $+\infty$ or $-\infty$, or (ii) $\liminf_{s \uparrow t} \varphi(s) < \limsup_{s \uparrow t} \varphi(s)$. Fix some integer $n > t$. Since $\varphi$ is bounded on bounded intervals, there is some $B_n \in (0, \infty)$ such that $|\varphi(s)| < B_n$ for all $s \in Q_+ \cap [0, n]$. This rules out the occurrence of (i), and thus (ii) must occur. Thus, we can find rationals $a, b$ with $a < b$ such that

$$\liminf_{s \uparrow t} \varphi(s) < a < b < \limsup_{s \uparrow t} \varphi(s).$$

Hence, from Proposition 4.2.6 (b), there exist infinite sequences $\{s_k\}$ and $\{\tau_k\}$ in $[0, t) \cap Q_+$, converging to $t$ from the left, such that $\varphi(s_k) < a$ and $\varphi(\tau_k) > b$ for all $k = 0, 1, 2, \ldots$. From these sequences we can clearly find an infinite sequence of pairs $\{(\sigma_k, \tau_k)\}$ of rational numbers $\sigma_k, \tau_k$ such that $0 \leq \sigma_k < \tau_k < \sigma_{k+1} < \tau_{k+1} < t$ and $\varphi(\sigma_k) < a$, $\varphi(\tau_k) > b$ for all $k = 0, 1, 2, \ldots$. It follows that $U_{[0,n] \cap Q_+} (a, b; \varphi) = \infty$ which contradicts (4.8). This establishes that $\varphi$ has a finite left limit at each $t \in (0, \infty)$. An exactly analogous argument shows that $\varphi$ must also have a finite right limit at each $t \in [0, \infty)$.

**Remark 4.2.10** Proposition 4.2.9 has a converse, namely if $\varphi : Q_+ \to IR$ is a regulated function then it is bounded on bounded intervals and (4.8) holds for all integers $n = 1, 2, \ldots$ and rationals $a, b$ with $a < b$. We shall not prove this converse assertion since we make no use of it in the sequel.

We are now going to look at the sample paths of a submartingale $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ with a view to showing that the mappings $s \to X_s(\omega) : Q_+ \to IR$ (i.e. the restriction of the sample paths to $Q_+$) are regulated functions. To this end it is useful to have a more explicit characterization of upcrossings than that given in Remark 4.2.8. This characterization is developed next:

**Remark 4.2.11** Fix some $T \in (0, \infty)$ and let $\{r_n, n = 0, 1, 2, \ldots\}$ be any fixed enumeration of the countable set of numbers $[0, T] \cap Q_+$. For each $N = 1, 2, \ldots$ let $\pi_N \overset{\Delta}{=} \{s_0^N, s_1^N, \ldots, s_N^N\}$ be the sequence obtained by re-arranging the set of numbers $\{r_0, r_1, \ldots, r_N\}$ in increasing order, so that $s_j^N < s_k^N$ when $j < k$. Clearly $\pi_N$ is a finite-length sequence of numbers in $[0, T] \cap Q_+$ with $\pi_N \subseteq \pi_{N+1}$. Moreover, it is easily seen that

$$\bigcup_{1 \leq N < \infty} \pi_N = [0, T] \cap Q_+,$$

and therefore

$$U_{[0,T] \cap Q_+} (a, b; \varphi) = \lim_{N \to \infty} U_{\pi_N} (a, b; \varphi)$$

where the left-hand quantity is defined by (4.7) and the convergence on the right-hand side is monotonic increasing.

**Remark 4.2.12** Fix some $T \in (0, \infty)$ and suppose that $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ is an adapted process on the probability space $(\Omega, \mathcal{F}, P)$. If $\pi_N \overset{\Delta}{=} \{s_0^N, s_1^N, \ldots, s_N^N\}$ is the finite ordered sequence constructed according to the recipe in Remark 4.2.11, then $\{(X_s, \mathcal{F}_s); s \in \pi_N\}$ is clearly an adapted sequence. We
shall use \( U_{\pi_N}(a, b; X(\cdot, \omega)) \) to denote the number of upcrossings of the interval \([a, b]\) by the finite sequence of real numbers \( \{X_s(\omega), s \in \pi_N\} \). From Remark 2.5.3 we see that the mapping \( \omega \to U_{\pi_N}(a, b; X(\cdot, \omega)) \) is \( F_N \)-measurable and hence \( F_T \)-measurable since \( s_N \leq T \). If we let \( U_{[0,T]\cap Q}(a, b; X(\cdot, \omega)) \) be the quantity on the left hand side of (4.7) when \( \varphi \) is identified with the restricted sample path \( s \to X(s, \omega) : Q_+ \to \mathbb{R} \), then we see from comparison with (4.9) that

\[
U_{[0,T]\cap Q}(a, b; X(\cdot, \omega)) = \lim_{N \to \infty} U_{\pi_N}(a, b; X(\cdot, \omega))
\]

for each \( \omega \). It follows that the quantity on the left-hand side of (4.10), viewed as a function of \( \omega \), is \( F_T \)-measurable. Next, define

\[
\Omega_1^* \triangleq \bigcap_{1 \leq n < \infty} \bigcap_{a, b} \{\omega : U_{[0,n]\cap Q}(a, b; X(\cdot, \omega)) < \infty\}
\]

where the intersections are over all positive integers \( n \) and rational numbers \( a \) and \( b \) with \( a < b \). We see, from the \( F_T \)-measurability of the function of \( \omega \) in (4.10), that \( \Omega_1^* \in F_\infty \). Also, define

\[
\Omega_2^* \triangleq \bigcap_{1 \leq n < \infty} \{\omega : \sup_{s \in Q_+ \cap [0,n]} |X(s, \omega)| < \infty\}
\]

Since the function of \( \omega \) deriving from the supremum in (4.12) is \( F_n \)-measurable (the supremum is taken over only countably many \( F_n \)-measurable random variables, since \( Q_+ \cap [0,n] \) is a countable set) we see that \( \Omega_2^* \in F_\infty \). Thus \( \Omega_1^* \cap \Omega_2^* \in F_\infty \), hence has a well defined probability. In view of Proposition 4.2.9 we see that if \( \omega \in \Omega_1^* \cap \Omega_2^* \) then the mapping \( s \to X(s, \omega) : Q_+ \to \mathbb{R} \) is a regulated function. We shall now show that \( P(\Omega_1^* \cap \Omega_2^*) = 1 \) when \( \{(X_t, F_t); t \in [0, \infty)\} \) is a submartingale.

**Proposition 4.2.13** Suppose that \( \{(X_t, F_t); t \in [0, \infty)\} \) is a submartingale on the probability space \((\Omega, F, P)\). Then there is some event \( \Omega^* \in F_\infty \) such that \( P(\Omega^*) = 1 \), and the mapping \( s \to X_s(\omega) : Q_+ \to \mathbb{R} \) is a regulated function for each \( \omega \in \Omega^* \). Moreover, \( \Omega^* = \Omega_1^* \cap \Omega_2^* \) where \( \Omega_1^* \) and \( \Omega_2^* \) are defined by (4.11) and (4.12) respectively.

Proof: Fix some arbitrary \( T \in (0, \infty) \). For each \( N = 1, 2, \ldots \) let \( \pi_N \triangleq \{s_N^0, s_N^1, \ldots s_N^N\} \) be the finite increasing sequence in \([0, T] \cap Q_+\) constructed in Remark 4.2.11, and put

\[
A_N(x) \triangleq \{\omega : \max_{s \in \pi_N} |X_s(\omega)| > x\} \quad \text{and} \quad A(x) \triangleq \{\omega : \sup_{s \in [0,T]\cap Q_+} |X_s(\omega)| > x\}
\]

for each \( x \in (0, \infty) \). Fix such an \( x \), and observe from Remark 4.2.11 that

\[
A_N(x) \subset A_{N+1}(x) \subset A(x) \quad \text{and} \quad A(x) = \bigcup_{1 \leq N < \infty} A_N(x)
\]

Thus, from the Monotone Convergence Theorem 1.2.12,

\[
P[A(x)] = E[1_{A(x)}] = \lim_{N \to \infty} E[1_{A_N(x)}] = \lim_{N \to \infty} P[A_N(x)].
\]
Now \( \{(X_s, \mathcal{F}_s) ; s \in \pi_N \} \) is a discrete-parameter submartingale, hence Theorem 2.4.1 shows that

\[
(4.14) \quad x \cdot P[A_N(x)] \leq 2E[X_{s_N}^+] + E[|X_0|] \leq 2E[X_T^+] + E[|X_0|]
\]

where the second inequality follows since \( \{(X_t^+, \mathcal{F}_t) , t \in [0, \infty) \} \) is a submartingale (use Proposition 4.1.12 with \( c(x) \triangleq \max\{x, 0\} \)) and thus \( E[X_{s_N}^+] \leq E[X_T^+] \) on account of \( s_N \leq T \). From (4.13) and (4.14) (with \( x \equiv n \)):

\[
P \left[ \sup_{s \in [0,T] \cap Q^+} |X_s| > n \right] \leq \frac{2E[X_T^+] + E[|X_0|]}{n}
\]

for all \( n = 1, 2, \ldots \), hence \( \lim_{n \to \infty} P[\sup_{s \in [0,T] \cap Q^+} |X_s| > n] = 0 \). Since

\[
\sup_{s \in [0,T] \cap Q^+} |X_s| = \infty = \bigcap_{1 \leq n < \infty} \left[ \sup_{s \in [0,T] \cap Q^+} |X_s| > n \right]
\]

we see that

\[
P \left[ \sup_{s \in [0,T] \cap Q^+} |X_s| = \infty \right] = 0
\]

for each \( T \in (0, \infty) \). In view of (4.12) we find that \( P(\Omega_2^*) = 1 \). We next show that \( P(\Omega_1^*) = 1 \). Fix arbitrary \( T \in (0, \infty) \) and rational numbers \( a \) and \( b \) such that \( a < b \). Since \( \{(X_s, \mathcal{F}_s) ; s \in \pi_N \} \) is a discrete-parameter submartingale, it follows from Theorem 2.5.4 that

\[
(4.15) \quad E[\mathcal{U}_{\pi_N}(a, b; X)] \leq \frac{E[(X_{s_N} - a)^+]}{b - a} \leq \frac{E[(X_T - a)^+]}{b - a}
\]

where the second inequality is a consequence of the fact that \( \{(X_t - a)^+, \mathcal{F}_t) , t \in [0, \infty) \} \) is a submartingale (take \( c(x) \triangleq \max\{x - a, 0\} \) in Proposition 4.1.12) and \( s_N \leq T \). Now, from (4.10), we see that

\[
\mathcal{U}_{[0,T] \cap Q^+}(a, b; X(\cdot, \omega)) = \lim_{n \to \infty} \mathcal{U}_{\pi_N}(a, b; X(\cdot, \omega))
\]

the convergence being monotonic increasing for each \( \omega \). Thus, by the Monotone convergence theorem 1.2.12 and (4.15):

\[
(4.16) \quad E[\mathcal{U}_{[0,T] \cap Q^+}(a, b; X)] \leq \frac{E[(X_T - a)^+]}{b - a} < \infty.
\]

It follows that

\[
P[\mathcal{U}_{[0,T] \cap Q^+}(a, b; X) = +\infty] = 0.
\]

Since this holds for arbitrary \( T \in (0, \infty) \) and rational numbers \( a \) and \( b \) with \( a < b \), we see from (4.11) that \( P(\Omega_1^*) = 1 \). The proposition follows upon taking \( \Omega^* \triangleq \Omega_1^* \cap \Omega_2^* \).

**Remark 4.2.14** An especially important class of regulated functions \( \varphi : [0, \infty) \to \mathbb{R} \) are those which have the property that \( \varphi(t) = \varphi(t+) \) for all \( t \in [0, \infty) \). Such functions are thus right-continuous at each \( t \in [0, \infty) \) and have finite left limits at each \( t \in (0, \infty) \). These mappings are known variously as Skorokhod functions, corloll functions (for continuous-on-right, finite-limits-on-left), cadlag
functions (from the French continu à la droite et pourvu des limites finies à la gauche), and r.c.l.l.
functions (for right-continuous-with-left-limits). We shall always use the term corlol functions.
The point about such functions is that they are extremely well behaved. Corlol functions are clearly
more regular than right-continuous functions, since they also have left limits at each \( t \in (0, \infty) \), and
are almost as nicely behaved as continuous functions while having the extra generality of allowing
for simple discontinuities. The next proposition shows how a given regulated function \( \varphi : Q_+ \to \mathbb{R} \)
determines a corlol function:

**Proposition 4.2.15** Suppose that \( \varphi : Q_+ \to \mathbb{R} \) is a regulated function, and put \( \psi(t) \triangleq \varphi(t+) \) for each
\( t \in [0, \infty) \). Then the mapping \( \psi : [0, \infty) \to \mathbb{R} \) is a corlol function.

The proof of this proposition is a simple exercise in real analysis and is left to the reader.

**Remark 4.2.16** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a submartingale on the probability space
\( (\Omega, \mathcal{F}, P) \), and fix an arbitrary \( \omega \) in \( \Omega^* \), the event identified in Proposition 4.2.13. Then the mapping
\( s \to X_s(\omega) : Q_+ \to \mathbb{R} \) has a finite left limit at each \( t \in (0, \infty) \) and a finite right limit at each
\( t \in [0, \infty) \); we denote these limits by \( X_{t-}(\omega) \) and \( X_{t+}(\omega) \) respectively. For typographical convenience,
we shall sometimes also use the notations \( X(t-, \omega) \) and \( X(t+, \omega) \) for the left and right limits. For
\( \omega \notin \Omega^* \) put \( X_{t-}(\omega) \triangleq 0 \) and \( X_{t+}(\omega) \triangleq 0 \ \forall t \in [0, \infty) \). In view of Proposition 4.2.15 we see that
\( t \to X_{t+}(\omega) : [0, \infty) \to \mathbb{R} \) is a corlol function for each \( \omega \in \Omega \).

In the next proposition we shall see how to derive a submartingale with corlol sample paths from a
given submartingale:

**Proposition 4.2.17** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a submartingale on the probability space
\( (\Omega, \mathcal{F}, P) \) such that \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \). Then, with respect to the notation in Remark
4.2.16 and Definition 3.1.26, the following hold true:
(a) \( E|X_{t+}| < \infty \) and \( E[X_{t+} | \mathcal{F}_t] \geq X_t \) a.s. for each \( t \in [0, \infty) \);
(b) \( E|X_{t-}| < \infty \) and \( E[X_t | \mathcal{F}_{t-}] \geq X_{t-} \) a.s. for each \( t \in (0, \infty) \);
(c) \( \{(X_{t+}, \mathcal{F}_{t+}); t \in [0, \infty)\} \) is a submartingale whose sample paths are corlol functions.

Proof: (a) Fix some arbitrary \( t \in [0, \infty) \) and let \( \{s_k, k = 0, -1, -2 \ldots\} \) be some negatively indexed
sequence in \( Q_+ \) such that \( t < s_{k-1} < s_k \) and \( \lim_{k \to -\infty} s_k = t \). From Remark 4.2.16 and Proposition
4.2.13 we see that
\[
(4.17) \quad \lim_{k \to -\infty} X(s_k, \omega) = X(t+, \omega)
\]
for each \( \omega \in \Omega^* \), and, since \( P(\Omega^*) = 1 \), the convergence in (4.17) is a.s. Since \( \{(X_{s_k}, \mathcal{F}_{s_k}); k = 0, -1, -2 \ldots\} \) is a negatively indexed discrete-parameter submartingale with \( -\infty < E[X_t] \leq E[X_{s_k}] \) for
all \( k = 0, -1, -2 \ldots \), Theorem 2.7.7 says that the collection of random variables \( \{X_{s_k}, k = 0, -1, -2 \ldots\} \)
is uniformly integrable, and hence Theorem 2.6.10 along with a.s. convergence in (4.17), implies that
\[ E|X_{t+}| < \infty \] and
\[ E|X_{s_k} - X_{t+}| = 0. \]  
(4.18)

Fix some arbitrary event \( A \in \mathcal{F}_t \). Then, clearly, \( |E[X_{s_k}; A] - E[X_{t+}; A]| \leq E|X_{s_k} - X_{t+}| \) hence from (4.18) we obtain:
\[ \lim_{k \to \infty} E[X_{s_k}; A] = E[X_{t+}; A]. \]  
(4.19)

Now, by the submartingale property, \( E[X_{s_k}; A] \geq E[X_t; A] \) for all \( k = 0, -1, -2 \ldots \), thus in view of (4.19) we find that:
\[ E[X_{t+}; A] \geq E[X_t; A]. \]  
(4.20)

Part (a) follows, since (4.20) holds for arbitrary \( A \in \mathcal{F}_t \) and \( X_t \) is \( \mathcal{F}_t \)-measurable.

(b) Fix some arbitrary \( t \in (0, \infty) \) and let \( \{s_k, k = 0, 1, 2 \ldots\} \) be some sequence in \( Q_+ \) such that \( s_k < s_{k+1} < t \) and \( \lim_{k \to \infty} s_k = t \). By the submartingale property
\[ E[X_t | \mathcal{F}_{s_k}] \geq X_{s_k} \] a.s.
for all \( k = 0, 1, 2 \ldots \). Now, from Remark 4.2.16 and Proposition 4.2.13,
\[ \lim_{k \to \infty} X(s_k, \omega) = X(t-, \omega) \] for each \( \omega \in \Omega^* \), and since \( P(\Omega^*) = 1 \) the convergence in (4.22) is a.s. In view of Proposition 3.1.30 (iii) and Theorem 2.7.6, we also obtain
\[ \lim_{k \to \infty} E[X_t | \mathcal{F}_{s_k}] = E[X_t | \mathcal{F}_{t-}] \] a.s.
(4.23)

Putting (4.21), (4.22) and (4.23) together gives \( E[X_t | \mathcal{F}_{t-}] \geq X_{t-} \) a.s. Finally, to see that \( X_{t-} \) is integrable, observe from Proposition 4.1.12 that \( \{(X_t^+, \mathcal{F}_t); t \in [0, \infty)\} \) is a submartingale hence
\[ \sup_k E[X_{s_k}^+] \leq E[X_{t+}] < \infty. \] Since \( \{(X_{s_k}, \mathcal{F}_{s_k}); k = 0, 1, 2 \ldots\} \) is a discrete-parameter submartingale we see from Theorem 2.5.7 and the a.s. convergence in (4.22) that \( E|X_{t-}| < \infty \).

(c) We have observed in Remark 4.2.16 that \( t \mapsto X_{t+}(\omega) \) defines a coroll function for each \( \omega \in \Omega \). To establish that \( \{(X_{t+}, \mathcal{F}_{t+}); t \in [0, \infty)\} \) is a submartingale, first observe from (a) that \( E|X_{t+}| < \infty \) for all \( t \in [0, \infty) \). Next, we must verify that \( X_{t+}(\omega) \) is \( \mathcal{F}_{t+} \)-measurable with respect to \( \omega \). Thus, fix some \( t \in [0, \infty) \) and let \( \{s_k, k = 0, -1, -2 \ldots\} \) be some negatively indexed sequence in \( Q_+ \) such that \( t < s_k < s_{k-1} \) and \( \lim_{k \to \infty} s_k = t \). From Proposition 4.2.13 and Remark 4.2.16:
\[ \lim_{k \to \infty} X_{s_k} = X_{t+} \] a.s.
(4.24)

Now put
\[ Y_t(\omega) \triangleq \lim_{k \to \infty} X_{s_k}(\omega) \]
for each $\omega \in \Omega$, and fix some $\epsilon > 0$. For all sufficiently negative $k$ one has $t < s_k < t + \epsilon$, and thus $X_{s_k}(\omega)$ is $F_{t+\epsilon}$-measurable with respect to $\omega$, hence $Y_t(\omega)$ is $F_{t+\epsilon}$-measurable with respect to $\omega$. By the arbitrary choice of $\epsilon > 0$ and the definition of $F_{t+}$ (see Definition 3.1.26) it follows that $Y_t(\omega)$ is $F_{t+}$-measurable in $\omega$. Moreover, from (4.29) and Theorem 2.6.10:

$$
\text{Now, by the submartingale property of}
$$

Proposition 4.2.17 (c) suggests the following question: what is the relation between the given submartingale $(X_t; t \in [0, \infty))$ and the corol submartingale $(X_{t+}; t \in [0, \infty))$ which is derived from it? The next theorem, which is the main result of the present section, answers this question:

**Theorem 4.2.18** Suppose that $(X_t; t \in [0, \infty))$ is a submartingale on the probability space $(\Omega, \mathcal{F}, P)$ and that $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a standard filtration (recall Definition 3.4.2). Then, $P[X_t = X_{t+}] = 1$ for all $t \in [0, \infty)$ if and only if the mapping $t \to E[X_t] : [0, \infty) \to \mathbb{R}$ is right-continuous.

Proof: Suppose that $t \to E[X_t]$ is right-continuous on $[0, \infty)$, and fix some arbitrary $t \in [0, \infty)$. Let $\{s_k, k = 0, -1, -2 \ldots\}$ be some negatively indexed sequence in $Q_+$ such that $t < s_{k-1} < s_k$ and $\lim_{k \to -\infty} s_k = t$. By the postulated right-continuity of $t \to E[X_t]$ it follows that

$$
\lim_{k \to -\infty} E[X_{s_k}] = E[X_t],
$$

and, from Proposition 4.2.13,

$$
\lim_{k \to -\infty} X_{s_k}(\omega) = X_{t+}(\omega) \quad \text{a.s.}
$$

Now, by the submartingale property of $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$, we see that $\{(X_{s_k}, \mathcal{F}_{s_k}); k = 0, -1, -2 \ldots\}$ is a negatively indexed discrete-parameter submartingale, and clearly $-\infty < E[X_t] \leq E[X_{s_k}]$ for all $k$. Thus Theorem 2.7.7 tells us that the collection $\{X_{s_k}, k = 0, -1, -2 \ldots\}$ is uniformly integrable, hence, from (4.29) and Theorem 2.6.10:

$$
\lim_{k \to -\infty} E[X_{s_k}] = E[X_{t+}].
$$

140
Now, from Proposition 4.2.17 (a), we have $E[X_{t+} | \mathcal{F}_t] \geq X_t$ a.s.; since $\mathcal{F}_{t+} = \mathcal{F}_t$ and $X_{t+}$ is $\mathcal{F}_{t+}$-measurable, this reduces to $X_{t+} \geq X_t$ a.s. Combining (4.28) and (4.30) we obtain

$$E[X_{t+} - X_t] = 0 \quad \text{and} \quad X_{t+} \geq X_t \text{ a.s.}$$

This, along with elementary measure and integration theory, shows that $P[X_t = X_{t+}] = 1$.

For the converse, suppose that $P[X_t = X_{t+}] = 1 \forall t \in [0, \infty)$. Fix some arbitrary $t \in [0, \infty)$, and let $\{t_k, k = 0, -1, -2 \ldots\}$ be some negatively indexed sequence in $\mathbb{R}$ such that $t \leq t_{k-1} \leq t_k$ and $\lim_{k \to -\infty} t_k = t$. Since $t \to X_{t+}(\omega)$ is right-continuous (see Remark 4.2.16) we clearly have

$$\lim_{k \to -\infty} X(t_k+, \omega) = X(t+, \omega) \quad (4.31)$$

for each $\omega \in \Omega$. Moreover, $\{(X_{t_k}, \mathcal{F}_{t_k}); k = 0, -1, -2 \ldots\}$ is a negatively indexed discrete-parameter submartingale with $-\infty < E[X_{t_k}] \leq E[X_{t_k}]$, hence (by Theorem 2.7.7) the collection of random variables $\{X_{t_k}, k = 0, -1, -2 \ldots\}$ is uniformly integrable. Since $P[X_{t_k} = X_{t_k+}] = 1$ for all $k$, we see from Remark 2.6.6 that $\{X_{t_k+}, k = 0, -1, -2 \ldots\}$ is also uniformly integrable. In view of (4.31) and Theorem 2.6.10 one obtains $\lim_{k \to -\infty} E[X(t_k+)] = E[X_{t+}]$. But $E[X(t_k+)] = E[X(t_k)]$ and $E[X_{t+}] = E[X_t]$ (since $P[X_t = X_{t+}] = 1 \forall t \in [0, \infty)$) hence $\lim_{k \to -\infty} E[X(t_k)] = E[X_t]$.

The following corollary is just a combination of Proposition 4.2.17(c) and Theorem 4.2.18:

**Corollary 4.2.19** Suppose that $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a standard filtration in the probability space $(\Omega, \mathcal{F}, P)$ and $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ is a submartingale. Then the following are equivalent:

(a) The mapping $t \to E[X_t] : [0, \infty) \to \mathbb{R}$ is right-continuous;

(b) $\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}$ has a corol modification, that is there exists a corol process $\{Y_t; t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that $\{(Y_t, \mathcal{F}_t); t \in [0, \infty)\}$ is a submartingale and $P[X_t = Y_t] = 1$ for each $t \in [0, \infty)$.

**Remark 4.2.20** In problems involving stochastic processes one can usually replace a given process with a modification of that process without altering the essential features of the problem. Corollary 4.2.19 then allows us to assume *a-priori* that the sample paths of a given submartingale are right-continuous (in fact, corol) provided that the associated filtration is a standard filtration and condition (a) of Corollary 4.2.19 holds. Continuous-parameter submartingales subject to these two restrictions are very common in applications, to the extent that some authors even incorporate the property of sample-path right-continuity into the very definition of submartingales, relying on Corollary 4.2.19 to justify this restriction (see e.g Stroock [29], page 28). Although we have not built this property into our Definition 4.1.1 of submartingales, we shall usually postulate right-continuity for the sample paths of the submartingales occurring in most of the theorems, lemmas etc. in the remainder of this chapter. Since Corollary 4.2.19 obviously applies not only to submartingales but also to supermartingales and martingales, the preceding remarks hold for these cases as well.
4.3 Martingale Inequalities

Our goal in this section is to establish continuous-parameter analogues of the basic martingale inequalities which were developed in a discrete-parameter context in Chapter 2. The simple proposition which follows will be used repeatedly. Its proof involves only simple real analysis and is left as an exercise.

**Proposition 4.3.1** Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a right-continuous mapping, and let \( S \) be some dense subset of \([0, \infty)\). Then, for each \( T \in [0, \infty) \), we have

\[
\sup_{t \in [0, T]} \varphi(t) = \sup_{s \in \Gamma_T} \varphi(s),
\]

for \( \Gamma_T \triangleq \{T\} \cup \{[0, T] \cap S\} \).

**Remark 4.3.2** In keeping with Remark 3.1.27 we shall call \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) a right-continuous martingale [supermartingale, submartingale] when \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale [supermartingale, submartingale] with right-continuous sample-paths. This terminology says nothing about \( \{\mathcal{F}_t, t \in [0, \infty)\} \), and, in particular, does not imply that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a right-continuous filtration.

We now extend the basic martingale inequalities developed in Theorems 2.4.1 and 2.4.4 to the continuous parameter case:

**Theorem 4.3.3** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous submartingale on \((\Omega, \mathcal{F}, P)\). Then the following inequalities hold:

(i) For each \( x \in [0, \infty) \) and \( T \in [0, \infty) \):

\[
x \cdot P \left[ \sup_{0 \leq t \leq T} X_t \geq x \right] \leq E[X_T^+] (4.32)
\]

and

\[
x \cdot P \left[ \inf_{0 \leq t \leq T} X_t \leq -x \right] \leq E[X_T^+] - E[X_0]. (4.33)
\]

(ii) For each \( p > 1 \) and \( T \in [0, \infty) \):

\[
\| \sup_{0 \leq t \leq T} X_t \|_p \leq q \cdot \|X_T\|_p (4.34)
\]

where \( \frac{1}{p} + \frac{1}{q} = 1 \), provided \( X_t \geq 0 \) for all \( t \in [0, \infty) \).
the sequence \( \{ s^N_0, s^N_1, \ldots, s^N_N, T \} \) obtained by adding the element \( T \) to the sequence \( \pi_N \). Then, using Theorem 2.4.1 on the discrete-parameter submartingale \( \{(X_s, F_s); s \in \nu_N\} \), we obtain

\[
(4.35) \quad x_n \cdot P \left[ \max_{s \in \nu_N} X_s > x_n \right] \leq E[X_T^+].
\]

Now clearly

\[
\sup_{s \in \Gamma_T} X_s > x_n = \bigcup_{1 \leq N < \infty} \max_{s \in \nu_N} X_s > x_n
\]

and the events in the set union on the right hand side are increasing as \( N \) increases. Thus:

\[
(4.36) \quad \lim_{N \to \infty} P \left[ \max_{s \in \nu_N} X_s > x_n \right] = P \left[ \sup_{s \in \Gamma_T} X_s > x_n \right] = P \left[ \sup_{0 \leq t \leq T} X_t > x_n \right],
\]

where the first equality follows from Theorem 1.2.5(c), and the second equality from Proposition 4.3.1 together with right-continuity of the mappings \( t \to X_t(\omega) \). From (4.35) and (4.36) we obtain:

\[
\lim_{n \to \infty} x_n \cdot P \left[ \sup_{0 \leq t \leq T} X_t > x_n \right] \leq E[X_T^+].
\]

Finally, taking \( n \to \infty \) gives (4.32). Using a similar argument, we obtain (4.33) from (2.33) and (4.34) from (2.46).

As an application of the bounds in Theorem 4.3.3 we next establish the following refinement of Proposition 4.1.13:

**Proposition 4.3.4** Assume the hypotheses of Proposition 4.1.13, but suppose in addition that \( p \in (1, \infty) \) and the \( L^p \)-martingales \( \{(X_n(t), F_t); t \in [0, \infty)\} \) are continuous for each \( n = 0, 1, 2, \ldots \). Then the limiting \( L^p \)-martingale \( \{(X(t), F_t); t \in [0, \infty)\} \) in Proposition 4.1.13 has a continuous modification. I.e. there exists a continuous process \( \{Y(t); t \in [0, \infty)\} \) on \( (\Omega, F, P) \) such that \( \{(Y(t), F_t); t \in [0, \infty)\} \) is a martingale and \( P[X(t) = Y(t)] = 1, \forall t \in [0, \infty) \).

**Proof:** Fix real numbers \( \alpha \in (1, \infty) \) and \( T \in [0, \infty) \). For arbitrary positive integers \( m, n \), we see that \( \{(X_n(t) - X_m(t), F_t); t \in [0, \infty)\} \) is a continuous \( L^p \)-martingale, and thus

\[
(4.37) \quad P \left[ \sup_{0 \leq t \leq T} |X_n(t) - X_m(t)| > r^{-\alpha} \right] \leq r^{\alpha p} E \left[ \sup_{0 \leq t \leq T} |X_n(t) - X_m(t)|^p \right]
\]

\[
\leq r^{\alpha p} \left( \frac{p}{p-1} \right)^p E \left[ \|X_n(T) - X_m(T)\|^p \right].
\]

Here we have used the Markov inequality (Theorem 1.2.23) to get the first bound, and the second bound follows from Theorem 4.3.3 (ii) along with the fact that \( \{|X_n(t) - X_m(t)|^p, F_t\}; t \in [0, \infty)\} \) is a submartingale with right-continuous (in fact, continuous) sample paths. Now \( \{X_n(T), n = 0, 1, 2, \ldots\} \) is a converging (hence Cauchy) sequence in \( L^p(\Omega, F, P) \), and thus

\[
(4.38) \quad \lim_{m \to \infty} \left( \sup_{n>m} E \left[ \|X_n(T) - X_m(T)\|^p \right] \right) = 0.
\]
Fix some $\beta > 1 + \alpha p$. In view of (4.38), for each integer $r = 1, 2 \ldots$ there is some integer $m_r$ such that

$$\sup_{n > m_r} E[|X_n(T) - X_{m_r}(T)|^p] < r^{-\beta}.$$ 

Furthermore, without loss of generality we can choose the $m_r$ such that $m_{r+1} > m_r$. Thus

$$E[|X_{m_{r+1}}(T) - X_{m_r}(T)|^p] < r^{-\beta},$$

hence, from (4.37),

$$P \left[ \sup_{0 \leq t \leq T} |X_{m_{r+1}}(t) - X_{m_r}(t)| > r^{-\alpha} \right] \leq r^{-(\beta - \alpha p)} \left( \frac{p}{p-1} \right)^p$$ 

for all $r = 1, 2 \ldots$ Now put $A_r \triangleq \{ \sup_{0 \leq t \leq T} |X_{m_{r+1}}(t) - X_{m_r}(t)| > r^{-\alpha} \}$ and $\Omega(T) \triangleq \limsup_{r \to \infty} A_r^c$. Since $(\beta - \alpha p) > 1$ we see from (4.39) and the Borel-Cantelli Theorem 1.2.5(f) that $P[\Omega(T)] = 1$. Now fix some arbitrary $\omega \in \Omega(T)$. Then, there is some integer $R(\omega)$ such that $\omega \notin A_r$ for all $r \geq R(\omega)$, hence (since $\alpha > 1$) we have

$$\sum_{r \geq 1} \sup_{0 \leq t \leq T} |X_{m_{r+1}}(t, \omega) - X_{m_r}(t, \omega)| < \infty.$$ 

In view of this bound and the Cauchy criterion it follows that $\{X_{m_r}(\cdot, \omega), r = 1, 2 \ldots \}$ is a Cauchy sequence in the complete metric space $C[0, T]$, hence there is some mapping $Y^T(\cdot, \omega) \in C[0, T]$ such that

$$\lim_{r \to \infty} \sup_{0 \leq t \leq T} |X_{m_r}(t, \omega) - Y^T(t, \omega)| = 0.$$ 

Finally, for $\omega \notin \Omega(T)$, put $Y^T(t, \omega) \triangleq 0$, $\forall t \in [0, T]$. We have now constructed a mapping $Y^T : [0, T] \otimes \Omega \to \mathcal{R}$ such that (4.40) holds for each $\omega \in \Omega(T)$ and $t \to Y^T(t, \omega)$ is continuous on $[0, T]$ for each $\omega \in \Omega$. Fix some arbitrary $t \in [0, T]$. Since $P[\Omega(T)] = 1$ we see from (4.40) that $\lim_{r \to \infty} X_{m_r}(t) = Y^T(t)$ a.s. and since $\{X_{m_r}(t), r = 1, 2 \ldots \}$ is $L^p$-converging to $X(t)$ it follows that

$$P[X(t) = Y^T(t)] = 1$$ 

for each $t \in [0, T]$. Now fix some sequence of numbers $\{T_k\}$ in $[0, \infty)$ such that $0 < T_k < T_{k+1}$ and $\lim_{k \to \infty} T_k = +\infty$, put

$$\tilde{\Omega} \triangleq \bigcap_{1 \leq k < \infty} \Omega(T_k),$$

and observe that $P[\tilde{\Omega}] = 1$, since $P[\Omega(T_k)] = 1$ for all $k = 1, 2 \ldots$. Choose some arbitrary $t \in [0, \infty)$ and let $k_1$ and $k_2$ be such that $0 \leq t \leq T_{k_1} < T_{k_2}$. Then, for each $\omega \in \tilde{\Omega}$, we see that (4.40) holds for both $T \triangleq T_{k_1}$ and $T \triangleq T_{k_2}$. It follows that $Y^{T_{k_1}}(t, \omega) = Y^{T_{k_2}}(t, \omega)$ for all $\omega \in \tilde{\Omega}$ and $t \in [0, T_{k_1}]$. We can thus define a mapping $Y : [0, T] \otimes \Omega \to \mathcal{R}$ as follows: for $\omega \notin \tilde{\Omega}$ put $Y(t, \omega) \triangleq 0$, $\forall t \in [0, \infty)$, and for each $\omega \in \tilde{\Omega}$ and $t \in [0, \infty)$ put $Y(t, \omega) \triangleq Y^{T_{k_2}}(t, \omega)$ for any $T_k > t$. Clearly $t \to Y(t, \omega)$ is continuous on $[0, \infty)$ for each $\omega \in \Omega$, and, since $P[\tilde{\Omega}] = 1$, it follows that $P[X(t) = Y(t)] = 1$ for each $t \in [0, \infty)$. The process $\{Y(t), t \in [0, \infty)\}$ is therefore continuous as well as being a modification of $\{X(t), t \in [0, \infty)\}$. Since $\mathcal{F}_0$ includes all $P$-null events in $\mathcal{F}$ it follows from Proposition 3.1.18 that $\{(Y_t, \mathcal{F}_t); t \in [0, \infty)\}$ is an adapted process which is clearly a martingale.

$\blacksquare$
4.4 Martingale Convergence Theorems

In this section we shall extend the martingale convergence theorems of § 2.5 to a continuous-parameter setting. To this end we need some basic facts and ideas from elementary real analysis. For the sake of completeness we recall these notions in the next few remarks.

Remark 4.4.1 A mapping \( \varphi : [0, \infty) \to \mathbb{R} \) converges to a finite limit as \( t \to \infty \) when there exists some \( \xi \in \mathbb{R} \) such that the following holds: corresponding to each \( \epsilon > 0 \) there is some \( T(\epsilon) \in [0, \infty) \) such that \( |\varphi(t) - \xi| < \epsilon \) for all \( t \in [T(\epsilon), \infty) \). Under these conditions \( \xi \) is called the limit of \( \varphi \) as \( t \to \infty \).

On the other hand, \( \varphi \) is said to converge to \( +\infty \) \([-\infty] \) when, corresponding to each \( c > 0 \) there exists some \( T(c) \in [0, \infty) \) such that \( \varphi(t) > c \) \([\varphi(t) < -c] \) for all \( t \in [T(c), \infty) \). We use the notations
\[
\lim_{t \to \infty} \varphi(t) = \xi \quad \text{or} \quad \lim_{t \to \infty} \varphi(t) = \pm \infty
\]
to indicate convergence of \( \varphi \) to a finite limit \( \xi \) or \( \pm \infty \) as \( t \to \infty \), and say that \( \varphi \) converges to a limit in \( \mathbb{R} \) as \( t \to \infty \) when any one of these cases occurs.

The next theorem, from elementary real analysis, identifies a particularly important class of mappings on \([0, \infty)\) which converge to a limit as \( t \to \infty \):

Theorem 4.4.2 Suppose that the mapping \( \varphi : [0, \infty) \to \mathbb{R} \) is either non-increasing or non-decreasing. Then \( \varphi \) converges to a limit in \( \mathbb{R} \) as \( t \to \infty \).

We next formulate a continuous-parameter analogue of the notions of superior and inferior limits for a sequence of real numbers. To this end, let \( \varphi : [0, \infty) \to \mathbb{R} \) be a given mapping, and for each \( t \in [0, \infty) \) put
\[
\underline{\psi}(t) \triangleq \inf_{u \geq t} \varphi(u) \quad \text{and} \quad \overline{\psi}(t) \triangleq \sup_{u \geq t} \varphi(u).
\]
In this way we get mappings \( \underline{\psi}, \overline{\psi} : [0, \infty) \to \mathbb{R} \) which are clearly non-decreasing and non-increasing respectively, and thus, from Theorem 4.4.2, converge to limits in \( \mathbb{R} \) as \( t \to \infty \). These limits are called the inferior and superior limits of \( \varphi \) as \( t \to \infty \) and are denoted by \( \liminf_{t \to \infty} \varphi(t) \) and \( \limsup_{t \to \infty} \varphi(t) \) respectively. Thus
\[
\liminf_{t \to \infty} \varphi(t) \triangleq \lim_{t \to \infty} \left[ \inf_{u \geq t} \varphi(u) \right] \quad \text{and} \quad \limsup_{t \to \infty} \varphi(t) \triangleq \lim_{t \to \infty} \left[ \sup_{u \geq t} \varphi(u) \right]
\]

The next proposition summarizes some basic properties of inferior and superior limits. The proof is just an easy exercise in real analysis and is left to the reader:

Proposition 4.4.3 Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a given mapping. Then
(a) \( \liminf_{t \to \infty} \varphi(t) \leq \limsup_{t \to \infty} \varphi(t) \);
(b) if \( \liminf_{t \to \infty} \varphi(t) < a < b \leq \limsup_{t \to \infty} \varphi(t) \) for some real numbers \( a \) and \( b \) then there exist infinite sequences \( \{t_k, k = 0, 1, 2 \ldots\} \) and \( \{\bar{t}_k, k = 0, 1, 2 \ldots\} \) in \([0, \infty)\) such that \( \lim_{k \to \infty} t_k = \lim_{k \to \infty} \bar{t}_k = +\infty \) and \( \varphi(t_k) < a, \ \varphi(\bar{t}_k) > b \), for all \( k = 0, 1, 2 \ldots \);
(c) \( \varphi \) converges to some limit \( \xi \in \mathbb{R} \) if and only if \( \liminf_{t \to \infty} \varphi(t) = \limsup_{t \to \infty} \varphi(t) = \xi \).
Remark 4.4.4 Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a given mapping and \( a \) and \( b \) are real numbers such that \( a < b \). By considering the restriction of \( \varphi \) to \( Q_+ \), the set of non-negative rational numbers, we can define \( U_{[0,T] \cap Q_+} (a, b; \varphi) \) as in Remark 4.2.8 for all \( T \in [0, \infty) \). Now put

\[
(4.41) \quad U_{[0,\infty) \cap Q_+} (a, b; \varphi) \triangleq \lim_{n \to \infty} U_{[0,n] \cap Q_+} (a, b; \varphi)
\]

It is clear that the convergence in this definition is monotonic and the limit exists in \([0, \infty]\). The quantity defined in (4.41) can be used to determine when a given right-continuous mapping converges to some limit in \( \overline{\mathbb{R}} \), as shown in the next proposition:

Proposition 4.4.5 Suppose that \( \varphi : [0, \infty) \to \mathbb{R} \) is a given right-continuous mapping. If

\[
U_{[0,\infty) \cap Q_+} (a, b; \varphi) < \infty
\]

for all rational numbers \( a \) and \( b \) such that \( a < b \) then \( \varphi \) converges to a limit in \( \overline{\mathbb{R}} \) as \( t \to \infty \).

Proof: We argue by contradiction and suppose that \( \varphi \) fails to converge to a limit in \( \overline{\mathbb{R}} \). Then, in view of Proposition 4.4.3, there are rational numbers \( a \) and \( b \) such that

\[
\liminf_{t \to \infty} \varphi(t) < a < b < \limsup_{t \to \infty} \varphi(t),
\]

and thus there are infinite sequences \( \{ t_k, k = 0, 1, 2 \ldots \} \) and \( \{ \overline{t}_k, k = 0, 1, 2 \ldots \} \) in \([0, \infty)\) such that

\[
\lim_{k \to \infty} t_k = \lim_{k \to \infty} \overline{t}_k = +\infty \quad \text{and} \quad \varphi(t_k) < a, \, \varphi(\overline{t}_k) > b \quad \text{for all} \quad k = 0, 1, 2 \ldots .
\]

Now, since \( \varphi \) is right-continuous, corresponding to each \( k \) we can find \( \underline{s}_k, \overline{s}_k \in Q_+ \) such that \( t_k \leq \underline{s}_k, \overline{t}_k \leq \overline{s}_k \), and \( \varphi(\underline{s}_k) < a, \varphi(\overline{s}_k) > b \) for all \( k \). It follows that \( U_{[0,\infty) \cap Q_+} (a, b; \varphi) = +\infty \).

Remark 4.4.6 Suppose that \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty)\} \) is some \( \mathbb{R} \)-valued adapted process on the probability space \( (\Omega, \mathcal{F}, P) \) such that the sample paths \( t \to X_t(\omega) \) are right-continuous on \([0, \infty)\) for each \( \omega \in \Omega \). Fixing real numbers \( a \) and \( b \) such that \( a < b \), we recall from Remark 4.2.12 that the mapping \( \omega \to U_{[0,T] \cap Q_+} (a, b; X(\cdot, \omega)) \) is \( \mathcal{F}_T \)-measurable for each \( T \in [0, \infty) \), and hence (see Remark 4.4.4) the mapping \( \omega \to U_{[0,\infty) \cap Q_+} (a, b; X(\cdot, \omega)) \triangleq \lim_{n \to \infty} U_{[0,n] \cap Q_+} (a, b; X(\cdot, \omega)) \) is \( \mathcal{F}_\infty \)-measurable. Put

\[
(4.42) \quad \Omega_3^* \triangleq \bigcap_{a,b} \{ \omega : U_{[0,\infty) \cap Q_+} (a, b; X(\cdot, \omega)) < +\infty \}
\]

the intersection being over all pairs of rational numbers \( a \) and \( b \) such that \( a < b \). It follows that \( \Omega_3^* \in \mathcal{F}_\infty \), and, from Proposition 4.4.5, one sees that if \( \omega \in \Omega_3^* \) then the mapping \( t \to X_t(\omega) \) converges to some limit in \( \overline{\mathbb{R}} \) as \( t \to \infty \).

We are now ready to establish the continuous-parameter analogue of Theorem 2.5.7:
Theorem 4.4.7 Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous submartingale on \((\Omega, \mathcal{F}, P)\), such that
\[
(4.43) \quad \sup_{t \in [0, \infty)} E[X_t^+] \overset{\triangle}{=} C < \infty.
\]
Then there exists some a.s. unique random variable \( X_\infty : \Omega \to \mathbb{R} \), measurable with respect to \( \mathcal{F}_\infty \), such that \( E|X_\infty| < \infty \) and \( \lim_{t \to \infty} X_t = X_\infty \) a.s.

Proof: Fix arbitrary rational numbers \( a \) and \( b \) such that \( a < b \) and let \( n \) be some positive integer. Then:
\[
(4.44) \quad E[\mathcal{U}_{[0,n]}(a,b;X)] \leq \frac{E[(X_n - a)^+] + C + |a|}{b-a} \leq \frac{C + |a|}{b-a}.
\]

Here, the first inequality follows by an argument identical to that which led to (4.16), and the second inequality is immediate consequence of (4.43). Taking \( n \to \infty \) in (4.44), recalling Remark 4.4.4, and using the Monotone Convergence Theorem 1.2.12, gives
\[
E[\mathcal{U}_{[0,\infty)}(a,b;X)] \leq \frac{C + |a|}{b-a} < \infty
\]
and hence \( P[\mathcal{U}_{[0,\infty)}(a,b;X)] < \infty \] = 1. Since there are only countably many pairs of rational numbers \( a \) and \( b \) we see that \( P[\Omega_3^*] = 1 \), for \( \Omega_3^* \) defined in (4.42). Now define
\[
(4.45) \quad X_\infty(\omega) \overset{\triangle}{=} \lim_{n \to \infty} X_n(\omega)
\]
An argument involving the use of Fatou’s Theorem 1.2.15 identical to that for (2.70) shows that \( E|X_\infty| < \infty \), and hence \( P[|X_\infty| = +\infty] = 0 \). Now define \( X_\infty(\omega) \overset{\triangle}{=} X_\infty(\omega) \) for all \( \omega \) such that \( |X_\infty(\omega)| < \infty \), and define \( X_\infty(\omega) \overset{\triangle}{=} 0 \) for all \( \omega \) such that \( |X_\infty(\omega)| = \infty \). We see that \( X_\infty \) is \( \mathbb{R} \)-valued and \( \mathcal{F}_\infty \)-measurable, that \( E|X_\infty| = E|X_\infty| < \infty \), and that \( \lim_{t \to \infty} X_t(\omega) = X_\infty(\omega) \) when \( \omega \in \Omega_3^* \cap \{|X_\infty| < \infty\} \). Since \( P[\Omega_3^* \cap \{|X_\infty| < \infty\}] = 1 \) the theorem is established. \( \square \)

Remark 4.4.8 When \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous submartingale [supermartingale] with \( X_t \leq 0 [X_t \geq 0] \) a.s. \( \forall t \in [0, \infty) \) then, exactly as in Remark 2.5.8, it follows there exists an a.s. unique random variable \( X_\infty \) with the properties in Theorem 4.4.7.

We next formulate the continuous-parameter analogue of a closed supermartingale (recall Definition 2.3.4):

Definition 4.4.9 The collection of pairs \( \{(X_t, \mathcal{F}_t); t \in [0, \infty]\} \) is a right-continuous closed supermartingale when:

(a) \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous supermartingale on a probability space \((\Omega, \mathcal{F}, P)\);

(b) \( \mathcal{F}_\infty \) is given by (3.3) and \( X_\infty : \Omega \to \mathbb{R} \) is a \( \mathcal{F}_\infty \)-measurable random variable with
\[
(4.46) \quad E|X_\infty| < \infty \quad \text{and} \quad X_t \geq E[X_\infty | \mathcal{F}_t] \quad \text{a.s.} \quad \forall t \in [0, \infty).
\]

Likewise, the collection of pairs \( \{(X_t, \mathcal{F}_t); t \in [0, \infty]\} \) is a right-continuous closed submartingale when \( \{(-X_t, \mathcal{F}_t); t \in [0, +\infty)\} \) is a right-continuous closed supermartingale, and it is a right-continuous closed martingale when it is both a right-continuous closed supermartingale and submartingale.
Remark 4.4.10 If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous supermartingale [submartingale, martingale] and \( X_\infty \) is some \( \mathcal{F}_\infty \)-measurable random variable such that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous closed supermartingale [submartingale, martingale] then we say that the random variable \( X_\infty \) closes the supermartingale [submartingale, martingale] \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \).

The next result, which is a continuous-parameter analogue of Theorem 2.7.1, enables one to determine when a right-continuous supermartingale is closed by a random variable:

**Theorem 4.4.11** (a) Suppose \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable and right-continuous supermartingale on \( (\Omega, \mathcal{F}, P) \). Then there exists an a.s. unique \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty : \Omega \to \mathbb{R} \) such that \( E|X_\infty| < \infty \) and the following hold:

\[
\lim_{t \to \infty} X_t = X_\infty \quad \text{a.s.}
\]

and

\[
\lim_{t \to \infty} E|X_t - X_\infty| = 0.
\]

Moreover

\[
X_t \geq E[X_\infty | \mathcal{F}_t] \quad \text{a.s.}
\]

for each \( t \in [0, \infty) \), so that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous closed supermartingale.

(b) If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable right-continuous martingale on \( (\Omega, \mathcal{F}, P) \), then the preceding assertions continue to hold, except that in (4.49) \( \geq \) is replaced with \( = \). In particular, \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous closed martingale.

Proof: Since the collection of random variables \( \{X_t; t \in [0, \infty)\} \) is uniformly integrable we see from Remark 2.6.3 that \( \sup_{t \in [0, \infty)} E|X_t| < \infty \) and hence (4.43) is verified. Thus Theorem 4.4.7 ensures the existence of a \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty \) such that \( E|X_\infty| < \infty \) and (4.47) holds a.s. Now the sequence \( \{X_n; n = 0, 1, 2 \ldots\} \), obtained by “sampling” the process \( \{X_t; t \in [0, \infty)\} \) at integer values, is clearly uniformly integrable, thus we see that

\[
\lim_{n \to \infty} X_n = X_\infty \quad (L^1\text{sense}).
\]

Now fix some arbitrary \( t \in [0, \infty) \) and observe that

\[
X_t \geq E[X_n | \mathcal{F}_t] \quad \text{a.s.}
\]

for all integers \( n > t \). Taking \( n \to \infty \) in (4.51) and using (4.50) along with Proposition 1.4.21(b), we see that (4.49) holds. It remains to establish (4.48). Suppose the contrary. Then there must exist some sequence \( \{t_k, k = 0, 1, 2 \ldots\} \) in \([0, \infty)\) such that \( t_{k+1} > t_k \), \( \lim_{k \to \infty} t_k = +\infty \), and some \( \delta > 0 \) such that

\[
E|X_{t_k} - X_\infty| > \delta
\]
for all \( k = 0,1,2 \ldots \). But, \( \lim_{k \to \infty} X_{t_k} = X_\infty \) a.s. (since (4.47) has been shown to hold a.s.) and \( \{ X_{t_k}, \; k = 0,1,2 \ldots \} \) is uniformly integrable. Application of Theorem 2.6.10 then gives \( \lim_{k \to \infty} E|X_{t_k} - X_\infty| = 0 \), in contradiction to (4.52), and (4.48) follows.

When \( \{(X_t, \mathcal{F}_t); \; t \in [0, \infty)\} \) is an \( L^p \)-bounded martingale for some \( p \in (1, \infty) \) then we can sharpen the previous result to obtain the following continuous-parameter analogue of Theorem 2.7.4:

**Theorem 4.4.12** Suppose \( p \in (1, \infty) \) and \( \{(X_t, \mathcal{F}_t); \; t \in [0, \infty)\} \) is a right-continuous \( L^p \)-bounded martingale on a probability space \((\Omega, \mathcal{F}, P)\). Then there exists an a.s. unique \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty : \Omega \to \mathbb{R} \) such that \( E|X_\infty|^p < \infty \) and the following hold:

\[
\begin{align*}
\text{(4.53)} & \quad \lim_{t \to \infty} X_t = X_\infty \quad \text{a.s.} \\
\text{(4.54)} & \quad \lim_{t \to \infty} E|X_t - X_\infty|^p = 0; \\
\text{(4.55)} & \quad E\left[ \sup_{t \in [0, \infty)} |X_t|^p \right] \leq \left( \frac{p}{p-1} \right)^p E|X_\infty|^p; \\
\text{(4.56)} & \quad X_t = E[X_\infty | \mathcal{F}_t] \quad \text{a.s.} \quad \forall \; t \in [0, \infty).
\end{align*}
\]

**In particular, \( \{(X_t, \mathcal{F}_t); \; t \in [0, \infty)\} \) is a right-continuous closed martingale.**

Proof: One establishes (4.53), (4.56) and (4.55) by obvious changes in the arguments which led respectively to parts (2.109), (2.111) and (2.112) of Theorem 2.7.4, but using Theorem 4.4.11 in place of Theorem 2.7.1 (to get (4.53) and (4.56)), and using Theorem 4.3.3(ii) in place of Theorem 2.4.4 (to get (4.55)). To obtain (4.54) fix some sequence \( \{t_n, \; n = 1,2,\ldots\} \) in \([0, \infty)\) with \( t_n \leq t_{n+1} < \infty \) and \( \lim_{n \to \infty} t_n = +\infty \). Then, clearly

\[
|X_{t_n} - X_\infty|^p \leq 2^p \sup_{t \in [0, \infty)} |X_t|^p, \quad \forall \; n = 1,2,\ldots
\]

In view of (4.53), the integrability of the right side of the preceding inequality (which is ensured by (4.55)), and the Dominated Convergence Theorem 1.2.21, we get \( \lim_{n \to \infty} E|X_{t_n} - X_\infty|^p = 0 \). Now (4.54) follows from the arbitrary choice of \( \{t_n\} \).

**Remark 4.4.13** Exactly as in Remark 2.7.5 one sees from (4.55) that, if \( \{(X_t, \mathcal{F}_t); \; t \in [0, \infty)\} \) is an \( L^p \)-bounded martingale for some \( p \in (1, \infty) \), then the collection of random variables \( \{X_t; \; t \in [0, \infty)\} \) is dominated by the \( L^p \)-bounded random variable \( \sup_{t \in [0, \infty)} |X_t| \).

**Remark 4.4.14** If \( \{(X_t, \mathcal{F}_t); \; t \in [0, \infty)\} \) is a uniformly integrable right-continuous supermartingale, \( X_\infty \) is the random variable whose existence is ensured by Theorem 4.4.11, and \( T \) is a \( \{\mathcal{F}_t\} \)-stopping time, then the mapping \( X_T : \Omega \to \mathbb{R} \) defined by (3.35) is \( \mathcal{F}_T \)-measurable.
4.5 The Optional Sampling Theorem

In this section our main goal is to establish a continuous parameter analogue of the optional sampling
Theorem 2.3.9. To this end we start with two preliminary results. The first of these is an obvious
extension of Lemma 2.3.13 to the continuous-parameter context, and gives a criterion for a progressively
measurable process to be a martingale. Exactly as in Remark 2.3.12 we call a stopping time \( T : \Omega \rightarrow [0, \infty] \) bounded when there is some constant \( C \in [0, \infty) \) such that \( T(\omega) \leq C, \forall \omega \in \Omega \).

**Lemma 4.5.1** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a real-valued progressively measurable process such that \( E|X_T| < \infty \) and \( E[X_T] = 0 \) for each bounded \( \{\mathcal{F}_t\} \)-stopping time \( T \). Then \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale.

**Remark 4.5.2** The proof of this lemma is identical to the proof of Lemma 2.3.13, except that one
now allows \( s, t \), in that proof to be real-valued rather than integer-valued. Observe that Proposition
3.3.14(b), together with the postulated progressive measurability of the process \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \)
and the fact that \( \{T < \infty\} \equiv \Omega \) (which is a consequence of the hypothesized boundedness of \( T \)), ensures
that \( X_T \) is well-defined everywhere on \( \Omega \) and is a \( \mathcal{F}_T \)-measurable (hence \( \mathcal{F} \)-measurable) mapping. Thus
the expectations \( E|X_T| \) and \( E[X_T] \) in the statement of Lemma 4.5.1 make sense.

Our second preliminary result shows how one can approximate a given stopping time \( T \) by a sequence
of stopping times \( \{T_n, n = 1, 2 \ldots\} \), where each \( T_n \) assumes a discrete set of values only:

**Proposition 4.5.3** Suppose that \( T \) is a stopping time with respect to some filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \)
in the probability space \( (\Omega, \mathcal{F}, P) \), let \( \mathcal{F}_\infty \) be given by (3.3), and define

\[
(4.57) \quad T_n(\omega) \triangleq 2^{-n}([2^n T(\omega)] + 1), \quad \forall \omega \in \Omega, \forall n = 1, 2, \ldots
\]

Then each \( T_n, n = 1, 2, \ldots \) is a \( \{\mathcal{F}_t\} \)-stopping time and we have
(a) \( T(\omega) \leq T_{n+1}(\omega) \leq T_n(\omega) \quad \forall \omega \in \Omega, \forall n = 1, 2 \ldots \);
(b) \( \lim_{n \to \infty} T_n(\omega) = T(\omega) \quad \forall \omega \in \Omega ; \)
(c) \( T_n(\omega) \in \{k2^{-n} : k = 0, 1, 2, \ldots \} \quad \forall \omega \in \Omega, \forall n = 1, 2 \ldots \)
(d) Fix integers \( m, n \) with \( 1 \leq m \leq n < \infty \) and put \( \mathcal{G}_k \triangleq \mathcal{F}_{k2^{-n}}, \forall k = 0, 1, 2, \ldots, \infty, U \triangleq 2^n T_m \). Then
\( U \) takes values in \( \{0, 1, 2, \ldots, \infty\} \) and is a discrete-parameter stopping time with respect to the discrete
filtration \( \{\mathcal{G}_k, \forall k = 0, 1, 2, \ldots, \infty\} \). Moreover, we have \( \mathcal{G}_U = \mathcal{F}_{T_m} \) for the usual pre-\( \sigma \)-algebras

\[
\mathcal{G}_U \triangleq \{A \subset \Omega : A \cap \{U \leq k\} \in \mathcal{G}_k, \forall k = 0, 1, 2, \ldots, \infty\},
\]

\[
\mathcal{F}_{T_m} \triangleq \{A \subset \Omega : A \cap \{T_m \leq t\} \in \mathcal{F}_t, \forall t \in [0, \infty]\}.
\]

Proof : Clearly, \( T_n = \infty \) on \( \{T = \infty\} \), and \( T_n \triangleq k2^{-n} \) on \( \{(k-1)2^{-n} \leq T < k2^{-n}\}, k = 1, 2 \ldots \) It
follows that \( T(\omega) \leq T_{n+1}(\omega) \leq T_n(\omega) \) and \( \lim_{n \to \infty} T_n(\omega) = T(\omega), \forall \omega \in \Omega \). Moreover, for arbitrary
Proof: (a) Let
\[ \forall \omega \]
Suppose that the analogue of Theorem 2.3.9:
We are now able to establish the main result of this section, namely the following continuous-parameter stopping times with respect to the filtration \( t \)
and, for \( t = 0 \), we have \( \{ T_n \leq t \} = \emptyset \in \mathcal{F}_0 \). Thus, \( T_n \) is a \( \{ \mathcal{F}_t \} \)-stopping time. As for (d), we have seen that \( T_m \in \{ k2^{-m}, k = 0, 1, 2, \ldots, \infty \} \), so that \( U \in \{ 0, 1, 2, \ldots, \infty \} \), since \( m \leq n \). Moreover, \( \{ U \leq k \} = \{ T_m \leq 2^{-n}k \} \in \mathcal{F}_{k2^{-n}} = \mathcal{G}_k, \forall \ k = 0, 1, 2, \ldots, \infty \), therefore \( U \) is a stopping time with respect to \( \{ \mathcal{G}_k, \ k = 0, 1, 2, \ldots, \infty \} \). Next, fix some \( A \in \mathcal{F}_{T_m} \). Then \( A \cap \{ U \leq k \} = A \cap \{ T_m \leq k2^{-n} \} \in \mathcal{F}_{k2^{-n}} = \mathcal{G}_k, \forall \ k = 0, 1, 2, \ldots, \infty \). For the opposite set inclusion fix \( A \in \mathcal{G}_U \) and \( t \in [0, \infty) \), and let \( k \) be the non-negative integer such that \( k2^{-n} \leq t < (k + 1)2^{-n} \). Thus \( A \cap \{ T_m \leq t \} = A \cap \{ U \leq 2^nt \} = A \cap \{ U \leq k \} \in \mathcal{G}_k = \mathcal{F}_{k2^{-n}} \subset \mathcal{F}_t \) where the second set equality follows since \( U \) takes integer values only. Moreover, it is clear that \( A \cap \{ T_m \leq \infty \} = A \cap \{ U \leq \infty \} \in \mathcal{G}_\infty = \mathcal{F}_\infty \), thus \( A \in \mathcal{F}_{T_m} \).

We are now able to establish the main result of this section, namely the following continuous-parameter analogue of Theorem 2.3.9:

**Theorem 4.5.4** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty] \} \) is a right-continuous closed supermartingale on the probability space \( (\Omega, \mathcal{F}, P) \). If \( S, T : \Omega \to [0, \infty) \) are \( \{ \mathcal{F}_t \} \)-stopping times with \( S(\omega) \leq T(\omega) \), \( \forall \ \omega \in \Omega \), then:

(a) \( E[X_T] < \infty \);

(b) \( E[X_T | \mathcal{F}_S] \leq X_S \) a.s.

Proof: (a) Let \( T_n \) be the \( \{ \mathcal{F}_t \} \)-stopping time defined by (4.57) and fix some integer \( n \geq 2 \). Put \( Z_k \triangleq X_{k2^{-n}} \) and \( \mathcal{G}_k \triangleq \mathcal{F}_{k2^{-n}}, \forall \ k = 0, 1, 2, \ldots, \infty \). Then \( \{(Z_k, \mathcal{G}_k); k = 0, 1, 2, \ldots + \infty \} \) is clearly a (discrete-parameter) closed supermartingale. Proposition 4.5.3(d) ensures that \( U \triangleq 2^nT_n \) and \( V \triangleq 2^nT_{n-1} \) are discrete-parameter stopping times with respect to the filtration \( \{ \mathcal{G}_k, \ k = 0, 1, 2, \ldots, \infty \} \), and clearly \( U \leq V \). Then, Theorem 2.3.9 establishes that \( E[Z_U] < \infty \), \( E[Z_V] < \infty \), and \( E[Z_V \mid \mathcal{G}_U] \leq Z_U \) a.s. Clearly, \( Z_V = X_{T_{n-1}}, Z_U = X_{T_n} \), and Proposition 4.5.3(d) says that \( \mathcal{G}_U = \mathcal{F}_{T_n} \), whence

\[ E \left[ X_{T_{n-1}} \mid \mathcal{F}_{T_n} \right] \leq X_{T_n} \text{ a.s. } \forall \ n = 2, 3, \ldots \quad (4.58) \]

Next, put \( \mathcal{H}_k \triangleq \mathcal{F}_{T_{1-k}}, W_k \triangleq X_{T_{1-k}}, \forall \ k = 0, -1, -2, \ldots \). Then, in view of (4.58), we see that \( \{(W_k, \mathcal{H}_k); k = 0, -1, -2, \ldots \} \) is a negatively indexed supermartingale, and clearly, by Theorem 2.3.9 and the fact that \( T_{1-k} \geq 0 \), we have \( E[W_k] = E[X_{T_{1-k}}] \leq E[X_0] < \infty, \forall \ k = 0, -1, -2, \ldots \). Thus, by the obvious adaptation of Theorem 2.7.7 to the case of supermartingales, we see that the collection of random variables \( \{W_k, k = 0, -1, -2, \ldots \} \equiv \{X_{T_n}, n = 1, 2, \ldots \} \) is uniformly integrable. Now Proposition 4.5.3(a,b), together with the postulated right-continuity of the sample-paths \( t \to X_t(\omega) \), ensures
that \( \lim_{n \to \infty} X_{T_n} = X_T, \forall \omega \in \Omega \), thus from Theorem 2.6.10 and the preceding uniform integrability, we have

\[
(4.59) \quad E|X_T| < \infty \quad \text{and} \quad \lim_{n \to \infty} E|X_{T_n} - X_T| = 0.
\]

(b) Put \( S_n \triangleq 2^{-n}(2^n S + 1), \forall n = 1, 2, \ldots \) Then, in view of Proposition 4.5.3, each \( S_n \) is a \( \{\mathcal{F}_t\} \)-stopping time taking values in \( \{k2^{-n} : k = 0, 1, 2, \ldots\} \), and it is clear that \( S_n \leq T_n \), since \( S \leq T \). An argument which is identical to that used in (a) then establishes that \( E[X_{T_n} | \mathcal{F}_{S_n}] \leq X_{S_n} \) a.s. In particular, fixing \( A \in \mathcal{F}_S \), it follows from \( S \leq S_n \) that \( A \in \mathcal{F}_{S_n}, \forall n = 1, 2, \ldots \), and so

\[
(4.60) \quad E[X_{T_n}; A] \leq E[X_{S_n}; A], \quad \forall n = 1, 2, \ldots
\]

Now \( |E[X_{T_n}; A] - E[X_T; A]| \leq E[X_{T_n} - X_T] \), thus the convergence in (4.59) establishes

\[
\lim_{n \to \infty} E[X_{T_n}; A] = E[X_T; A].
\]

An identical limit, but with \( S_n, S \) substituted for \( T_n, T \) clearly holds as well, whence taking \( n \to \infty \) in (4.60), we get \( E[X_T; A] \leq E[X_S; A], \forall A \in \mathcal{F}_S \), as required.

Corollary 4.5.5 Suppose that \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty]\} \) is a right-continuous closed supermartingale, with \( X_t \leq 0 \) a.s. for each \( t \in [0, \infty] \). If \( \mathcal{U} \) denotes the set of all \( \{\mathcal{F}_t\} \)-stopping times then the collection of random variables \( \{X_T : T \in \mathcal{U}\} \) is uniformly integrable.

Proof : It follows at once from Theorem 4.5.4 that

\[
(4.61) \quad 0 \leq -X_T \leq E[-X_{\infty} | \mathcal{F}_T] \quad \text{a.s.}
\]

for each \( T \in \mathcal{U} \). Since \( E|X_{\infty}| < \infty \), Theorem 2.6.7 ensures that the collection of random variables \( \{E[-X_{\infty} | \mathcal{F}_T] ; T \in \mathcal{U}\} \) are uniformly integrable, and this fact together with (4.61) establishes the result.

The next result, which is a continuous-parameter analogue of Theorem 2.3.3, postulates a supermartingale which need not be closed, but in return insists that the stopping times be bounded. The easy proof of this result is left to the reader:

Corollary 4.5.6 Suppose that \( \{(X_t, \mathcal{F}_t) ; t \in [0, \infty]\} \) is a right-continuous supermartingale and \( S \) and \( T \) are \( \{\mathcal{F}_t\} \)-stopping times on \( (\Omega, \mathcal{F}, P) \) such that \( S(\omega) \leq T(\omega) \leq a, \forall \omega \in \Omega \), for a constant \( a \in [0, \infty) \). Then:

(a) \( E|X_T| < \infty \);

(b) \( E[X_T | \mathcal{F}_S] \leq X_S \) a.s.
Corollary 4.5.7 Suppose that \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a right-continuous uniformly integrable martingale, and let \( \mathcal{U} \) denote the set of all \( \{ \mathcal{F}_t \} \)-stopping times. Then the collection of random variables \( \{ X_T : T \in \mathcal{U} \} \) is uniformly integrable.

Proof: In view of Theorem 4.4.11, there exists some \( \mathcal{F}_\infty \)-measurable random variable \( X_\infty : \Omega \to \mathbb{R} \) such that \( E|X_\infty| < \infty \) and \( X_t = E[X_\infty | \mathcal{F}_t] \) a.s. for each \( t \in [0, \infty) \). Then Theorem 4.5.4 establishes that \( X_T = E[X_\infty | \mathcal{T}_T] \) a.s. for each \( T \in \mathcal{U} \), and uniform integrability of \( \{ X_T : T \in \mathcal{U} \} \) is an immediate consequence of Theorem 2.6.7. \( \square \)

We end this section with the following continuous-parameter analogue of Theorem 2.3.14. This result is central to the technique of localization, and will be essential for our development of stochastic calculus in later chapters:

Corollary 4.5.8 Suppose that \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a right-continuous martingale and \( T \) is an \( \{ \mathcal{F}_t \} \)-stopping time defined on \( (\Omega, \mathcal{F}, P) \). Then \( \{ (X_{t\wedge T}, \mathcal{F}_t); t \in [0, \infty) \} \) is a right-continuous martingale.

Proof: Proposition 3.3.14(a) ensures that \( \{ (X_{t\wedge T}, \mathcal{F}_t); t \in [0, \infty) \} \) is a progressively measurable process, whose sample-paths \( t \mapsto X_{t\wedge T}(\omega) : [0, \infty) \to \mathbb{R} \) are clearly right-continuous for each \( \omega \in \Omega \). Put \( Y_t \overset{\Delta}{=} X_{t\wedge T} - X_0, \forall \, t \in [0, \infty) \), and let \( S \) be a bounded \( \{ \mathcal{F}_t \} \)-stopping time. Then \( S \wedge T \) is also a bounded \( \{ \mathcal{F}_t \} \)-stopping time, so that Corollary 4.5.6 ensures that \( E|X_{S\wedge T}| < \infty \) and \( E[X_{S\wedge T}] = E[X_0] \). Thus, \( E|Y_S| < \infty \) and \( E[Y_S] = 0 \) for arbitrary bounded \( \{ \mathcal{F}_t \} \)-stopping times \( S \), whence, from Lemma 4.5.1, we see that \( \{ (Y_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a martingale, as required to establish the result. \( \square \)

Remark 4.5.9 Suppose that \( \{ X_t; t \in [0, \infty) \} \) is some process and \( T \) is a random time on the probability space \( (\Omega, \mathcal{F}, P) \). Henceforth we put

\[
X^T_t(\omega) \overset{\Delta}{=} X_{t\wedge T}(\omega),
\]

for all \( \omega \in \Omega \) and \( t \in [0, \infty) \), and call \( \{ X^T_t; t \in [0, \infty) \} \) the process obtained by stopping \( \{ X_t; t \in [0, \infty) \} \) at the random time \( T \). The essence of Corollary 4.5.8 is that, if \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a right-continuous martingale and \( T \) is a \( \{ \mathcal{F}_t \} \)-stopping time, then the process \( \{ (X^T_t, \mathcal{F}_t); t \in [0, \infty) \} \) obtained by “stopping” at \( T \) is still a right-continuous martingale.

4.6 Local Martingales

When we study stochastic integration in the next chapter we shall often be dealing with processes which are not martingales in the full sense of Definition 4.1.1 but which possess a partial martingale-like property in a “local sense”. Our goal in this section is to precisely articulate this notion of a “partial martingale” and to look at some of its basic properties:
Definition 4.6.1 A real-valued, right-continuous, adapted process \( \{(X_t, \mathcal{F}_t); t \in [0, \infty]\} \) on a probability space \((\Omega, \mathcal{F}, P)\) is a **local martingale** when there exists some sequence of \(\mathcal{F}_t\)-stopping times \(\{T_n, n = 1, 2, \ldots\}\) such that

(i) \(0 \leq T_n(\omega) \leq T_{n+1}(\omega), \ \forall \omega \in \Omega, \ \forall n = 1, 2, \ldots;\)

(ii) \(P[\lim_{n \to \infty} T_n = \infty] = 1;\)

(iii) \(\{(X_{t \wedge T_n}, \mathcal{F}_t); t \in [0, \infty)\}\) is a martingale for each \(n = 1, 2, \ldots\).

We have formulated the notion of a local martingale in the continuous-parameter setting only. One can also define discrete-parameter local martingales, but we have refrained from doing so since this idea is most useful in the continuous-parameter context.

Remark 4.6.2 The sequence of stopping times \(\{T_n, n = 1, 2, \ldots\}\) in Definition 4.6.1 is called a **localizing sequence** for the process \(\{(X_t, \mathcal{F}_t); t \in [0, \infty]\}\). Observe that \(\{S_n, n = 1, 2, \ldots\}\) defined by \(S_n \overset{\Delta}{=} n \wedge T_n\) is also a localizing sequence for \(\{(X_t, \mathcal{F}_t); t \in [0, \infty]\}\), since (i) and (ii) in Definition 4.6.1 clearly hold with \(S_n\) in place of \(T_n\), and Corollary 4.5.8, together with Definition 4.6.1(iii), says that \(\{(X_{t \wedge S_n}, \mathcal{F}_t); t \in [0, \infty)\}\) is a martingale for each \(n = 1, 2, \ldots\). Moreover, by Definition 4.6.1(iii) and Corollary 4.5.6, we have

\[
X_{t \wedge S_n} = E[X_{t \wedge S_n} | \mathcal{F}_{t \wedge S_n}] \quad \text{a.s.} \quad \forall t \in [0, \infty),
\]

for each \(n = 1, 2, \ldots\), thus Theorem 2.6.7 ensures that \(\{X_{t \wedge S_n}; t \in [0, \infty)\}\) is uniformly integrable. In view of the preceding, without loss of generality we can assume that a local martingale always has a localizing sequence \(\{S_n, n = 1, 2, \ldots\}\) such that \(S_n(\omega) \leq n, \ \forall \omega \in \Omega, \ \text{and} \ \{(X_{t \wedge S_n}, \mathcal{F}_t); t \in [0, \infty)\}\) is a **uniformly integrable** martingale. These additional properties are often useful, and in fact they are included as part of the definition of a localizing sequence in some books. (see e.g. Definition 6 on page 69 of Liptser and Shiryayev [22]).

Remark 4.6.3 Suppose that \(\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}\) is a local martingale with a localizing sequence \(\{T_n, n = 1, 2, \ldots\}\), and \(S\) is some \(\mathcal{F}_t\)-stopping time. Then \(\{(X_{t \wedge S}, \mathcal{F}_t); t \in [0, \infty)\}\) is a local martingale also with localizing sequence \(\{T_n, n = 1, 2, \ldots\}\), i.e. a local martingale stopped at a stopping time is a local martingale. Indeed, Definition 4.6.1(iii) says that \(\{(X_{t \wedge T_n}, \mathcal{F}_t); t \in [0, \infty)\}\) is a right-continuous martingale, \(\forall \ n = 1, 2, \ldots\), hence Corollary 4.5.8 shows that \(\{(X_{t \wedge S \wedge T_n}, \mathcal{F}_t); t \in [0, \infty)\}\) is a martingale, \(\forall \ n = 1, 2, \ldots\), so that \(\{(X_{t \wedge S}, \mathcal{F}_t); t \in [0, \infty)\}\) is indeed a local martingale, with a localizing sequence \(\{T_n, n = 1, 2, \ldots\}\).

Remark 4.6.4 It is clear from Corollary 4.5.8 that, if \(\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}\) and \(\{(Y_t, \mathcal{F}_t); t \in [0, \infty)\}\) are local martingales with localizing sequences \(\{S_n, n = 1, 2, \ldots\}\) and \(\{T_n, n = 1, 2, \ldots\}\) respectively, then \(\{S_n \wedge T_n, n = 1, 2, \ldots\}\) is a common localizing sequence for both local martingales. It then follows at once that \(\{(\alpha X_t + \beta Y_t, \mathcal{F}_t); t \in [0, \infty)\}\) is a local martingale for constants \(\alpha, \beta \in \mathbb{R}\), that is, the set of local martingales relative to a fixed filtration \(\mathcal{F}_t, t \in [0, \infty)\) is a real vector space with respect to the usual notion of vector addition and scalar multiplication.
Remark 4.6.5 If \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\), we shall use \( M_{\text{loc}}(\{ \mathcal{F}_t \}, P) \) to denote the set of all processes \( \{ X_t; t \in [0, \infty) \} \) on \((\Omega, \mathcal{F}, P)\) such that \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a local martingale, and write \( X \in M_{\text{loc}}(\{ \mathcal{F}_t \}, P) \) to indicate that \( \{ X(t); t \in [0, \infty) \} \) is an element of \( M_{\text{loc}}(\{ \mathcal{F}_t \}, P) \). Of particular interest in later chapters will be local martingales with continuous sample paths, and we denote the set of these by \( M_{\text{loc}}^0(\{ \mathcal{F}_t \}, P) \). It follows from Remark 4.6.4 that \( M_{\text{loc}}(\{ \mathcal{F}_t \}, P) \) and \( M_{\text{loc}}^0(\{ \mathcal{F}_t \}, P) \) are real vector spaces. Finally, we denote by \( M_{\text{loc}}^0(\{ \mathcal{F}_t \}, P) \) the real vector space of local martingales [continuous local martingales] which are null at the origin.

Remark 4.6.6 It is clear that, if \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a right-continuous martingale, then it is also a local martingale, an obvious choice for the localizing sequence being \( \{ T_n \equiv +\infty, n = 1, 2 \ldots \} \). The question naturally arises: can one give examples of local martingales which fail to be martingales? Such examples would show that the set of local martingales is strictly larger than the set of martingales. Examples of this kind are best discussed when we have available the notion of stochastic integration, and will be deferred to the next chapter. Of more immediate importance are criteria which enable one to recognize when a given local martingale has genuine martingale-like properties. The next two propositions deal with this question:

Proposition 4.6.7 Suppose that \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a local martingale on the probability space \((\Omega, \mathcal{F}, P)\) with \( X_t \geq 0 \) a.s. for each \( t \in [0, \infty) \). Then \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \) is a supermartingale.

Proof: Let \( \{ T_n, n = 1, 2, \ldots \} \) be a localizing sequence for \( \{ (X_t, \mathcal{F}_t); t \in [0, \infty) \} \), and fix \( 0 \leq s < t < \infty \). Then, since \( \{ (X_{t\wedge T_n}, \mathcal{F}_t); t \in [0, \infty) \} \) is a martingale for each \( n = 1, 2 \ldots \), we have that \( X(t \wedge T_n) \) and \( X(0) \) are integrable, and

\[
E[X(t \wedge T_n)] = E[X(0)] < \infty \quad \text{and} \quad E[X(t \wedge T_n) | \mathcal{F}_s] = X(s \wedge T_n) \text{ a.s.} \tag{4.62}
\]

for each \( n = 1, 2, \ldots \). Now \( \lim_{n \to \infty} T_n = \infty \) a.s. thus

\[
\liminf_{n \to \infty} X(s \wedge T_n) = X(s) \text{ a.s.} \quad \text{and} \quad \liminf_{n \to \infty} X(t \wedge T_n) = X(t) \text{ a.s.} \tag{4.63}
\]

From Fatou’s theorem (see Theorem 1.2.15), the first relation of (4.62), and (4.63),

\[
E[X(t)] = E[\liminf_{n \to \infty} X(t \wedge T_n)] \leq \liminf_{n \to \infty} E[X(t \wedge T_n)] = E[X(0)] < \infty,
\]

which shows that \( X(t) \) is integrable for each \( t \in [0, \infty) \). Next, from Fatou’s theorem for conditional expectations (see Theorem 1.4.19(b)),

\[
E[X(t) | \mathcal{F}_s] = E\left[ \liminf_{n \to \infty} X(t \wedge T_n) \bigg| \mathcal{F}_s \right] \leq \liminf_{n \to \infty} E[X(t \wedge T_n) | \mathcal{F}_s] \text{ a.s.} \tag{4.64}
\]

Taking \( \liminf_{n \to \infty} \) in the second relation of (4.62) and using (4.63), (4.64), we get \( E[X(t) | \mathcal{F}_s] \leq X(s) \) a.s. as required. \( \blacksquare \)
Proposition 4.6.8 Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a local martingale with a localizing sequence \( \{T_n, n = 1, 2, \ldots\} \). If the sequence of random variables \( \{X_{t \wedge T_n}, n = 1, 2, \ldots\} \) is uniformly integrable for each \( t \in [0, \infty) \), then \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale.

Proof: From Definition 4.6.1 we know that \( \{(X_{t \wedge T_n}, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale and, for each \( t \in [0, \infty) \), we clearly have

\[
\lim_{n \to \infty} X_{t \wedge T_n} = X_t \quad \text{a.s.}
\]

In view of the postulated uniform integrability of \( \{X_{t \wedge T_n}, n = 1, 2, \ldots\} \) together with Theorem 2.6.10, the convergence in (4.65) holds in the \( L^1 \)-sense as well, and \( E[|X_t|] < \infty \). To verify the remaining conditions in the definition of a martingale, we note that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted process (by the definition of a local martingale). Lastly, fixing arbitrary \( s, t \in [0, \infty) \) with \( s < t \), we see from Definition 4.6.1(iii) that

\[
E[X_{t \wedge T_n} \mid \mathcal{F}_s] = X_{s \wedge T_n} \quad \text{a.s.}
\]

Since the convergence in (4.65) is in \( L^1 \), we observe from Proposition 1.4.21(b) that \( \lim_{n \to \infty} E[X_{t \wedge T_n} \mid \mathcal{F}_s] = E[X_t \mid \mathcal{F}_s] \) (in \( L^1 \)-sense), and of course we have \( \lim_{n \to \infty} X_{s \wedge T_n} = X_s \) a.s., (and, by uniform integrability, in \( L^1 \) as well). Taking \( n \to \infty \) in (4.66) then gives \( E[X_t \mid \mathcal{F}_s] = X_s \) a.s.

Given a local martingale, it is often necessary to construct a specific instance of a localizing sequence. In general this is a difficult task. However, when the sample-paths of the local martingale are continuous then the construction of a localizing sequence becomes easy:

Proposition 4.6.9 Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous local martingale on \((\Omega, \mathcal{F}, P)\). Put

\[
S_k(\omega) \triangleq \inf\{t \in [0, \infty) : |X_t(\omega)| \geq k\},
\]

\( \forall \omega \in \Omega, \forall k = 1, 2, \ldots \) Then \( \{S_k, k = 1, 2, \ldots\} \) is a localizing sequence for \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) and the martingale \( \{(X_{t \wedge S_k}, \mathcal{F}_t); t \in [0, \infty)\} \) is uniformly bounded: \( |X_{t \wedge S_k}(\omega)| \leq k, \forall \omega \in \Omega, \forall t \in [0, \infty) \).

Proof: In view of Remark 3.3.8 we see that each \( S_k \) is a \( \{\mathcal{F}_t\} \)-stopping time such that \( S_k(\omega) \leq S_{k+1}(\omega), \forall \omega \in \Omega, \forall k = 1, 2, \ldots, \) and \( \lim_{k \to \infty} S_k(\omega) = \infty, \forall \omega \in \Omega. \) Thus parts (i) and (ii) in Definition 4.6.1 are verified. To verify Definition 4.6.1(iii), let \( k \) be some arbitrary positive integer. Since \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a local martingale, it must have some localizing sequence \( \{T_n, n = 1, 2, \ldots\} \), whence Remark 4.6.3 establishes that \( \{(X_{t \wedge S_k}, \mathcal{F}_t); t \in [0, \infty)\} \) is a local martingale with the same localizing sequence \( \{T_n, n = 1, 2, \ldots\} \). But, by the definition of \( S_k \), it follows that \( \sup_{t, \omega, n} |X_{t \wedge S_k \wedge T_n}(\omega)| \leq k \), hence trivially the collection of random variables \( \{X_{t \wedge S_k \wedge T_n}, n = 1, 2, \ldots\} \) is uniformly integrable for each \( t \in [0, \infty) \). It follows from Proposition 4.6.8 that \( \{(X_{t \wedge S_k}, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale, as required.
4.7 The Quadratic Variation Process

In this section our goal is to introduce the notion of a quadratic variation process, a concept which is ubiquitous in the basic theory of martingales. As we shall see in the next chapter, the quadratic variation process will allow us to completely characterize the class of processes which are stochastically integrable with respect to a continuous local martingale, as well as to write down Itô’s celebrated formula on stochastic integrals. The quadratic variation process is usually defined on the basis of the Doob-Meyer decomposition theorem for general submartingales, a result which is well outside the scope of the present elementary introduction. Here we shall take full advantage of the fact that we really need the idea of quadratic variations only for local martingales with continuous sample-paths, in which case one can avoid use of a Doob-Meyer decomposition. Our first task is to clearly formulate the notion of a quadratic variation process:

**Definition 4.7.1** Suppose that \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\) is a continuous local martingale on the probability space \((\Omega, \mathcal{F}, P)\). An \(\mathbb{R}\)-valued process \(\{A(t); t \in [0, \infty)\}\) on \((\Omega, \mathcal{F}, P)\) is a quadratic variation process corresponding to \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\) when the following conditions hold:

(a) \(\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}\) is an adapted process;

(b) The mappings \(t \rightarrow A(t, \omega)\) are continuous and non-decreasing on \([0, \infty)\) for all \(\omega \in \Omega\);

(c) \(A(0) \equiv 0\);

(d) \(\{(X^2(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\}\) is a continuous local martingale.

**Remark 4.7.2** For typographical convenience we shall, throughout this section, use the notational convention of Remark 3.1.5.

Having formulated the notion of a quadratic variation process we must now demonstrate that such a process actually exists. For this purpose we shall need the following result which really pertains to discrete-parameter martingales:

**Lemma 4.7.3** Suppose that \(\{(Z_n, \mathcal{F}_n); n = 0, 1, 2, \ldots\}\) is a discrete-parameter martingale on the probability space \((\Omega, \mathcal{F}, P)\) such that \(|Z_n(\omega)| \leq C, \forall \omega \in \Omega, \forall n = 0, 1, 2, \ldots\), for some constant \(C \in [0, \infty)\). Then

\[
E \left[ \left( \sum_{n=0}^{N-1} (Z_{n+1} - Z_n)^2 \right)^2 \right] \leq 6C^4
\]

for all \(N = 1, 2, 3\ldots\)

Proof: Fix some arbitrary integer \(N \geq 1\) and put \(U_n \triangleq (Z_{n+1} - Z_n)^2\) for all \(n = 0, 1, 2\ldots\) Then

\[
\left( \sum_{n=0}^{N-1} U_n \right)^2 = \left( \sum_{n=0}^{N-1} U_n \right) \cdot \left( \sum_{m=0}^{N-1} U_m \right) = \sum_{n=0}^{N-1} U_n^2 + 2 \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} U_m U_n
\]
Now clearly $U_n \leq 4C^2$ for all $n$, thus

$$(4.70) \quad \sum_{n=0}^{N-1} U_n^2 \leq (\max_{0 \leq i < N} U_i) \cdot \sum_{n=0}^{N-1} U_n \leq 4C^2 \sum_{n=0}^{N-1} U_n. $$

Moreover, since $U_m$ is $\mathcal{F}_{m+1}$-measurable, we obtain from the composition rule and linearity of conditional expectations that

$$(4.71) \quad E \left\{ \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} U_m U_n \right\} = E \left\{ \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} E[U_m U_n | \mathcal{F}_{m+1}] \right\} = E \left\{ \sum_{m=0}^{N-2} U_m \left( \sum_{n=m+1}^{N-1} E[U_n | \mathcal{F}_{m+1}] \right) \right\}. $$

Now, we have

$$(4.72) \quad E[U_n | \mathcal{F}_{m+1}] = E[(Z_{n+1} - Z_n)^2 | \mathcal{F}_{m+1}] = E[Z_{n+1}^2 + Z_n^2 | \mathcal{F}_{m+1}] - 2E[Z_n Z_{n+1} | \mathcal{F}_{m+1}]$$

for all $n > m$, and from the composition rule for conditional expectations,

$$(4.73) \quad E[Z_n Z_{n+1} | \mathcal{F}_{m+1}] = E[E[Z_n Z_{n+1} | \mathcal{F}_n] | \mathcal{F}_{m+1}] = E[Z_n E[Z_{n+1} | \mathcal{F}_n] | \mathcal{F}_{m+1}] = E[Z_n^2 | \mathcal{F}_{m+1}],$$

for all $n > m$. Now, by (4.72) and (4.73):

$$(4.74) \quad \sum_{n=m+1}^{N-1} E[U_n | \mathcal{F}_{m+1}] = \sum_{n=m+1}^{N-1} E[Z_{n+1}^2 - Z_n^2 | \mathcal{F}_{m+1}] = E[Z_N^2 | \mathcal{F}_{m+1}] - E[Z_{m+1}^2 | \mathcal{F}_{m+1}] \leq C^2.$$

Combining (4.74) with (4.71) gives

$$(4.75) \quad E \left\{ \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} U_m U_n \right\} \leq C^2 E \left\{ \sum_{m=0}^{N-2} U_m \right\}. $$

Now, taking expectations in (4.69), and using (4.70) together with (4.75), we obtain

$$(4.76) \quad E \left\{ \left( \sum_{n=0}^{N-1} U_n \right)^2 \right\} \leq 6C^2 E \left\{ \sum_{n=0}^{N-1} U_n \right\}. $$

Since $U_n = Z_{n+1}^2 - 2Z_n Z_{n+1} + Z_n^2$ and 

$$E[Z_n Z_{n+1}] = E[E[Z_n Z_{n+1} | \mathcal{F}_n]] = E[Z_n E[Z_{n+1} | \mathcal{F}_n]] = E[Z_n^2]$$

158
it follows that $E[U_n] = E[Z_{n+1}^2 - Z_n^2]$, hence

\[(4.77) \quad E \left\{ \sum_{n=0}^{N-1} U_n \right\} = \sum_{n=0}^{N-1} \{E[Z_{n+1}^2] - E[Z_n^2]\} \leq E[Z_N^2]. \]

Now (4.68) follows from (4.76) and (4.77).

Using this lemma we first prove existence of a quadratic variation process for a rather more special structure than a continuous local martingale, namely a uniformly bounded continuous martingale, and show how this process arises as an $L^2$-limit of a sequence of “simple” processes. To establish existence of a quadratic variation process we shall need to use Propositions 4.1.13 and 4.3.4, and thus we postulate the following

**Condition 4.7.4** \(\{\mathcal{F}_t, t \in [0, \infty)\}\) is a given filtration in a probability space \((\Omega, \mathcal{F}, P)\), where \(\mathcal{F}_0\) includes all \(P\)-null events of \(\mathcal{F}\).

**Lemma 4.7.5** Suppose that Condition 4.7.4 holds, \(X \in M^c(\{\mathcal{F}_t\}, P)\), and \(|X(t, \omega)| \leq C, \forall \omega \in \Omega, \forall t \in [0, \infty)\), for some constant \(C \in [0, \infty)\). Put \(\tau^n_k \triangleq k2^{-n}, \forall k, n = 0, 1, 2\ldots\) and define

\[(4.78) \quad A_n(t, \omega) \triangleq \sum_{0 \leq k < \infty} [X(t \wedge \tau^n_k, \omega) - X(t \wedge \tau^n_{k+1}, \omega)]^2 \]

\(\forall t \in [0, \infty), \forall \omega \in \Omega, \forall n = 0, 1, 2\ldots\) Then the sequence of random variables \(\{A_n(t); n = 0, 1, 2\ldots\}\) converges in \(L^2(\Omega, \mathcal{F}, P)\) to some limit \(A(t)\) for each \(t \in [0, \infty)\), and the process \(\{A(t); t \in [0, \infty)\}\) has the following properties:

(a) \(\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}\) is an adapted process;

(b) The mappings \(t \to A(t, \omega)\) are continuous and non-decreasing on \([0, \infty)\) for all \(\omega \in \Omega\);

(c) \(A(0) \equiv 0\) and \(E[A^2(t)] < \infty, \forall t \in [0, \infty)\);

(d) \(\{(X^2(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\}\) is a continuous martingale.

In particular, \(\{A(t); t \in [0, \infty)\}\) is a quadratic variation process for \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\).

Proof: Notice that, for each \(t \in [0, \infty)\), there are only finitely many non-zero terms in the summation of (4.78). From Definition 4.7.1 it follows that a quadratic variation process for the martingale \(\{(X(t) - X(0), \mathcal{F}_t); t \in [0, \infty)\}\) is also a quadratic variation process for the martingale \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\).

We shall therefore assume, without loss of generality, that \(X(0) \equiv 0\). Now define

\[(4.79) \quad M_n(t) \triangleq \sum_{0 \leq k < \infty} X(t \wedge \tau^n_k)[X(t \wedge \tau^n_{k+1}) - X(t \wedge \tau^n_k)] \]

\(\forall t \in [0, \infty), \forall n = 0, 1, 2\ldots\) It easily follows from (4.79) and (4.78) that

\[(4.80) \quad X^2(t, \omega) = 2M_n(t, \omega) + A_n(t, \omega) \]

159
\[ \forall \omega \in \Omega, \forall t \in [0, \infty), \forall n = 0, 1, 2 \ldots \text{ Moreover, since } t \to X(t, \omega) \text{ is continuous on } [0, \infty), \text{ from (4.78)} \]

we see that \( t \to A_n(t, \omega) \) is continuous on \( [0, \infty) \) with \( A_n(0) \equiv 0, \) for each \( n = 1, 2, \ldots, \) and from (4.79) it is easily verified that \( \{(M_n(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale. Now fix integers \( n, m \) such that \( n > m > 0, \) and fix some \( t \in [0, \infty). \) We clearly have

\[
[X(t \wedge \tau_{k+1}^m) - X(t \wedge \tau_k^m)] = \sum_{0 \leq j < \infty} I_{[t \wedge \tau_k^m, t \wedge \tau_{k+1}^m]}(t \wedge \tau_j^n) \cdot [X(t \wedge \tau_{j+1}^n) - X(t \wedge \tau_j^n)]
\]

thus

\[
(4.81) \quad M_m(t) = \sum_{0 \leq k < \infty} \sum_{0 \leq j < \infty} X(t \wedge \tau_k^m) I_{[t \wedge \tau_k^m, t \wedge \tau_{k+1}^m]}(t \wedge \tau_j^n) [X(t \wedge \tau_{j+1}^n) - X(t \wedge \tau_j^n)]
\]

\[
= \sum_{0 \leq j < \infty} U_j[X(t \wedge \tau_{j+1}^n) - X(t \wedge \tau_j^n)]
\]

for

\[
(4.82) \quad U_j \triangleq \sum_{0 \leq k < \infty} X(t \wedge \tau_k^m) I_{[t \wedge \tau_k^m, t \wedge \tau_{k+1}^m]}(t \wedge \tau_j^n)
\]

(changing order of summation at the second equality of (4.81) is permissible since, for each \( t \in [0, \infty), \) there are only finitely many non-zero terms in the double summation). Now, for each fixed \( j, \) there is clearly only one value of \( k \) which contributes to the summation on the right side of (4.82), namely that \( k \) such that \( \tau_k^m \leq \tau_j^n < \tau_{k+1}^m, \) or equivalently, such that \( j \in [k2^{n-m}, (k+1)2^{n-m}). \) We denote this value of \( k \) by \( k(j), \) and it follows at once that

\[
(4.83) \quad k(j)2^{-m} \leq j2^{-n} < (k(j) + 1)2^{-m}.
\]

From (4.82) we get that \( U_j = X(t \wedge \tau_{k(j)}^m) \) hence \( U_j \) is \( \mathcal{F}_{t \wedge \tau_j^m} \)-measurable since, from (4.83), we have \( t \wedge \tau_{k(j)}^m \leq t \wedge \tau_j^m. \) Now, from (4.81), (4.82), and (4.79) we find that

\[
(4.84) \quad M_n(t) - M_m(t) = \sum_{0 \leq j < \infty} [X(t \wedge \tau_j^n) - U_j][X(t \wedge \tau_{j+1}^n) - X(t \wedge \tau_j^n)].
\]

Put \( V_j \triangleq [X(t \wedge \tau_j^n) - U_j] \) and \( \Delta_j \triangleq [X(t \wedge \tau_{j+1}^n) - X(t \wedge \tau_j^n)]. \) Then, by the \( \mathcal{F}_{t \wedge \tau_j^n} \)-measurability of \( U_j \) noted earlier, and the fact that \( \Delta_j \) is \( \mathcal{F}_{t \wedge \tau_{j+1}^n} \)-measurable, for all \( j < k \) we find that

\[
E[(V_j \Delta_j)(V_k \Delta_k)] = E[E [V_j \Delta_j V_k \Delta_k | \mathcal{F}_{t \wedge \tau_j^n}]] = E[V_j \Delta_j V_k E [\Delta_k | \mathcal{F}_{t \wedge \tau_k^m}]] = 0
\]

where we have use the fact that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale at the last equality. Thus all terms in the summation of (4.84) are orthogonal, and hence

\[
(4.85) \quad E\{[M_n(t) - M_m(t)]^2\} = E\left\{ \sum_{0 \leq j < \infty} V_j^2 \Delta_j^2 \right\} \leq E\left\{ W_m(t) \sum_{0 \leq j < \infty} \Delta_j^2 \right\}
\]
$$W^m(t, \omega) \triangleq \max_{0 \leq s - u \leq 2^{-m}} [X(s, \omega) - X(u, \omega)]^2.$$  

At the last inequality in (4.85) we have used the fact that $\tau_{k(j)}^n \leq \tau_j^m < \tau_{k(j)+1}^m$ to see that $0 \leq \tau_j^m - \tau_{k(j)}^m \leq 2^{-m}$ which in turn implies that $V_j^2(\omega) \leq W^m(t, \omega)$. From (4.85) and the Cauchy-Schwarz inequality we obtain

$$E\{[M_n(t) - M_m(t)]^2\} \leq E^{1/2}\left\{ \sum_{0 \leq j < \infty} \Delta_j^2 \right\} \cdot E^{1/2}\{[W^m(t)]^2\} \leq \sqrt{6C^2}E^{1/2}\{[W^m(t)]^2\},$$

for all $t \in [0, \infty)$ and all integers $n > m$. To get the last inequality of (4.87) we used Lemma 4.7.3 along with the fact that $\{(X(t \wedge \tau_j^n), \mathcal{F}_{t \wedge \tau_j^n}) ; j = 0, 1, 2 \ldots\}$ is a discrete-parameter martingale uniformly bounded by $C$. Now, for each $\omega$, the mapping $s \rightarrow X(s, \omega)$ is continuous on $[0, \infty)$, hence uniformly continuous on the closed and bounded interval $[0, t]$ for each $t \in [0, \infty)$. It follows that

$$\lim_{m \to \infty} W^m(t, \omega) = 0$$

for all $\omega \in \Omega$, all $t \in [0, \infty)$. Now, in view of (4.86), we observe that $|W^m(t, \omega)| \leq 4C^2$ for all $\omega \in \Omega$, all $t \in [0, \infty)$, hence the Lebesgue Dominated Convergence Theorem 1.2.21 together with (4.88) says that

$$\lim_{m \to \infty} E^{1/2}\{[W^m(t)]^2\} = 0$$

for each $t \in [0, \infty)$. Since the upper bound in (4.87) holds for all $n > m$ and $t \in [0, \infty)$ we observe from (4.89) that

$$\lim_{m \to \infty} \sup_{n > m} E\{[M_n(t) - M_m(t)]^2\} = 0.$$

Hence $\{M_n(t), n = 0, 1, 2 \ldots\}$ is a Cauchy sequence in the complete metric space $L^2(\Omega, \mathcal{F}, P)$ and therefore converges in $L^2$ to some limit $M^*(t) \in L^2(\Omega, \mathcal{F}, P)$. Now this convergence holds for each $t \in [0, \infty)$ and $\{(M_n(t), \mathcal{F}_t) ; t \in [0, \infty)\}$ is a continuous $L^2$-martingale for each $n = 0, 1, 2 \ldots$. Since $\mathcal{F}_0$ includes all $P$-null events in $\mathcal{F}$, we can use Proposition 4.1.13 and Proposition 4.3.4 to see that $\{(M^*(t), \mathcal{F}_t) ; t \in [0, \infty)\}$ is an $L^2$-martingale with a modification which is a continuous martingale. We shall denote this continuous modification by $\{(M(t), \mathcal{F}_t) ; t \in [0, \infty)\}$. Now define

$$B(t, \omega) \triangleq X^2(t, \omega) - 2M(t, \omega)$$

for all $\omega \in \Omega$, $t \in [0, \infty)$. It follows at once that $\{(B(t), \mathcal{F}_t) ; t \in [0, \infty)\}$ is a continuous adapted process, and $\{(X^2(t) - B(t), \mathcal{F}_t) ; t \in [0, \infty)\}$ is a continuous martingale. We next show that the mappings $t \rightarrow B(t, \omega)$ are non-decreasing on $[0, \infty)$ for all $\omega$ outside some $P$-null event. Fix arbitrary $s, t$ in the set of dyadic rationals

$$D_+ \triangleq \{k2^{-n} : k, n = 0, 1, 2, \ldots\},$$

161
with \( s < t \). Since \( M(t) = M^*(t) \) a.s. we see that \( M(t) \) is an \( L^2 \)-limit of the sequence \( \{M_n(t), n = 0, 1, 2 \ldots \} \), and in view of (4.91) and (4.80) it follows that \( B(t) \) is an \( L^2 \)-limit of the sequence \( \{A_n(t), n = 0, 1, 2 \ldots \} \). The Markov inequality (Theorem 1.2.23) then implies that \( \{A_n(t), n = 0, 1, 2 \ldots \} \) converges in probability to \( B(t) \), hence Theorem 1.2.11(b) says that there is some subsequence \( \{A_{n_k}(t), k = 0, 1, 2 \ldots \} \) of the sequence \( \{A_n(t), n = 0, 1, 2 \ldots \} \) and some \( N_t \in \mathcal{F} \) such that \( P(N_t) = 0 \) and

\[
\lim_{k \to \infty} A_{n_k}(t, \omega) = B(t, \omega)
\]

for all \( \omega \not\in N_t \). Since \( B(s) \) is likewise an \( L^2 \)-limit of \( \{A_n(s), n = 0, 1, 2 \ldots \} \), it is clearly also an \( L^2 \)-limit of the sequence \( \{A_{n_k}(s), k = 0, 1, 2 \ldots \} \). In view of the Markov inequality and Theorem 1.2.11(a), we see that there is some subsequence \( \{A_{r_n}(s), r = 0, 1, 2 \ldots \} \) of the sequence \( \{A_{n_k}(s), k = 0, 1, 2 \ldots \} \) and some \( N_s \in \mathcal{F} \) such that \( P(N_s) = 0 \) and \( \lim_{r \to \infty} A_{r_n}(s, \omega) = B(s, \omega) \) \( \forall \omega \not\in N_s \). Since \( s,t \in D_+ \) and \( \lim_{r \to \infty} n_r = +\infty \), there is some integer \( r_0 \), depending on \( s,t \) such that, for each \( r \geq r_0 \), we can write \( s = k(r, s)2^{-nr} \) and \( t = k(r, t)2^{-nr} \) for non-negative integers \( k(r, s) \) and \( k(r, t) \). From (4.78) we see that \( A_{r_n}(t, \omega) \geq A_{r_n}(s, \omega) \), \( \forall \omega \in \Omega \), and integers \( r \geq r_0 \), and since (4.92) implies, trivially, that \( \lim_{r \to \infty} A_{r_n}(t, \omega) = B(t, \omega) \), \( \forall \omega \not\in N_t \), we see that \( B(t, \omega) \geq B(s, \omega) \) for all \( \omega \not\in N_{s,t} \overset{\triangle}{=} N_s \cup N_t \). Now put

\[
N \overset{\triangle}{=} \bigcup_{s,t} N_{s,t}
\]

the set union being taken over all dyadic rational \( s,t \in D_+ \) with \( s < t \). Clearly \( N \in \mathcal{F} \), \( P(N) = 0 \), and the mappings \( u \to B(u, \omega) : D_+ \to \mathbb{R} \) are non-decreasing for each \( \omega \not\in N \). Since \( B(\cdot, \omega) \) is continuous on \( [0, \infty) \) and \( D_+ \) is dense in \( [0, \infty) \), this in turn implies that \( B(\cdot, \omega) \) is non-decreasing on all of \( [0, \infty) \) for each \( \omega \not\in N \). Finally, for each \( t \in [0, \infty) \) we define \( A(t, \omega) \overset{\triangle}{=} B(t, \omega) \) for all \( \omega \not\in N \), and \( A(t, \omega) \overset{\triangle}{=} 0 \) for all \( \omega \in N \). Since \( \mathcal{F}_0 \) includes all \( P \)-null events of \( \mathcal{F} \), we see from Proposition 3.1.18 that \( \{(A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous adapted process such that \( \{(X^2(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale and each \( t \to A(t, \omega) \) is non-decreasing on \( [0, \infty) \). Moreover, it is clear that \( A(t) \) is an \( L^2 \)-limit of the sequence \( \{A_n(t), n = 0, 1, 2 \ldots \} \) for each \( t \in [0, \infty) \), hence \( E|A(t)| \leq E[A^2(t)] < \infty \). Lastly, we easily observe from (4.79) and \( X(0) \equiv 0 \) that \( A(0) \equiv 0 \).

In the preceding lemma we constructed a quadratic variation process for a uniformly bounded continuous martingale as an \( L^2 \)-limit of the sums in (4.78). The question immediately arises: is it possible for there to exist more than one quadratic variation process for this martingale? We are going to show that the quadratic variation process is unique in the sense that any two such processes are necessarily indistinguishable. To this end, we first digress to recall the notion of the total variation of a function of a real variable. Since total variation also plays a significant role in several later chapters, we shall define this concept in full and state its most basic properties:

**Definition 4.7.6** The mapping \( A : [0, \infty) \to \mathbb{R} \) is said to be of bounded variation on the interval
[s, t] ⊂ [0, ∞) if there exists some \( c \in [0, \infty) \), generally depending on \( s \) and \( t \), such that

\[
\sum_{k=1}^{n} |A(t_k) - A(t_{k-1})| \leq c,
\]

for each and every finite partition \( s = t_0 < t_1 < \ldots < t_n = t \) of \([s, t]\). The mapping is said to be of unbounded variation on the interval \([s, t]\) when such a finite constant \( c \) fails to exist, that is when the supremum of the quantity on the left hand side of (4.93), taken over all finite partitions of \([s, t]\), is \(+\infty\). The mapping \( A \) is said to be of locally bounded variation when it is of bounded variation on the interval \([0, t]\) for each \( t \in [0, \infty) \).

**Definition 4.7.7** For a given mapping \( A : [0, \infty) \to \mathbb{R} \), the total variation of \( A \) on the interval \([s, t]\) \( \subset [0, \infty) \), denoted by \( V[A; s, t] \), is the quantity

\[
V[A; s, t] \triangleq \sup \sum_{k=1}^{n} |A(t_k) - A(t_{k-1})|,
\]

the supremum being taken over all finite partitions \( s = t_0 < t_1 < \ldots < t_n = t \) of \([s, t]\). The mapping \( t \to V[A; 0, t] : [0, \infty) \to [0, \infty] \) is called the total variation function of \( A \).

From the preceding definitions we see that the mapping \( A : [0, \infty) \to \mathbb{R} \) is of locally bounded variation if and only if the total variation function of \( A \) is finite-valued i.e. takes its values in \([0, \infty)\).

**Example 4.7.8**

(a) From Definition 4.7.7 we see that the function \( A : [0, \infty) \to \mathbb{R} \) assumes some constant value if and only if \( V[A; 0, t] = 0 \) for all \( t \in [0, \infty) \).

(b) Suppose that \( A : [0, \infty) \to \mathbb{R} \) is either non-increasing or non-decreasing. Then, clearly,

\[
V[A; s, t] = |A(t) - A(s)|, \quad \forall 0 \leq s < t < \infty,
\]

thus \( A \) is of locally bounded variation.

(c) Suppose that \( A(0) \triangleq 0 \) and \( A(t) \triangleq t \sin(\pi/t) \quad \forall t \in (0, \infty) \). Then \( A \) is of unbounded variation on any interval \([0, t]\) where \( t \in (0, \infty) \). Indeed, fix some positive integer \( n \) and consider the partition of \([0, t]\) by the points \( t_0 \triangleq 0 \) and \( t_k \triangleq t/(2n - 2k + 1), \forall k = 1, 2 \ldots n \). Then it is easily seen that

\[
\sum_{k=1}^{n} |A(t_k) - A(t_{k-1})| > \sum_{k=1}^{n} k^{-1}
\]

whence \( V[A; 0, t] = +\infty \).

The next proposition lists various basic properties of total variation. The proof, which involves only simple real analysis, is left to the reader:
Proposition 4.7.9 Suppose the mappings $A, B : [0, \infty) \to \mathbb{R}$ are of locally bounded variation. Then
(a) $A + B$ is of locally bounded variation, and $V[A + B; s, t] \leq V[A; s, t] + V[B; s, t], \forall 0 \leq s < t < \infty$.
(b) $cA$ is of locally bounded variation for every $c \in \mathbb{R}$, and $V[cA; s, t] = |c|V[A; s, t], \forall 0 \leq s < t < \infty$.
(c) $V[A; s, u] = V[A; s, t] + V[A; t, u], \forall 0 \leq s < t < u < \infty$.
(d) The mapping $t \mapsto V[A; 0, t]$ is non-decreasing on $[0, \infty)$ and $V[A; s, t] = 0$ when $s = t$.

Remark 4.7.10 For a given mapping $A : [0, \infty) \to \mathbb{R}$ of locally bounded variation, we see from Proposition 4.7.9(d) that the limit
$$V[A; 0, \infty] \overset{\Delta}{=} \lim_{n \to \infty} V[A; 0, n]$$
exists in $[0, \infty]$. We call $V[A; 0, \infty]$ the total variation of $A$ on $[0, \infty)$, and say that $A$ is of bounded variation on $[0, \infty)$ when $V[A; 0, \infty] < \infty$.

Remark 4.7.11 It follows at once from Proposition 4.7.9(a,b) together with Example 4.7.8(b) that any mapping which is the difference of two real-valued non-decreasing functions on $[0, \infty)$ is of locally bounded variation. We shall see later in Lemma 5.1.3 that this property actually characterizes mappings of locally bounded variation.

Having disposed of these preliminaries, we are now ready to return to the problem of uniqueness of the quadratic variation process in Lemma 4.7.5. The basic tool for establishing this uniqueness is the following lemma:

Lemma 4.7.12 Suppose that Condition 4.7.4 holds, $X \in \mathcal{M}^{00}(\{\mathcal{F}_t\}, P)$, and $|X(t, \omega)| \leq C, \forall \omega \in \Omega, \forall t \in [0, \infty)$, for some constant $C \in [0, \infty)$. If

$$V[X(\cdot, \omega); 0, t] < \infty \quad \forall \omega \in \Omega, \forall t \in [0, \infty), \quad (4.95)$$

then $P[\omega : X(t, \omega) = 0, \forall t \in [0, \infty]] = 1$.

Proof: Put $\tau^n_k \overset{\Delta}{=} k2^{-n}, \forall n, k = 0, 1, 2 \ldots$ and define $A_n(t, \omega)$ as in (4.78), $\forall t \in [0, \infty), \forall \omega \in \Omega$. Fix some arbitrary $t \in [0, \infty)$. Then we see from (4.78) and the definition of total variation that

$$A_n(t, \omega) \leq \max_{0 \leq k < \infty} |X(t \wedge \tau^n_{k+1}, \omega) - X(t \wedge \tau^n_k, \omega)| \cdot V[X(\cdot, \omega); 0, t]. \quad (4.96)$$

Since $X(\cdot, \omega)$ is uniformly continuous on the closed bounded interval $[0, t]$ we have

$$\lim_{n \to \infty} \max_{0 \leq k < \infty} |X(t \wedge \tau^n_{k+1}, \omega) - X(t \wedge \tau^n_k, \omega)| = 0 \quad \forall \omega \in \Omega.$$

In view of (4.95) and (4.96) it follows that

$$\lim_{n \to \infty} A_n(t, \omega) = 0 \quad \forall \omega \in \Omega. \quad (4.97)$$
Thus, if \( \{A(t), \ t \in [0, \infty)\} \) is the process constructed in Lemma 4.7.5, then \( A(t) \) is an \( L^2 \)-limit of the sequence \( \{A_n(t), \ n = 0, 1, 2 \ldots\} \), hence we observe from (4.97) that \( A(t) = 0 \) a.s. It follows from Lemma 4.7.5 that \( \{(X^2(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is a martingale, and hence \( E[X^2(t)] = E[X^2(0)] = 0 \) for all \( t \in [0, \infty) \). Thus, if \( N_t \triangleq \{\omega : X(t, \omega) \neq 0\} \), then \( P(N_t) = 0 \). Now put \( N \triangleq \bigcup_{t \in \mathbb{Q}_+} N_t \). Clearly \( P(N) = 0 \) and, since \( X(\cdot, \omega) \) is continuous on \([0, \infty)\) for each \( \omega \) and \( \mathbb{Q}_+ \) is dense in \([0, \infty)\), it follows that

\[
N^c = \{\omega : X(t, \omega) = 0 \ \forall t \in [0, \infty)\}.
\]

Since \( P(N^c) = 1 \) the lemma is established.

For later applications we must generalize the preceding lemma to do away with the hypothesis of uniform boundedness:

**Lemma 4.7.13** Suppose that Condition 4.7.4 holds, and \( M \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P) \). If the mapping \( t \to M(t, \omega) \) is of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \) then \( P[M(t) = 0, \ \forall t \in [0, \infty)] = 1 \).

**Proof:** For each \( n = 1, 2 \ldots \) and \( \omega \in \Omega \) put

\[
T_n(\omega) \triangleq \inf\{u \in [0, \infty) : |M(u, \omega)| \geq n\}.
\]

In view of Proposition 4.6.9 we know that \( \{T_n, \ n = 1, 2 \ldots\} \) is a localizing sequence of \( \{\mathcal{F}_t\} \)-stopping times, thus defining \( M_n(t, \omega) \triangleq M(t \wedge T_n(\omega), \omega) \), we see that \( \{(M_n(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is a continuous martingale with \( |M_n(t, \omega)| \leq n, \ \forall \omega \in \Omega, \ \forall t \in [0, \infty), \ \forall n = 1, 2 \ldots \). Fix some arbitrary \( \omega \in \Omega \) and positive integer \( n \). Since the mapping \( t \to M(t, \omega) \) is of locally bounded variation on \([0, \infty)\), so obviously is the mapping \( t \to M_n(t, \omega) \). Thus, we can apply Lemma 4.7.12 to the uniformly bounded continuous martingale \( \{(M_n(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) to conclude that

\[
P[\omega : \ M_n(t, \omega) = 0 \ \forall t \in [0, \infty)] = 1.
\]

Since this holds for all \( n = 1, 2 \ldots \), and \( \lim_{n \to \infty} T_n(\omega) = +\infty, \ \forall \omega \in \Omega \), implies

\[
[\omega : \ M(t, \omega) = 0 \ \forall t \in [0, \infty)] = \bigcap_{1 \leq n < \infty} [\omega : \ M_n(t, \omega) = 0 \ \forall t \in [0, \infty)],
\]

the lemma is established.

**Lemma 4.7.5** shows that a uniformly bounded continuous martingale has an associated quadratic variation process. Using Lemma 4.7.13 we show that this quadratic variation process is unique to within indistinguishability:

**Lemma 4.7.14** Suppose that Condition 4.7.4 holds, and \( X \in \mathcal{M}^c(\{\mathcal{F}_t\}, P) \) with \( |X(t, \omega)| \leq C, \ \forall \omega \in \Omega, \ \forall t \in [0, \infty) \), for some constant \( C \in [0, \infty) \). If \( \{A(t); t \in [0, \infty)\} \) and \( \{B(t); t \in [0, \infty)\} \) are any two quadratic variation processes for \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) then \( P[\omega : \ A(t, \omega) = B(t, \omega), \ \forall t \in [0, \infty)] = 1 \).
Theorem 4.7.15
Suppose that Condition 4.7.4 holds.

(a) Corresponding to $X \in \mathfrak{M}_{\text{loc}}^2(\{\mathcal{F}_t\}; P)$ is a quadratic variation process $\{A(t); t \in [0, \infty)\}$, and if $\{B(t); t \in [0, \infty)\}$ is another such quadratic variation process, then $\{A(t); t \in [0, \infty)\}$ and $\{B(t); t \in [0, \infty)\}$ are indistinguishable.

(b) If $X \in \mathfrak{M}_{\text{loc}}^2(\{\mathcal{F}_t\}, P)$ and $\{A(t); t \in [0, \infty)\}$ is a corresponding quadratic variation process, then $E|A(t)| < \infty$, $\forall t \in [0, \infty)$ and $\{(X^2(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous martingale.

Proof: (a) For each $\omega \in \Omega$, $n = 1, 2, \ldots$ define

$$T_n(\omega) \triangleq \inf\{t \in [0, \infty): |X(t, \omega)| \geq n\}.$$

In view of Proposition 4.6.9 we know that $\{T_n, n = 1, 2, \ldots\}$ is a localizing sequence of $\{\mathcal{F}_t\}$-stopping times, thus $\{(X(t \wedge T_n), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous martingale with $\sup_{(t, \omega)} |X(t \wedge T_n(\omega), \omega)| \leq n$, $\forall n = 1, 2, \ldots$. Then, for each $n = 1, 2, \ldots$, Lemma 4.7.5 says there exists some $\{A_{n+1}(t), t \in [0, \infty)\}$ which is a quadratic variation process for the martingale $\{(X(t \wedge T_{n+1}), \mathcal{F}_t); t \in [0, \infty)\}$, and, in particular, $\{(X^2(t \wedge T_{n+1}) - A_{n+1}(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous martingale. Thus, from Corollary 4.5.8, it follows that $\{(X^2(t \wedge T_{n+1}) - A_{n+1}(t \wedge T_n), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous martingale.

Since $X(t \wedge T_{n+1} \wedge T_n) \equiv X(t \wedge T_n)$, it then follows from Lemma 4.7.14 that the continuous processes $\{A_n(t), t \in [0, \infty)\}$ and $\{A_{n+1}(t \wedge T_n), t \in [0, \infty)\}$ are indistinguishable. Put

$$\Omega_n \triangleq \{\omega \in \Omega: A_n(t, \omega) = A_{n+1}(t \wedge T_n(\omega), \omega) \quad \forall t \in [0, \infty)\}, \quad \Omega^* \triangleq \bigcap_{1 \leq n < \infty} \Omega_n,$$

and observe that $P(\Omega_n) = 1$, $\forall n = 1, 2, \ldots$, hence $P(\Omega^*) = 1$. We now “piece together” the processes $\{A_n(t), t \in [0, \infty)\}$ into an overall quadratic variation process for $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$. To this end, define

$$A^*(t, \omega) \triangleq \limsup_{n \to \infty} A_n(t, \omega), \quad \forall \omega \in \Omega, \quad \forall t \in [0, \infty),$$

166
and observe from (4.99) that, when \( \omega \in \Omega^* \), then \( A^*(t, \omega) = A_n(t, \omega) \), \( \forall t \in [0, T_n(\omega)] \), \( n = 1, 2 \ldots \). Equivalently, for each \( \omega \in \Omega^* \), we have

\[
A^*(t \wedge T_n(\omega), \omega) = A_n(t \wedge T_n(\omega), \omega) \equiv A_n(t, \omega), \; \forall t \in [0, \infty), \; \forall n = 1, 2 \ldots.
\]

(4.101)

In view of (4.100), the fact that each \( \omega \to A_n(\omega, t) \) is \( \mathcal{F}_t \)-measurable, and Proposition 1.2.6, we find that \( \{(A^*(t), \mathcal{F}_t); \; t \in [0, \infty)\} \) is an adapted process, and it is clear from \( \lim_{n \to \infty} T_n(\omega) = +\infty \) and (4.101) that the mappings \( t \to A^*(t, \omega) \) are continuous and non-decreasing on \( [0, \infty) \) for each \( \omega \in \Omega^* \).

To regularize matters on the complement of \( \Omega^* \), we just put \( A(t, \omega) \overset{\Delta}{=} A^*(t, \omega) \), \( \forall \omega \in \Omega^* \), \( \forall t \in [0, \infty) \), and \( A(t, \omega) \overset{\Delta}{=} 0 \), \( \forall \omega \not\in \Omega^* \), \( \forall t \in [0, \infty) \). Then \( \{(A(t), \mathcal{F}_t); \; t \in [0, \infty)\} \) is a continuous adapted process (since \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \), thus \( \Omega^* \in \mathcal{F}_0 \subset \mathcal{F}_t \), \( \forall t \in [0, \infty) \)). Moreover, the mapping \( t \to A(t, \omega) \) is non-decreasing on \( [0, \infty) \) for each \( \omega \in \Omega \), and \( \{(X^2(t \wedge T_n) - A(t \wedge T_n), \mathcal{F}_t); \; t \in [0, \infty)\} \) is a continuous local martingale with localizing sequence \( \{T_n, \; n = 1, 2 \ldots\} \). It follows that the process \( \{A(t); \; t \in [0, \infty)\} \) is a quadratic variation process for \( \{(X(t), \mathcal{F}_t); \; t \in [0, \infty)\} \). As for uniqueness, if \( \{B(t); \; t \in [0, \infty)\} \) is another quadratic variation process, and \( M(t) \overset{\Delta}{=} A(t) - B(t) \), \( \forall t \in [0, \infty) \), then \( \{M(t); \; t \in [0, \infty)\} \) is clearly a continuous local martingale, null at the origin, such that \( t \to M(t, \omega) \) is of locally bounded variation on \( [0, \infty) \), \( \forall \omega \in \Omega \) (recall Remark 4.7.11). Thus Lemma 4.7.13 establishes \( P[M(t) = 0, \; \forall t \in [0, \infty)] = 1 \), as required.

(b) Fix some arbitrary \( t \in [0, \infty) \). Since \( \{(X^2(t \wedge T_n) - A(t \wedge T_n), \mathcal{F}_t); \; t \in [0, \infty)\} \) is a martingale, one has \( E[X^2(t \wedge T_n) - A(t \wedge T_n)] = E[X^2(0) - A(0)] \geq 0 \), or

\[
E[A(t \wedge T_n)] \leq E[X^2(t \wedge T_n)] \leq E[X^2(t)] < \infty
\]

(4.102)

for all \( n = 1, 2 \ldots \) Here, we obtained the second inequality of (4.102) by using Corollary 4.5.6 to see that \( X(t \wedge T_n) = E[X(t) \mid \mathcal{F}_{t \wedge T_n}] \); then Jensen’s inequality (see Theorem 1.4.20) gives

\[
X^2(t \wedge T_n) \leq E \left[ X^2(t) \mathcal{F}_{t \wedge T_n} \right],
\]

(4.103)

and taking expectations on each side of (4.103) gives the inequality in (4.102). Now, for each \( \omega \), the sequence \( \{A(t \wedge T_n(\omega), \omega), \; n = 1, 2 \ldots\} \) increases monotonically to \( A(t, \omega) \) as \( n \to \infty \), hence by the monotone convergence theorem 1.2.12 we find that \( E[A(t)] \leq E[X^2(t)] < \infty \). It remains to see that \( \{(X^2(t) - A(t), \mathcal{F}_t); \; t \in [0, \infty)\} \) is a martingale. Since \( t \to A(t, \omega) \) is non-decreasing on \( [0, \infty) \) for each \( \omega \in \Omega \), we find that

\[
|X^2(t \wedge T_n) - A(t \wedge T_n)| \leq X^2(t \wedge T_n) + A(t)
\]

(4.104)

for all \( n = 1, 2 \ldots \) In view of (4.103) and Theorem 2.6.7, it is clear that the collection of random variables \( \{X^2(t \wedge T_n), \; n = 1, 2 \ldots\} \) is uniformly integrable, and since we have seen that \( E[A(t)] < \infty \), it then follows from (4.104) that \( \{X^2(t \wedge T_n) - A(t \wedge T_n), \; n = 1, 2 \ldots\} \) is uniformly integrable. Since this
holds for each $t \in [0, \infty)$, Proposition 4.6.8 says that $\{(X^2(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a martingale.

Remark 4.7.16 We have shown in Theorem 4.7.15 that a continuous local martingale $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ always has a quadratic variation process, and that any two members in the class of quadratic variation processes for this local martingale are indistinguishable. Henceforth we shall use the notation $\{[X](t); t \in [0, \infty)\}$ to denote an arbitrarily chosen but fixed quadratic variation process for $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$.

Remark 4.7.17 Suppose that $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a scalar standard Wiener process on $(\Omega, \mathcal{F}, P)$. Then, one immediately sees from Theorem 4.7.15 and Proposition 4.1.8 that

\begin{equation}
(4.105) \quad P\{[W](t) = t, \ \forall \ t \in [0, \infty)\} = 1.
\end{equation}

Remark 4.7.18 When $X \in \mathcal{M}^{\text{loc}}_{\text{st}}(\mathcal{F}_t, P)$ and $\{[X](t); t \in [0, \infty)\}$ is an associated quadratic variation process, then we generally do not know that $[X](t)$ is integrable, or that $\{((X^2(t) - [X](t), \mathcal{F}_t); t \in [0, \infty)\}$ is a martingale (indeed, Definition 4.7.1 requires only that this latter process be local martingale). However, subject to the stronger hypothesis that $X \in \mathcal{M}^2_{\text{st}}(\mathcal{F}_t, P)$, we see from Theorem 4.7.15(b) that $E\{[X](t)\} < \infty, \ \forall \ t \in [0, \infty)$, and $\{((X^2(t) - [X](t), \mathcal{F}_t); t \in [0, \infty)\}$ is indeed a continuous martingale. Finally, when $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a uniformly bounded continuous martingale then $L^2$-convergence in Lemma 4.7.5 yields the further information that $[X](t)$ is square integrable for each $t \in [0, \infty)$.

Remark 4.7.19 Suppose that $X \in \mathcal{M}^{\text{loc}}_{\text{st}}(\mathcal{F}_t, P)$, $T$ is a $\{\mathcal{F}_t\}$-stopping time, and $\mathcal{F}_0$ includes all $P$-null events in $\mathcal{F}$. In view of Remark 4.6.3 we observe that the stopped process $\{(X^T(t), \mathcal{F}_t); t \in [0, \infty)\}$ is also a continuous local martingale, and hence has an associated quadratic variation process $\{[X^T](t); t \in [0, \infty)\}$. The next lemma relates this process to the stopped quadratic variation process of $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$:

Lemma 4.7.20 Suppose that Condition 4.7.4 holds, $X \in \mathcal{M}^{\text{loc}}_{\text{st}}(\mathcal{F}_t, P)$, and $T$ is a $\{\mathcal{F}_t\}$-stopping time. Then the processes $\{[X^T](t); t \in [0, \infty)\}$ and $\{[X](t \wedge T); t \in [0, \infty)\}$ are indistinguishable.

Proof: Since $\{(X^2(t) - [X](t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous local martingale, we know from Remark 4.6.3 that

\begin{equation}
(4.106) \quad \{(X^2(t \wedge T) - [X](t \wedge T), \mathcal{F}_t); t \in [0, \infty)\}
\end{equation}

is a continuous local martingale. Moreover, by Definition 4.7.1 and the fact that $X^T \in \mathcal{M}^{\text{loc}}_{\text{st}}(\mathcal{F}_t, P)$, we have

\begin{equation}
(4.107) \quad \{( (X^T(t))^2 - [X^T](t), \mathcal{F}_t); t \in [0, \infty)\}
\end{equation}
is a continuous local martingale. Now \( X^T(t) = X(t \wedge T) \), hence, writing \( C(t) \triangleq [X^T](t) - [X](t \wedge T) \), we see from taking the difference of (4.106) and (4.107) that \( \{(C(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous local martingale. But Remark 4.7.11 shows that the mapping \( t \to C(t, \omega) \) is of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \). Since \( C(0) \equiv 0 \), Lemma 4.7.13 says that \( \{[X^T](t); t \in [0, \infty)\} \) and \( \{[X](t \wedge T); t \in [0, \infty)\} \) are indistinguishable processes.

The notion of a quadratic variation process that we have formulated for a continuous local martingale may be viewed as a profound analogue of the familiar concept of variance for square-integrable random variables. In problems of basic probability which involve more than one random variable we often come across the covariance of a pair of square integrable random variables. A similar situation pertains to the present context where, given two continuous local martingales, we need some notion of a joint- or co-quadratic variation process for the two local martingales. The appropriate formulation of this concept is presented in the next definition, which should be contrasted with Definition 4.7.1:

**Definition 4.7.21** Suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) are continuous local martingales on the probability space \((\Omega, \mathcal{F}, P)\). A real-valued process \( \{A(t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) is called a **co-quadratic variation process** corresponding to the local martingales when the following conditions hold:

(a) \( \{(A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted process;

(b) The mappings \( t \to A(t, \omega) \) are continuous and of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \);

(c) \( A(0) \equiv 0 \);

(d) \( \{(X(t)Y(t) - A(t, \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous local martingale.

Notice that whereas the quadratic variation process of a given local martingale assumes non-negative values and has non-decreasing sample-paths, the co-quadratic variation process of a pair of local martingales is postulated to take values in \( \mathbb{R} \) and have sample-paths which are of locally bounded variation on \([0, \infty)\). Having formulated the concept of a co-quadratic variation process we must now establish existence and look at the issue of uniqueness, a task which is undertaken in the next theorem. Before doing so however, we draw attention to the following trivial identity which is frequently used for studying co-quadratic variation processes: for all \( x, y \in \mathbb{R} \) we have

\[
(4.108) \quad xy = \frac{1}{4} \{(x + y)^2 - (x - y)^2\}.
\]

**Theorem 4.7.22** Suppose that Condition 4.7.4 holds.

(a) Corresponding to a pair of continuous local martingales \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(Y_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a co-quadratic variation process \( \{A(t); t \in [0, \infty)\} \), which is given by

\[
(4.109) \quad A(t) \triangleq \frac{1}{4} \{[X + Y](t) - [X - Y](t)\}, \quad \forall t \in [0, \infty),
\]
and if \( \{B(t); t \in [0, \infty)\} \) is another such co-quadratic variation process, then \( \{A(t); t \in [0, \infty)\} \) and 
\( \{B(t); t \in [0, \infty)\} \) are indistinguishable.

(b) If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(Y_t, \mathcal{F}_t); t \in [0, \infty)\} \) are continuous \( L^2 \)-martingales and \( \{A(t); t \in [0, \infty)\} \) is a corresponding co-quadratic variation process then \( E|A(t)| < \infty, \forall t \in [0, \infty) \), and 
\( \{(X(t)Y(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale.

**Remark 4.7.23** In (4.109) the processes \( \{(X + Y)(t); t \in [0, \infty)\} \) and \( \{(X - Y)(t); t \in [0, \infty)\} \) are of course understood to mean arbitrary choices of quadratic variation processes for the continuous local martingales 
\( \{(X(t) + Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(X(t) - Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) respectively. The definition of \( A(t) \) in (4.109) is actually motivated by the expression for the covariance of two square integrable random variables in terms of the variances of their sums and differences (see (1.46)). The idea is that quadratic and co-quadratic variation processes for pairs of local martingales should stand in the same relation to each other as do variances and covariances for pairs of random variables.

Proof of Theorem 4.7.22: (a) From (4.109) and properties of the quadratic variation processes 
\( \{(X + Y)(t); t \in [0, \infty)\} \) and \( \{(X - Y)(t); t \in [0, \infty)\} \), we see that \( \{(A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted continuous process null at the origin. Since the mapping \( t \to A(t, \omega) \) is clearly the difference of two non-decreasing mappings it follows from Remark 4.7.11 that \( t \to A(t, \omega) \) is of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \). Finally, using the identity (4.108) together with the fact that 
\( \{(X+Y)^2(t) - [X+Y](t), \mathcal{F}_t); t \in [0, \infty)\} \) and 
\( \{(X-Y)^2(t) - [X-Y](t), \mathcal{F}_t); t \in [0, \infty)\} \) are continuous local martingales, we see that 
\( \{(X(t)Y(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous local martingale.

It follows that \( \{A(t); t \in [0, \infty)\} \) verifies all conditions of Definition 4.7.1 and is therefore a co-quadratic variation process for the local martingales 
\( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \). Next, suppose that \( \{B(t); t \in [0, \infty)\} \) is another co-quadratic variation process for the two given local martingales, and define \( C(t) \overset{\Delta}{=} A(t) - B(t) \). We see that \( \{C(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous local martingale, with \( C(0) \equiv 0 \), and in view of Proposition 4.7.9(a,b) it follows that the mapping \( t \to C(t, \omega) \) is of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \). Thus Lemma 4.7.13 implies that \( \{A(t); t \in [0, \infty)\} \) and 
\( \{B(t); t \in [0, \infty)\} \) are indistinguishable processes.

(b) When \( X, Y \in \mathbb{M}_2^\circ(\{\mathcal{F}_t\}, P) \) then (see Theorem 4.7.15(b)) we know that 
\( E\{(X + Y)(t)\} < \infty \) and 
\( E\{|X - Y|(y)\} < \infty, \forall t \in [0, \infty) \), while 
\( \{(X+Y)^2(t) - [X+Y](t), \mathcal{F}_t); t \in [0, \infty)\} \) and 
\( \{(X-Y)^2(t) - [X-Y](t), \mathcal{F}_t); t \in [0, \infty)\} \) are continuous martingales. It follows at once from (4.109) that 
\( E|A(t)| < \infty, \forall t \in [0, \infty) \), and 
\( \{(X(t)Y(t) - A(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale.

**Remark 4.7.24** If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) are continuous local martingales on a common probability space \((\Omega, \mathcal{F}, P)\), we shall use the notation \( \{(X,Y)(t); t \in [0, \infty)\} \) to denote an arbitrary but fixed co-quadratic variation process for the two martingales.
Remark 4.7.25 Suppose that \{(W(t), \mathcal{F}_t); t \in [0, \infty]\} is a d-dimensional Wiener process on \((\Omega, \mathcal{F}, \mathbb{P})\). Then \(W^k \in \mathcal{M}^c_2(\mathcal{F}_t), \forall k = 1, 2, \ldots, d\), and we see from Definition 4.7.21 together with the uniqueness assertion of Theorem 4.7.22 and Proposition 4.1.10, that
\[
P\{ [W^j, W^k](t) = t \delta_{j,k}, \ \forall t \in [0, \infty) \} = 1,
\]
for each \(j, k \in \{1, 2, \ldots d\}\).

In Lemma 4.7.5 we saw how the quadratic variation process of a continuous uniformly bounded martingale emerged as the \(L^2\)-limit of a sequence of sums defined by (4.78). We can extend this result to the case of continuous local martingales, but now we must be content with convergence in a somewhat weaker sense:

**Theorem 4.7.26** Suppose that Condition 4.7.4 holds, and \(X \in \mathcal{M}^c_{loc}(\mathcal{F}_t), \mathbb{P}\).

Put \(\tau^n_k \overset{\Delta}{=} k2^{-n}, \forall k, n = 0, 1, 2 \ldots\) and define
\[
A_n(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} \left[ X(t \wedge \tau^n_k, \omega) - X(t \wedge \tau^n_{k+1}, \omega) \right]^2
\]
for all \(t \in [0, \infty), \forall \omega \in \Omega\), \(\forall n = 0, 1, 2 \ldots\) Then the sequence of random variables \(\{A_n(t); n = 0, 1, 2 \ldots\}\) converges in probability to \([X](t)\) for each \(t \in [0, \infty)\).

Proof: For each \(\omega \in \Omega\) and \(N = 1, 2, \ldots\) put
\[
T_N(\omega) \overset{\Delta}{=} \inf\{t \in [0, \infty) : \ |X(t, \omega)| \geq N\}.
\]
Remark 3.3.8 says that each \(T_N\) is a \(\mathcal{F}_t\)-stopping time with \(\lim_{N \to \infty} T_N(\omega) = +\infty, \forall \omega \in \Omega\), which further implies that
\[
\lim_{N \to \infty} P\{T_N \leq t\} = 0
\]
for each \(t \in [0, \infty)\). Now we know from Corollary 4.5.8 that \(\{(X^T_N(t), \mathcal{F}_t); t \in [0, \infty)\}\) is a martingale which is clearly uniformly bounded and continuous, and from (4.111) it follows that
\[
A_n(t \wedge T_N(\omega), \omega) = \sum_{0 \leq k < \infty} \left[ X^T_N(t \wedge \tau^n_k, \omega) - X^T_N(t \wedge \tau^n_{k+1}, \omega) \right]^2
\]
for all \(t \in [0, \infty), \forall \omega \in \Omega\). Thus applying Lemma 4.7.5 to \(\{(X^T_N(t), \mathcal{F}_t); t \in [0, \infty)\}\), it follows that
\[
\lim_{n \to \infty} E|A_n(t \wedge T_N) - [X^T_N](t)|^2 = 0
\]
and hence, in view of Lemma 4.7.20, we get
\[
\lim_{n \to \infty} E|A_n(t \wedge T_N) - [X](t \wedge T_N)|^2 = 0
\]
for each \(t \in [0, \infty)\) and \(N = 1, 2, \ldots\) Now fix arbitrary \(t \in [0, \infty)\) and \(\epsilon > 0\). Clearly
\[
P\{|A_n(t) - [X](t)| > \epsilon\} \leq P\{|A_n(t) - A_n(t \wedge T_N)| > \epsilon/3\}
\]
for all \( N = 1, 2, \ldots \) Moreover, since \( \{|A_n(t)| > \epsilon/3 \} \subset \{T_N \leq t\}, \forall n = 1, 2, \ldots \) we get
\[
(4.117) \quad \sup_{n \geq 1} P\{|A_n(t) - A_n(t \wedge T_N)| > \epsilon/3 \} \leq P\{T_N \leq t\}
\]
for all \( N = 1, 2, \ldots \) Likewise, \( \{|[X](t \wedge T_N) - [X](t)| > \epsilon/3 \} \subset \{T_N \leq t\}, \) thus
\[
(4.118) \quad P\{|[X](t \wedge T_N) - [X](t)| > \epsilon/3 \} \leq P\{T_N \leq t\}.
\]
Now fix an arbitrary \( \delta > 0 \). Using (4.118), (4.117) and (4.113), we can find some integer \( N(\delta) \) such that
\[
(4.119) \quad \sup_{n \geq 1} P\{|A_n(t) - A_n(t \wedge T_{N(\delta)})| > \epsilon/3 \} < \delta/3
\]
and
\[
(4.120) \quad P\{|[X](t \wedge T_{N(\delta)}) - [X](t)| > \epsilon/3 \} < \delta/3.
\]
From (4.115) and the Markov inequality (Theorem 1.2.23) we can find some integer \( n(\delta) \) such that
\[
(4.121) \quad P\{|A_n(t \wedge T_{N(\delta)}) - [X](t \wedge T_{N(\delta)})| > \epsilon/3 \} < \delta/3
\]
for all integers \( n > n(\delta) \). Taking \( N = N(\delta) \) in (4.116), and using (4.121), (4.120) and (4.119) we find that
\[
P\{|A_n(t) - [X](t)| > \epsilon\} < \delta
\]
for all integers \( n > n(\delta) \). \( \blacksquare \)

We next formulate an analogue of Theorem 4.7.26 for co-quadratic variation processes:

**Theorem 4.7.27** Suppose that Condition 4.7.4 holds and \( X, Y \in M_{loc}^c(\{\mathcal{F}_t\}, P) \).
Put \( \tau^n_k \triangleq k2^{-n}, \forall k, n = 0, 1, 2, \ldots \) and define
\[
(4.122) \quad A_n(t, \omega) \triangleq \sum_{0 \leq k < \infty} [X(t \wedge \tau^n_{k+1}, \omega) - X(t \wedge \tau^n_k, \omega)][Y(t \wedge \tau^n_{k+1}, \omega) - Y(t \wedge \tau^n_k, \omega)]
\]
\( \forall t \in [0, \infty), \forall \omega \in \Omega, \forall n = 0, 1, 2, \ldots \) Then the sequence of random variables \( \{A_n(t); n = 0, 1, 2, \ldots\} \)
converges in probability to \([X, Y](t)\) for each \( t \in [0, \infty) \).

Proof: Apply Theorem 4.7.26 separately to the continuous local martingales \( \{(X(t) + Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(X(t) - Y(t), \mathcal{F}_t); t \in [0, \infty)\} \), and use (4.108) and (4.109). \( \blacksquare \)

The next lemma lists some useful properties of co-quadratic variation processes:
Lemma 4.7.28 Suppose that Condition 4.7.4 holds, and \( X, Y, Z \in M^c_{\text{loc}}(\{F_t\}, P) \). Then
(a) \( \{[X](t); \ t \in [0, \infty)\} \) and \( \{[X, X](t); \ t \in [0, \infty)\} \) are indistinguishable processes;
(b) \( \{[X, Y](t); \ t \in [0, \infty)\} \) and \( \{[Y, X](t); \ t \in [0, \infty)\} \) are indistinguishable processes;
(c) \( \{[\alpha X + \beta Y, Z](t); \ t \in [0, \infty)\} \) and \( \{[\alpha X, Z](t) + \beta[Y, Z](t); \ t \in [0, \infty)\} \) are indistinguishable processes for each \( \alpha, \beta \in \mathbb{R} \);
(d) There is some \( P \)-null event \( N \in \mathcal{F} \) such that, if \( \omega \notin N \), then
\[
|[X, Y](t, \omega) - [X, Y](s, \omega)| \leq \{[X](t, \omega) - [X](s, \omega)\}^{1/2}\{[Y](t, \omega) - [Y](s, \omega)\}^{1/2}
\]
for all \( s, t \in [0, \infty) \) with \( s < t \);
(e) If \( P[Y(t) = Y_0, \ \forall \ t \in [0, \infty)] = 1 \) then \( P[[X, Y](t) = 0, \ \forall \ t \in [0, \infty]] = 1 \).

Remark 4.7.29 Recall that, whenever we write down a quadratic or co-quadratic variation process, we are really specifying an arbitrary choice from within a class of processes, any two of whose members are indistinguishable. Strictly speaking then, the preceding lemma should have been stated in the following terms: Suppose that Condition 4.7.4 holds and \( X, Y, Z \in M^c_{\text{loc}}(\{F_t\}, P) \). Let \( \{[X](t); \ t \in [0, \infty)\}, \{[Y](t); \ t \in [0, \infty)\}, \{[X, X](t); \ t \in [0, \infty)\}, \{[X, Y](t); \ t \in [0, \infty)\}, \{[X, Z](t); \ t \in [0, \infty)\}, \{[Y, X](t); \ t \in [0, \infty)\}, \{[Y, Y](t); \ t \in [0, \infty)\}, \{[Y, Z](t); \ t \in [0, \infty)\} \) and \( \{[\alpha X + \beta Y, Z](t); \ t \in [0, \infty)\} \) be arbitrarily chosen quadratic or co-quadratic variation processes, for constants \( \alpha, \beta \in \mathbb{R} \). Then assertions (a) to (d) in Lemma 4.7.28 hold true. While this version of Lemma 4.7.28 leaves nothing to be desired from the point of view of accuracy, it is a very long-winded statement. In fact, it is customary to avoid the prologue about arbitrary choices of quadratic and co-quadratic variation processes, and to state results like Lemma 4.7.28 in exactly the way that we have done, it being understood that when we write down a quadratic or co-quadratic variation process what we really have in mind is an arbitrary but fixed choice of that process. This is quite similar to our use of the standard notation for conditional expectations, where we likewise write down a fixed but arbitrary choice from among a class of random variables, any two of which agree a.s. (see Remark 1.4.16).

Proof of Lemma 4.7.28: Parts (a), (b), (c) are immediate consequences of our definitions and the uniqueness to within indistinguishability ensured by Proposition 4.7.13. As for (d), fix arbitrary \( s, t \in [0, \infty) \) with \( s < t \), and let \( \lambda \in \mathbb{R} \). Since the sample-paths of a quadratic variation process are non-decreasing for each \( \omega \), we have
\[
0 \leq [X + \lambda Y](t, \omega) - [X + \lambda Y](s, \omega) \quad \forall \omega \in \Omega, \ \lambda \in \mathbb{R}.
\]
From (a) and (c) we get
\[
[X + \lambda Y](t) = [X + \lambda Y, X + \lambda Y](t) \quad \text{a.s.}
= [X](t) + 2\lambda[X, Y](t) + \lambda^2[Y](t) \quad \text{a.s.}
\]
Let $N_t$ denote the $P$-null event (in $\mathcal{F}_t$) at which the equalities in (4.124) fail to hold. Combining (4.123), (4.124) and the analogue of (4.124) which holds when $t$ is replaced with $s$, for each $\omega \not\in N_{s,t} \overset{\Delta}{=} N_s \cup N_t$ we have that

$$[X](t, \omega) - [X](s, \omega) + 2\lambda \{[X,Y](t, \omega) - [X,Y](s, \omega)\} + \lambda^2 \{[Y](t, \omega) - [Y](s, \omega)\} \geq 0 \quad \forall \lambda \in \mathbb{R}.$$  

The discriminant associated with the quadratic function of $\lambda$ in the preceding inequality must therefore be non-negative, whence

$$|[X,Y](t, \omega) - [X,Y](s, \omega)|^2 \leq \{[X](t, \omega) - [X](s, \omega)\}\{[Y](t, \omega) - [Y](s, \omega)\} \quad \forall \omega \not\in N_{s,t}.$$  

Now part (d) follows upon defining $N \overset{\Delta}{=} \bigcup N_{s,t}$, the union of events being taken over all rational $s,t \in \mathbb{R}$ with $s < t$, and the fact that the processes $\{(X(t); t \in [0,\infty))\}$, $\{(Y(t); t \in [0,\infty))\}$ and $\{(X,Y)(t); t \in [0,\infty))\}$ have continuous sample paths. As for part (e), since $Y(t)$ is $\mathcal{F}_0$-measurable for each $t \in [0,\infty)$, we see that $\{(X(t)Y(t), \mathcal{F}_t); t \in [0,\infty))\}$ is a continuous local martingale, hence the uniqueness part of Theorem 4.7.22(a) establishes that $[X,Y](t) = 0$, $\forall t \in [0,\infty)$, a.s.

Lemma 4.7.30 Suppose that Condition 4.7.4 holds and $X \in M^c_{loc}(\{\mathcal{F}_t\}, P)$. If

$$P[ [X](t) = 0, \forall t \in [0,\infty)] = 1$$

then

$$P[ X(t) = X(0), \forall t \in [0,\infty)] = 1.$$  

Proof: Suppose, to begin with, that $\{(X(t), \mathcal{F}_t); t \in [0,\infty))\}$ is a continuous $L^2$-martingale, and fix some arbitrary $t \in [0,\infty)$. Then we see that

$$E[(X(t) - X(0))^2] = E[X^2(t) - X^2(0)] = E([X](t)) = 0,$$

where the first equality follows from Proposition 4.1.5, the second equality follows since $\{(X^2(t) - [X](t), \mathcal{F}_t); t \in [0,\infty))\}$ is a martingale with $[X](0) \equiv 0$, and the third equality follows from (4.125). Now (4.127), together with the arbitrary choice of $t$ and sample-path continuity of $\{X(t); t \in [0,\infty)\}$, implies (4.126). In the general case where $\{(X(t), \mathcal{F}_t); t \in [0,\infty))\}$ is a continuous local martingale, let $\{S_k, n = 1, 2, \ldots\}$ be the localizing sequence for $\{(X(t), \mathcal{F}_t); t \in [0,\infty))\}$ given by Proposition 4.6.9. We then have

$$P[ X(t \land S_k) = X(0), \forall t \in [0,\infty)] = 1,$$

for all $k = 1, 2, \ldots$, and taking $k \to \infty$ gives (4.126).

Proposition 4.7.31 Suppose that Condition 4.7.4 holds, and $X, Y \in M^c_{loc}(\{\mathcal{F}_t\}, P)$. If $X(0) = Y(0)$ a.s. and, for each $U \in M^c_{loc}(\{\mathcal{F}_t\}, P)$, we have

$$P\{ [X,U](t) = [Y,U](t), \forall t \in [0,\infty)\} = 1,$$

then $\{X(t); t \in [0,\infty)\}$ and $\{Y(t); t \in [0,\infty)\}$ are indistinguishable processes.
Proof: Taking \( U \triangleq X - Y \) in (4.128), we see from Lemma 4.7.28(c) that \([X - Y](t) = 0, \forall t \in [0, \infty), \text{a.s.}\)
In view of Lemma 4.7.30 we obtain indistinguishability of \(\{X(t); t \in [0, \infty)\}\) and \(\{Y(t); t \in [0, \infty)\}\). \(\blacksquare\)

We next extend Lemma 4.7.20 to co-quadratic variations of continuous local martingales.

**Lemma 4.7.32** Suppose that Condition 4.7.4 holds and \(X, Y \in \mathbf{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, \mathbb{P}).\) If \(T\) is a \(\{\mathcal{F}_t\}\)-stopping time then the processes \(\{[X,Y](t \wedge T); t \in [0, \infty)\}, \{|X,Y|^2(t); t \in [0, \infty)\}, \{|X^T,Y^T|(t); t \in [0, \infty)\},\) and \(\{|X^T,Y^T|(t); t \in [0, \infty)\}\) are indistinguishable.

Proof: Suppose, first, that \(\{(X_t, \mathcal{F}_t); t \in [0, \infty)\}\) and \(\{(Y_t, \mathcal{F}_t); t \in [0, \infty)\}\) are continuous \(L^2\)-martingales. Put \(Z(t) \triangleq X(t)Y(t \wedge T) - X(t \wedge T)Y(t \wedge T)\) and let \(S\) be an arbitrary bounded \(\{\mathcal{F}_t\}\)-stopping time. i.e. \(S(\omega) \leq u, \forall \omega \in \Omega, u \in [0, \infty)\) being some constant. Then

\[
(Z(S)) |X(S)||Y(S \wedge T)| + |X(S \wedge T)||Y(S \wedge T)|.
\]

Now, by Corollary 4.5.6 and Jensen’s inequality for conditional expectations (Theorem 1.4.20) we get \(X^2(S) \leq E[X^2(u) | \mathcal{F}_S]\) a.s. thus \(E[X^2(S)] \leq E[X^2(u)] < \infty\). Likewise, one sees that \(E[X^2(S \wedge T)] < \infty\) and \(E[Y^2(S \wedge T)] < \infty\). Thus, from (4.129) and the Cauchy-Schwarz inequality, we obtain \(E|Z(S)| < \infty\). Moreover, by the composition rule for conditional expectations, clearly

\[
E[X(S)Y(S \wedge T)] = E[E[X(S)Y(S \wedge T) | \mathcal{F}_{S \wedge T}]]
\]

\[
= E[E[X(S) | \mathcal{F}_{S \wedge T}](S \wedge T)]
\]

\[
= E[X(S \wedge T)Y(S \wedge T)].
\]

Here we have used the \(\mathcal{F}_{S \wedge T}\)-measurability of \(Y(S \wedge T)\) guaranteed by Proposition 3.3.14(b) at the second equality, and Corollary 4.5.6 at the third equality in (4.130). Thus, we see that \(E|Z(S)| < \infty\) and \(E[Z(S)] = 0\) for each bounded \(\{\mathcal{F}_t\}\)-stopping time \(S\). Since \(\{(Z(t), \mathcal{F}_t); t \in [0, \infty)\}\) is clearly adapted and continuous it is progressively measurable (by Proposition 3.1.25), hence Lemma 4.5.1 says that

\[
\{(X(t)Y^T(t) - X^T(t)Y^T(t), \mathcal{F}_t); t \in [0, \infty)\}
\]

is a continuous martingale. By the definition of a co-quadratic variation process we see that

\[
\{(X(t)Y^T(t) - [X,Y]^2(t), \mathcal{F}_t); t \in [0, \infty)\}
\]

is a continuous martingale, and Corollary 4.5.8 applied to the martingale \(\{(X(t)Y(t) - [X,Y]^2(t), \mathcal{F}_t); t \in [0, \infty)\}\) shows that

\[
\{(X^T(t)Y^T(t) - [X,Y]((t \wedge T), \mathcal{F}_t); t \in [0, \infty)\}
\]

is a continuous martingale. Adding (4.131) and (4.133) and then subtracting (4.132) shows that \(\{([X(t)Y^T(t) - [X,Y](t \wedge T), \mathcal{F}_t); t \in [0, \infty)\}\) is a continuous martingale. Now the mappings \(t \rightarrow\)
\([X, YT](t, \omega) - [X,Y](t \wedge T(\omega), \omega)\) are clearly of locally bounded variation on \([0, \infty)\), thus Lemma 4.7.13 shows that \(\{[X,Y](t \wedge T(t); t \in [0, \infty)\}\) and \(\{[X,YT](t); t \in [0, \infty)\}\) are indistinguishable processes. Showing indistinguishability of the remaining processes involves just a trivial application of the preceding assertion together with Lemma 4.7.28(b), hence Lemma 4.7.32 is established when \(\{(X(t), F_t); t \in [0, \infty)\}\) and \(\{(Y(t), F_t); t \in [0, \infty)\}\) are continuous \(L^2\)-martingales. In the general case where these are continuous local martingales, we introduce a common localizing sequence for \(\{(X(t), F_t); t \in [0, \infty)\}\) and \(\{(Y(t), F_t); t \in [0, \infty)\}\) (see Remark 4.6.4) and apply the preceding argument to the resulting bounded martingales.

**Proposition 4.7.33** Suppose that Condition 4.7.4 holds and \(X \in M^c_{loc}(\{F_t\}, P)\). Then the limit

\[
[X](\infty, \omega) \overset{\Delta}{=} \lim_{t \to \infty} [X](t, \omega),
\]

exists in \([0, \infty]\) for each \(\omega \in \Omega\), and defines a \(\mathcal{F}_\infty\)-measurable function. Furthermore:

(a) \(X \in M^c_{2,b}(\{F_t\}, P)\) if and only if \(E|X^2(0)| < \infty\) and \(E\{[X](\infty)\} < \infty\); 

(b) If \(X \in M^c_{2,b}(\{F_t\}, P)\) then \(E\{X^2(\infty)\} = E\{[X](\infty)\} + E\{X^2(0)\}\).

*Proof:* Since \(t \to [X](t, \omega)\) is nondecreasing on \([0, \infty), \forall \omega \in \Omega\), the limit in (4.134) clearly exists and is \(\mathcal{F}_\infty\)-measurable. To establish (a), suppose that \(X \in M^c_{2,b}(\{F_t\}, P)\). Then, of course, \(E(X^2(0)) < \infty\), and Theorem 4.4.12(b) says that there is some a.s. unique \(\mathcal{F}_\infty\)-measurable random variable \(X(\infty) \in L^2(\Omega, \mathcal{F}, P)\) such that \(X(t) = E[X(\infty) | F_t]\) a.s. whence \(E(X^2(t)) \leq E(X^2(\infty)) < \infty\) for each \(t \in [0, \infty)\). Since \(\{(X^2(t) - [X](t), F_t); t \in [0, \infty)\}\) is a martingale and \([X](0) \equiv 0\), for each \(t \in [0, \infty)\) we have \(E(X^2(t) - [X](t)) = E(X^2(0))\), whence

\[
E([X](t)) \leq E(X^2(\infty)) - E(X^2(0)) < \infty.
\]

Taking \(t \to \infty\) and using the monotone convergence theorem then shows that \(E([X](\infty)) < \infty\). For the converse assertion, suppose that \(X \in M^c_{loc}(\{F_t\}, P)\) with \(E(X^2(0)) < \infty\) and \(E([X](\infty)) < \infty\). From Proposition 4.6.9 there is a localizing sequence \(\{T_n, n = 1, 2, \ldots\}\) of \(\{F_t\}\)-stopping times such that \(\{(X(t \wedge T_n), F_t); t \in [0, \infty)\}\) is a uniformly bounded martingale for each \(n = 1, 2, \ldots\), and thus, in view of Lemma 4.7.32, it follows that \(\{(X^2(t \wedge T_n) - [X](t \wedge T_n), F_t); t \in [0, \infty)\}\) is a martingale. Consequently, since \([X](0) \equiv 0\), for each \(n = 1, 2, \ldots\) we have

\[
E(X^2(t \wedge T_n) - [X](t \wedge T_n)) = E(X^2(0)), \quad \forall t \in [0, \infty).
\]

Since \(t \to [X](t)\) is non-decreasing, we see from (4.135) that \(E(X^2(t \wedge T_n)) \leq E(X^2(0)) + E([X](\infty))\), whence Fatou’s Theorem 1.2.15 establishes that \(E(X^2(t)) \leq E(X^2(0)) + E([X](\infty)) < \infty\) for each \(t \in [0, \infty)\), so that \(X \in M^c_{2,b}(\{F_t\}, P)\). To establish (b) suppose that \(X \in M^c_{2,b}(\{F_t\}, P)\). Then, as already observed, we have \(E(X^2(t)) = E(X^2(0)) + E([X](t))\), \(\forall t \in [0, \infty)\). Now Theorem 4.4.12 ensures that \(E[\sup_t X^2(t)] < \infty\), thus the Lebesgue Dominated Convergence Theorem shows that \(\lim_{t \to \infty} E(X^2(t)) = E(X^2(\infty))\), while the Monotone Convergence Theorem gives \(\lim_{t \to \infty} E([X](t)) = E([X](\infty))\).
4.8 Problems

In each of the following problems \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\), and \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \).

**Problem 4.8.1** Suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a progressively measurable \( \mathbb{R} \)-valued process and \( X_\infty \) is a \( \mathcal{F}_\infty \)-measurable random variable on \((\Omega, \mathcal{F}, P)\), such that \( E|X(T)| < \infty \) and \( EX(T) = 0 \) for each \( \{\mathcal{F}_t\} \)-stopping time \( T \). Show that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable martingale. Hint: Extend the argument used for Lemma 2.3.13.

**Problem 4.8.2** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous martingale on the probability space \((\Omega, \mathcal{F}, P)\). Show that \( \{(X_t, \mathcal{F}_{t+}); t \in [0, \infty)\} \) is also a martingale on \((\Omega, \mathcal{F}, P)\). Hint: Use Theorem 2.7.9.

**Problem 4.8.3** Suppose that \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale on \((\Omega, \mathcal{F}, P)\), and \( \{\mathcal{G}_t, t \in [0, \infty)\} \) is a filtration in \((\Omega, \mathcal{F}, P)\) with \( \mathcal{G}_t \subset \mathcal{F}_t, \forall t \in [0, \infty) \). Show that \( \{(Y_t, \mathcal{G}_t); t \in [0, \infty)\} \) is a martingale, for \( Y_t \triangleq E[X_t | \mathcal{G}_t], \forall t \in [0, \infty) \).

**Problem 4.8.4** Suppose that \( X \in \mathbb{M}^0_{loc}(\{\mathcal{F}_t\}, P) \). Show that \( X \in \mathbb{M}^0_{loc}(\{\mathcal{F}_t^X\}, P) \), for

\[ \mathcal{F}_t^X \triangleq \sigma\{X_u, u \in [0, t]\} \]

**Problem 4.8.5** Suppose \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a right-continuous martingale such that, for each \( (t, \omega) \in [0, \infty) \otimes \Omega, X(t, \omega) \) takes values in the two-member set \( \{a, b\} \), for constants \( a < b \). Show that

\[ P\{X(t) = X(0), \forall t \in [0, \infty)\} = 1 \]

that is, the process \( \{X(t); t \in [0, \infty)\} \) is a.s. constant. Hint: Affinely transform \( \{X(t)\} \) to a process \( \{Y(t)\} \) which takes values in the two-member set \( \{0, 1\} \).

**Problem 4.8.6** Suppose that \( \{Z_t; t \in [0, \infty)\} \) is a jointly measurable \( \mathbb{R} \)-valued process on a complete probability space \((\Omega, \mathcal{F}, P)\) with \( E|Z_t| < \infty, \forall t \in [0, \infty) \), and \( \{\mathcal{G}_t, t \in [0, \infty)\} \) is some filtration in \((\Omega, \mathcal{F}, P)\) such that \( \mathcal{G}_0 \) includes all \( P \)-null events in \( \mathcal{F} \). A fundamental result of Dellacherie and Meyer shows that there exists a \( \{\mathcal{G}_t\} \)-progressively measurable process \( \{(\hat{Z}_t, \mathcal{G}_t); t \in [0, \infty)\} \) such that

\[ \hat{Z}_t = E[Z_t | \mathcal{G}_t] \quad \text{a.s.} \]

for each \( t \in [0, \infty) \). This process is called an optional projection of the process \( \{Z_t; t \in [0, \infty)\} \) onto the filtration \( \{\mathcal{G}_t, t \in [0, \infty)\} \). Now do the following: Suppose that (i) \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a filtration
in \((\Omega, \mathcal{F}, P)\) such that \(\mathcal{G}_t \subset \mathcal{F}_t, \forall \ t \in [0, \infty)\); (ii) \(\{(Y_t, \mathcal{F}_t); t \in [0, \infty)\}\) is a progressively measurable \(\mathbb{R}\)-valued process and there is a constant \(C \in [0, \infty)\) such that \(|Y_t(\omega)| \leq C < \infty, \forall (t, \omega) \in [0, \infty) \times \Omega\); (iii) \(X \in \mathbb{M}(\{\mathcal{F}_t\}, P)\), where \(\{X_t; t \in [0, \infty)\}\) is a jointly measurable and \(\mathbb{R}\)-valued process on \((\Omega, \mathcal{F}, P)\). If

\[
Z_t \triangleq X_t + \int_0^t Y_s \, ds, \quad \forall \ t \in [0, \infty),
\]

and \(\{\hat{Z}_t; t \in [0, \infty)\}, \{\hat{Y}_t; t \in [0, \infty)\}\), denote optional projections of the processes \(\{Z_t; t \in [0, \infty)\}\), \(\{Y_t; t \in [0, \infty)\}\), onto the filtration \(\{\mathcal{G}_t, t \in [0, \infty)\}\), then show that

\[
\hat{Z}_t - \int_0^t \hat{Y}_s \, ds, \quad \forall \ t \in [0, \infty),
\]

is a \(\{\mathcal{G}_t\}\)-martingale.

**Problem 4.8.7** Suppose that \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\) and \(\{(Y(t), \mathcal{F}_t); t \in [0, \infty)\}\) are \(L^2\)-martingales on \((\Omega, \mathcal{F}, P)\), and write

\[
\mathcal{F}^X \triangleq \sigma\{X(t), t \in [0, \infty)\}, \quad \mathcal{F}^Y \triangleq \sigma\{Y(t), t \in [0, \infty)\},
\]

and

\[
\mathcal{F}^{X,Y}_t \triangleq \sigma\{X(s), Y(s), s \in [0, t]\}, \quad \forall \ t \in [0, \infty).
\]

(a) If the \(\sigma\)-algebras \(\mathcal{F}_X\) and \(\mathcal{F}_Y\) are \(P\)-independent, show that \(\{(X(t)Y(t), \mathcal{F}^{X,Y}_t), t \in [0, \infty)\}\) is a martingale on \((\Omega, \mathcal{F}, P)\).

Hint: use Theorem 1.4.15(f).

(b) If, in addition, the martingales \(\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}\) and \(\{(Y(t), \mathcal{F}_t); t \in [0, \infty)\}\) are continuous, use the result from (a) to prove that

\[
P\{[X, Y](t) = 0, \forall t \in [0, \infty)\} = 1.
\]

**Problem 4.8.8** Suppose that \(\{(X(t), \mathcal{G}_t); t \in [0, \infty)\}\) is a right-continuous supermartingale on \((\Omega, \mathcal{F}, P)\).

(a) Show that \(\{(X(t), \mathcal{G}_t); t \in [0, \infty)\}\) is a right-continuous supermartingale, for \(\mathcal{G}_t \triangleq \mathcal{F}^{X,Y}_{t+}, \forall t \in [0, \infty)\).

Hint: Use Theorem 2.7.9.

(b) Suppose, in addition, that \(X(t, \omega) \in [0, \infty), \forall t \in [0, \infty), \forall \omega \in \Omega, \) and put

\[
S_1(\omega) \triangleq \inf\{t \in [0, \infty) : X(t, \omega) = 0\}, \forall \omega \in \Omega,
\]

\[
S_2(\omega) \triangleq \inf\{t \in [0, \infty) : X(t-, \omega) = 0\}, \forall \omega \in \Omega,
\]

and \(T(\omega) \triangleq S_1(\omega) \wedge S_2(\omega), \forall \omega \in \Omega.\) Show that

\[
P[\omega : T(\omega) < \infty, \text{ and } X(t, \omega) = 0 \ \forall t \geq T(\omega)] = P[T < \infty].
\]
Hint: Observe that the $T_n$ defined by $$T_n(\omega) \triangleq \inf\{t \in [0, \infty) : X(t, \omega) < n^{-1}\}, \ \forall \omega \in \Omega, \ \forall n = 1, 2, \ldots$$ are $\{G_t\}$-stopping times, and use the result established in (a).

**Problem 4.8.9** (a) Suppose that $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a standard scalar Wiener process on $(\Omega, \mathcal{F}, P)$. For each $\lambda \in \mathbb{R}$ define

$$X^\lambda_t \triangleq \exp\left\{ \lambda W(t) - \frac{\lambda^2 t}{2} \right\}, \ \forall t \in [0, \infty). \tag{4.136}$$

Show that $\{(X^\lambda_t, \mathcal{F}_t); t \in [0, \infty)\}$ is a martingale for each $\lambda \in \mathbb{R}$.

Hint: use the moment-generating function for a Gaussian-distributed random variable.

(b) Using (a) and Theorem 4.3.3 show that

$$P\left[ \max_{0 \leq s \leq t} W(s) \geq \alpha t \right] \leq \exp\left( -\alpha^2 t \right),$$

for all $\alpha, t \in [0, \infty)$.

**Problem 4.8.10** Suppose that $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a scalar standard Wiener process on probability space $(\Omega, \mathcal{F}, P)$. For each $x \in \mathbb{R}$ define the process $\{W^x(t), t \in [0, \infty)\}$ by

$$W^x(t, \omega) \triangleq x + W(t, \omega), \ \forall t \in [0, \infty), \ \forall \omega \in \Omega. \tag{4.137}$$

If $a, b \in \mathbb{R}$ are constants with $a < b$, then, for each $x \in \mathbb{R}$, define

$$T^x(\omega) \triangleq \inf\{t \in [0, \infty) : W^x(t, \omega) \notin (a, b)\}, \ \forall \omega \in \Omega.$$

(a) Show that $E[T^x] < \infty$ for each $x \in \mathbb{R}$.

(b) Using (a) show that

$$E[T^x] = (b - x)(x - a), \ \forall x \in (a, b).$$

**Problem 4.8.11** Suppose that $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a non-negative continuous martingale such that

$$\lim_{t \to \infty} X(t) = 0, \ \text{a.s.}$$

Show that for each $x \in (0, \infty)$

$$P\left[ \sup_{s \in [0, \infty)} X(s) \geq x \ \bigg| X(0) \right] = \min\left[ 1, \frac{X(0)}{x} \right], \ \text{a.s.}$$

Hint: Stop $\{X(t)\}$ at the $\{\mathcal{F}_t\}$-stopping time given by

$$T(\omega) \triangleq \inf\{t \in [0, \infty) : X(t, \omega) \geq x\},$$

and use Corollary 4.5.8.
Problem 4.8.12 Suppose that $X \in \mathcal{M}_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$, $S$ is a $\{\mathcal{F}_t\}$-stopping time, and there is a constant $a \in [0, \infty)$ such that

$$E[X^2(S \wedge T)] \leq a,$$

for each $\{\mathcal{F}_t\}$-stopping time $T$ for which $P[T < \infty] = 1$. Show that $\{(X(t \wedge S), \mathcal{F}_t); t \in [0, \infty)\}$ is an $L^2$-bounded martingale.

Problem 4.8.13 (a) Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space, $\mathcal{G} \subset \mathcal{F}$ is a sub $\sigma$-algebra, and $A \in \mathcal{G}$. The trace of $\mathcal{G}$ on the event $A$ is defined by

$$A \cap \mathcal{G} \triangleq \{ A \cap G | G \in \mathcal{G} \}.$$

Now suppose that $T$ and $U$ are $\{\mathcal{F}_t\}$-stopping times. Show that

$$\{T \leq U\} \cap \mathcal{F}_T = \{T \leq U\} \cap \mathcal{F}_{T \wedge U},$$

that is, $\mathcal{F}_T$ and $\mathcal{F}_{T \wedge U}$ have identical trace on the event $\{T \leq U\}$.

(b) Suppose that $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a right-continuous martingale, and $T$ and $U$ are $\{\mathcal{F}_t\}$-stopping times, with $U$ uniformly bounded i.e. $U(\omega) \leq B$, all $\omega \in \Omega$, for some constant $B \in [0, \infty)$. Show that

$$E[X(U) | \mathcal{F}_T] = X(T \wedge U), \quad a.s.$$

Hint: Use the result in (a) together with the optional sampling theorem for ordered stopping times (see Corollary 4.5.6) and the localization property of conditional expectations established in Problem 1.6.11(a)(b).

(c) Again with $T$ a $\{\mathcal{F}_t\}$-stopping time, put

$$Y(t) \triangleq \xi[X(t) - X(t \wedge T)], \quad t \in [0, \infty),$$

where $\xi$ is uniformly bounded and $\mathcal{F}_T$-measurable. Show that $\{(Y(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a martingale. Hint: use Lemma 4.5.1 and the result established in (b).

Problem 4.8.14 Suppose that $X \in \mathcal{M}_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$.

(a) Show that

$$\{ \sup_{t \in [0, \infty)} X_t < \infty \} \subset \{ \lim_{t \to \infty} X_t \text{ exists in } \mathcal{R} \} \quad a.s.$$

Hint: define the $\{\mathcal{F}_t\}$-stopping times $T_n$ by

$$T_n \triangleq \inf\{t \in [0, \infty) : X_t \geq n\}, \quad \forall \ n = 1, 2, \ldots$$
and use the fact that \( \{ (n - X_{t \wedge T_{n}}), \mathcal{F}_{t} \}; \ t \in [0, \infty) \) is a continuous non-negative local martingale.

(b) Next, show that

\[
\{ \lim_{t \to \infty} X_{t} \text{ exists in } \mathbb{R} \} = \{ \sup_{t \in [0, \infty)} X_{t} < \infty \} \cup \{ \inf_{t \in [0, \infty)} X_{t} > -\infty \} \quad \text{a.s.}
\]

(c) Finally, conclude that

\[
P[\lim_{t \to \infty} X_{t} = +\infty] = P[\lim_{t \to \infty} X_{t} = -\infty] = 0,
\]

and

\[
\{ \lim_{t \to \infty} X_{t} \text{ exists in } \mathbb{R} \} = \{ \sup_{t \in [0, \infty)} |X_{t}| < \infty \} \quad \text{a.s.}
\]

Note: Recall the notation in Remark 1.4.17.

**Problem 4.8.15** Suppose that \( X \in \mathbb{M}_{loc}^{c,0}(\{\mathcal{F}_{t}\}, P) \).

(a) Define the \( \{\mathcal{F}_{t}\}\)-stopping time \( S \) by

\[
S \overset{\Delta}{=} \inf \{ t \in [0, \infty) : [X](t) \geq a \},
\]

for some constant \( a \in (0, \infty) \), and show that \( \{(X(t \wedge S), \mathcal{F}_{t}); \ t \in [0, \infty)\} \) is an \( L^{2} \)-bounded continuous martingale. Hint: use Problem 4.8.12.

(b) Next, show that

\[
\{[X](\infty) < \infty\} \subset \{ \lim_{t \to \infty} X_{t} \text{ exists in } \mathbb{R} \} \quad \text{a.s.}
\]

Hint: Use the result from (a) together with the \( \{\mathcal{F}_{t}\}\)-stopping times

\[
S_{n} \overset{\Delta}{=} \inf \{ t \in [0, \infty) : [X](t) \geq n \}, \quad \forall \ n = 1, 2, \ldots
\]

(c) Next, conclude that

\[
\{[X](\infty) < \infty\} = \{ \lim_{t \to \infty} X_{t} \text{ exists in } \mathbb{R} \} \quad \text{a.s.}
\]

Hint: Use the \( \{\mathcal{F}_{t}\}\)-stopping times

\[
T_{n} \overset{\Delta}{=} \inf \{ t \in [0, \infty) : |X_{t}| \geq n \}, \quad \forall \ n = 1, 2, \ldots
\]

to obtain the a.s. set-inclusion opposite to that established in (b).

Note: Recall the notation in Remark 1.4.17.

**Problem 4.8.16** In this problem \( \{\mathcal{F}_{t}, \ t \in [0, \infty)\} \) is a given *right-continuous* filtration on the probability space \((\Omega, \mathcal{F}, P)\), that is

\[
\mathcal{F}_{t} = \bigcap_{s \geq t} \mathcal{F}_{s}, \quad \text{for all } t \in [0, \infty).
\]
(a) Suppose that $T_n, n = 1, 2, \ldots$, are $\{\mathcal{F}_t\}$-stopping times. Show that

$$T \triangleq \inf_n T_n$$

is a $\{\mathcal{F}_t\}$-optional time,

then conclude that it is a $\{\mathcal{F}_t\}$-stopping time.

Hint: Use Definition 3.3.1 and Proposition 3.3.3.

(b) Suppose that $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a right-continuous adapted process on $(\Omega, \mathcal{F}, P)$, and that $T_n, n = 1, 2, \ldots$, are $\{\mathcal{F}_t\}$-stopping times such that

(i) $0 \leq T_n(\omega) \leq T_{n+1}(\omega)$ for all $\omega \in \Omega$ and $n = 1, 2, \ldots$

(ii) $P [\lim_{n \to \infty} T_n = \infty] = 1$.

Establish the following: $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a local martingale if and only if the stopped processes $\{(X(t \wedge T_n), \mathcal{F}_t), t \in [0, \infty)\}$ are local martingales for each $n = 1, 2, \ldots$

Hint: If $\{(X(t \wedge T_n), \mathcal{F}_t), t \in [0, \infty)\}$ is a local martingale then (see Remark 4.6.2) it has a localizing sequence $\{S_{n,m}, m = 1, 2, \ldots\}$ of $\{\mathcal{F}_t\}$-stopping times. Show that, for each $n = 1, 2, \ldots$, there is some positive integer $m(n)$ such that

$$P \left[ S_{n,m(n)} < n \wedge T_n \right] < 2^{-n}.$$

Then put $\tilde{T}_n \triangleq T_n \wedge S_{n,m(n)}$, show that $\lim_{n \to \infty} \tilde{T}_n = \infty$ a.s., and use Corollary 4.5.8.
Chapter 5

Stochastic Integration

We are now almost ready to address the main subject of these notes, namely stochastic integration, but there remain two preliminary items we must still discuss. In § 5.1 we shall look at a specialization of classical Lebesgue integration theory for functions of a real variable with respect to a signed measure, and arrive at the concept of a Lebesgue-Stieltjes integral. In § 5.2 we extend this notion to a setting in which the integrands and the signed measures with respect to which we integrate are allowed to be random, and we shall come up with the notion of a random Lebesgue-Stieltjes integral. These integrals, although not in themselves genuine stochastic integrals, are nevertheless essential for writing down the main results of stochastic integration theory such as Itô’s change-of-variables formula. In § 5.3, we present the definition and basic properties of stochastic integrals, in § 5.4 we derive Itô’s change-of-variables formula, and in § 5.7 we use this formula to establish an essential result of stochastic calculus called Girsanov’s theorem.

5.1 Lebesgue-Stieltjes Integration

In our study of stochastic integration we shall frequently need to write down Lebesgue integrals of real-valued functions on the half-line $[0, \infty)$ with respect to a signed measure generated by some given mapping of locally bounded variation. The resulting integrals are usually called Lebesgue-Stieltjes integrals. In view of the importance of these integrals for expressing the principal results of stochastic integration theory, we devote this section to a reasonably complete and self-contained account of Lebesgue-Stieltjes integration. In particular we shall develop an integration by parts formula and a generalization of the Cauchy-Schwarz inequality, both of which will serve us well when we start to look at the main ideas of stochastic integration in detail.

Before presenting the main features of Lebesgue-Stieltjes integration we must first return to the theme of functions of locally bounded variation and total variation, which were set forth in Definitions 4.7.6 and 4.7.7, and take these ideas somewhat further. We begin by expanding on Definition 4.7.7:
Definition 5.1.1 For a given mapping \( A : [0, \infty) \to IR \), the total variation, positive variation, and negative variation of \( A \) on the interval \([s, t] \subset [0, \infty)\), are, respectively, the non-negative numbers

\[
V[A; s, t] \triangleq \sup_n \sum_{k=1}^{n} |A(t_k) - A(t_{k-1})|, \\
V_+[A; s, t] \triangleq \sup_n \sum_{k=1}^{n} [A(t_k) - A(t_{k-1})]^+,
\]

(5.1)

\[
V_-[A; s, t] \triangleq \sup_n \sum_{k=1}^{n} [A(t_k) - A(t_{k-1})]^-, \\
\]

(5.2)

where, in each case, the supremum is over all finite partitions \( s = t_0 < t_1 < \ldots < t_n = t \) of \([s, t]\).

Remark 5.1.2 For a given mapping \( A : [0, \infty) \to IR \) the functions defined on \([0, \infty)\) by \( t \to V[A; 0, t], t \to V_+[A; 0, t] \) and \( t \to V_-[A; 0, t] \) are called respectively the total variation function, positive variation function and negative variation function of \( A \). To lighten the notation we shall usually denote the numbers \( V[A; 0, t], V_+[A; 0, t], \) and \( V_-[A; 0, t] \) by \( \hat{A}(t), \hat{A}_+(t) \) and \( \hat{A}_-(t) \) respectively. Observe that \( \hat{A}(\cdot), \hat{A}_+(\cdot) \) and \( \hat{A}_-(\cdot) \) are \([0, \infty]-valued non-decreasing functions on \([0, \infty)\), and if \( A(\cdot) \) is of locally bounded variation on \([0, \infty)\) then all of these functions are finitely-valued (i.e. take their values in \([0, \infty)\)). Also, it is clear from Definition 5.1.1 that \( \hat{A}, \hat{A}_+ \) and \( \hat{A}_- \) are null at the origin i.e. \( \hat{A}(0) = \hat{A}_+(0) = \hat{A}_-(0) = 0 \).

There is a simple relationship between a given mapping and its total, positive, and negative variation functions which is given by the next lemma. The proof involves just a simple manipulation of the preceding definitions and is left as an exercise for the reader:

Lemma 5.1.3 For a given mapping \( A : [0, \infty) \to IR \) we have

\[
V[A; s, t] = V_+[A; s, t] + V_-[A; s, t], \quad \forall 0 \leq s < t < \infty.
\]

(5.4)

Furthermore, if \( A \) is of locally bounded variation, then

\[
A(t) - A(s) = V_+[A; s, t] - V_-[A; s, t], \quad \forall 0 \leq s < t < \infty.
\]

(5.5)

Remark 5.1.4 For a mapping \( A : [0, \infty) \to IR \), Lemma 5.1.3 shows that

\[
\hat{A}(t) = \hat{A}_+(t) + \hat{A}_-(t), \quad \forall t \in [0, \infty),
\]

(5.6)

and, when \( A \) is of locally bounded variation, then

\[
A(t) - A(0) = \hat{A}_+(t) - \hat{A}_-(t), \quad \forall t \in [0, \infty).
\]

(5.7)

In view of Remark 5.1.2 we see that the difference on the right hand side of (5.7) is well defined (we never encounter the combination \( \infty - \infty \)).
In Remark 4.7.11 we saw that any mapping which is the difference of two non-decreasing functions is of locally bounded variation. Remark 5.1.4 asserts the converse, for (5.7) shows how a given mapping of locally bounded variation is always the difference of two non-decreasing functions. Of course there is nothing intrinsically unique about this decomposition. In fact, if \( D : [0, \infty) \to [0, \infty) \) is any non-decreasing function with \( D(0) = 0 \), and we put \( B(t) = \tilde{A}_+(t) + D(t) \) and \( C(t) = \tilde{A}_-(t) + D(t) \), \( \forall t \in [0, \infty) \), then \( B \) and \( C \) are also non-decreasing functions with \( B(0) = C(0) = 0 \), and clearly

\[
A(t) - A(0) = B(t) - C(t), \quad \forall t \in [0, \infty).
\]

Although uniqueness does not hold, the decomposition given by (5.7) is nevertheless minimal in the sense shown by the next lemma whose proof is just an elementary exercise:

**Lemma 5.1.5** Suppose that \( A : [0, \infty) \to \mathbb{R} \) is of locally bounded variation. If \( B, C : [0, \infty) \to [0, \infty) \) are non-decreasing mappings with \( B(0) = C(0) = 0 \), such that

\[
(5.8) \quad A(t) - A(0) = B(t) - C(t), \quad \forall t \in [0, \infty),
\]

then \( B(t) \geq \tilde{A}_+(t) \) and \( C(t) \geq \tilde{A}_-(t) \), \( \forall t \in [0, \infty) \). Moreover, if \( B \) and \( C \) in (5.8) are such that

\[
\tilde{A}(t) = B(t) + C(t), \quad \forall t \in [0, \infty),
\]

then \( B(t) = \tilde{A}_+(t) \) and \( C(t) = \tilde{A}_-(t) \), \( \forall t \in [0, \infty) \).

So far we have not postulated any hypotheses for the function \( A(\cdot) \), such as continuity or even Borel measurability. For later applications we shall need the following lemma whose easy proof is left to the reader:

**Lemma 5.1.6** Suppose that \( A : [0, \infty) \to \mathbb{R} \) is continuous and of locally bounded variation. Then the mappings \( \tilde{A}, \tilde{A}_+ \) and \( \tilde{A}_- \) are continuous and non-decreasing on \([0, \infty)\), assume their values in \([0, \infty)\) (i.e. are finitely-valued), and \( \tilde{A}(0) = \tilde{A}_+(0) = \tilde{A}_-(0) = 0 \). Moreover, for each \( t \in [0, \infty) \) we have

\[
\tilde{A}(t) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} |A(t \wedge \tau^n_{k+1}) - A(t \wedge \tau^n_k)|,
\]

\[
\tilde{A}_+(t) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} [A(t \wedge \tau^n_{k+1}) - A(t \wedge \tau^n_k)]^+,
\]

and

\[
\tilde{A}_-(t) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} [A(t \wedge \tau^n_{k+1}) - A(t \wedge \tau^n_k)]^-,
\]

where \( \tau^n_k = k2^{-n}, \forall k, n = 0, 1, 2, \ldots \)

Finally, to complete our preliminaries we need the following theorem which is a simple consequence of Theorem 1.2.3:
Theorem 5.1.7 Suppose that \( A : [0, \infty) \to \mathbb{R} \) is right-continuous and non-decreasing. Then there exists a unique measure \( \mu_A \) on the measurable space \( ([0, \infty), \mathcal{B}([0, \infty))) \) such that

\[
\mu_A(\{0\}) = 0 \quad \text{and} \quad \mu_A((s, t]) = A(t) - A(s), \quad \forall 0 \leq s < t < \infty.
\]

(5.9)

If, furthermore, \( A(\cdot) \) is continuous, then \( \mu_A(\{t\}) = 0, \forall t \in [0, \infty) \).

Remark 5.1.8 The measure \( \mu_A \) in Theorem 5.1.7 is usually called the Lebesgue-Stieltjes measure generated by \( A \).

Having disposed of the above preliminary notions we are now able to formalize the concept of a Lebesgue-Stieltjes integral. We do so in steps of successively increasing generality:

Step I : Suppose that \( A : [0, \infty) \to \mathbb{R} \) is continuous and non-decreasing, and \( \Phi : [0, \infty) \to [0, \infty] \) is \( \mathcal{B}([0, \infty)) \)-measurable. Then the Lebesgue integral

\[
\int_{[0, \infty)} \Phi d\mu_A
\]

is well defined, with a value in \([0, \infty]\). Henceforth, we are going to use an alternative notation for this integral and write

\[
\int_{[0, \infty)} \Phi(s) dA(s) \overset{\Delta}{=} \int_{[0, \infty)} \Phi d\mu_A.
\]

(5.10)

In keeping with the terminology established for the measure \( \mu_A \) in Remark 5.1.8 we call the quantity in (5.10) the Lebesgue-Stieltjes integral of \( \Phi \) with respect to \( A \). One may reasonably question why we have introduced this new notation, since the right hand side of (5.10) already gives a perfectly explicit symbolism for the integral concerned. The reason will become apparent when we develop integral relationships, such as the integration by parts formula, where it will be seen that the new notation has some significant advantages.

Step II : Next, suppose that \( A : [0, \infty) \to \mathbb{R} \) is continuous and non-decreasing, and \( \Phi : [0, \infty) \to [0, \infty] \) is \( \mathcal{B}([0, \infty)) \)-measurable. We define the Lebesgue-Stieltjes integral of \( \Phi \) with respect to \( A \) by

\[
\int_{[0, \infty)} \Phi(s) dA(s) \overset{\Delta}{=} \int_{[0, \infty)} \Phi_+(s) dA(s) - \int_{[0, \infty)} \Phi_-(s) dA(s)
\]

(5.11)

provided the two non-negative numbers on the right hand side are both finite, or equivalently, provided that

\[
\int_{[0, \infty)} |\Phi(s)| dA(s) < \infty.
\]

(5.12)

If the integral in (5.12) assumes the value \(+\infty\) then the Lebesgue-Stieltjes integral of the \( \mathbb{R} \)-valued function \( \Phi \) with respect to \( A \) is undefined.
Step III : Now suppose $A : [0, \infty) \rightarrow \mathbb{R}$ is continuous and of locally bounded variation, and $\Phi : [0, \infty) \rightarrow \mathbb{R}$ is $\mathcal{B}([0, \infty))$-measurable. We define the Lebesgue-Stieltjes integral of $\Phi$ with respect to $A$ by
\[ \int_0^\infty \Phi(s) \, dA(s) \triangleq \int_0^\infty \Phi(s) \, d\bar{A}_+(s) - \int_0^\infty \Phi(s) \, d\bar{A}_-(s) \tag{5.13} \]
provided the two integrals on the right hand side (which are formulated in Step II) are themselves defined, that is provided that
\[ \int_0^\infty |\Phi(s)| \, d\bar{A}_+(s) < \infty \quad \text{and} \quad \int_0^\infty |\Phi(s)| \, d\bar{A}_-(s) < \infty. \tag{5.14} \]
If either of the two conditions in (5.14) fails to hold then the Lebesgue-Stieltjes integral of $\Phi$ with respect to $A$ is again undefined. Thus, when the integral in (5.13) is defined, it must be $\mathbb{R}$-valued.

There is another way of stating the pair of conditions in (5.14) which will be useful for later analysis. Let $\mu_{\bar{A}}, \mu_{\bar{A}_+}$ and $\mu_{\bar{A}_-}$ be the Lebesgue-Stieltjes measures generated by the continuous non-decreasing mappings $\bar{A}, \bar{A}_+$ and $\bar{A}_-$ respectively (see Theorem 5.1.7). In view of (5.6) and (5.9) it is clear that
\[ \mu_{\bar{A}}((s, t]) = \mu_{\bar{A}_+}((s, t]) + \mu_{\bar{A}_-}((s, t]), \quad \forall \ 0 \leq s < t < \infty, \]
and in view of the uniqueness asserted in Theorem 5.1.7 it then follows that
\[ \mu_{\bar{A}}(\Gamma) = \mu_{\bar{A}_+}(\Gamma) + \mu_{\bar{A}_-}(\Gamma), \quad \forall \ \Gamma \in \mathcal{B}([0, \infty)). \]
Thus, we see at once that
\[ \int_{[0, \infty)} |\Phi| \, d\mu_{\bar{A}} = \int_{[0, \infty)} |\Phi| \, d\mu_{\bar{A}_+} + \int_{[0, \infty)} |\Phi| \, d\mu_{\bar{A}_-} \]
or, in our Lebesgue-Stieltjes notation,
\[ \int_0^\infty |\Phi(s)| \, d\bar{A}(s) = \int_0^\infty |\Phi(s)| \, d\bar{A}_+(s) + \int_0^\infty |\Phi(s)| \, d\bar{A}_-(s). \tag{5.15} \]
Therefore, the two conditions in (5.14) hold if and only if
\[ \int_0^\infty |\Phi(s)| \, d\bar{A}(s) < \infty, \tag{5.16} \]
and we only define the Lebesgue-Stieltjes integral of $\Phi$ with respect to $A$ when (5.16) holds, in which case the integral is real-valued and defined by (5.13).

We can easily extend the preceding ideas to arrive at the notion of the indefinite Lebesgue-Stieltjes integral. If $A : [0, \infty) \rightarrow \mathbb{R}$ is continuous and non-decreasing, and $\Phi : [0, \infty) \rightarrow [0, \infty]$ is $\mathcal{B}([0, \infty))$-measurable, then the indefinite Lebesgue-Stieltjes integral of $\Phi$ with respect to $A$ is the mapping from $[0, \infty)$ into $[0, \infty]$ defined for each $t \in [0, \infty)$ by
\[ \int_0^t \Phi(s) \, dA(s) \triangleq \int_0^\infty 1\{s \leq t\} \Phi(s) \, dA(s). \tag{5.17} \]
On the other hand, if $A : [0, \infty) \to \mathbb{R}$ is of locally bounded variation and $\Phi : [0, \infty) \to \mathbb{R}$ is $\mathcal{B}([0, \infty))$-measurable, then the indefinite Lebesgue-Stieltjes integral of $\Phi$ with respect to $A$ is the mapping from $[0, \infty)$ into $\mathbb{R}$ which is again given by (5.17). In view of the condition expressed by (5.16), this mapping is defined if and only if
\begin{equation}
\int_0^t |\Phi(s)| d\bar{A}(s) < \infty, \quad \forall t \in [0, \infty),
\end{equation}
and then it is easily seen that
\begin{equation}
\int_0^t \Phi(s) dA(s) = \left[ \int_0^t \Phi_+(s) d\bar{A}_+(s) - \int_0^t \Phi_-(s) d\bar{A}_+(s) \right]
- \left[ \int_0^t \Phi_+(s) d\bar{A}_-(s) - \int_0^t \Phi_-(s) d\bar{A}_-(s) \right]
\end{equation}
for all $t \in [0, \infty)$, where the four integrals on the right side of (5.19) always take their values in $[0, \infty)$. It is also clear that (5.19) holds when $t = +\infty$ with the four integrals on the right taking values in $[0, \infty)$ provided (5.16) holds.

**Remark 5.1.9** Suppose that $\Phi : [0, \infty) \to \mathbb{R}$ is continuous, $A : [0, \infty) \to \mathbb{R}$ is continuous and of locally bounded variation, and fix some arbitrary $t \in [0, \infty)$. Then there is some constant $M \in [0, \infty)$ (possibly depending on $t$) such that $|\Phi(s)| \leq M, \forall s \in [0, t]$, hence
\begin{equation}
\int_0^t |\Phi(s)| d\bar{A}(s) \leq M \bar{A}(t) < \infty.
\end{equation}
Thus condition (5.18) is verified, and hence the Lebesgue-Stieltjes integral $\int_0^t \Phi(s) dA(s)$ is defined, real-valued, and given by (5.19) for all $t \in [0, \infty)$.

The following results establish simple properties of Lebesgue-Stieltjes integrals. The proofs are elementary and we leave most of them as exercises for the reader:

**Lemma 5.1.10** Suppose $A : [0, \infty) \to \mathbb{R}$ is continuous and of locally bounded variation. If $T \in [0, \infty]$ is a constant and $A^T : [0, \infty) \to \mathbb{R}$ is defined by $A^T(t) \overset{\Delta}{=} A(t \wedge T), \forall t \in [0, \infty)$, then $A^T(\cdot)$ is continuous and of locally bounded variation on $[0, \infty)$. Furthermore, if $A^T(t) \overset{\Delta}{=} V[A^T; 0, t], \forall t \in [0, \infty)$, then
\begin{equation}
A^T(t) = \bar{A}(t \wedge T), \quad \forall t \in [0, \infty).
\end{equation}
If $\Phi : [0, \infty) \to \mathbb{R}$ is $\mathcal{B}([0, \infty))$-measurable, and
\begin{equation}
\int_0^t |\Phi(s)| d\bar{A}(s) < \infty
\end{equation}
for each $t \in [0, \infty)$, then
\begin{equation}
\int_0^t |\Phi(s)| d\bar{A}^T(s) < \infty
\end{equation}
for each $t \in [0, \infty)$, and

$$
(5.22) \quad \int_0^t \Phi(s)dA_T(s) = \int_0^{t\wedge T} \Phi(s)dA(s)
$$

for each $t \in [0, \infty)$. Moreover, if (5.20) holds when $t = +\infty$ then (5.21) and (5.22) also hold when $t = +\infty$.

**Lemma 5.1.11** Suppose $A, B : [0, \infty) \to IR$ are continuous and of locally bounded variation, and let $C(t) \triangleq A(t) + B(t)$, $\forall t \in [0, \infty)$. Then $C(\cdot)$ is continuous and of locally bounded variation on $[0, \infty)$. If $\Phi : [0, \infty) \to IR$ is $B([0, \infty))$-measurable, with

$$
(5.23) \quad \int_0^t |\Phi(s)|d\tilde{A}(s) < \infty \quad \text{and} \quad \int_0^t |\Phi(s)|d\tilde{B}(s) < \infty,
$$

for each $t \in [0, \infty)$, then

$$
(5.24) \quad \int_0^t |\Phi(s)|d\tilde{C}(s) < \infty
$$

and

$$
(5.25) \quad \int_0^t \Phi(s)d\tilde{C}(s) = \int_0^t \Phi(s)dA(s) + \int_0^t \Phi(s)dB(s)
$$

for each $t \in [0, \infty)$. Moreover, if (5.23) holds when $t = +\infty$ then so also do (5.24) and (5.25).

**Lemma 5.1.12** Suppose that $A : [0, \infty) \to IR$ is continuous and of locally bounded variation, $\Phi : [0, \infty) \to IR$ is $B([0, \infty))$-measurable, and

$$
(5.26) \quad \int_0^t |\Phi(s)|d\tilde{A}(s) < \infty, \quad \forall t \in [0, \infty).
$$

Then the mapping $\tilde{B}$ defined by

$$
(5.27) \quad \tilde{B}(t) \triangleq \int_0^t \Phi(s)dA(s), \quad \forall t \in [0, \infty),
$$

is finitely-valued (i.e. $\tilde{B}(t) \in IR, \forall t \in [0, \infty)$), continuous, and of locally bounded variation on $[0, \infty)$. Moreover, for each $t \in [0, \infty)$ we have

$$
(5.28) \quad \tilde{B}(t) = \int_0^t |\Phi(s)|d\tilde{A}(s),
$$

$$
(5.29) \quad \tilde{B}_+(t) = \int_0^t \Phi_+(s)d\tilde{A}_+(s) + \int_0^t \Phi_-(s)d\tilde{A}_-(s),
$$

$$
(5.30) \quad \tilde{B}_-(t) = \int_0^t \Phi_+(s)d\tilde{A}_-(s) + \int_0^t \Phi_-(s)d\tilde{A}_+(s).
$$

Proof : Condition (5.26) ensures that each of the four integrals on the right hand side of (5.19), viewed as functions of $t \in [0, \infty)$, take values in $[0, \infty)$ and are non-decreasing, hence we see that $\tilde{B}$ is finitely-valued and of locally bounded variation. In view of (5.19) we can, when establishing continuity of $\tilde{B}$,
assume without loss of generality that $A$ is non-decreasing and $\Phi$ takes values in $[0, \infty]$. Now fix some $T \in (0, \infty)$. Then $B(T) < \infty$, hence
$$\int_{[0,T]} \Phi d\mu_A < \infty,$$
and so it follows from Theorem 1.2.20 that, corresponding to each $\epsilon > 0$, there exists some $\eta(\epsilon) > 0$ such that
$$\int_{[s,t]} \Phi d\mu_A < \epsilon$$
for all $0 \leq s < t \leq T$ with $\mu_A([s,t]) < \eta(\epsilon)$. Equivalently, from (5.9) and (5.27), we have that $|B(t) - B(s)| < \epsilon$ for all $0 \leq s < t \leq T$ such that $|A(t) - A(s)| < \eta(\epsilon)$. But $A$ is continuous on $[0, \infty)$, hence uniformly continuous on the closed and bounded interval $[0, T]$, and so there exists some $\delta(\epsilon) > 0$ such that $|A(t) - A(s)| < \eta(\epsilon)$ for all $0 \leq s < t \leq T$ subject to $t - s < \delta(\epsilon)$. It follows that the mapping $B$ is uniformly continuous on $[0, T]$, and, in view of the arbitrary choice of $T \in (0, \infty)$, it must be continuous on $[0, \infty)$. The remaining assertions follow from Lemma 5.1.5.

Our next result is a chain-rule for Lebesgue-Stieltjes integrals:

**Theorem 5.1.13** (a) Suppose $A : [0, \infty) \to \mathbb{R}$ is continuous and non-decreasing, $\Phi : [0, \infty) \to [0, \infty]$ is $\mathcal{B}([0, \infty))$-measurable, and the mapping $B(\cdot)$ defined by

$$B(t) \triangleq \int_0^t \Phi(s) dA(s), \quad \forall t \in [0, \infty),$$

is finitely-valued, that is, $B(t) \in [0, \infty)$, $\forall t \in [0, \infty)$. Then $B(\cdot)$ is continuous and non-decreasing on $[0, \infty)$, and for each $\mathcal{B}([0, \infty))$-measurable function $\Psi : [0, \infty) \to [0, \infty]$ we have

$$\int_0^t \Psi(s) dB(s) = \int_0^t \Psi(s) \Phi(s) dA(s)$$

for all $t \in [0, \infty]$.

(b) Suppose $A : [0, \infty) \to \mathbb{R}$ is continuous and of locally bounded variation and $\Phi : [0, \infty) \to \mathbb{R}$ is $\mathcal{B}([0, \infty))$-measurable such that

$$\int_0^t |\Phi(s)| d\bar{A}(s) < \infty, \quad \forall t \in [0, \infty).$$

Then the mapping $B(\cdot)$ defined by (5.31) is continuous and of locally bounded variation on $[0, \infty)$. Moreover, if $\Psi : [0, \infty) \to \mathbb{R}$ is $\mathcal{B}([0, \infty))$-measurable, and

$$\int_0^t |\Psi(s)| dB(s) < \infty$$

for all $t \in [0, \infty)$, then

$$\int_0^t |\Psi(s)||\Phi(s)| d\bar{A}(s) < \infty,$$

and (5.32) holds, for all $t \in [0, \infty)$. Finally, if (5.34) holds when $t = +\infty$ then (5.35) and (5.32) hold when $t = +\infty$. 190
Proof: (a) We see from Lemma 5.1.12 that $B(\cdot)$ defined by (5.31) is continuous and non-decreasing on $[0, \infty)$. If $\mu_B$ is the Lebesgue-Stieltjes measure on $B([0, \infty))$ generated by $B(\cdot)$ then the uniqueness assertion of Theorem 5.1.7 ensures that

$$
\mu_B(\Gamma) = \int_{\Gamma} \Phi d\mu_A, \quad \forall \ \Gamma \in B([0, \infty)).
$$

Thus, by Corollary 1.2.13, we have

$$
\int_{[0, \infty)} \Psi d\mu_B = \int_{[0, \infty)} \Psi \Phi d\mu_A,
$$

which, re-written in the notation of Lebesgue-Stieltjes integrals, is just (5.32) when $t = +\infty$. To establish (5.32) for some $t \in [0, \infty)$ just replace $\Psi$ by $\Psi I_{[0,t]}$ in the preceding.

(b) From Lemma 5.1.12 and (5.33) we see that $B(\cdot)$ defined by (5.31) is continuous and of locally bounded variation on $[0, \infty)$. Now suppose that (5.34) holds for each $t \in [0, \infty)$. Then the Lebesgue-Stieltjes integral on the left hand side of (5.32) is defined and $\mathcal{R}$-valued for each $t \in [0, \infty)$. Moreover, since

$$
\breve{B}(t) \triangleq \int_{0}^{t} |\Phi(s)| d\breve{A}(s), \quad \forall \ t \in [0, \infty),
$$

(Lemma 5.1.12) we see from (a) along with (5.34) that

$$
\int_{0}^{t} |\Psi(s)| |\Phi(s)| d\breve{A}(s) \leq \int_{0}^{t} |\Psi(s)| d\breve{B}(s) < \infty, \quad \forall \ t \in [0, \infty).
$$

Thus (5.35) holds, so that the Lebesgue-Stieltjes integral on the right hand side of (5.32) is defined and $\mathcal{R}$-valued for each $t \in [0, \infty)$. It remains to show that the two sides of (5.32) are equal. This just involves expanding each side of (5.32) in the form given by (5.19) and using the positive and negative variations of $B(\cdot)$ which are given by Lemma 5.1.12. We leave the easy computations to the reader. If the case where (5.34) holds for $t = +\infty$ we follow an obvious repetition of the preceding argument.

We next establish an integration-by-parts formula for Lebesgue-Stieltjes integrals. To this end we first show that a Lebesgue-Stieltjes integral has a particularly simple representation as a limit of Riemann sums when the integrand is continuous:

Theorem 5.1.14 Suppose $\Phi : [0, \infty) \to \mathbb{R}$ is continuous and $A : [0, \infty) \to \mathbb{R}$ is continuous and of locally bounded variation. Then, for each $t \in [0, \infty)$, we have

$$
\int_{0}^{t} \Phi(s)dA(s) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} \Phi(t \wedge \tau_k^n)[A(t \wedge \tau_k^n) - A(t \wedge \tau_k^n)],
$$

where $\tau_k^n = k2^{-n}, \forall \ k, n = 0, 1, 2, \ldots$

Proof: For each $n = 0, 1, 2, \ldots$ define $\Phi_n : [0, \infty) \to \mathbb{R}$ by $\Phi_n(s) \triangleq \Phi(\tau_k^n)$ for all $s \in (\tau_k^n, \tau_{k+1}^n]$, $k = 0, 1, 2, \ldots$ and $\Phi_n(0) \triangleq \Phi(0)$. Then clearly, for each $t \in [0, \infty)$, we have

$$
\int_{[0,t]} \Phi_n d\mu_A = \sum_{0 \leq k < \infty} \Phi(t \wedge \tau_k^n)\mu_{A_+}((t \wedge \tau_k^n, t \wedge \tau_{k+1}^n]).
$$
or equivalently:

\[
\int_0^t \Phi_n(s)d\tilde{A}_+(s) = \sum_{0 \leq k < \infty} \Phi(t \wedge \tau_k^n)[\tilde{A}_+(t \wedge \tau_{k+1}^n) - \tilde{A}_+(t \wedge \tau_k^n)]
\]

Since \( \Phi \) is continuous we see that \( \lim_{n \to \infty} \Phi_n(s) = \Phi(s) \) for each \( s \in [0, t] \), and since \( \Phi \) is uniformly bounded on \( [0, t] \), there is some constant \( M \in [0, \infty) \) such that \( |\Phi_n(s)| \leq M, \forall \ s \in [0, t], \forall \ n = 0, 1, 2, \ldots \)

Thus, from the dominated convergence theorem, we have \( \lim_{n \to \infty} \int_0^t \Phi_n(s)d\tilde{A}_+(s) = \int_0^t \Phi(s)d\tilde{A}_+(s) \)

hence, from (5.38),

\[
\int_0^t \Phi(s)d\tilde{A}_+(s) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} \Phi(t \wedge \tau_k^n)[\tilde{A}_+(t \wedge \tau_{k+1}^n) - \tilde{A}_+(t \wedge \tau_k^n)].
\]

Similarly,

\[
\int_0^t \Phi(s)d\tilde{A}_-(s) = \lim_{n \to \infty} \sum_{0 \leq k < \infty} \Phi(t \wedge \tau_k^n)[\tilde{A}_-(t \wedge \tau_{k+1}^n) - \tilde{A}_-(t \wedge \tau_k^n)].
\]

Now, by (5.7),

\[
A(t \wedge \tau_{k+1}^n) - A(t \wedge \tau_k^n) = [\tilde{A}_+(t \wedge \tau_{k+1}^n) - \tilde{A}_+(t \wedge \tau_k^n)] - [\tilde{A}_-(t \wedge \tau_{k+1}^n) - \tilde{A}_-(t \wedge \tau_k^n)].
\]

Putting (5.39), (5.40) and (5.41) together gives (5.37). 

We are now ready to establish the integration by parts formula:

**Theorem 5.1.15** Suppose \( A, B : [0, \infty) \to \mathbb{R} \) are continuous and of locally bounded variation. Then, for each \( t \in [0, \infty) \),

\[
A(t)B(t) = A(0)B(0) + \int_0^t A(s)dB(s) + \int_0^t B(s)dA(s).
\]

Proof: Fix some \( t \in [0, \infty) \). In view of Theorem (5.1.14),

\[
\int_0^t A(s)dB(s) + \int_0^t B(s)dA(s)
\]

\[
= \lim_{n \to \infty} \left\{ \sum_{0 \leq k < \infty} A(t \wedge \tau_k^n)[B(t \wedge \tau_{k+1}^n) - B(t \wedge \tau_k^n)]
+ \sum_{0 \leq k < \infty} B(t \wedge \tau_k^n)[A(t \wedge \tau_{k+1}^n) - A(t \wedge \tau_k^n)]
+ \sum_{0 \leq k < \infty} [A(t \wedge \tau_k^n) - A(t \wedge \tau_{k+1}^n)][B(t \wedge \tau_{k+1}^n) - B(t \wedge \tau_k^n)] \right\}
\]

\[
= \lim_{n \to \infty} \left\{ \sum_{0 \leq k < \infty} [A(t \wedge \tau_k^n)B(t \wedge \tau_{k+1}^n) - A(t \wedge \tau_k^n)B(t \wedge \tau_k^n)]
+ \sum_{0 \leq k < \infty} [A(t \wedge \tau_k^n) - A(t \wedge \tau_{k+1}^n)][B(t \wedge \tau_{k+1}^n) - B(t \wedge \tau_k^n)] \right\}.
\]
Now
\( \sum_{0 \leq k < \infty} [A(t \land \tau^n_k)B(t \land \tau^n_{k+1}) - A(t \land \tau^n_k)B(t \land \tau^n_k)] = A(t)B(t) - A(0)B(0), \)
and clearly
\( \left| \sum_{0 \leq k < \infty} [A(t \land \tau^n_k) - A(t \land \tau^n_k)] [B(t \land \tau^n_{k+1}) - B(t \land \tau^n_k)] \right| \)
\( \leq \max_{0 \leq k < \infty} |A(t \land \tau^n_{k+1}) - A(t \land \tau^n_k)| \frac{\hat{B}(t)}{\hat{B}(t)}. \)

Since \( A \) is uniformly continuous on the closed and bounded interval \([0, t]\), we have
\( \lim_{n \to \infty} \max_{0 \leq k < \infty} |A(t \land \tau^n_{k+1}) - A(t \land \tau^n_k)| = 0, \)
while the fact that \( B \) is of locally bounded variation ensures that \( \hat{B}(t) < \infty \). It follows from (5.45) that
\( \lim_{n \to \infty} \left| \sum_{0 \leq k < \infty} [A(t \land \tau^n_k) - A(t \land \tau^n_k)] [B(t \land \tau^n_{k+1}) - B(t \land \tau^n_k)] \right| = 0. \)

Now (5.42) follows from (5.43), (5.44) and (5.46).

The final item in this section is an inequality which will play a crucial role in our later study of stochastic calculus:

**Lemma 5.1.16** Suppose \( A : [0, \infty) \to \mathbb{R} \) is continuous and of locally bounded variation, and the mappings \( B, C : [0, \infty) \to \mathbb{R} \) are continuous and non-decreasing, with
\( [A(t) - A(s)]^2 \leq [B(t) - B(s)][C(t) - C(s)], \quad \forall 0 \leq s < t < \infty. \)

Then
\( [\bar{A}(t) - \bar{A}(s)]^2 \leq [B(t) - B(s)][C(t) - C(s)], \quad \forall 0 \leq s < t < \infty, \)
and
\( [\mu_A(\Gamma)]^2 \leq [\mu_B(\Gamma)]\mu_C(\Gamma), \quad \forall \Gamma \in \mathcal{B}([0, \infty)). \)

**Proof:** Fix \( 0 \leq s < t < \infty \) and let \( s = t_0 < t_1 < \ldots t_n = t \) be a partition of \([s, t]\). Then, from (5.47) and the Cauchy-Schwarz inequality,
\( \sum_{k=0}^{n-1} |A(t_{k+1}) - A(t_k)| \)
\( \leq \sum_{k=0}^{n-1} ([B(t_{k+1}) - B(t_k)][C(t_{k+1}) - C(t_k)])^{1/2} \)
\( \leq \left\{ \sum_{k=0}^{n-1} [B(t_{k+1}) - B(t_k)] \right\}^{1/2} \left\{ \sum_{k=0}^{n-1} [C(t_{k+1}) - C(t_k)] \right\}^{1/2} \)
\( \leq [B(t) - B(s)]^{1/2}[C(t) - C(s)]^{1/2}. \)
Taking the supremum over all finite partitions of \([s,t]\) then gives
\[
(5.50) \quad V[A; s, t] \leq [B(t) - B(s)]^{1/2}[C(t) - C(s)]^{1/2}.
\]

Now, Proposition 4.7.9 implies
\[
(5.51) \quad V[A; s, t] = A(t) - A(s).
\]

Upon combining (5.50) and (5.51) we obtain (5.48).

It remains to prove (5.49). Fix some positive integer \(n\), and write \(B([0,n)) \overset{\Delta}{=} \{ \Gamma \cap [0,n) : \Gamma \in B([0,\infty)) \} \). We show first that (5.49) holds for all \(\Gamma \in B([0,n))\). Let \(\mathcal{A}\) denote the collection of all subsets of \([0,n)\) each of which is some finite union of disjoint intervals \([t,u)\) where \(0 \leq t \leq u \leq n\). Clearly \(\mathcal{A}\) is an algebra of sets over \([0,n)\), \(\mathcal{A} \subset B([0,n))\), and \(\sigma\{\mathcal{A}\} = B([0,n)) \). Now define
\[
(5.52) \quad \mathcal{M} \overset{\Delta}{=} \{ \Gamma \in B([0,n)) : [\mu_{\mathcal{A}}(\Gamma)]^{2} \leq \mu_{B}(\Gamma)\mu_{C}(\Gamma) \}.
\]

We observe that \(\mathcal{M}\) is a monotone class of sets in \([0,n)\). In fact, suppose that \(\{\Gamma_k, \ k = 1,2,\ldots\}\) is a sequence of members of \(\mathcal{M}\) with \(\Gamma_{k+1} \subset \Gamma_k\), \(\forall k = 1,2,\ldots\) Then we have
\[
(5.53) \quad [\mu_{\mathcal{A}}(\Gamma_k)]^{2} \leq \mu_{B}(\Gamma_k)\mu_{C}(\Gamma_k), \quad \forall k = 1,2,\ldots
\]

Since \(\mu_{\mathcal{A}}\) is a measure on \(B([0,n))\), and \(\mu_{\mathcal{A}}(\Gamma_1) \leq \mu_{\mathcal{A}}([0,n)) = A(n) - A(0) < \infty\), from Theorem 1.2.5(d) we see that \(\mu_{\mathcal{A}}(\lim_{k \to \infty} \Gamma_k) = \lim_{k \to \infty} \mu_{\mathcal{A}}(\Gamma_k)\). Similar limits hold with \(\mu_{B}\) and \(\mu_{C}\) in place of \(\mu_{\mathcal{A}}\), thus taking \(k \to \infty\) in (5.53) gives
\[
(5.54) \quad [\mu_{\mathcal{A}}(\lim_{k \to \infty} \Gamma_k)]^{2} \leq \mu_{B}(\lim_{k \to \infty} \Gamma_k)\mu_{C}(\lim_{k \to \infty} \Gamma_k).
\]

It follows that \(\lim_{k \to \infty} \Gamma_k\) is a member of \(\mathcal{M}\). An identical argument holds when \(\{\Gamma_k, \ k = 1,2,\ldots\}\) is an increasing sequence of sets in \(\mathcal{M}\), hence \(\mathcal{M}\) is a monotone class of sets in \([0,n)\). Now fix some \(\Gamma \in \mathcal{A}\). Then \(\Gamma\) can be written as
\[
\Gamma = \bigcup_{i=1}^{r} [t_i, u_i)
\]
for some pairwise disjoint sets \([t_i, u_i), \ i = 1,2,\ldots r\), with \(0 \leq t_i \leq u_i \leq n\). Then
\[
\mu_{\mathcal{A}}(\Gamma) = \sum_{i=1}^{r} \mu_{\mathcal{A}}([t_i, u_i)) = \sum_{i=1}^{r} [A(u_i) - A(t_i)]
\]
\[
\leq \sum_{i=1}^{r} [B(u_i) - B(t_i)]^{1/2}[C(u_i) - C(t_i)]^{1/2}
\]
\[
\leq \left[ \sum_{i=1}^{r} [B(u_i) - B(t_i)] \right]^{1/2} \left[ \sum_{i=1}^{r} [C(u_i) - C(t_i)] \right]^{1/2}
\]
\[
= [\mu_{B}(\Gamma)]^{1/2} [\mu_{C}(\Gamma)]^{1/2}.
\]

Here we have used (5.48) at the second line and the Cauchy-Schwarz inequality for discrete sums at the third line. By the arbitrary choice of \(\Gamma\) it follows that \(\mathcal{A} \subset \mathcal{M}\). Thus Theorem 1.5.9 says that
\( \sigma \{ A \} \subset M \), and since we have seen that \( \sigma \{ A \} = B([0, n)) \), we get \( B([0, n)) \subset M \subset B([0, n)) \) or \( M = B([0, n)) \). In view of (5.52), this set equality tells us that (5.49) holds for all \( \Gamma \in B([0, n)) \).

Next, fix an arbitrary \( \Gamma \in B([0, \infty)) \), and put \( \Gamma_n = \Gamma \cap [0, n), n = 1, 2, \ldots \) Clearly \( \Gamma_n \in B([0, n)) \) hence, by what we have just proved,

\[
(5.55) \quad [\mu_A(\Gamma_n)]^2 \leq \mu_B(\Gamma_n)\mu_C(\Gamma_n), \quad \forall \ n = 1, 2, \ldots
\]

Since \( \{ \Gamma_n, n = 1, 2, \ldots \} \) is an increasing sequence of sets in \( B([0, \infty)) \) with limit \( \Gamma \), Theorem 1.2.5(c) shows that

\[
\mu_A(\Gamma) = \lim_{n \to \infty} \mu_A(\Gamma_n).
\]

Similar limits holds with \( \mu_B \) and \( \mu_C \) in place of \( \mu_A \) hence taking \( n \to \infty \) in (5.55) gives (5.49).

**Theorem 5.1.17** Suppose that

(a) The mappings \( \Phi, \Psi : [0, \infty) \to \mathbb{R} \) are \( B([0, \infty)) \)-measurable;

(b) \( A : [0, \infty) \to \mathbb{R} \) is continuous and of locally bounded variation, and the mappings \( B, C : [0, \infty) \to \mathbb{R} \) are continuous and non-decreasing;

(c) \( [A(t) - A(s)]^2 \leq [B(t) - B(s)][C(t) - C(s)], \quad \forall \ 0 \leq s < t < \infty. \)

Then

\[
(5.56) \quad \left[ \int_0^\infty |\Phi(s)|\Psi(s)|dA(s)\right] \leq \left[ \int_0^\infty |\Phi(s)|^2dB(s) \right]^{1/2} \left[ \int_0^\infty |\Psi(s)|^2dC(s) \right]^{1/2}.
\]

Proof: We assume that the right hand side of (5.56) is finite, since the inequality holds trivially when the right hand side has the value \(+\infty\). First suppose that \( \Phi \) and \( \Psi \) are simple functions of the form

\[
(5.57) \quad \Phi(s) = \sum_{i=1}^m \alpha_i I_{\Gamma_i}(s) \quad \text{and} \quad \Psi(s) = \sum_{j=1}^n \beta_j I_{\Delta_j}(s),
\]

where \( \Gamma_i \) and \( \Delta_j \) are sets in \( B([0, \infty)) \) and \( \alpha_i, \beta_j \in \mathbb{R} \) are constants. By appropriately intersecting the sets \( \Gamma_i \) with \( \Delta_j \) and fine-tuning the choice of the constants \( \alpha_i \) and \( \beta_j \), we may suppose without loss of generality that \( m = n, \Gamma_i = \Delta_i, \) and the sets \( \{ \Gamma_1, \Gamma_2, \ldots, \Gamma_n \} \) are pairwise disjoint. Then,

\[
(5.58) \quad \int_0^\infty |\Phi(s)||\Psi(s)|dA(s) = \sum_{i=1}^n |\alpha_i\beta_i|\mu_A(\Gamma_i)
\]

\[
\leq \sum_{i=1}^n |\alpha_i\beta_i|\left[ \mu_B(\Gamma_i) \right]^{1/2} \left[ \mu_C(\Gamma_i) \right]^{1/2}
\]

\[
\leq \left[ \sum_{i=1}^n \alpha_i^2\mu_B(\Gamma_i) \right]^{1/2} \left[ \sum_{i=1}^n \beta_i^2\mu_C(\Gamma_i) \right]^{1/2}
\]

\[
= \left[ \int_0^\infty |\Phi(s)|^2dB(s) \right]^{1/2} \left[ \int_0^\infty |\Psi(s)|^2dC(s) \right]^{1/2},
\]

where we have used (5.49) at the second line and the Cauchy-Schwarz inequality for discrete sums at the third line of (5.58). We have therefore established (5.56) when \( \Phi \) and \( \Psi \) are \( B([0, \infty)) \)-measurable

195
simple functions, and it remains to consider the general case. Clearly there is no loss of generality if we suppose that $\Phi$ and $\Psi$ are non-negative. Then, by Proposition 1.2.9, there are sequences $\{\Phi_n, n = 1, 2, \ldots\}$ and $\{\Psi_n, n = 1, 2, \ldots\}$ of measurable simple functions such that $\{\Phi_n(s), n = 1, 2, \ldots\}$ and $\{\Psi_n(s), n = 1, 2, \ldots\}$ increase monotonically to $\Phi(s)$ and $\Psi(s)$ as $n \to \infty$ for each $s \in [0, \infty)$. Now, by what has already been proved, we have

$$\left[ \int_0^\infty |\Phi_n(s)\Psi_n(s)|d\mathcal{A}(s) \right] \leq \left[ \int_0^\infty |\Phi_n(s)|^2dB(s) \right]^{1/2} \left[ \int_0^\infty |\Psi_n(s)|^2dC(s) \right]^{1/2}$$

for all $n = 1, 2, \ldots$ hence taking $n \to \infty$ and using the monotone convergence theorem 1.2.12 gives the desired result. $\blacksquare$

### 5.2 Random Lebesgue-Stieltjes Integration

In this section our goal is to look at Lebesgue-Stieltjes integrals

$$(5.59) \quad \int_0^t \Phi(s, \omega)dA(s, \omega)$$

where $\{\Phi(t); t \in [0, \infty)\}$ and $\{A(t); t \in [0, \infty)\}$ are processes on the probability space $(\Omega, \mathcal{F}, P)$, and the mapping $t \to A(t, \omega)$ is continuous and of locally bounded variation on $[0, \infty)$ for each $\omega \in \Omega$. Fixing $t \in [0, \infty]$, and supposing that the integral in (5.59) exists for each and every $\omega$, we clearly have a function defined on $\Omega$, and an issue of obvious importance concerns its measurability:

**Proposition 5.2.1** Suppose $\{A(t); t \in [0, \infty)\}$ is a process on the probability space $(\Omega, \mathcal{F}, P)$ with $A(\cdot, \omega)$ being continuous and non-decreasing on $[0, \infty)$ for each $\omega \in \Omega$. If, for some fixed $t \in [0, \infty)$, the mapping $H : [0, t] \otimes \Omega \to [0, \infty]$ is $\mathcal{B}([0, t]) \otimes \mathcal{F}$-measurable, then the mapping

$$(5.60) \quad \omega \to U(H)(\omega) \triangleq \int_0^t H(s, \omega)dA(s, \omega) : \Omega \to [0, \infty]$$

is $\mathcal{F}$-measurable.

Proof: The result is established by an argument which is quite similar to that used for proving the measurability assertion of the Tonelli Theorem 1.2.40, but here one has the additional feature that the measure with respect to which one integrates (the Lebesgue-Stieltjes measure generated by $A(\cdot, \omega)$) also depends on $\omega$. Let $M^+(\mathcal{B}([0, t]) \otimes \mathcal{F})$ denote the set of all $[0, \infty]$-valued $\mathcal{B}([0, t]) \otimes \mathcal{F}$-measurable functions on $[0, t] \otimes \Omega$, and let $\mathcal{H}$ be the set of all functions $H : [0, t] \otimes \Omega \to [0, \infty]$ in $M^+(\mathcal{B}([0, t]) \otimes \mathcal{F})$ with the further property that the mapping in (5.60) is $\mathcal{F}$-measurable. It is easily verified that $\mathcal{H}$ is a $\mathcal{L}$-class of functions over $[0, t] \otimes \Omega$ (see Definition 1.5.10). Indeed, when $H \equiv 1$ then

$$U(H)(\omega) = A(t, \omega) - A(0, \omega)$$
which is $\mathcal{F}$-measurable, whence $1 \in \mathcal{H}$. As for Definition 1.5.10(ii), fix $H_1, H_2 \in \mathcal{H}$ such that $0 \leq H_1(s, \omega) \leq H_2(s, \omega) \leq B < \infty$, $\forall (s, \omega) \in [0, t] \otimes \Omega$. Since $t$ is finite and $H_1, H_2,$ are uniformly bounded by the constant $B < \infty$, one sees that the mappings $s \to H_i(s, \omega) : [0, t] \to [0, \infty), i = 1, 2,$ are integrable with respect to the Lebesgue-Stieltjes measure on $\mathcal{B}[0, t]$ generated by $A(\cdot, \omega)$, for each $\omega \in \Omega$, thus Theorem 1.2.18 says that

$$U(H_2 - H_1)(\omega) = U(H_2)(\omega) - U(H_1)(\omega), \quad \forall \omega \in \Omega.$$  

Since $H_1, H_2 \in \mathcal{H}$, the functions of $\omega$ on the right hand side are $\mathcal{F}$-measurable, thus Proposition 1.2.7 ensures that $\omega \to U(H_2 - H_1)(\omega)$ is $\mathcal{F}$-measurable, whence $H_2 - H_1 \in \mathcal{H}$, and Definition 1.5.10(ii) is established for $\mathcal{H}$. A similar argument shows that $\mathcal{H}$ verifies Definition 1.5.10(iii), and Definition 1.5.10(iv) is checked by an obvious application of Theorem 1.2.12 and Proposition 1.2.6. We see that $\mathcal{H}$ is a $\mathcal{L}$-class of functions over $[0, t] \otimes \Omega$. Now clearly $\mathcal{B}([0, t]) = \sigma \{C\}$ for $C \triangleq \{[u, v] : 0 \leq u \leq v \leq t\}$ hence, if $\mathcal{D} \triangleq \{[u, v] \otimes \Gamma : 0 \leq u \leq v \leq t \text{ and } \Gamma \in \mathcal{F}\}$, then $\mathcal{B}([0, t]) \otimes \mathcal{F} = \sigma \{\mathcal{D}\}$, and it is of course clear that $\mathcal{D}$ is a $\pi$-class of sets over $[0, t] \otimes \Omega$. Finally, if $H(\cdot, \cdot)$ is the indicator function of some member of $\mathcal{D}$ with the form $[u, v] \otimes \Gamma$, for $0 \leq u \leq v \leq t$, $\Gamma \in \mathcal{F}$, then the right hand side of (5.60) is $I_{\Gamma}(\omega)[A(v, \omega) - A(u, \omega)]$, which is clearly a $\mathcal{F}$-measurable function of $\omega$. It follows that $I_D \in \mathcal{H}$ for each $D \in \mathcal{D}$. By Theorem 1.5.11 we get $M^+(\mathcal{B}([0, t]) \otimes \mathcal{F}_t) = \mathcal{H}$, as required. 

**Proposition 5.2.2** Suppose that $\{A(t); t \in [0, \infty)\}$ and $\{\Phi(t); t \in [0, \infty)\}$ are processes on $(\Omega, \mathcal{F}, P)$ such that

(a) $A(\cdot, \omega)$ is continuous and non-decreasing on $[0, \infty)$, $\forall \omega \in \Omega$;

(b) $\{\Phi(t); t \in [0, \infty)\}$ is jointly measurable and $\Phi(t, \omega) \in [0, \infty], \forall \omega \in \Omega, \forall t \in [0, \infty)$.

Then the mapping

$$\omega \to \int_0^t \Phi(s, \omega)dA(s, \omega) : \Omega \to [0, \infty]$$  

is $\mathcal{F}$-measurable for each $t \in [0, \infty]$.

Proof: The result follows at once from Proposition 5.2.1 when $t \in [0, \infty)$. In the case where $t = +\infty$, fix some sequence $\{t_n, \ n = 1, 2, \ldots\} \subset [0, \infty)$ with $t_n < t_{n+1}$ and $\lim_{n \to \infty} t_n = +\infty$. Then Theorem 1.2.12 ensures that $\int_0^\infty \Phi(s, \omega)dA(s, \omega) = \lim_{n \to \infty} \int_0^{t_n} \Phi(s, \omega)dA(s, \omega), \forall \omega \in \Omega$. In view of Proposition 5.2.1, this is a limit of $\mathcal{F}$-measurable functions, and the result follows by Proposition 1.2.6.

Next, we introduce a filtration in $(\Omega, \mathcal{F}, P)$ with respect to which $\{A(t); t \in [0, \infty)\}$ and $\{\Phi(t); t \in [0, \infty)\}$ are adapted, and want to secure conditions which ensure that the integral in (5.59) is also adapted to this filtration. This matter is effectively settled by the next result, which closely resembles Proposition 5.2.1:

**Proposition 5.2.3** Suppose $\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}$ is an adapted process on the probability space $(\Omega, \mathcal{F}, P)$ with $A(\cdot, \omega)$ being continuous and non-decreasing on $[0, \infty)$ for each $\omega \in \Omega$. If, for some fixed
$t \in [0, \infty)$, the mapping $H : [0, t] \otimes \Omega \rightarrow [0, \infty]$ is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$-measurable, then

\begin{equation}
\omega \mapsto \int_0^t H(s, \omega) dA(s, \omega) : \Omega \rightarrow [0, \infty]
\end{equation}

is $\mathcal{F}_t$-measurable.

Proof: The proof is identical to the proof of Proposition 5.2.2, with $\mathcal{F}_t$ everywhere replacing $\mathcal{F}$. □

The next corollary, which resembles Proposition 5.2.2, is an immediate consequence of Proposition 5.2.3 and the definition of progressive measurability:

**Corollary 5.2.4** Suppose that $\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}$ and $\{((\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ are progressively measurable processes on $(\Omega, \mathcal{F}, P)$ such that

(a) $A(\cdot, \omega)$ is continuous and non-decreasing on $[0, \infty)$, $\forall \omega \in \Omega$;

(b) $\Phi(t, \omega) \in [0, \infty]$, $\forall \omega \in \Omega, \forall t \in [0, \infty)$.

Then the mapping

\begin{equation}
\omega \mapsto \int_0^t \Phi(s, \omega) dA(s, \omega) : \Omega \rightarrow [0, \infty]
\end{equation}

is $\mathcal{F}_t$-measurable for each $t \in [0, \infty)$ (see Remark 3.1.15 for the definition of $\mathcal{F}_\infty$.)

**Remark 5.2.5** Our next task is to generalize Corollary 5.2.4 to the case where $\Phi(t)$ is $\mathbb{R}$-valued rather than non-negative, and the sample-paths of $\{A(t), t \in [0, \infty)\}$ are of locally bounded variation rather than just non-decreasing. We shall proceed in a manner which is very reminiscent of Step III in § 5.1. For these as well as later developments it is useful to have the following notation: Let $\{\mathcal{F}_t, t \in [0, \infty)\}$ be a filtration in the probability space $(\Omega, \mathcal{F}, P)$, let $\mathbf{FV}^c\{\mathcal{F}_t\}$ denote the set of all $\mathbb{R}$-valued processes $\{A(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that $\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}$ is adapted and the mappings $t \mapsto A(t, \omega)$ are continuous and of locally bounded variation on $[0, \infty)$ for each $\omega \in \Omega$, and let $\mathbf{FV}^{c,0}\{\mathcal{F}_t\}$ denote the set of all members of $\mathbf{FV}^c\{\mathcal{F}_t\}$ which are null at the origin. We shall use the notation $A \in \mathbf{FV}^c\{\mathcal{F}_t\}$ to indicate that $\{A(t); t \in [0, \infty)\}$ is an element of $\mathbf{FV}^c\{\mathcal{F}_t\}$, and write $A \in \mathbf{FV}^{c,0}\{\mathcal{F}_t\}$ with an obviously similar interpretation. It is at once clear that $\mathbf{FV}^c\{\mathcal{F}_t\}$ and $\mathbf{FV}^{c,0}\{\mathcal{F}_t\}$ are real vector spaces with the obvious notions of vector addition and scalar multiplication. Furthermore, if $T$ is an $\{\mathcal{F}_t\}$-stopping time and $A \in \mathbf{FV}^c\{\mathcal{F}_t\}$, then Proposition 3.3.14 ensures that $A_T \in \mathbf{FV}^c\{\mathcal{F}_t\}$.

Suppose $A \in \mathbf{FV}^c\{\mathcal{F}_t\}$. If we denote the total variation, positive variation, and negative variation functions of the mapping $A(\cdot, \omega)$ by $A(\cdot, \omega)$, $A^+ (\cdot, \omega)$ and $A^- (\cdot, \omega)$ respectively, then the following result is an immediate consequence of Lemma 5.1.6 and Proposition 1.2.6:

**Lemma 5.2.6** Suppose $A \in \mathbf{FV}^c\{\mathcal{F}_t\}$. Then $\{(A(t), \mathcal{F}_t); t \in [0, \infty)\}$, $\{(A^+(t), \mathcal{F}_t); t \in [0, \infty)\}$ and $\{(A^-(t), \mathcal{F}_t); t \in [0, \infty)\}$ are adapted processes. Moreover, the mappings $A(\cdot, \omega)$, $A^+(\cdot, \omega)$ and $A^- (\cdot, \omega)$ are continuous and non-decreasing on $[0, \infty)$, take their values in $[0, \infty)$, and are null at the origin for each $\omega \in \Omega$.  

198
Remark 5.2.7 Suppose \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a progressively measurable process on \((\Omega, \mathcal{F}, P)\) and \( A \in \text{FV}^c\{\mathcal{F}_t\} \) is such that, for each \( \omega \in \Omega \), we have

\[
\int_0^t |\Phi(s, \omega)| d\tilde{A}(s, \omega) < \infty, \quad \forall \ t \in [0, \infty).
\]

We see from Step III of §5.1 that the indefinite Lebesgue-Stieltjes integral of \( \Phi(\cdot, \omega) \) with respect to \( A(\cdot, \omega) \) is defined and \( \mathbb{R} \)-valued for each \( \omega \in \Omega \), with

\[
\int_0^t \Phi(s, \omega) dA(s, \omega) \triangleq \left[ \int_0^t \Phi_+(s, \omega) d\tilde{A}_+(s, \omega) - \int_0^t \Phi_-(s, \omega) d\tilde{A}_+(s, \omega) \right] - \left[ \int_0^t \Phi_+(s, \omega) d\tilde{A}_-(s) - \int_0^t \Phi_-(s, \omega) d\tilde{A}_-(s, \omega) \right], \quad \forall \ t \in [0, \infty).
\]

Lemma 5.1.12 tells us that, for each \( \omega \in \Omega \), the function of the variable \( t \in [0, \infty) \) in (5.65) is continuous and of locally bounded variation on \([0, \infty)\), and gives its total, positive and negative variation functions as well (see (5.28) etc.). Trivially we observe that \( \{(\Phi_+(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) and \( \{(\Phi_-(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) are progressively measurable processes, thus we can use Corollary 5.2.4 on each of the right-hand terms of (5.65) to see that, for each fixed \( t \in [0, \infty) \), the function of \( \omega \) on the left-hand-side of (5.65) is \( F_t \)-measurable. Subject to (5.64) we have thus obtained a process

\[
\left\{ \left( \int_0^t \Phi(s) dA(s), \mathcal{F}_t \right); \ t \in [0, \infty) \right\},
\]

which is an element of \( \text{FV}^{c,0}\{\mathcal{F}_t\} \), by “path-wise” Lebesgue-Stieltjes integration.

Remark 5.2.8 In the special case where \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous adapted process and \( A \in \text{FV}^c\{\mathcal{F}_t\} \), we see from Remark 5.1.9 that (5.64) holds for each \( \omega \in \Omega \) and therefore the process (5.66) is defined and a member of \( \text{FV}^{c,0}\{\mathcal{F}_t\} \).

Remark 5.2.9 In our later study of stochastic integration we shall typically have to deal with a progressively measurable process \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) and a process \( A \in \text{FV}^c\{\mathcal{F}_t\} \) on \((\Omega, \mathcal{F}, P)\), for which it is not generally true that (5.64) holds for each and every \( \omega \in \Omega \), but where we have the somewhat less precise information that

\[
P \left[ \int_0^t |\Phi(s)| d\tilde{A}(s) < \infty \right] = 1, \quad \forall \ t \in [0, \infty).
\]

Again, we would like to write down the integral given by (5.65), but now we must be careful to avoid the undefined combination \( \infty - \infty \) possibly occurring on the set of \( P \)-measure zero allowed by (5.67). To this end, put

\[
N_t \overset{\Delta}{=} \{ \omega : \int_0^t |\Phi(s, \omega)| d\tilde{A}(s, \omega) = +\infty \}, \quad \forall \ t \in [0, \infty),
\]
observe from Corollary 5.2.4 that \( N_t \in F_t \), and furthermore, by (5.67), that \( P(N_t) = 0, \forall t \in [0, \infty) \).

Now let
\[
N \triangleq \bigcup_{1 \leq n < \infty} N_n
\]
(the set union being taken over all positive integers \( n \)) so that \( N \in F_\infty \) and \( P(N) = 0 \). From (5.68) we have \( N_{t_1} \subset N_{t_2}, \forall 0 \leq t_1 < t_2 < \infty \), thus \( N_t \subset N, \forall t \in [0, \infty) \), and hence (5.64) holds for each \( \omega \not\in N \). In particular we see that \( \Phi(\cdot, \omega) \) is integrable with respect to \( A(\cdot, \omega) \) and
\[
\int_0^t \Phi(s, \omega) dA(s, \omega)
\]
is \( \mathbb{IR} \)-valued and given by the right-hand-side of (5.65) for each \( \omega \not\in N \). When \( \omega \in N \) we shall simply put
\[
\int_0^t \Phi(s, \omega) dA(s, \omega) \triangleq 0, \forall t \in [0, \infty).
\]
Thus, for each \( t \in [0, \infty) \), we have a well-defined mapping
\[
\omega \rightarrow \int_0^t \Phi(s, \omega) dA(s, \omega) : \Omega \rightarrow \mathbb{IR}
\]
which (since \( N \in F_\infty \subset F_t \)) is easily seen to be \( F_t \)-measurable. There are many applications where we want this mapping to be \( F_t \)-measurable, for each \( t \in [0, \infty) \), (for example, to ensure that the process in (5.66) is adapted), and for this stronger statement to hold we must impose some conditions on the underlying filtration \( \{F_t, t \in [0, \infty)\} \). If we suppose that \( F_0 \) includes all \( P \)-null events in \( F \) then, since \( P(N) = 0 \), we have \( N \in F_0 \subset F_t, \forall t \in [0, \infty) \), and it now follows that the mapping in (5.70) is \( F_t \)-measurable for each \( t \in [0, \infty) \).

**Remark 5.2.10** To conclude, from now on we shall write down the process (5.66) only when
(a) \( A \in FV^c_0\{F_t\} \);
(b) \( \{(\Phi(t), F_t); t \in [0, \infty)\} \) is progressively measurable;
(c) (5.67) holds and the filtration \( \{F_t, t \in [0, \infty)\} \) is such that \( F_0 \) includes all \( P \)-null events in \( F \).

Under these conditions the process in (5.66) is properly defined and is a member of \( FV^c_0\{F_t\} \). In many applications we will in fact be able to verify that
\[
E\left[ \int_0^t |\Phi(s)| d\tilde{A}(s) \right] < \infty, \quad \forall t \in [0, \infty),
\]
or even that
\[
E\left[ \int_0^\infty |\Phi(s)| d\tilde{A}(s) \right] < \infty,
\]
either of which implies (5.67).

**Remark 5.2.11** Suppose that \( \{A(t); t \in [0, \infty)\} \) and \( \{\Phi(t); t \in [0, \infty)\} \) are processes on \( (\Omega, \mathcal{F}, P) \) such that \( A(\cdot, \omega) \) is continuous and of locally bounded variation on \([0, \infty)\) for each \( \omega \in \Omega \), and \( \{\Phi(t); t \in [0, \infty)\} \) is jointly measurable with
\[
E\left[ \int_0^\infty |\Phi(s)| d\tilde{A}(s) \right] < \infty
\]
(observe that Proposition 5.2.2 ensures the function of $\omega$ within the square brackets in (5.71) is $\mathcal{F}$-measurable). Put
\[ N_\infty \triangleq \{ \omega : \int_0^\infty |\Phi(s, \omega)|d\tilde{A}(s, \omega) = \infty \}. \]

Then (5.71) ensures that $P(N_\infty) = 0$. For each $\omega \not\in N_\infty$ we follow Step III of § 5.1 and define
\[
\tag{5.72}
\int_0^\infty \Phi(s, \omega)dA(s, \omega) \triangleq \left[ \int_0^\infty \Phi_+(s, \omega)d\tilde{A}_+(s, \omega) - \int_0^\infty \Phi_-(s, \omega)d\tilde{A}_+(s, \omega) \right] \\
- \left[ \int_0^\infty \Phi_+(s, \omega)d\tilde{A}_-(s, \omega) - \int_0^\infty \Phi_-(s, \omega)d\tilde{A}_-(s, \omega) \right],
\]
and, for $\omega \in N_\infty$, we put
\[
\tag{5.73}
\int_0^\infty \Phi(s, \omega)dA(s, \omega) \triangleq 0.
\]

By Proposition 5.2.2 each of the four mappings
\[
\omega \to \int_0^\infty \Phi_{\pm}d\tilde{A}_{\pm}(s, \omega) : \Omega \to [0, \infty]
\]
is $\mathcal{F}$-measurable, and since $N_\infty \in \mathcal{F}$, one sees that $\omega \to \int_0^\infty \Phi(s, \omega)dA(s, \omega)$ is an $\mathcal{F}$-measurable $\mathbb{R}$-valued mapping on $\Omega$. Moreover, (5.71) ensures that this mapping is integrable with
\[
E\left( \int_0^\infty \Phi(s)dA(s) \right) = \left[ E\left( \int_0^\infty \Phi_+(s)d\tilde{A}_+(s) \right) - E\left( \int_0^\infty \Phi_-(s)d\tilde{A}_+(s) \right) \right] \\
- \left[ E\left( \int_0^\infty \Phi_+(s)d\tilde{A}_-(s) \right) - E\left( \int_0^\infty \Phi_-(s)d\tilde{A}_-(s) \right) \right].
\]

### 5.3 Stochastic Integration

The random Lebesgue-Stieltjes integrals defined in the preceding section may be regarded as “junior grade” stochastic integrals, and thus constitute a first step in our development of a systematic stochastic calculus. It turns out, however, that stochastic calculus acquires real power and a broad range of applications only when we are able to define and manipulate integrals of the form
\[
\tag{5.74}
\int_0^t \Phi(s) \, dX(s),
\]
in which $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous local martingale. Assuming that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable, one might imagine that it is possible to proceed by analogy with Section 5.2, and construct this integral in a Lebesgue-Stieltjes sense individually for each $\omega$. Here, however, we encounter an obvious difficulty: in order for this plan to work, we need the mappings $t \to X(t, \omega)$ to be of locally bounded variation on $[0, \infty)$, but Proposition 4.7.13 says that this happens only when $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is “trivial” in the sense that its sample paths are constant a.s. Since the only interesting continuous local martingales are those whose sample paths are of unbounded
variation, and there appears to be no natural extension of the ideas of Section 5.2 which allows us to define (5.74) individually for each \( \omega \) when this is the case, it seems that an essentially new approach to the whole issue of integration is needed. Such a new approach was pioneered by Itô\cite{15}, \cite{16}, and subsequently greatly expanded in scope by Kunita and Watanabe \cite{21}. As we shall see, the essential idea involved in this approach is to exploit the martingale structure of \( \{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) to define (5.74) \emph{simultaneously} for \( P \)-almost all \( \omega \) rather than “pathwise” for individually chosen \( \omega \).

Our goal in this section is to present the definition and basic properties of stochastic integrals with respect to continuous local martingales. Before undertaking this task, it is perhaps worthwhile to recall the general manner in which we formulated another basic idea of probability theory, namely conditional expectation. In Section 1.4.5 we defined the notion of conditional expectation on the basis of Theorem 1.4.13, which establishes the existence and uniqueness of a random variable having the properties that we require in any meaningful concept of conditional expectation. Our approach to stochastic integration will be in much the same spirit, namely we shall first establish a theorem (see Theorem 5.3.14) which asserts the existence and uniqueness of a certain stochastic process having the sort of properties that one requires of a useful concept of stochastic integration, and then we shall call this desirable process a stochastic integral. In order to establish this theorem, we shall make extensive use of quadratic variation processes of continuous martingales and random Lebesgue-Stieltjes integrals. We recall that, in order to develop most of those notions, we needed to assume our filtration \( \{\mathcal{F}_t, \ t \in [0, \infty)\} \) was such that \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \). Accordingly, for the remainder of these notes, we shall adopt the following:

**Condition 5.3.1** \( \{\mathcal{F}_t, \ t \in [0, \infty)\} \) is a given filtration in a probability space \( (\Omega, \mathcal{F}, P) \) such that \( \mathcal{F}_0 \) includes all events in \( \mathcal{N} \triangleq \{N \in \mathcal{F} : P(N) = 0\} \).

We intend to define stochastic integrals of the form (5.74) when \( \{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is a continuous local martingale and \( \{((\Phi(t), \mathcal{F}_t); \ t \in [0, \infty)) \) is a stochastic “integrand”, the basic properties of which will shortly be made clear. It turns out that the problems associated with formulating this definition are most easily overcome when we impose more restrictive boundedness conditions on \( \{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) and suppose that it is a continuous \( L^2 \)-bounded martingale (that is, \( X \in \mathcal{M}^{2,b}({\mathcal{F}_t}, P) \), in the notation of Remark 4.1.6(c)). Once we have defined stochastic integrals in this special case we shall then easily be able to generalize the definition to include the case of \( X \in \mathcal{M}^{2,b}_\text{loc}({\mathcal{F}_t}, P) \) by introducing stopping times and systematically using the technique of localization based upon Corollary 4.5.8. The advantage of first restricting ourselves to the case of \( X \in \mathcal{M}^{2,b}_\text{loc}({\mathcal{F}_t}, P) \) lies in the fact that we can easily define a Hilbert space structure in \( \mathcal{M}^{2,b}_\text{loc}({\mathcal{F}_t}, P) \), and this in turn makes it possible to use the Riesz Representation Theorem 1.1.2 as the main device for establishing the afore-mentioned existence and uniqueness Theorem 5.3.14 as the main device for establishing the afore-mentioned existence and uniqueness Theorem 5.3.14 upon which the definition of stochastic integration is based. Our next goal is to define this Hilbert space structure.
Remark 5.3.2 Recall from Remark 4.1.7 that $M_{2,b}^c(\{\mathcal{F}_t\}, P)$ is a real vector space when vector addition and scalar multiplication are defined in the usual pointwise sense. For each $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$ define
\begin{equation}
\|X\|_M \triangleq E^{1/2}[|X(\infty)|^2],
\end{equation}
where $X(\infty)$ is the last element of $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, whose existence, a.s. uniqueness and basic properties are guaranteed by Theorem 4.4.12. It is readily verified that the mapping in (5.75) defines a pseudo-norm on the real vector space $M_{2,b}^c(\{\mathcal{F}_t\}, P)$ (see Section 1.1). In order for this mapping to constitute a norm, we must adjust our interpretation of the vector space structure of $M_{2,b}^c(\{\mathcal{F}_t\}, P)$, much as we had to do in Section 1.2.5. From (4.55) of Theorem 4.4.12 we see that, when $\|X - Y\|_M = 0$ for $X, Y \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, then $\{X_t; t \in [0, \infty)\}$ and $\{Y_t; t \in [0, \infty)\}$ are indistinguishable processes. Hence, if we relax the sense of equality in $M_{2,b}^c(\{\mathcal{F}_t\}, P)$ and declare any two indistinguishable members of this space to be equal, then $M_{2,b}^c(\{\mathcal{F}_t\}, P)$ becomes a normed vector space with norm defined by (5.75).

Theorem 5.3.3 Suppose Condition 5.3.1. Then the normed vector space $M_{2,b}^c(\{\mathcal{F}_t\}, P)$, with norm defined by (5.75), is a Banach space.

Proof: Fix a Cauchy sequence $\{X_n, n = 1, 2, \ldots\}$ in $M_{2,b}^c(\{\mathcal{F}_t\}, P)$. Then it follows from (5.75) that $\{X_n(\infty), n = 1, 2, \ldots\}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, P)$ which $L^2$-converges to some limit $Z \in L^2(\Omega, \mathcal{F}, P)$ (see Theorem 1.2.27). Now define $X(t) \overset{\Delta}{=} E[Z | \mathcal{F}_t], \forall t \in [0, \infty)$. Proposition 1.4.21(b) shows that $\{X_n(t), n = 1, 2, \ldots\}$ converges in $L^2$ to $X(t)$ for each $t \in [0, \infty)$, and then the result follows from Proposition 4.3.4. \hfill \Box

Remark 5.3.4 It is now easy to introduce a Hilbert space structure into $M_{2,b}^c(\{\mathcal{F}_t\}, P)$. Put
\begin{equation}
(X,Y)_M \overset{\Delta}{=} E[X(\infty)Y(\infty)],
\end{equation}
for $X, Y \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$. Then it follows at once that (5.76) defines an inner product on the vector space $M_{2,b}^c(\{\mathcal{F}_t\}, P)$, and in view of Theorem 5.3.3 we have:

Corollary 5.3.5 Suppose Condition 5.3.1. Then the inner product space $M_{2,b}^c(\{\mathcal{F}_t\}, P)$, with inner product defined by (5.76), is a Hilbert space.

We observe that $M_{2,b}^{c,0}(\{\mathcal{F}_t\}, P)$ is a closed subspace of $M_{2,b}^c(\{\mathcal{F}_t\}, P)$ and is therefore also a Hilbert space with inner product defined by (5.76).

Having established a Hilbert space structure in the space of martingales $M_{2,b}^c(\{\mathcal{F}_t\}, P)$, our next task is to formulate an appropriate class of integrands $\{\Phi(t); t \in [0, \infty)\}$ for the stochastic integral. For this purpose we need the following simple result:
Proposition 5.3.6 Suppose Condition 5.3.1, and fix some \( X \in M_{2,0}^T(\mathcal{F}_t, P) \).

(a) For each \( \Gamma \in \mathcal{B}([0, \infty)) \otimes \mathcal{F} \), define

\[
\nu_X(\Gamma) \triangleq E \left[ \int_0^\infty I_\Gamma(s, \omega) \, d[X](s, \omega) \right].
\]

Then \( \nu_X \) is a finite measure on the measurable space \( ([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}) \);

(b) If \( \Phi : [0, \infty) \otimes \Omega \to [0, \infty] \) is \( \mathcal{B}([0, \infty)) \otimes \mathcal{F} \)-measurable, then

\[
\int_{[0, \infty) \otimes \Omega} \Phi \, d\nu_X = E \left[ \int_0^\infty \Phi(s) \, d[X](s) \right];
\]

(c) If \( \Phi, \Psi \) are elements of \( L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X) \) then

\[
E \left[ \int_0^\infty |\Phi(s)\Psi(s)| \, d[X](s) \right] < \infty,
\]

and

\[
\int_{[0, \infty) \otimes \Omega} \Phi \Psi \, d\nu_X = E \left[ \int_0^\infty \Phi(s)\Psi(s)d[X](s) \right].
\]

Proof: (a) Proposition 5.2.2 ensures that the expectation in (5.77) exists. Fix some sequence of sets \( \{\Gamma_n, n = 1, 2, \ldots\} \) in \( \mathcal{B}([0, \infty)) \otimes \mathcal{F} \), with \( \bigcap_{n=1}^\infty \Gamma_n \) if and only if \( n \neq m \). From the Monotone Convergence Theorem 1.2.12 one sees that

\[
\lim_{n \to \infty} \int_{[0, \infty) \otimes \Omega} I_{\bigcup_{k=1}^n \Gamma_k}(s, \omega) \, d[X](s, \omega) = \int_0^\infty I_{\bigcup_{1 \leq k < \infty} \Gamma_k}(s, \omega) \, d[X](s, \omega), \quad \forall \omega \in \Omega,
\]

where the convergence on the left hand side is monotonic non-decreasing. In view of this, together with a second application of Theorem 1.2.12, one sees that \( \nu_X(\cdot) \) is a measure on \( \mathcal{B}([0, \infty)) \otimes \mathcal{F} \). To see that it is a finite measure, observe from Proposition 4.7.33 that

\[
\nu_X([0, \infty) \otimes \Omega) = E[[X](\infty)] < \infty.
\]

(b) In view of (a), clearly (5.78) holds when \( \Phi \) is the indicator function of a set \( \Gamma \in \mathcal{B}([0, \infty)) \otimes \mathcal{F} \), hence also holds when \( \Phi \) is a non-negative \( \mathcal{B}([0, \infty)) \otimes \mathcal{F} \)-measurable simple function. The general case follows by a standard application of the Monotone Convergence Theorem 1.2.12 and Proposition 1.2.9.

(c) In view of (b) and the Cauchy-Schwarz inequality,

\[
E \left[ \int_0^\infty |\Phi(s)\Psi(s)| \, d[X](s) \right] = \int_{[0, \infty) \otimes \Omega} |\Phi\Psi| \, d\nu_X
\]

\[
\leq \left\{ \int_{[0, \infty) \otimes \Omega} |\Phi|^2 \, d\nu_X \right\}^{1/2} \left\{ \int_{[0, \infty) \otimes \Omega} |\Psi|^2 \, d\nu_X \right\}^{1/2} < \infty.
\]

As for (5.80), write

\[
\Phi \Psi = \Phi_+ \Psi_+ + \Phi_+ \Psi_- + \Phi_- \Psi_+ + \Phi_- \Psi_-,
\]
and then apply (5.78) to each of the terms on the right hand side.

Clearly, any measurable mapping $\Phi : ([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}) \to \mathbb{R}$ is synonymous with a \textit{jointly measurable} $\mathbb{R}$-valued process $\{\Phi(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ (recall Definition 3.1.12). Thus, for each $X \in \mathcal{M}_{2, b}^c(\{\mathcal{F}_t\}, P)$, we can regard the Lebesgue space $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ as the real vector space of jointly measurable processes $\{\Phi(t); t \in [0, \infty)\}$ such that

$$\int_{[0, \infty) \otimes \Omega} |\Phi|^2 \, d\nu_X < \infty.$$

In view of Section 1.2.5 and Proposition 5.3.6(b), this real vector space is a normed vector space with norm defined by

$$\|\Phi\|_X \triangleq \left\{ E \left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] \right\}^{1/2},$$

provided we understand equality in the following sense: two jointly measurable processes $\{\Phi(t); t \in [0, \infty)\}$ and $\{\Psi(t); t \in [0, \infty)\}$ in $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ are defined to be \textit{equal} when $\|\Phi - \Psi\|_X = 0$, that is, when

$$\nu_X\{ (t, \omega) \in [0, \infty) \otimes \Omega : \Phi(t, \omega) \neq \Psi(t, \omega) \} = 0.$$

It then follows from Theorem 1.2.27 that $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ is a Banach space. In fact, the norm in (5.81) clearly arises from the inner product on $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ defined by

$$(\Phi, \Psi)_X \triangleq E \left[ \int_0^\infty \Phi(s)\Psi(s)\, d[X](s) \right],$$

(see Proposition 5.3.6(c)) and hence the Banach space $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ is actually a Hilbert space.

\textbf{Remark 5.3.7} Suppose that Condition 5.3.1 holds. For a fixed martingale $X \in \mathcal{M}_{2, b}^c(\{\mathcal{F}_t\}, P)$ we are going to define the stochastic integral (5.74) over some appropriate class of integrands $\{\Phi(t); t \in [0, \infty)\}$. In the first instance, we shall limit ourselves to integrands which are \textit{"bounded enough"} to be members of $L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$, since this rather strong boundedness will be essential to our use of the Riesz Representation Theorem 1.1.2 when we establish the basic existence and uniqueness Theorem 5.3.14. Once this step is accomplished we shall significantly relax the boundedness conditions on the integrand by a systematic use of stopping times. As far as measurability of the integrand is concerned, we shall soon see important technical reasons why joint measurability is not enough, and that in order to make our definition of stochastic integration work, we will need to suppose that the integrand is a progressively measurable process. All of this motivates the next definition in which we formulate the appropriate class of stochastic integrands:
Remark 5.3.9 When \( (5.83) \)

\[
E \left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] < \infty.
\]

Remark 5.3.11 In view of this proposition it follows that there is some \( \Psi \) specifically measurable (rather than just jointly measurable). Thus, suppose Condition 5.3.1. Let We next establish an inequality which will be essential for our definition of stochastic integrals:

inner product defined by \( (5.82) \).

Proposition 5.3.10 Suppose that Condition 5.3.1 holds, and \( X \in M_{\infty,d}(\{F_t\}) \). Then \( L^2(X, \{F_t\}) \) is a linear subspace of \( L^2([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F, \nu_X) \). The next result elucidates this relationship more precisely:

Proposition 5.3.10 Suppose that Condition 5.3.1 holds, and \( X \in M_{\infty,d}(\{F_t\}) \). Then \( L^2(X, \{F_t\}) \) is a closed linear subspace of the Hilbert space \( L^2([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F, \nu_X) \) whose inner product is defined by \( (5.82) \).

Proof: In view of Remark 3.1.24, if \( \{(\Phi(t), F_t); t \in [0, \infty)\} \) is progressively measurable then \( \{(\Phi(t); t \in [0, \infty)\} \) is jointly measurable, hence \( L^2(X, \{F_t\}) \) is a linear subspace of \( \Phi \in L^2([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F, \nu_X) \), and it remains to show that it is a closed linear subspace. Suppose \( \Phi_n, n = 1, 2, \ldots \) is a Cauchy sequence in \( L^2(X, \{F_t\}) \). We must show that \( \lim_{n \to \infty} \|\Phi_n - \Phi\|_X = 0 \) for some \( \Phi \in L^2(X, \{F_t\}) \). By the completeness of \( L^2([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F, \nu_X) \) (see Proposition 1.2.27), there is some \( \Psi \in L^2([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F, \nu_X) \) such that

\[
\lim_{n \to \infty} \|\Phi_n - \Psi\|_X = 0.
\]

Thus, in view of the Markov inequality (Theorem 1.2.23), the sequence \( \Phi_n, n = 1, 2, \ldots \) of measurable functions on \( ([0, \infty) \otimes \Omega, B([0, \infty)) \otimes F) \) converges in \( \nu_X \)-measure to \( \Psi \), hence by Theorem 1.2.11(b) there is some subsequence \( \Phi_{n_r}, r = 1, 2, \ldots \) such that \( \lim_{r \to \infty} \Phi_{n_r} = \Psi \) \( \nu_X \)-a.e. Now put

\[
\Phi(t, \omega) \triangleq \limsup_{r \to \infty} \Phi_{n_r}(t, \omega), \quad \forall \omega \in \Omega, \forall t \in [0, \infty),
\]

We observe that \( \Phi = \Psi \) \( \nu_X \)-a.e., and Propositions 1.2.6 and 3.1.22 imply that \( \{(\Phi(t), F_t); t \in [0, \infty)\} \) is progressively measurable. We thus have \( \|\Phi\|_X < \infty \), so that \( \Phi \in L^2(X, \{F_t\}, P) \), and, by \( (5.84) \), we also have \( \lim_{n \to \infty} \|\Phi_n - \Phi\|_X = 0 \), as required.

Remark 5.3.11 In view of this proposition it follows that \( L^2(X, \{F_t\}, P) \) is itself a Hilbert space with inner product defined by \( (5.82) \).

We next establish an inequality which will be essential for our definition of stochastic integrals:

Theorem 5.3.12 (Kunita-Watanabe) Suppose Condition 5.3.1. Let \( \{(X(t), F_t); t \in [0, \infty)\} \) and \( \{(Y(t), F_t); t \in [0, \infty)\} \) be continuous local martingales, let \( \{(\Phi(t); t \in [0, \infty)\}, \{(\Psi(t); t \in [0, \infty)\} \) be jointly measurable processes on \( (\Omega, F, P) \), and put

\[
||[X, Y]||(t, \omega) \triangleq V[ [X, Y](\cdot, \omega); 0, t], \quad \forall t \in [0, \infty), \forall \omega \in \Omega.
\]
Then
\begin{equation}
\left[ \int_0^t |\Phi(s)\Psi(s)| d||X,Y||(s) \right] \leq \left[ \int_0^t |\Phi(s)|^2 d|X|(s) \right]^{1/2} \left[ \int_0^t |\Psi(s)|^2 d|Y|(s) \right]^{1/2} \ a.s.
\end{equation}
and
\begin{equation}
E \left[ \int_0^t |\Phi(s)\Psi(s)| d||X,Y||(s) \right] \leq \left\{ E \left[ \int_0^t |\Phi(s)|^2 d|X|(s) \right] \right\}^{1/2} \left\{ E \left[ \int_0^t |\Psi(s)|^2 d|Y|(s) \right] \right\}^{1/2}
\end{equation}
for each \( t \in [0, \infty] \).

Proof: It is clearly enough to establish the assertions in the special case where \( t = +\infty \). Without loss of generality suppose that the right hand side of (5.87) is finite, since the inequality is trivially true when it has the value \(+\infty\). By the joint measurability of \( \{\Phi(t); t \in [0, \infty)\} \) and \( \{\Psi(t); t \in [0, \infty)\} \) we see that \( \Phi(\cdot, \omega) \) and \( \Phi(\cdot, \omega) \) are \( \mathcal{B}([0, \infty)) \)-measurable (Lemma 1.2.38), and by Lemma 4.7.28(d) there is some \( P \)-null event \( N \in \mathcal{F} \) such that, if \( \omega \notin N \), then
\[
||X,Y|(t,\omega) - [X,Y](s,\omega)||^2 \leq ([X](t,\omega) - [X](s,\omega))(Y(t,\omega) - Y(s,\omega))
\]
\( \forall \ 0 \leq s < t < \infty \). Now fix some \( \omega \notin N \) and take \( A(t) \overset{\Delta}{=} [X,Y](t,\omega) \), \( B(t) \overset{\Delta}{=} [X](t,\omega) \), and \( C(t) \overset{\Delta}{=} [Y](t,\omega) \), \( \forall \ t \in [0, \infty) \), in Theorem 5.1.17 to see that (5.87) holds for each \( \omega \notin N \), as required.

As for (5.88), Proposition 5.2.2 shows that the functions of \( \omega \) in \( E[...] \) are \( \mathcal{F} \)-measurable hence all expectations in (5.88) are defined (and may take the value \(+\infty\)). If the right hand side of (5.88) is \(+\infty\) then of course (5.88) holds trivially. If, on the other hand, this quantity is finite, then (5.88) is an immediate consequence of (5.87) and the Cauchy-Schwarz inequality.

\[ \blacksquare \]

**Remark 5.3.13** Suppose Condition 5.3.1 and take \( X, Z \in M_{2,k}^2(\{\mathcal{F}_t\}, P) \) and \( \Phi \in L^2(X, \{\mathcal{F}_t\}, P) \). By Theorem 5.3.12 and (5.81), we have
\begin{equation}
E \left[ \int_0^\infty |\Phi(s)| d||X,Z||(s) \right] \leq \left\{ E \left[ \int_0^\infty |\Phi(s)|^2 d|X|(s) \right] \right\}^{1/2} \left\{ E \left[ \int_0^\infty 1 d|Z|(s) \right] \right\}^{1/2}
\end{equation}
Now Proposition 4.7.33(a) ensures \( \{E(|Z|(\infty))\}^{1/2} < \infty \), thus it follows from (5.89) that
\[
E \left[ \int_0^\infty |\Phi(s)| d||X,Z||(s) \right] < \infty,
\]
and therefore, in view of Remark 5.2.10, we see that the process
\begin{equation}
\left\{ \int_0^t \Phi(s) d[X,Z](s); \ t \in [0, \infty) \right\}
\end{equation}
is well-defined and is a member of \( FV^{c,0}(\{\mathcal{F}_t\}) \). This fact will be needed in the following theorem, which is the essential step in formulating a useful concept of stochastic integration:
Theorem 5.3.14 (Kunita-Watanabe) Suppose Condition 5.3.1. If \( X \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \) and \( \Phi \in L^2(X, \{\mathcal{F}_t\}, P) \) then there exists some \( M \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \) such that a.s.:

\[
[M, Z](t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall \ t \in [0, \infty),
\]

for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \). Moreover, if \( \tilde{M} \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \) is such that a.s.:

\[
[M, Z](t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall \ t \in [0, \infty),
\]

for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \), then \( \{M(t); t \in [0, \infty)\} \) and \( \{\tilde{M}(t); t \in [0, \infty)\} \) are indistinguishable.

Proof: Recall from Corollary 5.3.5 that \( M^c_{2,b}(\{\mathcal{F}_t\}, P) \) is a Hilbert space with inner product defined by (5.76). Exactly as in Remark 5.3.13 we know that

\[
E \left[ \int_0^\infty |\Phi(s)| \, d|[X, Z]](s) \right] < \infty,
\]

for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \). Thus

\[
\phi(Z) \triangleq E \left[ \int_0^\infty \Phi(s) \, d[X, Z](s) \right], \quad \forall \ Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P),
\]

defines a mapping \( \phi : M^c_{2,b}(\{\mathcal{F}_t\}, P) \to \mathbb{R} \), which, in view of Lemma 4.7.28(c) and Lemma 5.1.11, is easily seen to be a linear functional on \( M^c_{2,b}(\{\mathcal{F}_t\}, P) \) (recall the definition of linear functional in Section 1.1). As in Remark 5.3.13 (see (5.89)), for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \) we have

\[
E \left[ \int_0^\infty |\Phi(s)| \, d|[X, Z]](s) \right] \leq \|\Phi\|_X \cdot \{E([Z](\infty))\}^{1/2}.
\]

But Proposition 4.7.33(b) gives

\[
E\{Z^2(\infty)\} = E\{[Z](\infty)\},
\]

when \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \), and thus from (5.94) with the defining relation (5.75), we get

\[
E \left[ \int_0^\infty |\Phi(s)| \, d|[X, Z]](s) \right] \leq \|\Phi\|_X \cdot \|Z\|_M,
\]

for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \). Hence, from (5.93) and (5.95), we see that

\[
|\phi(Z)| = \left| E \left[ \int_0^\infty \Phi(s) \, d[X, Z](s) \right] \right| \leq E \left[ \int_0^\infty |\Phi(s)| \, d|[X, Z]](s) \right] \leq \|\Phi\|_X \cdot \|Z\|_M, \quad \forall \ Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P).
\]
Since \( \| \Phi \|_X < \infty \), it follows from (5.96) that \( \phi(\cdot) \) given by (5.93) is a bounded linear functional on \( M^{0}_{2,b}(\{F_t\}, P) \). By the Riesz Representation Theorem 1.1.2, there exists some \( M \in M^{0}_{2,b}(\{F_t\}, P) \) such that \( \phi(Z) = (M, Z)_M \), or equivalently,

\[
(5.97) \quad E \left[ \int_{0}^{\infty} \Phi(s) \, d[X, Z](s) \right] = E[M(\infty)Z(\infty)], \quad \forall \, Z \in M^{0}_{2,b}(\{F_t\}, P).
\]

We shall now show that (5.91) holds for this martingale \( M \) and each \( Z \in M^{c}_{2,b}(\{F_t\}, P) \). To begin with fix some arbitrary \( Z \in M^{c}_{2,b}(\{F_t\}, P) \), and put

\[
(5.98) \quad U(t) \triangleq M(t)Z(t) - \int_{0}^{t} \Phi(s) \, d[X, Z](s), \quad \forall \, t \in [0, \infty).
\]

If it can be shown that \( \{(U(t), F_t); \, t \in [0, \infty)\} \) is a continuous martingale then the uniqueness part of Theorem 4.7.22 and Remark 5.3.13 implies that \( \{M, Z\}(t) \, t \in [0, \infty) \) and \( \{\int_{0}^{t} \Phi(s) \, d[X, Z](s), \, t \in [0, \infty)\} \) are indistinguishable processes, as required for (5.91). By Remark 5.3.13 it follows that \( \{(U(t), F_t); \, t \in [0, \infty)\} \) is a continuous adapted process, and hence, from Proposition 3.1.25, it is progressively measurable. Let \( T \) be any bounded \( \{F_t\}\)-stopping time, i.e. \( T(\omega) \leq C < \infty, \, \forall \, \omega \in \Omega \). Then it is easily verified that \( E|U(T)| < \infty \). By Corollary 4.5.4 we have \( Z^T \in M^{c}_{2,b}(\{F_t\}, P) \), and hence by (5.97), Lemma 4.7.32 and Lemma 5.1.10 we get

\[
(5.99) \quad E \left[ \int_{0}^{T} \Phi(s) \, d[X, Z](s) \right] = E[M(\infty)Z^T(\infty)] = E[M(\infty)Z(T)].
\]

Moreover, by Theorem 4.5.4 we have \( M(T) = E[M(\infty) \, | \, F_T] \) a.s. and since \( Z(T) \) is \( F_T \)-measurable (recall Remark 4.4.14), we get

\[
(5.100) \quad E[M(T)Z(T)] = E\left[E \left[ M(\infty)Z(T) \mid F_T \right] \right] = E[M(\infty)Z(T)].
\]

Combining (5.100), (5.99), and (5.98) we find that \( E[U(T)] = 0 \), and since this holds for arbitrary bounded \( \{F_t\}\)-stopping times \( T \), Lemma 4.5.1 says that \( \{(U(t), F_t); \, t \in [0, \infty)\} \) is a continuous martingale. We have thus established that (5.91) holds for each \( Z \in M^{c}_{2,b}(\{F_t\}, P) \). Now fix some \( Z \in M^{c}_{2,b}(\{F_t\}, P) \), and write \( \bar{Z}(t) \triangleq Z(t) - Z_0, \, \forall \, t \in [0, \infty) \). By Lemma 4.7.28(e), \( [M, Z](t) = [M, \bar{Z}](t) \) and \( [X, Z](t) = [X, \bar{Z}](t), \, \forall \, t \in [0, \infty) \), a.s., and since \( \bar{Z} \in M^{c}_{2,b}(\{F_t\}, P) \) we see that (5.91) holds for all \( Z \in M^{c}_{2,b}(\{F_t\}, P) \). Existence of the asserted element \( M \in M^{c}_{2,b}(\{F_t\}, P) \) is thus established, and it remains to prove uniqueness to within indistinguishability. Taking the difference of (5.91) and (5.92), we find \( [M, Z](t) = [M, \bar{M}](t), \, \forall \, t \in [0, \infty) \), a.s. and taking \( Z \triangleq M - \bar{M} \), we see from Lemma 4.7.28 that \( [M - \bar{M}](t) = 0, \, \forall \, t \in [0, \infty) \), a.s. In view of Proposition 4.7.30 we obtain indistinguishability of \( \{M(t); \, t \in [0, \infty)\} \) and \( \{\bar{M}(t); \, t \in [0, \infty)\} \).

Theorem 5.3.14 establishes that the set of martingales \( M \in M^{c}_{2,b}(\{F_t\}, P) \) having the property that (5.91) holds for each \( Z \in M^{c}_{2,b}(\{F_t\}, P) \), is non-empty, and asserts furthermore that any two members of this set are indistinguishable. Theorem 5.3.14 provides the basis for the definition of stochastic integrals:
Definition 5.3.15 Suppose Condition 5.3.1. Given $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$ and $\Phi \in L^2(X, \{\mathcal{F}_t\}, P)$ we shall call any martingale $M \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, which is such that (5.91) holds for each $Z \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, an Itô stochastic integral of the progressively measurable process $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ with respect to the countinuous $L^2$-bounded martingale $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$. We shall henceforth write $\{(\Phi \cdot X)(t), \mathcal{F}_t\}; t \in [0, \infty)\}$ or, more compactly, $\Phi \cdot X$, to denote some arbitrarily chosen but fixed martingale $M$ having this property. We call $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$ the stochastic integrator and $\Phi \in L^2(X, \{\mathcal{F}_t\}, P)$ the stochastic integrand.

We see that Itô stochastic integrals have this much in common with conditional expectations and quadratic variations: they are unique to within indistinguishability, and when we write down an Itô stochastic integral $\Phi \cdot X$ we are really indicating some fixed but arbitrary choice from within a class of choices, any two members of which are indistinguishable. For this reason we use the indefinite article and speak of “an Itô stochastic integral” rather than “the Itô stochastic integral”.

There is another notation for Itô stochastic integrals which is suggested by Leibnitz’ notation for integrals in ordinary calculus and to which we have already alluded: we indicate the random variable $(\Phi \cdot X)(t)$ by the notation (5.74). Generally speaking, this latter notation is most useful for stochastic differential equations and formal manipulations involving Itô’s formula, while the notation $\Phi \cdot X$ will usually be the more convenient one for writing out the basic rules of stochastic integration.

Remark 5.3.16 Notice the essential role played by the postulated progressive measurability of $\{(\Phi(t); t \in [0, \infty)\}$ in the proof of Theorem 5.3.14, for this ensures that $\{(U(t), \mathcal{F}_t); t \in [0, \infty)\}$ defined by (5.98) is progressively measurable, from which we conclude that $U(T)$ is a random variable and are able to use Lemma 4.5.1. It turns out that if one strengthens the hypotheses on the stochastic integrator $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, and postulates that the mappings $[X](\cdot, \omega)$ are absolutely continuous with respect to Lebesgue measure, then in return one can slightly extend the class of stochastic integrands and define $\Phi \cdot X$ for $\Phi \in L^2([0, \infty) \otimes \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F}, \nu_X)$ which are only $\{\mathcal{F}_t\}$-adapted (rather than progressively measurable). Interesting though this extension might be from a theoretical point of view, it turns out that it in no way enlarges scope of applications of the stochastic calculus. For this reason we refrain from further discussion of this matter.

In ordinary calculus the term integration connotes a generalized summation, and one is therefore led to ask how Itô stochastic integrals relate to summation. We shall soon answer this question, but first we must establish some basic rules for manipulating stochastic integrals. Our first result gives an explicit formula for the co-quadratic variation of two stochastic integrals:

Theorem 5.3.17 Suppose Condition 5.3.1. If $X,Y \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$, $\Phi \in L^2(X, \{\mathcal{F}_t\}, P)$, and $\Psi \in L^2(Y, \{\mathcal{F}_t\}, P)$, then

\begin{equation}
P \left[ \int_0^\infty |\Phi(s)\Psi(s)| \, d|[X,Y]|(s) < \infty \right] = 1,
\end{equation}

210
and we have a.s:

\[(\Phi \bullet X, \Psi \bullet Y)(t) = \int_0^t \Phi(s)\Psi(s) \, d[X, Y](s), \quad \forall \ t \in [0, \infty].\tag{5.102}\]

Proof: By Theorem 5.3.12 we get

\[E\left[\int_0^\infty |\Phi(s)| |\Psi(s)| \, d|[X, Y]|(s)\right] < \infty,
\]

so (5.101) follows, and thus the Lebesgue-Stieltjes integral in (5.102) is properly defined (see Remark 5.2.10). By Definition 5.3.15 we have, a.s.:

\[(\Psi \bullet Y, \Phi \bullet X)(t) = \int_0^t \Psi(s) \, d[Y, \Phi \bullet X](s), \quad \forall \ t \in [0, \infty),\tag{5.103}\]

where, in view of Remark 5.3.13, we have

\[E\left[\int_0^\infty |\Psi(s)| \, d|[Y, \Phi \bullet X]|(s)\right] < \infty,\tag{5.104}\]

so that the Lebesgue-Stieltjes integral on the right hand side of (5.103) is properly defined. Again, by Definition 5.3.15, we have, a.s.:

\[(\Phi \bullet X, Y)(t) = \int_0^t \Phi(s) \, d[X, Y](s), \quad \forall \ t \in [0, \infty),\tag{5.105}\]

where, in view of Remark 5.3.13,

\[E\left[\int_0^\infty |\Phi(s)| \, d|[X, Y]|(s)\right] < \infty.\tag{5.106}\]

By Theorem 5.1.13(b), together with (5.103), (5.104), (5.105) and (5.106), we conclude that (5.102) holds a.s.

\[\square\]

Corollary 5.3.18 Suppose Condition 5.3.1. If \(X \in \mathcal{M}^c_{2,b}(\mathcal{F}_t, P)\) and \(\Phi \in L^2(X, \mathcal{F}_t, P)\) then we have a.s.:

\[(\Phi \bullet X)(t) = \int_0^t |\Phi(s)|^2 \, d[X](s), \quad \forall \ t \in [0, \infty).\tag{5.107}\]

Since \(\Phi \bullet X \in \mathcal{M}^c_{2,b}(\mathcal{F}_t, P)\), we see from Theorem 4.4.12 that it has a last element \((\Phi \bullet X)(\infty)\) which is a square integrable random variable. The next result computes the \(L^2\)-magnitude of this random variable and is usually called the Itô isometry:

Proposition 5.3.19 Suppose Condition 5.3.1. If \(X \in \mathcal{M}^c_{2,b}(\mathcal{F}_t, P)\) and \(\Phi \in L^2(X, \mathcal{F}_t, P)\) then

\[E[(\Phi \bullet X)(\infty)]^2 = \|\Phi\|_X^2.\tag{5.108}\]
Proof: By Corollary 5.3.18 we have

\[ E[\Phi \mathcal{O} X](t) = E \left[ \int_0^t |\Phi(s)|^2 \, d[X](s) \right], \quad \forall t \in [0, \infty). \]

Taking \( t \to \infty \) along the positive integers and using the Monotone Convergence Theorem 1.2.12, we find

\[ E[\Phi \mathcal{O} X](\infty) = E \left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right]. \tag{5.109} \]

By Proposition 4.7.33, \( E[\Phi \mathcal{O} X](\infty) = E[(\Phi \mathcal{O} X)(\infty)]^2 \). Combining this with (5.109) then gives (5.108).

The next result is a chain rule for stochastic integration:

**Theorem 5.3.20** Suppose Condition 5.3.1. If \( X \in \mathcal{M}^{2,b}_{2,0}(\{\mathcal{F}_t\}, P) \), \( \Phi \in \mathbf{L}^2(X, \{\mathcal{F}_t\}, P) \), and \( \Psi \in \mathbf{L}^2(\Phi \mathcal{O} X, \{\mathcal{F}_t\}, P) \), then \( (\Psi \Phi) \in \mathbf{L}^2(X, \{\mathcal{F}_t\}, P) \) and, a.s.:

\[ ((\Psi \Phi) \mathcal{O} X)(t) = (\Psi \mathcal{O} (\Phi \mathcal{O} X))(t), \quad \forall t \in [0, \infty). \tag{5.110} \]

Proof: Since \( \Psi \in \mathbf{L}^2(\Phi \mathcal{O} X, \{\mathcal{F}_t\}, P) \) we have

\[ E \left[ \int_0^\infty |\Psi(s)|^2 \, d[\Phi \mathcal{O} X](s) \right] < \infty. \tag{5.111} \]

Again, since \( \Phi \in \mathbf{L}^2(X, \{\mathcal{F}_t\}, P) \), it is clear that

\[ P \left\{ \int_0^\infty |\Phi(s)|^2 \, d[X](s) < \infty \right\} = 1, \tag{5.112} \]

and, in view of Corollary 5.3.18, we have a.s. that

\[ \Phi \mathcal{O} X(t) = \int_0^t |\Phi(s)|^2 \, d[X](s), \quad \forall t \in [0, \infty). \tag{5.113} \]

From (5.112), (5.113), and Theorem 5.1.13(a) we get, a.s.:

\[ \int_0^\infty |\Psi(s)|^2 |\Phi(s)|^2 \, d[X](s) = \int_0^\infty |\Psi(s)|^2 \, d[\Phi \mathcal{O} X](s). \tag{5.114} \]

From (5.111) and (5.114) we find that \( \Psi \Phi \in \mathbf{L}^2(X, \{\mathcal{F}_t\}, P) \). It remain to prove (5.110). Fix some arbitrary \( Z \in \mathcal{M}^{2,b}(\{\mathcal{F}_t\}, P) \). By Definition 5.3.15 and the fact that \( \Psi \in \mathbf{L}^2(\Phi \mathcal{O} X, \{\mathcal{F}_t\}, P) \), we get, a.s.:

\[ \Psi \mathcal{O} (\Phi \mathcal{O} X, Z)(t) = \int_0^t \Psi(s) \, d[\Phi \mathcal{O} X, Z](s), \quad \forall t \in [0, \infty), \tag{5.115} \]

where Remark 5.3.13 ensures that

\[ P \left[ \int_0^\infty |\Psi(s)| \, d[\Phi \mathcal{O} X, Z](s) < \infty \right] = 1, \tag{5.116} \]

212
so that the right hand side of (5.115) is well defined. Likewise, since \( \Phi \in L^2(X, \{\mathcal{F}_t\}, P) \), for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \) we get, a.s.:

\[
(5.117) \quad [\Phi \cdot X, Z](t) = \int_0^t \Phi(s) d[X, Z](s), \quad \forall \, t \in [0, \infty),
\]

and, in view of Remark 5.3.13,

\[
(5.118) \quad P \left[ \int_0^\infty |\Phi(s)| d|[X, Z]|(s) < \infty \right] = 1.
\]

Combining (5.115), (5.116), (5.117) and (5.118), and using Theorem 5.1.13 we get, a.s.:

\[
(5.119) \quad [\Psi \cdot (\Phi \cdot X), Z](t) = \int_0^t \Psi(s) \Phi(s) d[X, Z](s), \quad \forall \, t \in [0, \infty),
\]

for each \( Z \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \). Now (5.110) follows from (5.119) and the uniqueness to within indistinguishability asserted by Theorem 5.3.14.

The next result asserts that stochastic integration is a linear mapping from \( L^2(X, \{\mathcal{F}_t\}, P) \) into \( M^c_{2,b}(\{\mathcal{F}_t\}, P) \) for a given stochastic integrator \( X \). The proof involves just a straightforward application of the uniqueness part of Theorem 5.3.14.

**Theorem 5.3.21** Suppose Condition 5.3.1. If \( X, \Phi, \Psi \in L^2(X, \{\mathcal{F}_t\}, P) \), and \( \alpha, \beta \in \mathbb{R} \), then \( \alpha \Phi + \beta \Psi \in L^2(X, \{\mathcal{F}_t\}, P) \), and we have a.s.:

\[
(5.120) \quad ((\alpha \Phi + \beta \Psi) \cdot X)(t) = \alpha(\Phi \cdot X)(t) + \beta(\Psi \cdot X)(t), \quad \forall \, t \in [0, \infty).
\]

The next theorem, whose proof is also just a straightforward application of the uniqueness assertion of Theorem 5.3.14, asserts linearity of stochastic integration with respect to the stochastic integrator for a given stochastic integrand:

**Theorem 5.3.22** Suppose Condition 5.3.1. If \( X, Y \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \), \( \Phi \in L^2(X, \{\mathcal{F}_t\}, P) \cap L^2(Y, \{\mathcal{F}_t\}, P) \), then \( \alpha X + \beta Y \in M^c_{2,b}(\{\mathcal{F}_t\}, P) \), and we have, a.s.:

\[
(5.121) \quad (\Phi \cdot (\alpha X + \beta Y))(t) = \alpha(\Phi \cdot X)(t) + \beta(\Phi \cdot Y)(t), \quad \forall \, t \in [0, \infty).
\]

**Remark 5.3.23** We next look at properties of stochastic integrals under stopping. The results we shall develop will be essential when we extend the concept of stochastic integration to stochastic integrators which are continuous local martingales, and remove the restrictive boundedness hypothesis (5.83) on the stochastic integrand. Let \( T \) be some \( \{\mathcal{F}_t\}\)-stopping time. Then \( \{(I_{[0,T]}(t), \{\mathcal{F}_t\}; \, t \in [0, \infty)\} \) is an adapted process with left-continuous sample paths, hence progressively measurable (Proposition 3.1.25). If \( \{(\Phi(t), \mathcal{F}_t); \, t \in [0, \infty)\} \) is a progressively measurable process then it follows at once that
\{(\Phi(t)I_{[0,T]}(t), \mathcal{F}_t); \ t \in [0,\infty)\} is progressively measurable. We denote this process by \Phi[0,T] for short, and call it the \textbf{truncation} of \Phi at the stopping time T. Obviously, if \(X \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\) and \(\Phi \in \mathbb{L}^2(\varphi, \mathcal{F}_t, P)\) then \(\Phi[0,T] \in \mathbb{L}^2(X, \varphi, \mathcal{F}_t, P)\). Moreover, Corollary 4.5.8 ensures that \(X^T \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\) (using the notation of Remark 4.5.9) and from Lemmas 4.7.20 and 5.1.10 we see that \(\Phi \in \mathbb{L}^2(X^T, \varphi, \mathcal{F}_t, P)\).

**Theorem 5.3.24** Suppose that Condition 5.3.1 holds, take \(X \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\) and \(\Phi \in \mathbb{L}^2(X, \mathcal{F}_t, P)\), and let \(T \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\) (using the notation of Remark 4.5.9) and from Lemmas 4.7.20 and 5.1.10 we see that \(\Phi \in \mathbb{L}^2(X^T, \mathcal{F}_t, P)\).

\[ (5.122) \quad (\Phi \bullet X^T)(t) = (\Phi[0,T] \bullet X)(t) = (\Phi \bullet X)(t \wedge T), \quad \forall \ t \in [0,\infty). \]

Proof: Fix an arbitrary \(Z \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\). By Definition 5.3.15, we have a.s.:

\[ (5.123) \quad [\Phi \bullet X, Z|[t \wedge T) = \int_0^{t \wedge T} \Phi(s) d[X, Z|(s), \quad \forall \ t \in [0,\infty). \]

Now, from Lemma 4.7.32 we have a.s.:

\[ (5.124) \quad [\Phi \bullet X^T, Z|[t \wedge T) = [(\Phi \bullet X), Z|[t \wedge T), \quad \forall \ t \in [0,\infty), \]

and

\[ (5.125) \quad [X^T, Z|[s) = [X, Z|[s \wedge T), \quad \forall \ t \in [0,\infty). \]

By (5.123), (5.124), (5.125), and Lemma 5.1.10 we get a.s.:

\[ (5.126) \quad [(\Phi \bullet X)^T, Z|[t) = \int_0^t \Phi(s) d[X^T, Z|[s), \quad \forall \ t \in [0,\infty). \]

Since (5.126) holds for each \(Z \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\), the uniqueness part of Theorem 5.3.14 ensures that \\{(\Phi \bullet X^T)(t), \ t \in [0,\infty)\} and \\{(\Phi \bullet X)^T(t), \ t \in [0,\infty)\} are indistinguishable. Now, from (5.126) and Lemma 5.1.10 we have, a.s.:

\[ (5.127) \quad [(\Phi \bullet X)^T, Z|[t) = \int_0^t \Phi[0,T|[s) d[X^T, Z|[s), \quad \forall \ t \in [0,\infty), \]

for each \(Z \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\), whence, again by the uniqueness assertion of Theorem 5.3.14, we see that \\{(\Phi \bullet X)^T(t), \ t \in [0,\infty)\} and \\{(\Phi[0,T] \bullet X)(t), \ t \in [0,\infty)\} are indistinguishable.

**Corollary 5.3.25** Suppose that Condition 5.3.1 holds, take \(X \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\), and let \(T \in \mathbb{M}_{2,b}^c(\mathcal{F}_t, P)\) (using the notation of Remark 4.5.9) and from Lemmas 4.7.20 and 5.1.10 we see that \(\Phi \in \mathbb{L}^2(X^T, \mathcal{F}_t, P)\).

\[ (5.127) \quad (I_{[0,T]} \bullet X)(t) = X^T(t), \quad \forall \ t \in [0,\infty). \]
Remark 5.3.26 Thus far, we have defined an Itô stochastic integral $\Phi \cdot X$ for stochastic integrators $X \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$ and stochastic integrands $\Phi \in L^2(X, \{\mathcal{F}_t\}, P)$. While these conditions permit the application of Hilbert space ideas for establishing Theorem 5.3.14, which forms the whole basis for our definition of stochastic integration, they are nevertheless too restrictive for most applications. Our next task is therefore to generalize the notion of stochastic integration to a larger class of stochastic integrators and integrands. Using Theorem 5.3.24 as the basic tool, we shall extend the stochastic integrators from continuous $L^2$-bounded martingales to continuous local martingales, and correspondingly relax the boundedness restriction we have hitherto imposed on the stochastic integrands. The next definition formulates the appropriate class of stochastic integrands when the stochastic integrator is a continuous local martingale:

Definition 5.3.27 Suppose that Condition 5.3.1 holds, and let $X \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$. We shall write $L_{loc}^2(X, \{\mathcal{F}_t\}, P)$ for the set of all processes $\{\Phi(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable and

$$P \left[ \int_0^t |\Phi(s)|^2 \, d[X](s) < \infty \right] = 1, \quad \forall \ t \in [0, \infty).$$

We shall write $\Phi \in L_{loc}^2(X, \{\mathcal{F}_t\}, P)$ to indicate that $\{\Phi(t); t \in [0, \infty)\}$ is a member of $L_{loc}^2(X, \{\mathcal{F}_t\}, P)$.

A useful alternative characterization of $L_{loc}^2(X, \{\mathcal{F}_t\}, P)$ is as follows:

Proposition 5.3.28 Suppose that Condition 5.3.1 holds, and let $X \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$. Then $\Phi \in L_{loc}^2(X, \{\mathcal{F}_t\}, P)$ if and only if

(a) $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable, and

(b) there is some sequence of $\{\mathcal{F}_t\}$-stopping times $\{T_n, n = 1, 2, \ldots\}$ such that

(i) $T_n(\omega) \leq T_{n+1}(\omega), \ \forall \ \omega \in \Omega, \ \forall \ n = 1, 2, \ldots$;

(ii) $P[\lim_{n \to \infty} T_n = +\infty] = 1$;

(iii) $X^{T_n} \in M_{2,b}^c(\{\mathcal{F}_t\}, P), \ \forall \ n = 1, 2, \ldots$;

(iv) $\Phi \in L^2(X^{T_n}, \{\mathcal{F}_t\}, P), \ \forall \ n = 1, 2, \ldots$.

Proof: Suppose that $\Phi \in L_{loc}^2(X, \{\mathcal{F}_t\}, P)$, namely $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable, and (5.128) holds. It remains to establish (b)(i-iv). Since $X \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$, we can use Proposition 4.6.9 to find a localizing sequence $\{S_n, n = 1, 2, \ldots\}$ of $\{\mathcal{F}_t\}$-stopping times for $\{X(t); t \in [0, \infty)\}$ such that $|X^{S_n}(t, \omega)| \leq n, \ \forall \ \omega \in \Omega, \ \forall \ t \in [0, \infty)$, whence obviously $X^{S_n} \in M_{2,b}^c(\{\mathcal{F}_t\}, P)$. Also, put

$$R_n \triangleq \inf\{t \in [0, \infty) : \int_0^t |\Phi(s)|^2 \, d[X](s) \geq n\},$$

and observe from Proposition 3.3.7(b), (5.128), and Condition 5.3.1, that each $R_n$ is an $\{\mathcal{F}_t\}$-stopping time. It follows that the $T_n \triangleq R_n \wedge S_n$ are $\{\mathcal{F}_t\}$-stopping times for which conditions (b)(i)-(iv) hold. Conversely, suppose that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is such that (a) and (b) hold, and fix some arbitrary
Since \( t \in [0, \infty) \). If \( \{T_n, n = 1, 2, \ldots\} \) is a sequence of \( \{\mathcal{F}_t\} \)-stopping times associated with \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) (see (b)), then clearly

\[
E \left[ \int_0^{T_n} |\Phi(s)|^2 \, d[X](s) \right] < \infty, \quad \forall n = 1, 2,\ldots
\]

(5.129)

Put \( N \triangleq \bigcup_{1 \leq n < \infty} N_n \) for \( N_n \) defined by

\[
N_n \triangleq \left\{ \int_0^{T_n} |\Phi(s)|^2 \, d[X](s) = +\infty \right\} \quad \forall n = 1, 2,\ldots
\]

By (5.129) we have \( P(N_n) = 0, \forall n = 1, 2,\ldots \) whence \( P(N) = 0 \), and, for each \( \omega \notin N \),

\[
\int_0^{T_n} |\Phi(s)|^2 \, d[X](s) < \infty, \quad \forall n = 1, 2,\ldots
\]

(5.130)

Since \( \lim_{n \to \infty} T_n = +\infty \) a.s., it follows from (5.130) that (5.128) holds.

Remark 5.3.29 Since \( (x+y)^2 \leq 2x^2 + 2y^2 \), \( \forall x, y \in \mathbb{R} \), an immediate consequence of Definition 5.3.27 is that \( L^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P) \) is a real vector space whose zero element is the process \( \{\Phi(t); t \in [0, \infty)\} \) with \( \Phi \equiv 0 \).

Remark 5.3.30 Suppose Condition 5.3.1, take \( X, Z \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) and \( \Phi \in L^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P) \). By Definition 5.3.27 together with Theorem 5.3.12 (see (5.87)) we get

\[
\int_0^t |\Phi(s)| \, d|[X, Z]|(s) \leq \left( \int_0^t |\Phi(s)|^2 \, d[X](s) \right)^{1/2} ([Z](t))^{1/2} < \infty, \quad \text{a.s.}
\]

for all \( t \in [0, \infty) \). In view of Remark 5.2.10, we see that the process

\[
\left\{ \int_0^t \Phi(s) \, d[X, Z](s), \; t \in [0, \infty) \right\}
\]

is a member of \( \text{FV}^{c,0}[\mathcal{F}_t] \). This fact will be needed in the following theorem, which will be used to extend the definition of Itô’s stochastic integral to include stochastic integrators which are only continuous local martingales.

Theorem 5.3.31 Suppose Condition 5.3.1. If \( X \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) and \( \Phi \in L^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P) \), then there exists some \( M \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) such that a.s.

\[
[M, Z](t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall t \in [0, \infty),
\]

(5.131)

for each \( Z \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \). Moreover, if \( \tilde{M} \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) is such that a.s.

\[
[\tilde{M}, Z](t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall t \in [0, \infty),
\]

(5.132)

for each \( Z \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \), then \( \{M(t); t \in [0, \infty)\} \) and \( \{\tilde{M}(t); t \in [0, \infty)\} \) are indistinguishable.
Proof: Since $\Phi \in L_{\text{loc}}^2(X, \{\mathcal{F}_t\}, P)$ we can find a sequence $\{T_n, \ n = 1, 2, \ldots \}$ of $\{\mathcal{F}_t\}$-stopping times such that (b)(i-iv) of Proposition 5.3.28 hold. Write $M_n(t) \triangleq (\Phi \cdot X^T_n)(t) \in M^{c,0}_{2,b}(\{\mathcal{F}_t\}, P)$, where $\Phi \cdot X^T_n$ is an arbitrarily chosen but fixed Itô stochastic integral of $\Phi \in L^2(X^T_n, \{\mathcal{F}_t\}, P)$ with respect to $X^T_n \in M^{c,0}_{2,b}(\{\mathcal{F}_t\}, P)$. It follows from Theorem 5.3.24 and $T_n \leq T_{n+1}$, that the continuous processes $\{M_n(t), \ t \in [0, \infty)\}$ and $\{M_{n+1}(t \land T_n), \ t \in [0, \infty)\}$ are indistinguishable. Put

\[(5.133) \quad \Omega_n \triangleq \{\omega \in \Omega : M_n(t, \omega) = M_{n+1}(t \land T_n(\omega), \omega) \ \forall t \in [0, \infty)\}, \quad \Omega^* \triangleq \bigcap_{1 \leq n < \infty} \Omega_n,\]

and observe that $P(\Omega_n) = 1$, $\forall n = 1, 2, \ldots$, hence $P(\Omega^*) = 1$. We now “piece together” the processes $\{M_n(t), \ t \in [0, \infty)\}$ into an overall local martingale $M \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P)$ for which (5.131) holds, using an argument which is very reminiscent of that found in the proof of Theorem 4.7.15. For each $\omega \in \Omega$ and $t \in [0, \infty)$ define

\[(5.134) \quad M^*(t, \omega) \triangleq \limsup_{n \to \infty} M_n(t, \omega).\]

Exactly as in the proof of Theorem 4.7.15 we see that $\{(M^*(t, \mathcal{F}_t); \ t \in [0, \infty))\}$ is an adapted process, and for each $\omega \in \Omega^*$ the mapping $t \to M^*(t, \omega)$ is continuous with $M^*(t \land T_n(\omega)) = M_n(t, \omega)$, $\forall t \in [0, \infty)$. Define $\{M(t); t \in [0, \infty)\}$ as follows: for each $\omega \in \Omega^*$, $M(t, \omega) \triangleq M^*(t, \omega)$, $\forall t \in [0, \infty)$; and for each $\omega \notin \Omega^*$, $M(t, \omega) \triangleq 0$, $\forall t \in [0, \infty)$. By Condition 5.3.1 we see that $\{(M(t, \mathcal{F}_t); t \in [0, \infty))\}$ is a continuous adapted process, and since

\[(5.135) \quad M_n(t) = M(t \land T_n), \quad \forall t \in [0, \infty),\]

for each $\omega \in \Omega^*$, it also follows that $M^T_n \in M^{c,0}_{2,b}(\{\mathcal{F}_t\}, P)$ for all $n = 1, 2, \ldots$, whence $M \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P)$. Now, for each $Z \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P)$, we see from the definition of $M_n$ that, a.s.:

\[(5.136) \quad [M_n, Z](t) = \int_0^t \Phi(s) \ d[X^T_n, Z](s), \quad \forall t \in [0, \infty),\]

thus, in view of (5.135) and Lemmas 4.7.32 and 5.1.10, it follows, a.s.:

\[(5.137) \quad [M, Z](t \land T_n) = \int_0^{t \land T_n} \Phi(s) \ d[X, Z](s), \quad \forall t \in [0, \infty).\]

Taking $n \to \infty$ then establishes (5.131) for each $Z \in M^{c,0}_{2,b}(\{\mathcal{F}_t\}, P)$. To show that (5.131) continues to hold for an arbitrary $Z \in M^{c}_{\text{loc}}(\{\mathcal{F}_t\}, P)$ is just a routine exercise involving another application of Lemma 4.7.32. It remains to establish uniqueness to within indistinguishability. The argument is identical to that used in the corresponding part of the proof of Theorem 5.3.14.

**Definition 5.3.32** Suppose Condition 5.3.1. Given $X \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P)$ and $\Phi \in L_{\text{loc}}^2(X, \{\mathcal{F}_t\}, P)$, we call any local martingale $M \in M^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P)$, such that (5.131) holds for each $Z \in M^{c}_{\text{loc}}(\{\mathcal{F}_t\}, P)$, an **Itô stochastic integral** of $\Phi$ with respect to $X$. As in Definition 5.3.15 we write $\Phi \cdot X$ to indicate
some arbitrary choice of this local martingale, and denote the random variable \((\Phi \circ X)(t)\) alternatively by
\[
\int_0^t \Phi(s) \, dX(s).
\]
It follows from Theorem 5.3.31 that any two choices of Itô stochastic integrals of \(\Phi\) with respect to \(X\) must be indistinguishable.

It is easily seen that the proof used to establish Theorem 5.3.24 carries over essentially without change to our more general concept of stochastic integration and we have the following:

**Theorem 5.3.33** Suppose Condition 5.3.1. Take \(X \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\) and \(\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\), and let \(T\) be an \(\{\mathcal{F}_t\}\)-stopping time. Then \(\Phi \in \mathbb{L}^2_{\text{loc}}(X^T, \{\mathcal{F}_t\}, P)\), \(\Phi[0,T] \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\), and (5.122) (5.127) hold a.s.

It is also easily verified that Theorem 5.3.17, Corollary 5.3.18, and Theorems 5.3.20, and 5.3.21 continue to hold for stochastic integrals in the broader setting covered by Definition 5.3.32. For completeness we state these results next, leaving the routine details of proof to the reader.

**Theorem 5.3.34** Suppose Condition 5.3.1. If \(X, Y \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\), \(\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\), and \(\Psi \in \mathbb{L}^2_{\text{loc}}(Y, \{\mathcal{F}_t\}, P)\), then
\[
P \left[ \int_0^t |\Phi(s)\Psi(s)| \, d|[X,Y]|(s) < \infty \right] = 1, \quad \forall \ t \in [0, \infty),
\]
and (5.102) holds a.s.

**Corollary 5.3.35** Suppose \(X \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\) and \(\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\). Then (5.107) holds a.s.

**Theorem 5.3.36** Suppose Condition 5.3.1. If \(X \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\), \(\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\), and \(\Psi \in \mathbb{L}^2_{\text{loc}}(\Phi \cdot X, \{\mathcal{F}_t\}, P)\), then \(\Psi\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\) and (5.110) holds a.s.

**Theorem 5.3.37** Suppose Condition 5.3.1. If \(X \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\), \(\Phi, \Psi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)\), and \(\alpha, \beta \in \mathbb{R}\), then (5.120) holds a.s.

**Theorem 5.3.38** Suppose Condition 5.3.1. If \(X, Y \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\), \(\alpha, \beta \in \mathbb{R}\), and \(\Phi \in \mathbb{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P) \cap \mathbb{L}^2_{\text{loc}}(Y, \{\mathcal{F}_t\}, P)\), then \(\alpha X + \beta Y \in \mathbb{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)\), \(\Phi \in \mathbb{L}^2_{\text{loc}}(\alpha X + \beta Y, \{\mathcal{F}_t\}, P)\) and (5.121) holds a.s.

Under the rather strong boundedness conditions of Theorem 5.3.14, the Itô stochastic integral is a continuous \(L^2\)-bounded martingale, and the Itô isometry given by Proposition 5.3.19 holds. When the less restrictive conditions of Theorem 5.3.31 are in force then the most we can generally assert is that the stochastic integral is a continuous local martingale, and now there is generally no Itô isometry. The next result shows that when the stochastic integrator is just a continuous local martingale, the
stochastic integral can nevertheless still be an $L^2$-bounded martingale provided that the integrand is sufficiently bounded, and in this case there is an Itô isometry. First we trivially extend the notation of Definition 5.3.8 to the case of continuous local martingales:

**Definition 5.3.39** Suppose that Condition 5.3.1 holds, and let $X \in \mathcal{M}_{\text{loc}}^c(\{\mathcal{F}_t\}, P)$. Then $L^2(X, \{\mathcal{F}_t\}, P)$ denotes the set of all $\mathbb{R}$-valued processes $\{\Phi(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable and

$$E\left[\int_0^\infty |\Phi(s)|^2 \, d[X](s)\right] < \infty. \quad (5.139)$$

In particular, if $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a scalar standard Wiener process on $(\Omega, \mathcal{F}, P)$ then $L^2(W, \{\mathcal{F}_t\}, P)$ denotes the set of all processes $\{\Phi(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is progressively measurable, and

$$E\left[\int_0^\infty |\Phi(s)|^2 \, ds\right] < \infty.$$

**Proposition 5.3.40** Suppose that Condition 5.3.1 holds, that $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a progressively measurable process, and that $X \in \mathcal{M}_{\text{loc}}^{c,0}(\{\mathcal{F}_t\}, P)$. If $\Phi \in L^2(X, \{\mathcal{F}_t\}, P)$ then $\Phi \cdot X \in \mathcal{M}_{2,b}^{c,0}(\{\mathcal{F}_t\}, P)$, and we have the Itô isometry

$$E[\{(\Phi \cdot X)(t)\}^2] = E\left[\int_0^t |\Phi(s)|^2 \, d[X](s)\right], \quad (5.140)$$

for each $t \in [0, \infty]$.

Proof: Fix $t_1 \in [0, \infty)$ and let $T$ be some $\{\mathcal{F}_t\}$-stopping time such that $0 \leq T \leq t_1$. Put

$$T_n \triangleq \inf\{t \in [0, \infty) : |X(t)| \geq n\},$$

for all $n = 1, 2, \ldots$. Then $T \land T_n$ is a $\{\mathcal{F}_t\}$-stopping time (by Remark 3.3.8 and Proposition 3.3.13(a)), and thus $Z_n \triangleq X^{T \land T_n} \in \mathcal{M}_{2,b}^{c,0}(\{\mathcal{F}_t\}, P)$ (by Proposition 4.6.9 we see that $X^{T_n}$ is a uniformly bounded martingale, and $Z_n$, being the martingale $X^{T_n}$ stopped at $T$, is also a uniformly bounded martingale, by Corollary 4.5.8). In view of (5.139) we see that $\Phi \in L^2(Z_n, \{\mathcal{F}_t\}, P)$ (recall Definition 5.3.8), thus Proposition 5.3.19 gives

$$E[\{(\Phi \cdot Z_n)(\infty)\}^2] = E\left[\int_0^\infty |\Phi(s)|^2 \, d[Z_n](s)\right],$$

or equivalently, in view of the definition of $Z_n$,

$$E[\{(\Phi \cdot X)(T \land T_n)\}^2] = E\left[\int_0^{T \land T_n} |\Phi(s)|^2 \, d[X](s)\right], \quad (5.141)$$
for all \( n = 1, 2, \ldots \) Since the finite constant \( t_1 \) is an upper-bound for the stopping-time \( T \), we have \( \lim_{n \to \infty} T \land T_n = T, \forall \omega \). Thus from (5.139), (5.141) and Fatou’s Theorem 1.2.15, we get

\[(5.142) \quad E[((\Phi \bullet X)(T))^2] = E[\lim_{n \to \infty}((\Phi \bullet X)(T \land T_n))^2] \leq \lim_{n \to \infty} E[((\Phi \bullet X)(T \land T_n))^2] \leq \lim_{n \to \infty} E \left[ \int_0^{T \land T_n} |\Phi(s)|^2 \, d[X](s) \right] \leq E \left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] < \infty,
\]

for each \( \{\mathcal{F}_t\}\)-stopping time \( T \) that is uniformly bounded by a finite constant. Now \( X \in \mathcal{M}^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) so that \( \Phi \bullet X \in \mathcal{M}^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \). If \( \{S_k\} \) is some localizing sequence for \( \Phi \bullet X \) then \( T \triangleq S_k \land t_1 \) is a \( \{\mathcal{F}_t\}\)-stopping-time upper-bounded by the finite constant \( t_1 \), hence (5.142) gives

\[(5.143) \quad \sup_k E[((\Phi \bullet X)(S_k \land t_1))^2] < \infty.
\]

From (5.143) we see that the sequence of random variables \( \{(\Phi \bullet X)(t_1 \land S_k), k = 1, 2, \ldots\} \) is uniformly integrable, and, in view of the arbitrary choice of \( t_1 \in [0, \infty) \), it follows from Proposition 4.6.8 that \( \Phi \bullet X \in \mathcal{M}^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \). Moreover, since (5.142) holds for the trivial stopping-times \( T \triangleq t_1 \in [0, \infty) \), we get

\[E[((\Phi \bullet X)(t_1))^2] \leq E \left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] < \infty,
\]

for all \( t_1 \in [0, \infty) \), thus \( \Phi \bullet X \in \mathcal{M}^{c,0}_{2,b}(\{\mathcal{F}_t\}, P) \). It remains to establish the isometry (5.140). To this end fix some arbitrary \( t_1 \in [0, \infty) \). Then certainly

\[(5.144) \quad ((\Phi \bullet X)(t_1 \land T_n))^2 \leq \sup_{t \in [0, \infty)} ((\Phi \bullet X)(t))^2,
\]

and, since \( \Phi \bullet X \in \mathcal{M}^{c,0}_{2,b}(\{\mathcal{F}_t\}, P) \), we find from Theorem 4.4.12 that

\[(5.145) \quad E[\sup_{t \in [0, \infty)} ((\Phi \bullet X)(t))^2] < \infty.
\]

Now \( \lim_{n \to \infty} (\Phi \bullet X)(t_1 \land T_n) = (\Phi \bullet X)(t_1) \), thus (5.144), (5.145), and the Lebesgue Dominated Convergence Theorem give

\[(5.146) \quad \lim_{n \to \infty} E[((\Phi \bullet X)(t_1 \land T_n))^2] = E[((\Phi \bullet X)(t_1))^2].
\]

But, upon taking \( T \triangleq t_1 \land T_n \) in (5.141), we get

\[(5.147) \quad E[((\Phi \bullet X)(t_1 \land T_n))^2] = E \left[ \int_0^{t_1 \land T_n} |\Phi(s)|^2 \, d[X](s) \right],
\]

for all \( n = 1, 2, \ldots \) In view of (5.146), (5.147), and the Monotone Convergence Theorem, we find

\[(5.148) \quad E[((\Phi \bullet X)(t_1))^2] = E \left[ \int_0^{t_1} |\Phi(s)|^2 \, d[X](s) \right].
\]

220
for each \( t_1 \in [0, \infty) \). It remains to show that (5.140) holds when \( t = \infty \). Since we have shown \( \Phi \cdot X \in \mathcal{M}^c_{2,b}(\{\mathcal{F}_t\}, P) \), it follows from Theorem 4.4.12 that

\[
\lim_{n \to \infty} E[(\Phi \cdot X)(n))^2] = E[(\Phi \cdot X)(\infty))^2].
\]

Upon taking \( t_1 = n, n = 1, 2, \ldots \), in (5.148), and using (5.149) and the Monotone Convergence Theorem, we get (5.140) when \( t = \infty \).

In contrast to Proposition 5.3.40 the next result postulates less boundedness on the integrand, but in return delivers a correspondingly weaker result about the stochastic integral:

**Corollary 5.3.41** Suppose that Condition 5.3.1 holds, that \( \{\Phi(t), \mathcal{F}_t \}; t \in [0, \infty) \) is a progressively measurable process, and that \( X \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P) \). If

\[
E\left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] < \infty, \quad \forall \; t \in [0, \infty),
\]

then \( \Phi \cdot X \in \mathcal{M}^c_{2,b}(\{\mathcal{F}_t\}, P) \), and (5.140) holds for each \( t \in [0, \infty) \).

Proof: Fix some \( t_1 \in (0, \infty) \), and put \( \Psi(t) \triangleq I_{[0,t]} \Phi(t), \forall \; t \in [0, \infty) \). Then certainly

\[
E\left[ \int_0^\infty |\Psi(s)|^2 \, d[X](s) \right] < \infty,
\]

thus Proposition 5.3.40 gives \( \Psi \cdot X \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P) \) and

\[
E[(\Psi \cdot X)(t))^2] = E\left[ \int_0^t |\Psi(s)|^2 \, d[X](s) \right], \quad \forall \; t \in [0, \infty).
\]

Now the result follows from the definition of \( \Psi \).

One useful consequence of Proposition 5.3.40 are the following *Wald identities*:

**Corollary 5.3.42** Suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty) \} \) is a scalar standard Wiener process and \( T \) is a \( \mathcal{F}_t \)-stopping time such that \( E[T] < \infty \). Then

(a) \( E[X(T)] < \infty \) and \( E[X(T)] = 0 \);
(b) \( E[X^2(T)] = E[T] \).

Proof: Put \( \Phi(t) \triangleq I_{[0,T]}(t), \forall \; t \in [0, \infty) \). Then \( \Phi \cdot X = X^T \) (see Theorem 5.3.33). Since \( X \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P) \) with \( [X](t) = t \), it follows that

\[
E\left[ \int_0^\infty |\Phi(s)|^2 \, d[X](s) \right] = E\{[X](T)\} = E[T] < \infty.
\]

Then, according to Proposition 5.3.40, we have \( X^T \in \mathcal{M}^c_{2,b}(\{\mathcal{F}_t\}, P) \), and (a) follows from Theorem 4.4.12 together with the obvious fact that \( X^T(\infty) = X(T) \). As for (b), taking \( \Phi \triangleq I_{[0,T]} \) in the Itô
isometry (5.140), we get $E[X^2(T)] = E[(X^T(\infty))^2] = E\{[X](T)\} = E[T]$. 

We are now going to show how the stochastic integral we have defined relates to summation, thus justifying our use of the term “integral” for the concept we have formulated.

**Definition 5.3.43** Suppose that $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a filtration in the probability space $(\Omega, \mathcal{F}, P)$. A process $\{\Phi(t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ is called a $\{\mathcal{F}_t\}$-simple process when it is of the form

$$\Phi(t, \omega) = \xi_0(\omega)I_{[0)}(t) + \sum_{0 \leq k < \infty} \xi_k(\omega)I_{(t_k, t_{k+1}]}(t),$$

where: (i) $\{t_k, k = 0, 1, 2 \ldots\}$ is a sequence in $[0, \infty)$ with $t_0 = 0$, $t_k < t_{k+1}$, and $\lim_{k \to \infty} t_k = +\infty$, and (ii) $\{\xi_k, k = 0, 1, 2, \ldots\}$ is a sequence of random variables on $(\Omega, \mathcal{F}, P)$ such that $\xi_k$ is $\mathcal{F}_{t_k}$-measurable, $\forall k = 0, 1, 2, \ldots$ The set of all $\{\mathcal{F}_t\}$-simple processes on $(\Omega, \mathcal{F}, P)$ will be denoted by $\mathcal{S}\{\mathcal{F}_t\}$.

**Remark 5.3.44** We see at once that a $\{\mathcal{F}_t\}$-simple process is $\mathcal{F}_t$-adapted and left-continuous, whence, by Proposition 3.1.25, a simple process is always progressively measurable with respect to $\{\mathcal{F}_t, t \in [0, \infty)\}$. Next, suppose that $\Phi \in \mathcal{S}\{\mathcal{F}_t\}$ is uniformly bounded, that is there is a constant $C \in [0, \infty)$ such that

$$|\Phi(t, \omega)| \leq C, \quad \forall (t, \omega) \in [0, \infty) \otimes \Omega.$$

Then, view of Definition 5.3.8 and Proposition 4.7.33(a), it follows that $\Phi \in L^2(\mathcal{X}, \mathcal{F}_s, \mathcal{P})$, for each $X \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$. Hence $\Phi \cdot X \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$ is defined for each uniformly bounded $\Phi \in \mathcal{S}\{\mathcal{F}_t\}$ and $X \in M_{2,b}^c(\mathcal{F}_t, \mathcal{P})$. Likewise, from Definition 5.3.27, we see that $\mathcal{S}\{\mathcal{F}_t\} \subset L^2(\mathcal{X}, \mathcal{F}_s, \mathcal{P})$ and $X \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$, so that $\Phi \cdot X \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$ is defined for each $\Phi \in \mathcal{S}\{\mathcal{F}_t\}$ and $X \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$. We shall next write down explicit formulae for the stochastic integrals of simple processes (see Proposition 5.3.46) but first we need the following simple result:

**Lemma 5.3.45** Suppose that $X, Z \in M_{2,b}^c(\mathcal{F}_s, \mathcal{P})$ and $0 \leq s < t \leq u < v < \infty$ are constants. If $\xi_1$ and $\xi_2$ are bounded random variables, $\xi_1$ being $\mathcal{F}_s$-measurable and $\xi_2$ being $\mathcal{F}_u$-measurable then

$$E[\xi_1\{X(t) - X(s)\}\xi_2\{Z(v) - Z(u)\} \mid \mathcal{F}_s] = 0 \quad \text{a.s.}$$

and

$$E[\xi_1\{X(t) - X(s)\}\{Z(t) - Z(s)\} \mid \mathcal{F}_s] = E[\xi_1\{[X, Z](t) - [X, Z](s)\} \mid \mathcal{F}_s] \quad \text{a.s.}$$

Proof: Using the composition rule for conditional expectations we get

$$E[\xi_1(X(t) - X(s))\xi_2(Z(v) - Z(u)) \mid \mathcal{F}_s] = E[E[\xi_1(X(t) - X(s))\xi_2(Z(v) - Z(u)) \mid \mathcal{F}_u] \mid \mathcal{F}_s] \quad \text{a.s.}$$

$$= E[\xi_1(X(t) - X(s))\xi_2E[Z(v) - Z(u) \mid \mathcal{F}_u] \mid \mathcal{F}_s] = 0 \quad \text{a.s.}$$
as required for (5.153). As for (5.154), since \( E[X(t)Z(s) \mid \mathcal{F}_s] = E[X(s)Z(t) \mid \mathcal{F}_s] = X(s)Z(s) \) a.s. we find that
\[
E[\xi_1\{X(t) - X(s)\}\{Z(t) - Z(s)\} \mid \mathcal{F}_s] = \xi_1 E[[X, Z](t) - [X, Z](s) \mid \mathcal{F}_s] \quad \text{a.s.}
\]
as required for (5.154).

**Proposition 5.3.46** Suppose that Condition 5.3.1 holds, and let \( \Phi \in \mathcal{G}\{\mathcal{F}_t\} \) have the form (5.151).

(a) If \( X \in \mathcal{M}^{\mathbb{C},0}_{2b}(\{\mathcal{F}_t\}, P) \), \( \Phi \) is uniformly bounded, and the process \( \{M(t); t \in [0, \infty)\} \) is defined by
\[
M(t, \omega) \triangleq \sum_{0 \leq k < \infty} \xi_k(\omega) (X(t_{k+1} \wedge t, \omega) - X(t_k \wedge t, \omega)), \quad \forall t \in [0, \infty), \ \forall \omega \in \Omega.
\]
then \( M \in \mathcal{M}^{\mathbb{C},0}_{2b}(\{\mathcal{F}_t\}, P) \), and the processes \( \{M(t); t \in [0, \infty)\} \) and \( \{(\Phi \cdot X)(t), t \in [0, \infty)\} \) are indistinguishable.

(b) If \( X \in \mathcal{M}^{\mathbb{C},0}_{loc}(\{\mathcal{F}_t\}, P) \) and \( \{M(t); t \in [0, \infty)\} \) is defined by (5.155), then \( M \in \mathcal{M}^{\mathbb{C},0}_{loc}(\{\mathcal{F}_t\}, P) \) and the processes \( \{M(t); t \in [0, \infty)\} \) and \( \{(\Phi \cdot X)(t), t \in [0, \infty)\} \) are indistinguishable.

**Proof:** (a) Fix \( s, t \in [0, \infty) \) with \( s < t \). For integers \( k \) such that \( 0 \leq t_k < t_{k+1} \leq s \), we clearly have
\[
E[\xi_k\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \mid \mathcal{F}_s] = \xi_k\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \quad \text{a.s.}
\]
For integer \( k \) such that \( t_k \leq s < t_{k+1} \),
\[
E[\xi_k\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \mid \mathcal{F}_s] = \xi_k\{X(t_{k+1} \wedge s) - X(t_k \wedge s)\} \quad \text{a.s.}
\]
For each integer \( k \) such that \( s < t_k < t_{k+1} \),
\[
E[\xi_k\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \mid \mathcal{F}_s] = E[\xi_k E[\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \mid \mathcal{F}_{t_k}] \mid \mathcal{F}_s] = 0 \quad \text{a.s.}
\]
since \( E\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \mid \mathcal{F}_{t_k} = 0 \) a.s. when \( s < t_k < t_{k+1} \). Combining (5.156), (5.157) and (5.158), one obtains \( E[M(t) \mid \mathcal{F}_s] = M(s) \) a.s. whence \( \{(M(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale. Clearly \( M(0) \equiv 0 \) and \( t \to M(t, \omega) \) is continuous on \([0, \infty)\) for each \( \omega \in \Omega \). Put \( V_k(t) \triangleq \xi_k\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\} \). When \( j < k \) we have, by Lemma 5.3.45, that \( E[V_j(t)V_k(t)] = 0 \), thus
\[
E[M^2(t)] = \sum_{0 \leq k < \infty} E[V_k^2(t)].
\]
Moreover, since the \( \xi_k \) are uniformly bounded,
\[
E[V_k^2(t)] = E[\xi_k^2\{X(t_{k+1} \wedge t) - X(t_k \wedge t)\}^2] \leq C^2 E[X^2(t_{k+1} \wedge t) - X^2(t_k \wedge t)] \leq C^2 E[[X](t_{k+1} \wedge t) - [X](t_k \wedge t)]
\]

223
so that, by (5.159), we have \( E[M^2(t)] \leq C^2E[X](t) \). By Proposition 4.7.33 we have \( E[X](t) \leq E[X](\infty) < \infty \), \( \forall \ t \in [0, \infty) \), thus \( \sup_{t \in [0, \infty)} E[M^2(t)] < \infty \) and therefore \( M \in \mathcal{M}_{2,b}^c(\mathcal{F}_t, P) \). We have already observed that \( \Phi \in \mathcal{L}^2(X, \{ \mathcal{F}_t \}, P) \), thus it remains to prove that \( \{(\Phi \cdot X)(t), \ t \in [0, \infty)\} \) and \( \{M(t); \ t \in [0, \infty)\} \) are indistinguishable. Fix \( s, t \in [0, \infty) \) with \( s < t \), and let \( m \) and \( n \) be integers such that \( t_{m-1} < t \leq t_m \) and \( t_n \leq t < t_{n+1} \), and suppose that \( Z \) is an arbitrary member of \( \mathcal{M}_{2,b}^c(\mathcal{F}_t, P) \). Then, by repeated application of Lemma 5.3.45 together with the composition rule for conditional expectation, we get

\[
E[M(t)Z(t) - M(s)Z(s) | \mathcal{F}_s] = E[\{M(t) - M(s)\}{Z(t) - Z(s)} | \mathcal{F}_s] \quad \text{a.s.}
\]

\[
= E\left[ \left\{ \xi_{m-1}\{X(t_m) - X(s)\} + \sum_{k=m}^{n-1} \xi_k\{X(t_{k+1}) - X(t_k)\} + \xi_n\{X(t) - X(t_n)\} \right\} \right] \quad \text{a.s.}
\]

\[
= E\left[ \left\{ \xi_{m-1}\{X(t_m) - X(s)\}\{Z(t_m) - Z(s)\} \right\} \right] \quad \text{a.s.}
\]

\[
+ \sum_{k=m}^{n-1} E\left[ \left\{ \xi_k\{X(t_{k+1}) - X(t_k)\}\{Z(t_{k+1}) - Z(t_k)\} \right\} \right] \quad \text{a.s.}
\]

\[
+ E\left[ \left\{ \xi_n\{X(t) - X(t_n)\}\{Z(t) - Z(t_n)\} \right\} \right] \quad \text{a.s.}
\]

\[
= E\left[ \int_s^t \Phi(u)d[X, Z](u) | \mathcal{F}_s \right] \quad \text{a.s.}
\]

Thus we have shown that

\[
\left\{ \left( M(t)Z(t) - \int_0^t \Phi(u) \ d[X, Z](u), \mathcal{F}_t \right) : \ t \in [0, \infty) \right\}
\]

is a martingale, hence the uniqueness to within indistinguishability of co-quadratic variation processes ensures that, a.s.: 

\[
[M, Z](t) = \int_0^t \Phi(u) \ d[X, Z](u), \quad \forall \ t \in [0, \infty).
\]

The uniqueness assertion of Theorem 5.3.14 then ensures indistinguishability of \( \{(\Phi \cdot X)(t), \ t \in [0, \infty)\} \) and \( \{M(t); \ t \in [0, \infty)\} \).

(b) This follows from (a) by routine localization arguments which are left to the reader.

**Remark 5.3.47** In (5.155) we have formulated an “intuitively sensible” integral of a simple process \( \{\Phi(t); \ t \in [0, \infty)\} \) with respect to a local martingale \( \{X(t); \ t \in [0, \infty)\} \). Proposition 5.3.46 shows that
this definition coincides, to within indistinguishability, with the Itô stochastic integral which was defined on the basis of Theorem 5.3.14. We are next going to show that, for processes \( \{ \Phi(t); t \in [0, \infty) \} \) which are more general than simple processes, discrete sums of the form (5.155) are good approximations to Itô stochastic integrals. For this purpose we need to establish some preliminary results, the first of which is a “dominated convergence theorem” for stochastic integration:

**Proposition 5.3.48** Suppose that Condition 5.3.1 holds, take \( X \in M^2_{\mathcal{F}, P} \), and let \( \Phi, \Phi_n, n = 1, 2, \ldots \) be members of \( L^2(X, \mathcal{F}, P) \) such that

(a) \( |\Phi_n(t, \omega)| \leq |\Phi(t, \omega)|, \forall \omega \in \Omega, \forall t \in [0, \infty), \forall n = 1, 2, \ldots \)

(b) \( \lim_{n \to \infty} |\Phi_n(t, \omega)| = 0, \forall \omega \in \Omega, \forall t \in [0, \infty). \)

Then

\[
\lim_{n \to \infty} E \left[ \sup_{t \in [0, \infty)} |(\Phi_n \cdot X)(t)|^2 \right] = 0. 
\]

**Proof:** By Theorem 4.4.12, Proposition 5.3.19, and (5.78),

\[
E \left[ \sup_{t \in [0, \infty)} |(\Phi_n \cdot X)(t)|^2 \right] \leq 4E[|(\Phi_n \cdot X)(\infty)|^2]
\]

\[
\leq 4E \left[ \int_0^\infty |\Phi_n(s)|^2 \, d|X|(s) \right]
\]

\[
\leq 4 \int_{[0, \infty) \otimes \Omega} |\Phi_n(s)|^2 \, d\nu_X. 
\]

Next, in view of (a), (b) and the Dominated Convergence Theorem 1.2.21, we get

\[
\lim_{n \to \infty} \int_{[0, \infty) \otimes \Omega} |\Phi_n(s)|^2 \, d\nu_X = 0.
\]

Now (5.160) follows from (5.161) and (5.162).

**Proposition 5.3.49** Suppose that Condition 5.3.1 holds, take \( X \in M^2_{\mathcal{F}, P} \), and let \( \Phi \in L^2(X, \mathcal{F}, P) \) be such that

(a) \( |\Phi(t, \omega)| \leq C, \forall \omega \in \Omega, \forall t \in [0, \infty) \), and some constant \( C \in [0, \infty). \)

(b) \( t \to \Phi(t, \omega) \) is continuous on \([0, \infty)\) for each \( \omega \in \Omega \).

Put \( \tau^n_k \overset{\Delta}{=} k2^{-n}, \forall k, n = 0, 1, 2, \ldots \) and define

\[
M_n(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} \Phi(t \land \tau^n_k, \omega)[X(t \land \tau^n_k, \omega) - X(t \land \tau^n_{k+1}, \omega)]
\]

\[ \forall \omega \in \Omega, \forall t \in [0, \infty), \forall n = 0, 1, 2, \ldots \] Then:

\[
\lim_{n \to \infty} E \left[ \sup_{t \in [0, \infty)} \left| M_n(t) - (\Phi \cdot X)(t) \right|^2 \right] = 0.
\]
Proof: For each \( n = 0, 1, 2, \ldots, \omega \in \Omega \) and \( t \in [0, \infty) \), put
\[
\Phi_n(t, \omega) \triangleq \Phi(0, \omega)I_{[0]}(t) + \sum_{0 \leq k < \infty} \Phi(\tau_k^n, \omega)I_{(\tau_k^n, \tau_{k+1}^n]}(t).
\]
From (a) we see that
\[
|\Phi_n(t, \omega) - \Phi(t, \omega)| \leq 2C, \quad \forall \omega \in \Omega, \quad \forall t \in [0, \infty), \quad \forall n = 0, 1, 2, \ldots
\]
and from (b),
\[
\lim_{n \to \infty} |\Phi_n(t, \omega) - \Phi(t, \omega)| = 0, \quad \forall \omega \in \Omega, \quad t \in [0, \infty).
\]
In view of Proposition 5.3.48, (5.165) and (5.166), we get
\[
\lim_{n \to \infty} E \left[ \sup_{t \in [0, \infty)} |(\Phi_n \cdot X)(t) - (\Phi \cdot X)(t)|^2 \right] = 0.
\]
Now, by Proposition 5.3.46(a), we see that \( \{M_n(t), \ t \in [0, \infty)\} \) and \( \{(\Phi_n \cdot X)(t), \ t \in [0, \infty)\} \) are indistinguishable, thus (5.164) follows from (5.167).

We can now establish an analogue of Theorem 5.1.14, which shows that Itô stochastic integrals can be approximated by Riemann sums when the integrand is continuous:

**Theorem 5.3.50** Suppose that Condition 5.3.1 holds, take \( X \in \mathcal{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, \mathcal{P}) \) and let \( \{(\Phi(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) be a progressively measurable \( \mathcal{F} \)-valued process whose sample paths \( t \to \Phi(t, \omega) \) are continuous on \([0, \infty), \forall \omega \in \Omega \). Put \( \tau_k^n \triangleq k2^{-n}, \forall k, n = 0, 1, 2, \ldots \) and define \( \{M_n(t), \ t \in [0, \infty)\}, \forall n = 0, 1, 2, \ldots \) by (5.163). Then, for each \( t \in [0, \infty) \) and \( \epsilon \in (0, \infty) \), we have
\[
\lim_{n \to \infty} P\{|M_n(t) - (\Phi \cdot X)(t)| > \epsilon\} = 0.
\]

Proof: Define
\[
S_N(\omega) \triangleq \inf\{t \in [0, \infty): |\Phi(t, \omega)| \geq N\}, \quad \text{and} \quad R_N(\omega) \triangleq \inf\{t \in [0, \infty): |X(t, \omega)| \geq N\},
\]
\( \forall N = 1, 2, \ldots, \forall \omega \in \Omega \). In view of Remark 3.3.8, along with Proposition 3.3.13(a), we see that each \( T_N \triangleq R_N \wedge S_N \) is an \( \mathcal{F}_t \)- stopping time, and clearly \( \lim_{N \to \infty} T_N(\omega) = \infty, \forall \omega \in \Omega \). Recalling the notation of Remark 4.5.9, we observe that \( X^{T_N} \in \mathcal{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, \mathcal{P}) \) (Corollary 4.5.8), \( \Phi^{T_N} \) is progressively measurable (Proposition 3.3.14(a)) and \( \Phi^{T_N} \in L^2(X^{T_N}, \{\mathcal{F}_t\}, \mathcal{P}) \). Moreover, \( |\Phi^{T_N}(t, \omega)| \leq N, \forall \omega \in \Omega, \forall t \in [0, \infty) \), and \( t \to \Phi^{T_N}(t, \omega) \) is continuous on \([0, \infty) \) for each \( \omega \in \Omega \). From the definition of \( M_n(t, \omega) \) in (5.163), we get
\[
M_n(t \wedge T_N(\omega), \omega) \triangleq \sum_{0 \leq k < \infty} \Phi^{T_N}(t \wedge \tau_k^n, \omega)[X^{T_N}(t \wedge \tau_k^n, \omega) - X(t \wedge \tau_k^n, \omega)],
\]
and it follows from Proposition 5.3.49 that

\[
\lim_{n \to \infty} E|M_n(t \wedge T_N) - (\Phi^{T_N} \bullet X^{T_N})(t)|^2 = 0
\]

for each \( t \in [0, \infty), N = 1, 2, \ldots \). Plainly, \( \Phi^{T_N}[0, T_N] \equiv \Phi[0, T_N] \), hence we see from Theorem 5.3.33 and (5.170) that

\[
\lim_{n \to \infty} E|M_n(t \wedge T_N) - (\Phi \bullet X)(t \wedge T_N)|^2 = 0
\]

for each \( t \in [0, \infty), N = 1, 2, \ldots \). One now completes the proof by following an argument which is identical to that used for the proof of Theorem 4.7.26 from line (4.115) on.

5.4 Itô’s Change-of-Variables Formula

We are now ready to establish what is undoubtedly the most important result in these notes, namely Itô’s formula for the systematic manipulation of stochastic integrals. Itô’s formula has revolutionized the whole area of continuous-parameter stochastic processes and is now one of the most powerful tools available to the modern probabilist. As we know from the previous section, Itô stochastic integrals are defined by quite an elaborate procedure involving the use of Riesz’s Theorem 1.1.2 followed by application of the optional sampling theorem for extending the space of stochastic integrators and stochastic integrands. As yet, however, we do not know how to calculate stochastic integrals. Itô’s formula is extraordinarily useful because it is the device for doing actual concrete computations with stochastic integrals, and thus for bringing the power of the stochastic integral concept within the realm of applications. We shall soon see that Itô’s formula simplifies and unifies much of the theory of continuous-parameter stochastic processes, and is the key to providing short and elegant proofs of a number of profound results on stochastic processes, some of which were originally established without the benefit of stochastic calculus by extremely involved arguments. More importantly however, Itô’s formula is also the basic tool for obtaining further significant results which seemingly cannot be established by more classical methods.

Central to Itô’s formula is the notion of a semimartingale:

**Definition 5.4.1** Suppose that Condition 5.3.1 holds. An \( \mathbb{R} \)-valued process \( \{X(t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) is called a **continuous semimartingale** with respect to \( \{\mathcal{F}_t; t \in [0, \infty)\} \) when it can be written in the form

\[
X(t) = X_0 + M(t) + A(t), \quad \forall t \in [0, \infty),
\]

where (i) \( X_0 \) is an \( \mathcal{F}_0 \)-measurable random variable; (ii) \( M \in \mathbf{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P) \); and (iii) \( A \in \mathbf{FV}^{c,0}(\{\mathcal{F}_t\}) \).
\textbf{Remark 5.4.2} A continuous semimartingale with respect to \( \{F_t, \ t \in [0, \infty)\} \) is thus \( \{F_t\} \)-progressively measurable. Observe also, that the space of continuous semimartingales with respect to \( \{F_t, \ t \in [0, \infty)\} \) is a real vector space, with the obvious notions of vector addition and scalar multiplication. We shall denote this space by \( SM^c(\{F_t\}, P) \) and write \( X \in SM^c(\{F_t\}, P) \) to indicate that \( \{X(t); \ t \in [0, \infty)\} \) is a continuous semimartingale with respect to \( \{F_t, \ t \in [0, \infty)\} \). We note that \( M^c_{loc}(\{F_t\}, P) \subset SM^c(\{F_t\}, P) \) and \( FV^c\{F_t\} \subset SM^c(\{F_t\}, P) \). Clearly, if \( T \) is an \( \{F_t\} \)-stopping time, then \( X^T \in SM^c(\{F_t\}, P) \) whenever \( X \in SM^c(\{F_t\}, P) \) (recall the notation in Remark 4.5.9). Also, we let \( SM^{c,0}(\{F_t\}, P) \) be the subset of all members of \( SM^c(\{F_t\}, P) \) which are null at the origin.

\textbf{Remark 5.4.3} If \( X \in SM^c(\{F_t\}, P) \) is given by (5.172) then we call \( M \) the \textbf{local martingale part} of \( X \), and \( A \) the \textbf{bounded variation part} of \( X \). It is extremely important to note that the local martingale and bounded variation parts of a given continuous semimartingale are unique to within indistinguishability. Indeed, suppose \( X \in SM^c(\{F_t\}, P) \) has an alternative representation to (5.172), namely:

\begin{equation}
(5.173) \quad X(t) = X_0 + \overline{M}(t) + \overline{A}(t), \quad \forall t \in [0, \infty),
\end{equation}

for some \( \overline{M} \in M^c_{loc}(\{F_t\}, P) \) and \( \overline{A} \in FV^{c,0}\{F_t\} \). Taking the difference of (5.172) and (5.173) we find

\begin{equation}
M(t) - \overline{M}(t) = \overline{A}(t) - A(t), \quad \forall t \in [0, \infty),
\end{equation}

whence \( M - \overline{M} \in M^c_{loc}(\{F_t\}, P) \), and has continuous sample-paths of locally bounded variation. In the light of Proposition 4.7.13, we see that \( \{M(t); \ t \in [0, \infty)\} \) and \( \{
\overline{M}(t); \ t \in [0, \infty)\} \) are indistinguishable, hence of course \( \{A(t); \ t \in [0, \infty)\} \) and \( \{\overline{A}(t); \ t \in [0, \infty)\} \) are also indistinguishable.

\textbf{Theorem 5.4.4} Suppose \( X, Y \in SM^c(\{F_t\}, P) \), \( X \) having a local martingale part \( M \), and \( Y \) having a local martingale part \( N \). Put \( \tau^n_k \overset{\Delta}{=} k2^{-n}, \forall k, n = 0, 1, 2 \ldots \) and define

\begin{equation}
(5.174) \quad D_n(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} [X(t \wedge \tau^n_{k+1}, \omega) - X(t \wedge \tau^n_k, \omega)][Y(t \wedge \tau^n_{k+1}, \omega) - Y(t \wedge \tau^n_k, \omega)]
\end{equation}

\( \forall t \in [0, \infty), \forall \omega \in \Omega, \forall n = 0, 1, 2 \ldots \) Then the sequence of random variables \( \{D_n(t); \ n = 0, 1, 2 \ldots\} \) converges in probability to \([M, N](t)\) for each \( t \in [0, \infty) \).

Proof: Write \( X = X_0 + M + A \), and \( Y = Y_0 + N + B \) for \( A, B \in FV^{c,0}\{F_t\} \). By Theorem 4.7.27, the sequence of random variables \( \{D^1_n(t), \ n = 0, 1, 2, \ldots\} \) defined by

\begin{equation}
(5.175) \quad D^1_n(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} [M(t \wedge \tau^n_{k+1}, \omega) - M(t \wedge \tau^n_k, \omega)][N(t \wedge \tau^n_{k+1}, \omega) - N(t \wedge \tau^n_k, \omega)]
\end{equation}

converges in probability to \([M, N](t)\) for each \( t \in [0, \infty) \). Next, consider

\begin{equation}
(5.176) \quad D^2_n(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} [M(t \wedge \tau^n_{k+1}, \omega) - M(t \wedge \tau^n_k, \omega)][B(t \wedge \tau^n_{k+1}, \omega) - B(t \wedge \tau^n_k, \omega)]
\end{equation}
for some fixed \( t \in [0, \infty) \). Clearly,

\[
|D_n^2(t, \omega)| \leq \left\{ \sup_{0 \leq k < \infty} |M(t \wedge \tau_{k+1}^n, \omega) - M(t \wedge \tau_k^n, \omega)| \right\} . \dot{B}(t, \omega).
\]

By the uniform continuity of each \( M(\cdot, \omega) \) on \([0, t]\) we have

\[
\lim_{n \to \infty} \left\{ \sup_{0 \leq k < \infty} |M(t \wedge \tau_{k+1}^n, \omega) - M(t \wedge \tau_k^n, \omega)| \right\} = 0,
\]

and since \( \dot{B}(t, \omega) < \infty \), we see from (5.177) that

\[
\lim_{n \to \infty} |D_n^2(t, \omega)| = 0, \quad \forall \omega \in \Omega.
\]

Likewise, if we put

\[
D_n^3(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} [A(t \wedge \tau_{k+1}^n, \omega) - A(t \wedge \tau_k^n, \omega)][N(t \wedge \tau_{k+1}^n, \omega) - N(t \wedge \tau_k^n, \omega)]
\]

and

\[
D_n^4(t, \omega) \overset{\Delta}{=} \sum_{0 \leq k < \infty} [A(t \wedge \tau_{k+1}^n, \omega) - A(t \wedge \tau_k^n, \omega)][B(t \wedge \tau_{k+1}^n, \omega) - B(t \wedge \tau_k^n, \omega)]
\]

then, for each \( t \in [0, \infty) \), we have

\[
\lim_{n \to \infty} |D_n^3(t, \omega)| = \lim_{n \to \infty} |D_n^4(t, \omega)| = 0, \quad \forall \omega \in \Omega.
\]

Since

\[
D_n(t) = D_n^1(t) + D_n^2(t) + D_n^3(t) + D_n^4(t),
\]

the theorem follows.

**Remark 5.4.5** Suppose \( X, Y \in \text{SM}^c(\{\mathcal{F}_t\}, P) \), \( X \) having a local martingale part \( M \), and \( Y \) having a local martingale part \( N \). For later manipulations it is useful to extend the bracket notation already introduced for co-quadratic variations of continuous local martingales and put

\[
[X, Y](t) \overset{\Delta}{=} [M, N](t) \quad \text{and} \quad [X](t) \overset{\Delta}{=} [M](t), \quad \forall t \in [0, \infty).
\]

The next proposition lists the basic rules for using this notation with continuous semimartingales. The proof is an immediate consequence of the preceding remark along with Proposition 4.7.28 and Lemma 4.7.30.

**Proposition 5.4.6** Suppose that Condition 5.3.1 holds, and take \( X, Y, Z \in \text{SM}^c(\{\mathcal{F}_t\}, P) \) and \( \alpha, \beta \in \mathbb{R} \). Then, we have a.s.:

(a) \([X](t) = [X, X](t), \forall t \in [0, \infty);\]

(b) \([X, Y](t) = [Y, X](t), \forall t \in [0, \infty);\]

(c) \([\alpha X + \beta Y, Z](t) = \alpha[X, Z](t) + \beta[Y, Z](t), \forall t \in [0, \infty);\]

(d) \([X](t) = 0, \forall t \in [0, \infty) \) if and only if \( X \in \text{FV}^c(\{\mathcal{F}_t\});\]

(e) If either \( X \in \text{FV}^c(\{\mathcal{F}_t\}) \) or \( Y \in \text{FV}^c(\{\mathcal{F}_t\}) \), then \([X, Y](t) = 0, \forall t \in [0, \infty).\)

229
Remark 5.4.7 Suppose $X, Y \in \text{SM}^c(\{\mathcal{F}_t\}, P)$, with $X(0) = Y(0)$ a.s. and for each $U \in \text{SM}^c(\{\mathcal{F}_t\}, P)$ we have a.s.:

$$[X, U](t) = [Y, U](t) \quad \forall \ t \in [0, \infty),$$

Taking $U \triangleq X - Y$ and using Proposition 5.4.6(c,d) we conclude that the local martingale parts of $X$ and $Y$ are indistinguishable and hence that $X - Y \in \text{FV}^{c,0}\{\mathcal{F}_t\}$. In general we cannot conclude that $X$ and $Y$ are themselves indistinguishable, a situation which should be compared with Proposition 4.7.31.

We next formulate the notion of stochastic integration with respect to continuous semimartingales. To this end we first establish the appropriate class of stochastic integrands:

Definition 5.4.8 Suppose that Condition 5.3.1 holds, and let $X \in \text{SM}^c(\{\mathcal{F}_t\}, P)$ have the representation in (5.172). We shall use $L^2_{loc}(X, \{\mathcal{F}_t\}, P)$ to denote the set of all $\mathbb{R}$-valued progressively measurable processes $(\Phi(t), \mathcal{F}_t); t \in [0, \infty))$ on $(\Omega, \mathcal{F}, P)$ such that

$$P \left[ \int_0^t |\Phi(s)|^2 d|M|(s) \right] < \infty \quad \text{and} \quad P \left[ \int_0^t |\Phi(s)| d\hat{A}(s) \right] < \infty$$

for each $t \in [0, \infty)$.

Remark 5.4.9 In the special case where $X \in \text{SM}^c(\{\mathcal{F}_t\}, P)$ is actually a member of $\text{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P)$ (thus the bounded variation part of $X$ is just the zero-process $A(t, \omega) = 0, \forall (t, \omega) \in [0, \infty) \otimes \Omega$) one sees that $L^2_{loc}(X, \{\mathcal{F}_t\}, P)$ in the preceding definition is identical to the class of integrands formulated in Definition 5.3.27. More generally, if $X \in \text{SM}^c(\{\mathcal{F}_t\}, P)$ has the representation in (5.172) and $\Phi \in L^2_{loc}(X, \{\mathcal{F}_t\}, P)$, then one sees from the first condition in (5.178) that $\Phi \in L^2_{loc}(M, \{\mathcal{F}_t\}, P)$.

Definition 5.4.10 Suppose that Condition 5.3.1 holds, that $X \in \text{SM}^c(\{\mathcal{F}_t\}, P)$ has the representation in (5.172), and that $\Phi \in L^2_{loc}(X, \{\mathcal{F}_t\}, P)$. Put

$$\Phi \bullet X(t) \triangleq (\Phi \bullet M)(t) + \int_0^t \Phi(s) dA(s), \quad \forall \ t \in [0, \infty),$$

where $(\Phi \bullet M) \in \text{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P)$ is an Itô stochastic integral of $\Phi$ with respect to $M$ (see Definition 5.3.32), and the second part of (5.178), together with Remark 5.2.10, ensures the second term on the right hand side of (5.179) defines a process which is an element of $\text{FV}^{c,0}\{\mathcal{F}_t\}$. It follows that $\Phi \bullet X \in \text{SM}^{c,0}(\{\mathcal{F}_t\}, P)$, and the two terms on the right hand side of (5.179) give the local martingale and bounded variation parts of $\Phi \bullet X$. We call $\Phi \bullet X$ an Itô stochastic integral of $\Phi$ with respect to the continuous semimartingale $X$.

The next proposition is an immediate consequence of Definition 5.4.10 and Proposition 5.4.6:
Proposition 5.4.11 Suppose that Condition 5.3.1 holds, that \( \{ \Phi(t), \mathcal{F}_t ; t \in [0, \infty) \} \) is a progressively measurable process such that (5.178) holds for each \( t \in [0, \infty) \). Then, for each \( Z \in \text{SM}^c(\{\mathcal{F}_t\}, P) \), we have a.s.:

\[
(5.180) \quad [\Phi \circ X, Z](t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall \, t \in [0, \infty).
\]

Remark 5.4.12 With reference to Proposition 5.4.11 we observe that, if \( \tilde{M} \in \text{SM}^{c,0}(\{\mathcal{F}_t\}, P) \) be such that a.s.:

\[
(5.181) \quad \tilde{M}, Z(t) = \int_0^t \Phi(s) \, d[X, Z](s), \quad \forall \, t \in [0, \infty),
\]

for each \( Z \in \text{SM}^c(\{\mathcal{F}_t\}, P) \), then it does not necessarily follow that \( \Phi \circ X \) and \( \tilde{M} \) are indistinguishable. In fact, in the light of Remark 5.4.7, we can conclude only that \( \tilde{M} \) has the form \( \tilde{M} = \Phi \circ X + B \) for some \( B \in \text{FV}^{c,0}(\{\mathcal{F}_t\}) \). This should be contrasted with Theorem 5.3.31 which establishes such a determining property for stochastic integration with respect to continuous local martingales.

For a given \( X \in \text{SM}^{c,0}(\{\mathcal{F}_t\}, P) \) with representation (5.172), and a progressively measurable process \( \{ (\Phi(t), \mathcal{F}_t) ; t \in [0, \infty) \} \), the stochastic integral \( \Phi \circ X \) is defined only when the conditions in (5.178) hold. This dependence of the “stochastic integrability” of \( \Phi \) on the stochastic integrator \( X \) can be inconvenient, and we now specify a class of progressively measurable processes \( \{ (\Phi(t), \mathcal{F}_t) ; t \in [0, \infty) \} \) which is general enough to be useful in applications, but special enough to ensure that \( \Phi \circ X \) is defined for each and every \( X \in \text{SM}^{c,0}(\{\mathcal{F}_t\}, P) \):

Definition 5.4.13 An \( \mathbb{R} \)-valued progressively measurable process \( \{ (\Phi(t), \mathcal{F}_t) ; t \in [0, \infty) \} \) is said to be **locally bounded** when there exists a sequence of \( \{\mathcal{F}_t\} \)-stopping times \( \{T_n, n = 1, 2, \ldots\} \), and real constants \( \{C_n, n = 1, 2, \ldots\} \), such that

(i) \( T_n(\omega) \leq T_{n+1}(\omega), \forall \, \omega \in \Omega, \forall \, n = 1, 2, \ldots \);
(ii) \( P[\lim_{n \to \infty} T_n = +\infty] = 1 \);
(iii) \( 0 \leq C_n \leq C_{n+1} < \infty, \forall \, n = 1, 2, \ldots \);
(iv) \( |\Phi[0, T_n](t, \omega)| \leq C_n, \forall \, \omega \in \Omega, \forall \, t \in [0, \infty), \forall \, n = 1, 2, \ldots \).

Remark 5.4.14 It is clear from Definition 5.4.13 that the space of locally bounded processes is a real vector space, with the obvious notions of vector addition, scalar multiplication, and zero element. We shall denote this space by \( \text{B}_\text{loc}(\{\mathcal{F}_t\}) \), and write \( \Phi \in \text{B}_\text{loc}(\{\mathcal{F}_t\}) \) to indicate that \( \{ (\Phi(t), \mathcal{F}_t) ; t \in [0, \infty) \} \) is a locally bounded process.

The proof of the next proposition is straightforward and is left to the reader:

Proposition 5.4.15 We have

\[
\text{B}_\text{loc}(\{\mathcal{F}_t\}) \subset \text{L}^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P), \quad \forall \, X \in \text{SM}^c(\{\mathcal{F}_t\}, P).
\]
Remark 5.4.16 Suppose that Condition 5.3.1 holds, and take $\Phi \in B_{loc}\{\mathcal{F}_t\}$. It follows from Proposition 5.4.15 that $\Phi \bullet X$ is a member of $SM^c(\{\mathcal{F}_t\}, P)$ defined by (5.179), for each $X \in SM^c(\{\mathcal{F}_t\}, P)$ with the representation (5.172). Thus $\Phi \in B_{loc}\{\mathcal{F}_t\}$ is “stochastically integrable” with respect to each and every $X \in SM^c(\{\mathcal{F}_t\}, P)$.

Remark 5.4.17 Of crucial importance is the fact that any continuous adapted process $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a member of $B_{loc}\{\mathcal{F}_t\}$. Indeed, we need only put $C_n \overset{\Delta}{=} n$ and

$$T_n(\omega) \overset{\Delta}{=} \inf\{t \in [0, \infty) : |\Phi(t, \omega)| \geq n\}, \quad \omega \in \Omega,$$

for each $n = 1, 2, \ldots$, in order to check the conditions of Definition 5.4.13. Thus, if $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is continuous and adapted, then the stochastic integral $\Phi \bullet X$ is defined for each and every $X \in SM^c(\{\mathcal{F}_t\}, P)$. In particular, $X \bullet Y$ and $Y \bullet X$ are defined whenever $X, Y \in SM^c(\{\mathcal{F}_t\}, P)$.

The next result summarizes the basic rules for manipulating stochastic integrals of locally bounded processes with respect to continuous semimartingales:

Proposition 5.4.18 Suppose that Condition 5.3.1 holds, and take $X, Y \in SM^c(\{\mathcal{F}_t\}, P)$, $\Phi, \Psi \in B_{loc}\{\mathcal{F}_t\}$, and $\alpha, \beta \in \mathbb{R}$. Then:

(a) We have

$$P \left[ \int_0^t |\Phi(s)\Psi(s)| \, d|X, Y|(s) < \infty \right] = 1, \quad \forall t \in [0, \infty),$$

and, a.s.:

$$[\Phi \bullet X, \Psi \bullet Y](t) = \int_0^t \Phi(s)\Psi(s) \, d[X, Y](s), \quad \forall t \in [0, \infty).$$

(b) $\Phi \Psi \in B_{loc}\{\mathcal{F}_t\}$ and we have a.s.:

$$((\Phi\Psi) \bullet X)(t) = (\Psi \bullet (\Phi \bullet X))(t), \quad \forall t \in [0, \infty).$$

(c) $(\alpha\Phi + \beta\Psi) \in B_{loc}\{\mathcal{F}_t\}$ and we have a.s.:

$$((\alpha\Phi + \beta\Psi) \bullet X)(t) = \alpha(\Phi \bullet X)(t) + \beta(\Psi \bullet X)(t), \quad \forall t \in [0, \infty).$$

(d) $\alpha X + \beta Y \in SM^c(\{\mathcal{F}_t\}, P)$ and we have a.s.:

$$(\Phi \bullet (\alpha X + \beta Y))(t) = \alpha(\Phi \bullet X)(t) + \beta(\Phi \bullet Y)(t), \quad \forall t \in [0, \infty).$$

(e) For each $\{\mathcal{F}_t\}$-stopping time $T$ we have $X^T \in SM^c(\{\mathcal{F}_t\}, P)$ and, a.s.:

$$(\Phi \bullet X^T)(t) = (\Phi[0, T] \bullet X)(t) = (\Phi \bullet X)^T(t), \quad \forall t \in [0, \infty).$$

(f) If $\Phi$ is a simple process of the form (5.151) then $\Phi \in L^2_{loc}(X, \{\mathcal{F}_t\}, P)$ and we have a.s.:

$$(\Phi \bullet X)(t) = \sum_{0 \leq k < \infty} \xi_k [X(t \wedge t_{k+1}) - X(t \wedge t_k)] \quad \forall t \in [0, \infty).$$
Proof: (a) is a consequence of Definition 5.4.10, Remark 5.4.5, and Theorem 5.3.34. (b) follows from
Definition 5.4.10, Theorem 5.3.36 and Theorem 5.1.13. (c) results from Definition 5.4.10, Theorem
5.3.37 and the obvious linearity of Lebesgue-Stieltjes integration. Also, (d) follows from Definition
5.4.10, Theorem 5.3.38 and Lemma 5.1.11. (e) is a consequence of Definition 5.4.10, Theorem 5.3.33
and Lemma 5.1.10, and (f) follows by Proposition 5.3.46(b).

The next result is an integration-by-parts formula for stochastic integration which should be compared
with Theorem 5.1.15. It will soon be seen that this result is really the crucial step in our derivation of
Itô’s formula.

**Theorem 5.4.19** Suppose that Condition 5.3.1 holds, and take \( X, Y \in \text{SM}^c(\{\mathcal{F}_t\}, P) \). Then we have
a.s.:

\[
(5.182) X(t)Y(t) = X(0)Y(0) + \int_0^t X(s) \, dY(s) + \int_0^t Y(s) \, dX(s) + [X, Y](t), \quad \forall \, t \in [0, \infty);
\]

and, in particular,

\[
(5.183) \quad X^2(t) = X^2(0) + 2 \int_0^t X(s) \, dX(s) + [X](t), \quad \forall \, t \in [0, \infty).
\]

Proof: Write \( \tau^*_k \overset{\Delta}{=} k 2^{-n} \), \( \forall \, k, n = 1, 2, \ldots \) and fix some arbitrary \( t \in [0, \infty) \). Then, by simple
manipulations, we see that

\[
(5.184) \quad \sum_{0 \leq k < \infty} [X(t \wedge \tau^*_n \wedge \tau^*_{k+1}, \omega) - X(t \wedge \tau^*_n, \omega)] [Y(t \wedge \tau^*_n \wedge \tau^*_{k+1}, \omega) - Y(t \wedge \tau^*_n, \omega)]
\]

\[
= X(t)Y(t) - X(0)Y(0) - \sum_{0 \leq k < \infty} X(t \wedge \tau^*_n, \omega) [Y(t \wedge \tau^*_n \wedge \tau^*_{k+1}, \omega) - Y(t \wedge \tau^*_n, \omega)]
\]

\[
- \sum_{0 \leq k < \infty} Y(t \wedge \tau^*_n, \omega) [X(t \wedge \tau^*_n \wedge \tau^*_{k+1}, \omega) - X(t \wedge \tau^*_n, \omega)].
\]

Now, by Theorem 5.4.4 and Remark 5.4.5 the left hand side of (5.184) converges in probability to \([X, Y]\)
as \( n \to \infty \). Considering the third term on the right hand side of (5.184), put

\[
(5.185) \quad X_n(t, \omega) \overset{\Delta}{=} X(0, \omega)I_{\{0\}}(t) + \sum_{0 \leq k < \infty} X(\tau^*_n, \omega)I_{\{\tau^*_n, \tau^*_n\wedge \tau^*_{k+1}\}}(t),
\]

\( \forall \, \omega \in \Omega, \, \forall \, t \in [0, \infty), \, \forall \, n = 1, 2, \ldots \)

and suppose the continuous semimartingale \( Y \) has the representation

\[
Y(t) = Y_0 + N(t) + B(t)
\]

for some \( N \in \text{M}^{c,0}_{\text{loc}}(\{\mathcal{F}_t\}, P) \) and \( B \in \text{FV}^{c,0}{\mathcal{F}_t} \). By Proposition 5.4.18(f) we have, a.s.:

\[
(5.186) \quad \sum_{0 \leq k < \infty} X(t \wedge \tau^*_n, \omega) [Y(t \wedge \tau^*_n \wedge \tau^*_{k+1}, \omega) - Y(t \wedge \tau^*_n, \omega)] = (X_n \bullet N)(t) + \int_0^t X_n(s) \, dB(s).
\]
Now Theorem 5.3.50 says that the first term on the right hand side of (5.186) converges in probability to \((X \cdot N)(t)\) as \(n \to \infty\). As for the second term, since \(t \to X(t, \omega)\) is continuous and hence uniformly bounded on \([0, t]\), we see from the Dominated Convergence Theorem 1.2.21 that
\[
\lim_{n \to \infty} \int_0^t X_n(s, \omega) \, dB(s, \omega) = \int_0^t X(s, \omega) \, dB(s, \omega), \quad \forall \omega \in \Omega.
\]
Combining these facts with (5.186), we find that the third term on the right hand side of (5.184) converges in probability to \((X \cdot N)(t) + \int_0^t X(s) \, dB(s) \triangleq (X \cdot Y)(t)\).

In the same way, the fourth term on the right hand side of (5.184) converges in probability to \((Y \cdot X)(t)\) as \(n \to \infty\), and (5.182) follows. (5.183) is now a special case of (5.182).

**Remark 5.4.20** An immediate consequence of Theorem 5.4.19 is that, if \(X, Y \in \mathbf{SM}^c(\{\mathcal{F}_t\}, P)\), then the product of \(X\) and \(Y\), namely \(XY \triangleq \{(X(t)Y(t), \mathcal{F}_t); \ t \in [0, \infty)\}\), is a member of \(\mathbf{SM}^c(\{\mathcal{F}_t\}, P)\), the local martingale part of which is the sum of the local martingale parts in the second and third terms of (5.182), and the bounded variation part of which is the sum of the bounded variation parts in the second, third and fourth terms of (5.182).

**Remark 5.4.21** When either \(X \in \mathbf{FV}^c(\mathcal{F}_t)\) or \(Y \in \mathbf{FV}^c(\mathcal{F}_t)\) then, from Proposition 5.4.6(e), the integration-by-parts formula (5.182) takes the simpler form
\[
X(t)Y(t) = X(0)Y(0) + \int_0^t X(s) \, dY(s) + \int_0^t Y(s) \, dX(s), \quad \forall t \in [0, \infty).
\]
When \(X, Y \in \mathbf{FV}^c(\mathcal{F}_t)\) then this result is consistent with that which follows from application of Theorem 5.1.15 for each \(\omega\).

**Definition 5.4.22** A mapping \(F : \mathbb{R}^d \to \mathbb{R}\) is a polynomial on \(\mathbb{R}^d\) in the real variables \(x^i, \ i = 1, 2, \ldots, d\), when it is a linear combination of mappings \(\phi : \mathbb{R}^d \to \mathbb{R}\) which have the special form
\[
\phi(x^1, x^2, \ldots x^d) = \prod_{i=1}^d (x^i)^{m_i},
\]
in which the \(m_i\) are non-negative integers.

**Definition 5.4.23** An \(\mathbb{R}^d\)-valued process \(\{X(t); t \in [0, \infty)\}\) is a \(d\)-dimensional continuous vector semimartingale with respect to the filtration \(\{\mathcal{F}_t, t \in [0, \infty)\}\) if it can be expanded as a vector of continuous semimartingales \(X_i \in \mathbf{SM}^c(\{\mathcal{F}_t\}, P)\), namely
\[
X(t) = (X_1(t), X_2(t), \ldots X_d(t)), \quad \forall t \in [0, \infty).
\]
**Theorem 5.4.24** Suppose that Condition 5.3.1 holds. If the mapping \( F : \mathbb{R}^d \to \mathbb{R} \) is a polynomial and \( X = (X_1, X_2, \ldots, X_d) \) is a \( d \)-dimensional continuous vector semimartingale with respect to the filtration \( \{ \mathcal{F}_t, t \in [0, \infty) \} \), then we have a.s.:

\[
F(X(t)) = F(X(0)) + \sum_{i=1}^{d} \int_0^t \frac{\partial F}{\partial x^i}(X(s)) \, dX_i(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 F}{\partial x^i \partial x^j}(X(s)) \, d[X_i, X_j](s), \quad \forall t \in [0, \infty).
\]

**Remark 5.4.25** Clearly \( \{(\partial F(X(t))/\partial x^i, \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous adapted process for each \( i = 1, 2, \ldots, d \), hence, in view of Remark 5.4.17, a member of \( \mathbf{B}_{\text{loc}}\{\mathcal{F}_t\} \). Thus, the stochastic integrals in the second term on the right hand side of (5.187) are well defined and elements of \( \mathbf{SM}^{c,0}\{\mathcal{F}_t\}, P \). The third term on the right hand side of (5.187) is a sum of Lebesgue-Stieltjes integrals. For each \( \omega \in \Omega \) and \( t \in [0, \infty) \) there exists some \( C(t, \omega) \in [0, \infty) \) such that

\[
\left| \frac{\partial^2 F}{\partial x^i \partial x^j}(X(s, \omega)) \right| \leq C(t, \omega), \quad \forall s \in [0, t].
\]

Thus, Remark 5.2.10 ensures the third term on the right of (5.187) defines a process which is a member of \( \mathbf{FV}^{c,0}\{\mathcal{F}_t\} \). It follows that \( \{(F(X(t)), \mathcal{F}_t); t \in [0, \infty)\} \) is a member of \( \mathbf{SM}^{c}\{\mathcal{F}_t\}, P \), whose martingale and bounded variation parts clearly follow from the right hand side of (5.187).

**Proof of Theorem 5.4.24:** We see at once from Theorem 5.4.19 that (5.187) holds a.s. for polynomials \( F : \mathbb{R}^d \to \mathbb{R} \) with the special bilinear form \( F(x^i, x^j) \overset{\Delta}{=} x^i x^j \), for \( i, j = 1, 2, \ldots, d \). Now suppose that \( F : \mathbb{R}^d \to \mathbb{R} \) is some polynomial for which (5.187) holds a.s., and put \( G(x) \overset{\Delta}{=} x^k F(x) \) for some \( k \in \{1, 2, \ldots, d\} \). In view of Remark 5.4.25, the process \( \{F(X_k(t)), t \in [0, \infty)\} \) is a member of \( \mathbf{SM}^{c}\{\mathcal{F}_t\}, P \), hence we can use Theorem 5.4.19 to get, a.s.:

\[
G(X(t)) = X_k(t) F(X(t)) = X_k(0) F(X(0)) + \int_0^t X_k(s) \, dF(X(s)) + \int_0^t F(X(s)) \, dX_k(s) + [X_k, F(X)](t), \quad \forall t \in [0, \infty).
\]

By (5.187), Proposition 5.4.18(b), and Theorem 5.1.13(b) we have, a.s.:

\[
\int_0^t X_k(s) \, dF(X(s)) = \sum_{i=1}^{d} \int_0^t X_k(s) \frac{\partial F}{\partial x^i}(X(s)) \, dX_i(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t X_k(s) \frac{\partial^2 F}{\partial x^i \partial x^j}(X(s)) \, d[X_i, X_j](s), \quad \forall t \in [0, \infty).
\]

Furthermore, by (5.187), Proposition 5.4.6(c,e), and Proposition 5.4.18(a), we have a.s.:

\[
[X_k, F(X)](t) = \sum_{i=1}^{d} \int_0^t \frac{\partial F}{\partial x^i}(X(s)) \, d[X_k, X_i](s), \quad \forall t \in [0, \infty).
\]
Combining (5.188), (5.189), (5.190) and simplifying, we see that, a.s.:

\[ G(X(t)) = G(X(0)) + \sum_{i=1}^{d} \int_{0}^{t} \frac{\partial G}{\partial x^{i}}(X(s)) \, dX_{i}(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_{0}^{t} \frac{\partial^{2} G}{\partial x^{i} \partial x^{j}}(X(s)) \, d[X_{i}, X_{j}](s), \quad \forall \, t \in [0, \infty). \]

To summarize, we have shown that (5.187) holds for polynomials \( F \colon \mathbb{R}^{d} \to \mathbb{R} \) having the form \( F(x) = x^{i}x^{j} \), and that if (5.187) holds for some polynomial \( F \) on \( \mathbb{R}^{d} \) then it also holds with \( F \) replaced by a polynomial \( G \) of the form \( G(x) \triangleq x^{k}F(x) \). Finally, we observe trivially that if (5.187) holds for polynomials \( F_{1} \) and \( F_{2} \) on \( \mathbb{R}^{d} \) then it also hold for the polynomial given by \( F(x) \triangleq F_{1}(x) + F_{2}(x) \). Together, the preceding implies that (5.187) holds for every polynomial \( F \) on \( \mathbb{R}^{d} \). \( \square \)

Suppose that \( D \) is an open subset of \( \mathbb{R}^{d} \). We recall that a function \( F : D \to \mathbb{R} \) is called a \textbf{C²-function} (or twice continuously differentiable function) when \( F \) is continuous on \( D \), the partial derivatives \( \partial F(x)/\partial x^{i} \) and \( \partial^{2} F(x)/\partial x^{i}\partial x^{j} \) exist at each \( x \in D \), and the mappings \( x \to \partial F(x)/\partial x^{i} \) and \( x \to \partial^{2} F(x)/\partial x^{i}\partial x^{j} \) are continuous on \( D \) for all \( i, j = 1, 2, \ldots d \).

The following theorem is well known from real analysis and will be needed for establishing Itô’s formula:

**Theorem 5.4.26 (Weierstrass)** Suppose \( D \subset \mathbb{R}^{d} \) is an open set, \( F : D \to \mathbb{R} \) is a \textbf{C²-function}, and \( K \subset D \) is closed and bounded. Then there exists a sequence of polynomials \( F_{k} : \mathbb{R}^{d} \to \mathbb{R}, \, k = 1, 2, \ldots, \) such that

\[ \lim_{k \to \infty} \max_{x \in K} |F_{k}(x) - F(x)| = 0, \]

\[ \lim_{k \to \infty} \max_{x \in K} \left| \frac{\partial F_{k}}{\partial x^{i}}(x) - \frac{\partial F}{\partial x^{i}}(x) \right| = 0, \quad \forall \, i = 1, 2, \ldots d, \]

\[ \lim_{k \to \infty} \max_{x \in K} \left| \frac{\partial^{2} F_{k}}{\partial x^{i}\partial x^{j}}(x) - \frac{\partial^{2} F}{\partial x^{i}\partial x^{j}}(x) \right| = 0, \quad \forall \, i, j = 1, 2, \ldots d. \]

**Theorem 5.4.27 (Itô, Kunita-Watanabe)** Suppose that Condition 5.3.1 holds, that \( D \subset \mathbb{R}^{d} \) is an open set, and that \( F : D \to \mathbb{R} \) is a \textbf{C²-function}. If \( X = (X_{1}, X_{2}, \ldots, X_{d}) \) is a \( d \)-dimensional continuous vector semimartingale with respect to filtration \( \{\mathcal{F}_{t}, \, t \in [0, \infty)\} \) such that

\[ P[X(t) \in D \quad \forall \, t \in [0, \infty)) = 1, \]

then we have, a.s.:

\[ F(X(t)) = F(X(0)) + \sum_{i=1}^{d} \int_{0}^{t} \frac{\partial F}{\partial x^{i}}(X(s)) \, dX_{i}(s) + \frac{1}{2} \sum_{i,j=1}^{d} \int_{0}^{t} \frac{\partial^{2} F}{\partial x^{i}\partial x^{j}}(X(s)) \, d[X_{i}, X_{j}](s), \quad \forall \, t \in [0, \infty). \]
Remark 5.4.28 The comments in Remark 5.4.25 in regard to Theorem 5.4.24 pertain also to Theorem 5.4.27.

Proof of Theorem 5.4.27: Let $D_n$, $n = 1, 2, \ldots$ be open bounded sets in $\mathbb{R}^d$ such that $D_n \subset D_{n+1} \subset D$ and $\overline{D_n} \subset D$ for each $n = 1, 2, \ldots$, and $D = \bigcup_{1 \leq n < \infty} D_n$. For each $n = 1, 2, \ldots$ put

$$T_n(\omega) \triangleq \inf\{t \in [0, \infty) : X(t, \omega) \in D^*_n \} \land n.$$  

In view of Condition 5.3.1 we shall suppose, without loss of generality, that $X(t, \omega) \in D, \forall t \in [0, \infty)$, and $\lim_{n \to \infty} T_n(\omega) = +\infty$ for each $\omega \in \Omega$. Proposition 3.3.7(b) ensures that each $T_n$ is an $\{\mathcal{F}_t\}$-stopping time. Fix some positive integer $n$, put $K \triangleq \overline{D_n}$, and use Theorem 5.4.26 to find a sequence of polynomials $F_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, 2, \ldots$, such that (5.191), (5.192) and (5.193) hold. By Remark 5.4.2 we see that $X^{T_n} \triangleq (X_1^{T_n}, X_2^{T_n}, \ldots, X_d^{T_n})$ is a $d$-dimensional continuous vector semimartingale with respect to $\{\mathcal{F}_t, t \in [0, \infty)\}$, hence, in view of Theorem 5.4.24 we have, a.s.:

$$F_k(X^{T_n}(t)) = F_k(X^{T_n}(0)) + \sum_{i=1}^d \int_0^t \frac{\partial F_k}{\partial x^i}(X^{T_n}(s)) \, dX^T_i(s) + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 F_k}{\partial x^i \partial x^j}(X^{T_n}(s)) \, d[X^T_i(s), X^T_j(s)], \quad \forall t \in [0, \infty).$$

(5.195)

Fix some arbitrary $t \in [0, \infty)$. By (5.191) we have

$$\lim_{k \to \infty} |F_k(X^{T_n}(t)) - F(X^{T_n}(t))| = 0 \quad \text{and} \quad \lim_{k \to \infty} |F_k(X^{T_n}(0)) - F(X^{T_n}(0))| = 0$$

for each $\omega \in \Omega$. Considering the second term on the right hand side of (5.195), let $M_i$ and $A_i$ be the local martingale and bounded variation parts respectively of $X_i^{T_n}$. Since $\overline{D_n}$ is closed and bounded, one easily sees that $M_i \in \mathbb{M}_{2,b}([\mathcal{F}_t], \mathbb{P})$ and there is some constant $B_n \in [0, \infty)$ such that

$$\left| \frac{\partial F_k}{\partial x^i}(X^{T_n}(s)) - \frac{\partial F}{\partial x^i}(X^{T_n}(s)) \right| \leq B_n,$$

and

(5.197)

$$\left| \frac{\partial^2 F_k}{\partial x^i \partial x^j}(X^{T_n}(s)) - \frac{\partial^2 F}{\partial x^i \partial x^j}(X^{T_n}(s)) \right| \leq B_n,$$

(5.198)

$\forall \omega \in \Omega, \forall s \in [0, \infty), \forall k = 1, 2, \ldots$. Thus, from (5.192), (5.197) and Proposition 5.3.48 we have

$$\lim_{k \to \infty} E \left| \int_0^t \frac{\partial F_k}{\partial x^i}(X^{T_n}(s)) \, dM_i(s) - \int_0^t \frac{\partial F}{\partial x^i}(X^{T_n}(s)) \, dM_i(s) \right|^2 = 0,$$

(5.199)

while (5.192), (5.193), (5.197), (5.198) and the Dominated Convergence Theorem 1.2.21 shows that

$$\lim_{k \to \infty} \int_0^t \frac{\partial F_k}{\partial x^i}(X^{T_n}(s)) \, dA_i(s) = \int_0^t \frac{\partial F}{\partial x^i}(X^{T_n}(s)) \, dA_i(s),$$

(5.200)
and
\[
\lim_{k \to \infty} \int_0^t \frac{\partial^2 F_k}{\partial x^i \partial x^j}(X^T_k(s)) \, d[X^T_k, X^T_k](s) = \int_0^t \frac{\partial^2 F}{\partial x^i \partial x^j}(X^T(s)) \, d[X^T, X^T](s),
\]
for each \( \omega \in \Omega \). Combining (5.195), (5.196), (5.200), and (5.201), and using Proposition 5.4.18(e), we find that
\[
F(X(t \wedge T_n)) = F(X(0)) + \sum_{i=1}^d \int_0^{t \wedge T_n} \frac{\partial F}{\partial x^i}(X(s)) \, dX_i(s) + \frac{1}{2} \sum_{i,j=1}^d \int_0^{t \wedge T_n} \frac{\partial^2 F}{\partial x^i \partial x^j}(X(s)) \, d[X_i, X_j](s) \text{ a.s.}
\]
Taking \( n \to \infty \) establishes (5.194).

Remark 5.4.29 Itô’s formula shows that a \( C^2 \)-mapping of a given continuous vector semimartingale is a continuous semimartingale, and gives its local martingale and bounded variation parts.

Remark 5.4.30 Notice from Proposition 5.4.6(d) that, in the special case where \( X_i \in \text{FV}^c\{\mathcal{F}_t\} \) for \( i = 1, 2, \ldots d \), the relation (5.194) takes the form
\[
F(X(t)) = F(X(0)) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x^i}(X(s)) \, dX_i(s), \quad \forall \ t \in [0, \infty),
\]
a result familiar from ordinary calculus.

5.5 Exponential Semimartingales

Our first application of Itô’s formula concerns a special class of semimartingales which have a particular exponential structure. Semimartingales of this kind will be seen in § 5.7 to play an essential role in the so-called Girsanov transformation of continuous local martingales.

Definition 5.5.1 Suppose that Condition 5.3.1 holds, and take \( X \in \text{SM}^c(\mathcal{F}_t, P) \). Then the process \( \{ \mathcal{E}(X)(t), \ t \in [0, \infty) \} \) defined by
\[
(5.202) \quad \mathcal{E}(X)(t) \triangleq \exp \left\{ X(t) - \frac{1}{2}[X](t) \right\}
\]
is called the **Itô exponential** of the continuous semimartingale \( X \). We note that \( \{ (\mathcal{E}(X)(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a continuous adapted process hence, in view of Remark 5.4.17, it is an element of \( \text{B}_{loc} \{ \mathcal{F}_t \} \).

Thus the stochastic integral \( \mathcal{E}(X) \cdot Y \) is defined for each \( X, Y \in \text{SM}^c(\mathcal{F}_t, P) \), and is a member of \( \text{SM}^c(\mathcal{F}_t, P) \).

Using Itô’s formula we can establish the basic properties of Itô exponentials of continuous semimartingales:
Thus, in order to evaluate the third term on the right hand side of (5.194), we only need the double

\[ E \]

from Definition 5.3.32 that

\[ E \]

From Proposition 5.4.6(e) and the fact that

\[ Y \]

(b) We have a.s.:

\[ E \]

\[ E \]

This fact, together with easy manipulations of (5.202), yields (5.204).

(c) Since

\[ X \]

follows. As for (5.204), observe from Proposition 5.4.6 that a.s.:

\[ E \]

Proof : (a) Define

\[ F \]

and

\[ F \]

Clearly, \( \{(Y_1(t), Y_2(t)), \mathcal{F}_t) ; t \in [0, \infty) \} \) is a 2-dimensional continuous vector semimartingale. Since \( F \) is a \( C^2 \)-mapping on \( \mathbb{R}^2 \) we see from Remark 5.4.29 that \( E(X) \equiv F(Y) \) is an element of \( SM^c(\{\mathcal{F}_t\}, P) \).

(b) We shall use Itô’s formula to establish (5.203). With the notation in (a) observe that

\[ \frac{\partial F}{\partial y_1}(y_1, y_2) = F(y_1, y_2), \quad \text{and} \quad \frac{\partial F}{\partial y_2}(y_1, y_2) = \frac{1}{2} F(y_1, y_2). \]

From Proposition 5.4.6(e) and the fact that \( Y_2 \in FV^{c,0}(\mathcal{F}_t) \), we have a.s.:

\[ [Y_1, Y_2](t) = [Y_2, Y_1](t) = [Y_2](t) = 0, \quad \forall t \in [0, \infty). \]

Thus, in order to evaluate the third term on the right hand side of (5.194), we only need the double derivative

\[ \frac{\partial^2 F}{\partial y_1 \partial y_2}(y_1, y_2) = F(y_1, y_2). \]

Substituting into (5.194), by Proposition 5.4.6(e) we get a.s.:

\[ F(Y(t)) = F(Y(0)) + \int_0^t F(Y(s)) \, dY_1(s) \]

\[ \quad - \frac{1}{2} \int_0^t F(Y(s)) \, dY_2(s) + \frac{1}{2} \int_0^t F(Y(s)) \, d[Y_1](s) \quad \text{a.s.} \]

Since \( Y_2(t) = [Y_1](t), \forall t \in [0, \infty) \), we see that the third and fourth terms in (5.205) cancel, and (5.203) follows. As for (5.204), observe from Proposition 5.4.6 that a.s.:

\[ [X + Y + [X, Y]](t) = [X](t) + 2[X, Y](t) + [Y](t), \quad \forall t \in [0, \infty). \]

This fact, together with easy manipulations of (5.202), yields (5.204).

(c) Since \( E(X) \in B_{loc}^{c} \{\mathcal{F}_t\} \) it is clear that \( E(X) \in L^2_{loc}(X, \{\mathcal{F}_t\}, P) \). Since \( X \in M^c_{loc}(\{\mathcal{F}_t\}, P) \) we see from Definition 5.3.32 that \( E(X) \bullet X \in M^c_{loc}(\{\mathcal{F}_t\}, P) \). This fact, together with (5.203), shows that \( E(X) \in M^c_{loc}(\{\mathcal{F}_t\}, P) \).
Remark 5.5.3 Suppose that $X \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$. Then $\mathcal{E}(X) \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, P)$ (by Proposition 5.5.2(c)), thus, in view of Proposition 4.6.7, it follows that $\{(\mathcal{E}(X)(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a non-negative supermartingale. We therefore have

\begin{equation}
E[\mathcal{E}(X)(t)] \leq E[\mathcal{E}(X)(0)] = 1, \quad \forall \ t \in [0, \infty);
\end{equation}

moreover, from Remark 4.4.8, there exists some a.s. unique $\mathcal{F}_\infty$-measurable random variable $Z$ such that $E|Z| < \infty$ and

\begin{equation}
\lim_{t \to \infty} \mathcal{E}(X)(t) = Z \quad \text{a.s.}
\end{equation}

Since the $\mathcal{E}(X)(t)$ are non-negative, without loss of generality we may suppose that $Z(\omega) \in [0, \infty)$, $\forall \ \omega \in \Omega$. Suppose that $\{t_n\} \subset [0, \infty)$ is a sequence with $t_n \to \infty$. In view of (5.206), (5.207) and Fatou’s Theorem, one sees that

\begin{equation}
EZ \leq \liminf_{n \to \infty} E[\mathcal{E}(X)(t_n)] \leq 1.
\end{equation}

Henceforth, we shall write $\mathcal{E}(X)(\infty)$ to denote an arbitrarily chosen but fixed random variable $Z$, taking values in $[0, \infty)$, such that (5.207) and (5.208) hold.

Proposition 5.5.4 Suppose that $X \in M_{loc}^{c,0}(\{\mathcal{F}_t\}, \mathcal{P})$, and define the events

$$A \triangleq \{\mathcal{E}(X)(\infty) = 0\} \quad \text{and} \quad B \triangleq \{[X](\infty) = \infty\},$$

where $[X](\infty)$ is defined by (4.134). Then $A = B$ a.s. (recall Remark 1.4.17).

Remark 5.5.5 Proposition 5.5.4 essentially says that, except for a set of $\mathcal{P}$-measure zero, one has $\mathcal{E}(X)(\infty) = 0$ if and only if $[X](\infty) = \infty$. The exceptional set of $\mathcal{P}$-measure zero arises, of course, because $\mathcal{E}(X)(\infty)$ and $[X](\infty)$ are arbitrary choices of random variables which are uniquely defined to within a $\mathcal{P}$-null set only.

Proof of Proposition 5.5.4: From (5.202) one sees that

\begin{equation}
\mathcal{E}(X)(t) \mathcal{E}(-X)(t) = \exp\{-[X](t)\}, \quad \forall \ t \in [0, \infty).
\end{equation}

Define the events

$$\Omega_1 \triangleq \{\lim_{t \to \infty} \mathcal{E}(X)(t) = \mathcal{E}(X)(\infty)\}, \quad \Omega_2 \triangleq \{\lim_{t \to \infty} \mathcal{E}(-X)(t) = \mathcal{E}(-X)(\infty)\},$$

and observe from Remark 5.5.3 that $P(\Omega_1 \cap \Omega_2) = 1$. Now fix some $\omega \in A \cap \Omega_1 \cap \Omega_2$. Then $\lim_{t \to \infty} \mathcal{E}(X)(t, \omega) = \mathcal{E}(X)(\infty, \omega) = 0$ and $\lim_{t \to \infty} \mathcal{E}(-X)(t, \omega) = \mathcal{E}(-X)(\infty, \omega) < \infty$, whence, taking $t \to \infty$ in (5.209), one finds $\lim_{t \to \infty} \exp\{-[X](t, \omega)\} = 0$, and therefore $\omega \in B$. It follows that
\(A \cap \Omega_1 \cap \Omega_2 \subset B \cap \Omega_1 \cap \Omega_2\), whence \(A - B \subset (\Omega_1 \cap \Omega_2)^c\) which establishes \(P(A - B) = 0\). It remains to show that \(P(B - A) = 0\). To this end observe from (5.202) that

\[
(5.210) \quad \mathcal{E}(X)(t) = \left\{ \mathcal{E} \left( \frac{X}{2} \right)(t) \right\}^2 \exp \left\{ \frac{-[X](t)}{8} \right\}, \quad \forall t \in [0, \infty),
\]

define the event

\[
\Omega_3 \triangleq \{ \lim_{t \to \infty} \mathcal{E}(X/2)(t) = \mathcal{E}(X/2)(\infty) \},
\]

and observe from Remark 5.5.3 that \(P(\Omega_1 \cap \Omega_3) = 1\). Now fix some \(\omega \in B \cap \Omega_1 \cap \Omega_3\). Then \(\lim_{t \to \infty} \mathcal{E}(X/2)(t, \omega) = \mathcal{E}(X/2)(\infty, \omega) < \infty\) whence, taking \(t \to \infty\) in (5.210), clearly \(\lim_{t \to \infty} \mathcal{E}(X)(t, \omega) = 0\), thus \(\omega \in A\). It follows that \(B \cap \Omega_1 \cap \Omega_3 \subset A \cap \Omega_1 \cap \Omega_3\), giving \(P(B - A) = 0\).

**Definition 5.5.6** Given an \(\mathbb{R}\)-valued continuous process \(\{Y(t); t \in [0, \infty)\}\) and a probability measure \(P\) on the measure space \((\Omega, \mathcal{F})\), we say that \(\{Y(t); t \in [0, \infty)\}\) is **\(P\)-strictly positive** when

\[
P[\omega \in \Omega : Y(t, \omega) > 0, \quad \forall t \in [0, \infty]] = 1.
\]

**Remark 5.5.7** Suppose that Condition 5.3.1 holds, that \(\{(Y(t), \mathcal{F}_t); t \in [0, \infty)\}\) is a continuous adapted \(P\)-strictly positive process on \((\Omega, \mathcal{F}, P)\), and let \(N\) denote the \(P\)-null event in \(\mathcal{F}\) comprising all \(\omega\) at which one fails to get \(Y(t, \omega) > 0, \forall t \in [0, \infty)\). Under these conditions we shall abuse notation, much as we did in Remark 1.2.36, and write \(Y^{-1} \triangleq \{(Y^{-1}(t), \mathcal{F}_t); t \in [0, \infty)\}\) for the process defined by \(Y^{-1}(t, \omega) \triangleq \frac{1}{Y(t, \omega)}, \forall t \in [0, \infty)\), when \(\omega \notin N\); and \(Y^{-1}(t, \omega) \triangleq 0, \forall t \in [0, \infty)\), when \(\omega \in N\). Since Condition 5.3.1 ensures \(N \in \mathcal{F}_0 \subset \mathcal{F}_t, \forall t \in [0, \infty)\), it follows that \(Y^{-1}\) is a continuous adapted process, and hence by Remark 5.4.17 we have \(Y^{-1} \in \mathcal{B}_{loc}\{\mathcal{F}_t\}\). Thus, for any \(X \in \text{SM}^c(\{\mathcal{F}_t\}, P)\), the Itô stochastic integral \((Y^{-1} \bullet X)\) is properly defined (see Remark 5.4.16).

Of particular interest in applications are strictly positive continuous local martingales. As the next proposition shows, a strictly positive continuous local martingale is necessarily the Itô exponential of some uniquely defined continuous local martingale:

**Proposition 5.5.8** Suppose that Condition 5.3.1 holds, and let \(Y \in \mathcal{M}_{loc}^c(\{\mathcal{F}_t\}, P)\) be \(P\)-strictly positive. Then there exists some \(X \in \mathcal{M}_{loc}^c(\{\mathcal{F}_t\}, P)\) such that a.s.

\[
(5.211) \quad Y(t) = \mathcal{E}(X)(t), \quad \forall t \in [0, \infty).
\]

Moreover, if \(\tilde{X} \in \mathcal{M}_{loc}^c(\{\mathcal{F}_t\}, P)\) is such that a.s. \(Y(t) = \mathcal{E}(\tilde{X})(t), \forall t \in [0, \infty)\), then \(\{X(t); t \in [0, \infty)\}\) and \(\{\tilde{X}(t); t \in [0, \infty)\}\) are indistinguishable, and the continuous local martingale \(X\) in (5.211) is given by

\[
(5.212) \quad X(t) = \log Y(0) + \int_0^t Y^{-1}(s) \, dY(s), \quad \forall t \in [0, \infty).
\]
Thus, \( T \) conditions on non-trivial matter. We shall now give an important result which establishes rather general sufficient can make the stronger statement that
\[
\lim_{n \to \infty} E[Y_n(t)] = 1, \quad \forall n = 1, 2, \ldots.
\]

To establish (5.211) use Itô’s formula (Theorem 5.4.27) with \( D = (0, \infty) \) and \( F : D \to IR \) defined by \( F(x) = \log(x) \) to obtain a.s.:

\[
\log Y(t) = \log Y(0) + \int_0^t Y^{-1}(s) dY(s) - \frac{1}{2} \int_0^t Y^{-2} d[Y](s), \quad \forall t \in [0, \infty).
\]

Now, for \( X \) defined by (5.212), Proposition 5.4.18(a) ensures that a.s.:

\[
[X](t) = \int_0^t Y^{-2}(s) d[Y](s), \quad \forall t \in [0, \infty).
\]

Putting (5.212), (5.213) and (5.214) together we get a.s. \( \log Y(t) = X(t) - (1/2)[X](t), \forall t \in [0, \infty) \), which is (5.211). As for uniqueness to within indistinguishability, we observe that \( \mathcal{E}(X) = \mathcal{E}(X^\Delta)(t), \forall t \in [0, \infty) \), implies a.s.:

\[
X(t) - X^\Delta(t) = \frac{1}{2}([X](t) - [X^\Delta](t)), \quad \forall t \in [0, \infty).
\]

Thus, \( X - X^\Delta \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P) \cap \mathcal{FV}^c(\{\mathcal{F}_t\}) \), whence Lemma 4.7.30 establishes indistinguishability of \( X \) and \( X^\Delta \).

Proposition 5.5.2(c) says that if \( \{(X(t), \mathcal{F}_t) ; t \in [0, \infty]\} \) is a continuous local martingale then so also is \( \{(\mathcal{E}(X)(t), \mathcal{F}_t) ; t \in [0, \infty]\} \). In Section 5.7 it will be particularly important to know when we can make the stronger statement that \( \{(\mathcal{E}(X)(t), \mathcal{F}_t) ; t \in [0, \infty]\} \) is actually a martingale. Securing reasonably general conditions on \( \{(X(t), \mathcal{F}_t) ; t \in [0, \infty]\} \) which allow us to make this assertion is a non-trivial matter. We shall now give an important result which establishes rather general sufficient conditions on \( \{(X(t), \mathcal{F}_t) ; t \in [0, \infty]\} \) enabling one to conclude that \( \mathcal{E}(X) \in \mathcal{M}^c(\{\mathcal{F}_t\}, P) \). In order to establish this result we need the following simple lemma:

**Lemma 5.5.9** Suppose that Condition 5.3.1 holds. If \( X \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P) \) then, for each \( \{\mathcal{F}_t\} \)-stopping time \( T \), we have

\[
1 \geq E[\mathcal{E}(X)(T)] \geq E[\mathcal{E}(X)(\infty)] \geq 0.
\]

Moreover, \( \{(\mathcal{E}(X)(t), \mathcal{F}_t) ; t \in [0, \infty]\} \) is a uniformly integrable continuous martingale if and only if \( E[\mathcal{E}(X)(\infty)] = 1 \).

Proof: Let \( T \) be some \( \{\mathcal{F}_t\} \)-stopping time. Then, by Theorem 4.5.4, we have \( E[\mathcal{E}(X)(T)] \geq E[\mathcal{E}(X)(T + n)], \forall n = 1, 2, \ldots \), so that Fatou’s Theorem gives

\[
E[\mathcal{E}(X)(T)] \geq \lim_{n \to \infty} E[\mathcal{E}(X)(T + n)] \geq E[\mathcal{E}(X)(\infty)] \geq 0.
\]

Likewise, by Theorem 4.5.4 and (5.206), one also has \( 1 = E[\mathcal{E}(X)(0)] \geq E[\mathcal{E}(X)(T \wedge n)] \geq 0, \forall n = 1, 2, \ldots \), so that Fatou’s Theorem gives \( 1 \geq \liminf_{n \to \infty} E[\mathcal{E}(X)(T \wedge n)] \geq E[\mathcal{E}(X)(T)], \) hence (5.215) follows.
Next, suppose that $E[\mathcal{E}(X)(\infty)] = 1$. Then, in view of (5.215), one has $E[\mathcal{E}(X)(t)] = 1, \forall t \in [0, \infty]$, thus Remark 4.1.2 ensures that $\mathcal{E}(X) \in \mathbf{M}^c(\{\mathcal{F}_t\}, P)$. It remains to show that the random variables \{\mathcal{E}(X)(t), t \in [0, \infty)\} are a uniformly integrable collection. To this end, fix some sequence \{\{t_n, n = 1, 2, \ldots\} \subset [0, \infty)\) with $\lim_{n \to \infty} t_n = \infty$. Then $E[\mathcal{E}(X)(t_n)] = E[\mathcal{E}(X)(\infty)] = 1, \forall n = 1, 2, \ldots$, hence in view of the non-negativity of the random variables $\mathcal{E}(X)(t_n)$ and $\mathcal{E}(X)(\infty)$, we see from Corollary 1.2.22 that $\lim_{n \to \infty} E[\mathcal{E}(X)(t_n) - \mathcal{E}(X)(\infty)] = 0$. By the arbitrary choice of the sequence \{\{t_n\} it then follows that

$$\lim_{t \to \infty} E[\mathcal{E}(X)(t) - \mathcal{E}(X)(\infty)] = 0.$$  

Fixing some $\epsilon \in (0, \infty)$, we see from (5.216) that there is some $T(\epsilon) \in [0, \infty)$ such that $E[\mathcal{E}(X)(t) - \mathcal{E}(X)(\infty)] < \epsilon/2, \forall t \in [T(\epsilon), \infty)$. Thus, for arbitrary $A \in \mathcal{F}$, one obtains

$$E[\mathcal{E}(X)(t); A] \leq E[\mathcal{E}(X)(\infty); A] + E[\mathcal{E}(X)(t) - \mathcal{E}(\infty)] \leq E[\mathcal{E}(X)(\infty); A] + \epsilon/2, \forall t \in [T(\epsilon), \infty).$$

Since $\mathcal{E}(X)(\infty)$ is integrable, it follows from Theorem 1.2.20 and (5.217) that there is some $\delta(\epsilon) \in (0, \infty)$ such that $E[\mathcal{E}(X)(t); A] < \epsilon$, for all $t \in [T(\epsilon), \infty)$ and all $A \in \mathcal{F}$ with $P(A) < \delta(\epsilon)$. In order to conclude that \{\mathcal{E}(X)(t), t \in [0, \infty)\} is uniformly integrable it remains to show that the collection \{\mathcal{E}(X)(t), t \in [0, T(\epsilon)]\} is uniformly integrable. However, this follows from Theorem 2.6.7 since $\mathcal{E}(X)(t) = E[\mathcal{E}(X)(T(\epsilon)) | \mathcal{F}_t]$ a.s. $\forall t \in [0, T(\epsilon)]$.

The next result formulates a rather general sufficient condition on $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ which entails $\mathcal{E}(X) \in \mathbf{M}^c(\{\mathcal{F}_t\}, P)$:

**Theorem 5.5.10 (Novikov)** Suppose that $\{\mathcal{F}_t, t \in [0, \infty)\}$ is a filtration in the probability space $(\Omega, \mathcal{F}, P)$ subject to Condition 5.3.1. If $X \in \mathbf{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P)$, and

$$E \left[ \exp \left( \frac{1}{2} [X](\infty) \right) \right] < \infty,$$

then $\{(\mathcal{E}(X)(t), \mathcal{F}_t), t \in [0, \infty)\}$ is a uniformly integrable martingale.

**Proof:** In view of Lemma 5.5.9 it is enough to show that $E[\mathcal{E}(X)(\infty)] = 1$. Now clearly $x^2 \leq \exp(x/2), \forall x \in [0, \infty)$, thus (5.218) shows that $[X](\infty) \in L^2(\Omega, \mathcal{F}, P)$, and Proposition 4.7.33 then shows that $X \in \mathbf{M}^{c,0}_{2b}(\{\mathcal{F}_t\}, P)$. According to Theorem 4.4.12 there exists some a.s. unique $\mathcal{F}_\infty$-measurable random variable $X(\infty) \in L^2(\Omega, \mathcal{F}, P)$ which is the a.s. (and $L^2$) limit of $X(t)$ as $t \to \infty$, and

$$X(t) = E[X(\infty) | \mathcal{F}_t] \text{ a.s. \forall } t \in [0, \infty).$$

Now put

$$Z(t) \triangleq \exp \left( \frac{1}{2} X(t) \right), \forall t \in [0, \infty),$$

243
and observe that \{ (Z(t), \mathcal{F}_t); t \in [0, \infty) \} is a non-negative continuous uniformly integrable submartingale. Indeed, since \( X(\infty), [X](\infty) \) and \( \mathcal{E}(X)(\infty) \) are the a.s. limits (as \( t \to \infty \)) of \( X(t), [X](t) \) and \( \mathcal{E}(X)(t) \) respectively, we see from (5.202) that

\[
\mathcal{E}(X)(\infty) = \exp \left\{ X(\infty) - \frac{1}{2} [X](\infty) \right\} \text{ a.s.}
\]

whence

\[
(5.221) \quad \exp \left( \frac{1}{2} X(\infty) \right) = \left\{ \mathcal{E}(X)(\infty) \right\}^{1/2} \left\{ \exp \left( \frac{1}{2} [X](\infty) \right) \right\}^{1/2} \text{ a.s.}
\]

Then (5.218), (5.221) and the Cauchy-Schwarz inequality give

\[
(5.222) \quad E \left[ \exp \left( \frac{1}{2} X(\infty) \right) \right] \leq \left\{ \mathcal{E}(X)(\infty) \right\}^{1/2} \left\{ E \left[ \exp \left( \frac{1}{2} [X](\infty) \right) \right] \right\}^{1/2} < \infty
\]

(since \( 0 \leq \mathcal{E}(X)(\infty) \leq 1 \), by Remark 5.5.3). Now, by (5.220), (5.219) and Jensen’s inequality for conditional expectations (Theorem 1.4.20) one obtains

\[
(5.223) \quad 0 \leq Z(t) \leq E \left[ \exp \left( \frac{1}{2} X(\infty) \right) \mid \mathcal{F}_t \right] \text{ a.s. } \forall t \in [0, \infty),
\]

whence, from (5.222), (5.223), Theorem 2.6.7, and Proposition 4.1.12, one sees that \{ (Z(t), \mathcal{F}_t); t \in [0, \infty) \} is a continuous uniformly integrable submartingale. Then Theorem 4.4.11(a) ensures that there is some \( Z(\infty) \in L^1(\Omega, \mathcal{F}, P) \) such that \( \lim_{t \to \infty} Z(t) = Z(\infty) \) (a.s. and \( L^1 \)-sense) and \{ (Z(t), \mathcal{F}_t), t \in [0, \infty) \} is a closed submartingale. Now let \( \mathcal{U} \) denote the set of all \( \{ \mathcal{F}_t \} \)-stopping times (not necessarily bounded!). Then (see Theorem 4.5.4) we have

\[
(5.224) \quad Z(T) \leq E \left[ Z(\infty) \mid \mathcal{F}_T \right], \text{ a.s. } \forall t \in \mathcal{U},
\]

and, taking \( t \to \infty \) in (5.220), we find

\[
(5.225) \quad Z(\infty) = \exp \left( \frac{1}{2} X(\infty) \right), \text{ a.s.}
\]

Now define

\[
(5.226) \quad Z_\lambda(t, \omega) \triangleq \exp \left\{ \frac{\lambda X(t, \omega)}{1 + \lambda} \right\}, \quad \forall \lambda \in [0, 1), \ \forall \omega \in \Omega, \ \forall t \in [0, \infty],
\]

and fix some \( \lambda \in [0, 1) \). Observe that \( 0 \leq Z_\lambda(t, \omega) \leq 1 \) when \( X(t, \omega) \leq 0 \), and since the mapping \( \lambda \to \lambda/(1 + \lambda) \) is nondecreasing over the interval \( 0 \leq \lambda \leq 1 \), it also follows that \( 0 \leq Z_\lambda(t, \omega) \leq Z(t, \omega) \) when \( X(t, \omega) \geq 0 \). Thus,

\[
(5.227) \quad 0 \leq Z_\lambda(t, \omega) \leq 1 + Z(t, \omega), \quad \forall \lambda \in [0, 1), \ \forall \omega \in \Omega, \ \forall t \in [0, \infty],
\]

and, in view of (5.224), we get

\[
Z_\lambda(T) \leq 1 + E \left[ Z(\infty) \mid \mathcal{F}_T \right], \text{ a.s. } \forall \lambda \in [0, 1), \ T \in \mathcal{U}.
\]
Then, for each $\lambda \in [0,1)$ and $T \in \mathcal{U}$, we have

$$E[I_A Z_\lambda(T)] \leq P(A) + E[I_A Z(\infty)], \quad \forall A \in \mathcal{F}_T,$$

and therefore, since $\{E(\lambda X)(T) > c\} \in \mathcal{F}_T$, we find

\begin{align*}
(5.228) \quad E[Z_\lambda(T); \ E(\lambda X)(T) > c] & \leq P[E(\lambda X)(T) > c] + E[Z(\infty); \ E(\lambda X)(T) > c],
\end{align*}

for each $\lambda \in [0,1)$, each $T \in \mathcal{U}$, and each $c \in (0, \infty)$. Now Lemma 5.5.9 gives $E[E(\lambda X)(T)] \leq 1$ for each $T \in \mathcal{U}$ and $\lambda \in [0,1)$ so that the Markov inequality gives

\begin{align*}
(5.229) \quad P[E(\lambda X)(T) > c] & \leq \frac{1}{c}, \quad \forall \lambda \in [0,1), \ \forall T \in \mathcal{U}, \ \forall c \in (0, \infty).
\end{align*}

Since $E[Z(\infty)] < \infty$, it follows from Theorem 1.2.20 with (5.229) and (5.228), that

\begin{align*}
(5.230) \quad \lim_{c \to \infty} \left\{ \sup_{T \in \mathcal{U}} E[Z_\lambda(T); \ E(\lambda X)(T) > c] \right\} = 0,
\end{align*}

for each $\lambda \in [0,1)$. We shall now establish that $\{(E(\lambda X)(t), \mathcal{F}_t), \ t \in [0, \infty]\}$ is a uniformly integrable martingale for each $\lambda \in [0,1)$. From (5.226), the definition of Itô exponentials, and an easy calculation, we get

\begin{align*}
(5.231) \quad E(\lambda X)(t) &= [E(X)(t)]^{1/\lambda^2} [Z_\lambda(t)]^{1-\lambda^2}, \quad \forall \omega \in \Omega, \ \forall t \in [0, \infty).
\end{align*}

Now fix some $\lambda \in [0,1)$, some $A \in \mathcal{F}$ and some $T \in \mathcal{U}$, and observe from (5.231) that

$$I_A E(\lambda X)(T) = [E(X)(T)]^{1/\lambda^2} [I_A Z_\lambda(T)]^{1-\lambda^2}, \quad \forall \omega \in \Omega,$$

from which it easily follows that

\begin{align*}
(5.232) \quad E[I_A E(\lambda X)(T)] & \leq \{E[E(X)(T)]\}^{1/\lambda^2} \{E[I_A Z_\lambda(T)]\}^{1-\lambda^2} \leq \{E[I_A Z_\lambda(T)]\}^{1-\lambda^2}.
\end{align*}

Here we have used Hölder’s inequality (see Theorem 1.2.25) with conjugate exponents $p = 1/\lambda^2$, $q = 1/(1 - \lambda^2)$ at the first inequality of (5.232), and the second inequality follows since Lemma 5.5.9 shows that $0 \leq E[E(X)(T)] \leq 1$. In view of (5.232) and (5.230), for each $\lambda \in [0,1)$ we obtain

$$\lim_{c \to \infty} \left\{ \sup_{T \in \mathcal{U}} E[E(\lambda X)(T); \ E(\lambda X)(T) > c] \right\} = 0,$$

as required to see that $\{E(\lambda X)(T), \ T \in \mathcal{U}\}$ is uniformly integrable; since $E(\lambda X) \in \mathcal{M}^c_{loc}(\mathcal{F}_t, P)$ one then sees from Proposition 4.6.8 that $\{E(\lambda X)(t), \ t \in [0, \infty]\}$ is a martingale, which is obviously uniformly integrable. In particular, this implies that $E[E(\lambda X)(\infty)] = 1, \ \forall \lambda \in [0,1)$. To complete the proof of the theorem, let $A \triangleq \Omega$ and $T \triangleq \infty$ in (5.232), so that

\begin{align*}
(5.233) \quad 1 & \leq \{E[E(X)(\infty)]\}^{1/\lambda^2} \{E[Z_\lambda(\infty)]\}^{1-\lambda^2}, \quad \forall \lambda \in [0,1).
\end{align*}
Fix some sequence \( \{ \lambda_n, n = 1, 2, \ldots \} \subset [0, 1) \) such that \( \lim_{n \to \infty} \lambda_n = 1 \). In view of (5.226) and (5.225) we have \( \lim_{n \to \infty} Z_{\lambda_n}(\infty) = Z(\infty) \), \( \forall \omega \in \Omega \), hence (5.227), the integrability of \( Z(\infty) \) and the Lebesgue dominated convergence theorem yields \( \lim_{n \to \infty} E[Z_{\lambda_n}(\infty)] = E[Z(\infty)] \), which in turn gives
\[
\lim_{n \to \infty} \{E[Z_{\lambda_n}(\infty)]\}^{1 - \lambda_n^2} = 1.
\]
Now, upon taking \( \lambda \triangleq \lambda_n \) and \( n \to \infty \) in (5.233), we obtain
\[
E[E[Z_{\lambda_n}(\infty)]|] = 1, \quad \forall t \in [0, \infty).
\]
(5.234)

As an obvious consequence of Theorem 5.5.10 we have:

**Corollary 5.5.11** Suppose that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is a filtration in the probability space \((\Omega, \mathcal{F}, P)\) subject to Condition 5.3.1. If \( X \in \mathcal{M}_{\text{loc}}^0(\mathcal{F}_t, P) \), and \( E[\exp([X](t)/2)] < \infty \), \( \forall t \in [0, \infty) \), then \( \{(\mathcal{E}(X)(t), \mathcal{F}_t), t \in [0, \infty)\} \) is a martingale.

**Remark 5.5.12** In Example 2.5.9 we saw an example of a discrete-parameter martingale which converges a.s. to an integrable limiting random variable, but for which \( L^1 \)-convergence to the same limit fails. The results of this section can be used to suggest similar examples in a continuous-parameter context. Thus, suppose that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a standard scalar Wiener process on \((\Omega, \mathcal{F}, P)\). Then it follows at once from Corollary 5.5.11 and Remark 4.7.17 that \( \{(\mathcal{E}(X)(t), \mathcal{F}_t), t \in [0, \infty)\} \) is a martingale, and therefore
\[
(5.234) \quad E[\mathcal{E}(X)(t)] = 1, \quad \forall t \in [0, \infty).
\]
Now, from Proposition 5.5.4 and Remark 4.7.17, we see that \( P\{\mathcal{E}(X)(\infty) = 0\} = 1 \), thus in particular, \( E[\mathcal{E}(X)(\infty)] = 0 \). In view of (5.234) it follows that we cannot get \( L^1 \)-convergence of \( \mathcal{E}(X)(t) \) to \( \mathcal{E}(X)(\infty) \).

### 5.6 Levy’s Characterization of a Wiener Process

Remark 4.7.25 establishes the relation (4.110) for the co-quadratic variations of the scalar components of a \( d \)-dimensional standard Wiener process. A profound theorem of Levy asserts a partial converse, namely that any vector of continuous local martingales for which this relation holds is necessarily a Wiener process. In the present section our goal is to show how Itô’s formula can be used to give a surprisingly easy proof of Levy’s theorem. To this end we need the following simple result:

**Lemma 5.6.1** Suppose \( X \) is an \( \mathbb{R}^d \)-valued random vector on some probability space \((\Omega, \mathcal{F}, P)\) and \( \mathcal{G} \subset \mathcal{F} \) is a sub-\( \sigma \)-algebra of \( \mathcal{F} \). If \( Q \) is a \( d \) by \( d \) symmetric positive semidefinite matrix, and
\[
(5.235) \quad E[\exp\{i(\theta, X)\} | \mathcal{G}] = \exp\left\{-\frac{1}{2}(\theta, Q\theta)\right\} \quad \text{a.s.}
\]
for each \( \theta \in \mathbb{R}^d \), then we have
(i) \( X \sim N(0, Q) \), and
(ii) \( X \) is independent of \( \mathcal{G} \).
we have
\( \theta \)
or, equivalently,
\( t \\forall \)
(5.240)
Proof: Put
\( j, k \)
for each
(5.239)
\[0 \) \( \in \Omega \) adapted process which is null at the origin and defined on a probability space
\( X \)
Suppose that
Theorem 5.6.2 (Levy, Kunita-Watanabe)
The main result of this section is as follows:
Proof: Upon taking expectations of (5.235), and writing \( \phi_X : R^d \to C \) for the characteristic function of \( X \), we get
(5.236)
\[ \phi_X(\theta) = \exp \left\{ -\frac{1}{2}(\theta, Q\theta) \right\}, \quad \forall \theta \in R^d, \]
whence (Definition 1.4.10) \( X \sim N(0, Q) \). To establish independence of \( X \) and \( G \), let \( Z \) be an arbitrary \( G \)-measurable random variable. Now, for any \( \beta \in R \), we see that \( \exp(i\beta Z) \) is a \( C \)-valued \( G \)-measurable function of \( \omega \), hence, by Theorem 1.4.15(c) and (5.235) we have
(5.237)
\[ E [\exp\{i(\theta, X) + i\beta Z\} | G] = E [\exp\{i(\theta, X)\} | G] \exp(i\beta Z) \]
\[ = \exp \left\{ -\frac{1}{2}(\theta, Q\theta) \right\} \exp(i\beta Z), \quad \text{a.s.} \]
Taking expectations in (5.237) and using (5.236) we obtain
(5.238)
\[ E[\exp\{i(\theta, X) + i\beta Z\}] = E[\exp\{i(\theta, X)\}]E[\exp\{i\beta Z\}], \quad \forall \theta \in R^d, \quad \forall \beta \in R. \]
In view of Theorem 1.4.8 and (5.238) it follows that \( X \) is independent of \( Z \), and in particular \( X \) is independent of \( I_A, \forall A \in G \). Thus \( X \) is independent of \( G \).

The main result of this section is as follows:

**Theorem 5.6.2 (Levy, Kunita-Watanabe)** Suppose that \( \{(W(t), F_t); t \in [0, \infty)\} \) is an \( R^d \)-valued adapted process which is null at the origin and defined on a probability space \( (\Omega, F, P) \). If \( \{(W^k(t), F_t); t \in [0, \infty)\} \) is a continuous local martingale for each \( k = 1, 2, \ldots d, \) and we have a.s.
(5.239)
\[ [W^j, W^k](t) = t \delta_{j,k}, \quad \forall t \in [0, \infty), \]
for each \( j, k = 1, 2, \ldots d, \) then \( \{(W(t), F_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process.
Proof: Put
(5.240)
\[ Z^\theta(t, \omega) \triangleq \exp \left\{ i(\theta, W(t, \omega)) + \frac{1}{2}|\theta|^2t \right\}, \]
\( \forall t \in [0, \infty), \forall \omega \in \Omega, \forall \theta \in R^d \). Clearly \( \{(Z^\theta(t), F_t); t \in [0, \infty)\} \) is continuous adapted process for each \( \theta \in R^d \). If it can be shown that \( Z^\theta \in M^c(F_t, P) \) for each \( \theta \in R^d \) then, for \( s, t \in [0, \infty), s < t, \) we have
\[ E [\exp\{i(\theta, W(t) + (1/2)|\theta|^2t)\} | F_s] = \exp\{i(\theta, W(s)) + (1/2)|\theta|^2s\} \quad \text{a.s.} \]
or, equivalently,
(5.241)
\[ E [\exp\{i(\theta, W(t) - W(s))\} | F_s] = \exp\{-(1/2)|\theta|^2(t - s)\} \quad \text{a.s.} \]
for each \( \theta \in R^d \). Thus, from Lemma 5.6.1, we obtain \( W(t) - W(s) \sim N(0, (t - s)I_d) \) and \( W(t) - W(s) \) is independent of \( F_s \). We have thus verified all conditions of Definition 3.2.10, so that \( \{(W(t), F_t); t \in [0, \infty)\} \) is a \( d \)-dimensional standard Wiener process.

247
To show that $Z^\theta \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)$, $\forall \theta \in \mathbb{R}^d$, fix some arbitrary $\theta \in \mathbb{R}^d$ and put  

$$F(t, x) \triangleq \exp \left\{ i\langle \theta, x \rangle + \frac{1}{2} |\theta|^2 t \right\}, \quad \forall x \in \mathbb{R}^d, \forall t \in [0, \infty).$$  

Then $Z^\theta(t) = F(t, W(t))$ and we can use Itô’s formula to write a.s.:  

\begin{align}
5.242 \theta(t) &= 1 + \sum_{j=1}^d \int_0^t \frac{\partial F}{\partial x^j}(s, W(s)) dW^j(s) \\
&+ \int_0^t \frac{\partial F}{\partial t}(s, W(s)) ds + \frac{1}{2} \sum_{j,k=1}^d \int_0^t \frac{\partial^2 F}{\partial x^j \partial x^k}(s, W(s)) d[W^j, W^k](s), \quad \forall t \in [0, \infty).
\end{align}  

Using (5.239) it is easily verified that the third and fourth terms on the right hand side of (5.242) cancel. As for the second term, the integrand of each stochastic integral is clearly a member of $\mathcal{B}_{\text{loc}}\{\mathcal{F}_t\}$. Since $W^j \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)$ it follows that these stochastic integrals are members of $\mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)$, and thus $Z^\theta \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)$. Now it follows from (5.240) that  

$$|Z^\theta(t, \omega)| \leq \exp \left\{ \frac{1}{2} |\theta|^2 t \right\}, \quad \forall \omega \in \Omega, \forall t \in [0, \infty).$$  

Thus, if $\{T_n, \ n = 1, 2, \ldots\}$ is a localizing sequence for $Z^\theta \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)$, then the sequence of random variables $\{Z^\theta(t \wedge T_n), \ n = 1, 2, \ldots\}$ is uniformly bounded, hence uniformly integrable, for each $t \in [0, \infty)$. By Proposition 4.6.8 we get $Z^\theta \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)$, as required. \[\Box\]  

**Remark 5.6.3** The elegant proof we have just given is due to Kunita and Watanabe [21] and amply demonstrates the power of Itô’s formula for studying continuous-parameter stochastic processes. Earlier proofs of Theorem 5.6.2 were lengthy and involved, and usually relied on an application of the central limit theorem for martingales (see e.g. Theorem 11.9 in Chapter VII of Doob [7]).  

### 5.7 Changes of Measure and the Girsanov Transformation

In this section we shall use Itô’s formula to address a rather subtle question which turns out to have many ramifications for stochastic processes in general, as well as considerable relevance for specific applications in such diverse areas as nonlinear filtering, the pricing of derivative securities, and quantum mechanics. Suppose that $\{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\}$ is a continuous martingale on a probability space $(\Omega, \mathcal{F}, P)$. If one replaces the probability measure $P$ on $(\Omega, \mathcal{F})$ with some other probability measure $\tilde{P}$, then $\{(X(t), \mathcal{F}_t); \ t \in [0, \infty)\}$ is still a continuous adapted process on the new probability space $(\Omega, \mathcal{F}, \tilde{P})$, but in general it is no longer a martingale, since, for $s < t$, we have no reason at all to expect  

$$\tilde{E} [X_t | \mathcal{F}_s] = X_s \quad \tilde{P} - \text{a.s.}$$  

when the conditional expectation on the left hand side is calculated with respect to the new measure $\tilde{P}$. The same remarks continue to hold, of course, for continuous local martingales: if $X \in \mathcal{M}^c_{\text{loc}}(\{\mathcal{F}_t\}, P)$
then it is not generally the case that \( X \in \mathcal{M}^{\text{loc}}_{\text{c}}(\{F_t\}, \tilde{P}) \). Suppose now that the probability measure \( \tilde{P} \), instead of just being arbitrarily chosen, is equivalent to \( P \) on \( \mathcal{F} \) (recall Definition 1.2.35). If \( X \in \mathcal{M}^{\text{loc}}_{\text{c}}(\{F_t\}, P) \) then the main result of this section, namely the Girsanov theorem, asserts that \( X \in \mathcal{S}\mathcal{M}^{\text{c}}(\{F_t\}, \tilde{P}) \), and explicitly characterizes the bounded variation part of \( X \) as well. Equivalently, the Girsanov theorem determines some \( A \in \mathcal{F}^{\text{c}} \{F_t\} \) such that \( X - A \in \mathcal{M}^{\text{loc}}_{\text{c}}(\{F_t\}, \tilde{P}) \), that is the compensated process \( X - A \) is a continuous local martingale when \( \tilde{P} \) is the measure on \( (\Omega, \mathcal{F}) \). Thus, although \( X \) is generally not a member of \( \mathcal{M}^{\text{loc}}_{\text{c}}(\{F_t\}, \tilde{P}) \), if one subtracts from \( X \) a suitable continuous adapted process of locally bounded variation, then the result is indeed a member of \( \mathcal{M}^{\text{loc}}_{\text{c}}(\{F_t\}, \tilde{P}) \). This result has many implications, and Girsanov’s theorem, like Itô’s formula, is now indispensable to the modern probabilist. Our goal in this section is to establish Girsanov’s theorem, deferring to later chapters a study of some of its applications.

For much of the present section we shall need the following strengthened form of Condition 5.3.1:

**Condition 5.7.1** \( \{F_t, t \in [0, \infty)\} \) is a filtration in a probability space \( (\Omega, \mathcal{F}, P) \) such that \( F_0 \) includes all \( P \)-null events of \( \mathcal{F} \), and \( \tilde{P} \) is a probability measure on \( (\Omega, \mathcal{F}) \) with \( \tilde{P} \equiv P[\mathcal{F}] \).

**Remark 5.7.2** With Condition 5.7.1 in force we see that \( F_0 \) also includes all \( \tilde{P} \)-null events in \( \mathcal{F} \). A statement such as \( X = Y \) a.s. will, of course, be interpreted to hold with reference to both measures \( P \) and \( \tilde{P} \), and, likewise, \( X \) will be called a.s. strictly positive when it is both \( P \)-strictly positive and \( \tilde{P} \)-strictly positive. For a random variable \( X \) on \( (\Omega, \mathcal{F}) \) and sub-\( \sigma \)-algebra \( \mathcal{G} \subset \mathcal{F} \), we write \( E[X] \) and \( E[X | \mathcal{G}] \) to denote expectations and conditional expectations calculated with respect to measure \( P \), and use \( \tilde{E}[X] \) and \( \tilde{E}[X | \mathcal{G}] \) for the same quantities but calculated with reference to \( \tilde{P} \).

**Remark 5.7.3** In view of Condition 5.7.1, we can define the Radon-Nikodym derivatives

\[
(5.243) \quad \Lambda(t) \triangleq \left. \frac{d\tilde{P}}{dP} \right|_{\mathcal{F}_t} \quad \forall \ t \in [0, \infty].
\]

The process \( \{(\Lambda(t), \mathcal{F}_t); t \in [0, \infty)\} \) defined in this way is clearly adapted and is called the **density process** for the pair of probability measures \( (P, \tilde{P}) \). From Remark 1.4.18 it follows that

\[
(5.244) \quad \Lambda(t) = E[\Lambda(\infty) \mid \mathcal{F}_t] \quad \text{a.s.}
\]

for each \( t \in [0, \infty] \), thus \( \{(\Lambda(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable martingale on \( (\Omega, \mathcal{F}, P) \). If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an adapted process defined on \( (\Omega, \mathcal{F}, P) \), then, for each \( t \in [0, \infty) \), one sees that

\[
(5.245) \quad \tilde{E}[X(t)] < \infty \quad \text{if and only if} \quad E[|X(t)| \Lambda(t)] < \infty,
\]

in which case

\[
(5.246) \quad \tilde{E}[X(t) ; A] = E[X(t)\Lambda(t) ; A], \quad \forall \ A \in \mathcal{F}_t,
\]

(take \( (E, \mathcal{S}) \triangleq (\Omega, \mathcal{F}), \mu \triangleq P, \nu \triangleq \tilde{P}, \mathcal{H} \triangleq \mathcal{F}_t, \) and \( g \triangleq X(t) \) in (1.15) and (1.16) of Remark 1.2.34).
If we suppose that \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is right-continuous (Remark 3.1.30) then it becomes a standard filtration on \((\Omega, \mathcal{F}, \mathbb{P})\) (Definition 3.4.2) and hence we can use Corollary 4.2.19 to conclude that the density process \( \{(\Lambda(t), \mathcal{F}_t); t \in [0, \infty)\} \) has some modification with coroll sample-paths. Actually, rather than postulating right-continuity of \( \{\mathcal{F}_t, t \in [0, \infty)\} \) in order to ensure good path-wise properties, we shall instead impose a condition directly on the paths of the density process itself:

**Condition 5.7.4** There exists a continuous modification of the density process for the pair of probability measures \((\mathbb{P}, \tilde{\mathbb{P}})\) (see (5.243)).

**Remark 5.7.5** Despite the restrictive appearance of Condition 5.7.4 it turns out that the structure of the measure \( \tilde{\mathbb{P}} \) and filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is usually such that this condition holds. Henceforth, we shall use \( \{\Lambda(t); t \in [0, \infty)\} \) to denote the continuous modification of the density process postulated by Condition 5.7.4, so that \( \{(\Lambda(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous uniformly integrable martingale given by (5.243).

We are now able to show that the process \( \{\Lambda(t); t \in [0, \infty)\} \) is both \( \mathbb{P} \) and \( \tilde{\mathbb{P}} \)-strictly positive (recall Definition 5.5.6):

**Proposition 5.7.6** Suppose that Conditions 5.7.1 and 5.7.4 hold, and put

\[ \Omega_1 \triangleq \{ \omega : \Lambda(t, \omega) > 0, \quad \forall t \in [0, \infty) \}. \]

Then \( \mathbb{P}(\Omega_1) = \tilde{\mathbb{P}}(\Omega_1) = 1 \).

Proof: For each \( \omega \in \Omega \) and \( n = 1, 2, \ldots \) put

\[ T_n(\omega) \triangleq \inf \{ t \in [0, \infty) : \Lambda(t, \omega) \leq n^{-1} \}. \]  
(5.247)

In view of Proposition 3.3.7(b) we see that the \( T_n \) are \( \{\mathcal{F}_t\} \)-stopping times, and clearly \( T_n(\omega) \leq T_{n+1}(\omega) \), \( \forall n = 1, 2, \ldots \), \( \forall \omega \in \Omega \); define

\[ T(\omega) \triangleq \lim_{n \to \infty} T_n(\omega), \quad \forall \omega \in \Omega. \]  
(5.248)

Then clearly \( \{T < \infty\} \subset \{T_n < \infty\} \); hence

\[ 0 \leq E[\Lambda(\infty); T < \infty] \leq E[\Lambda(\infty); T_n < \infty], \quad \forall n = 1, 2, \ldots \]  
(5.249)

Now, from Theorem 4.5.7, we see that \( E[\Lambda(\infty) | \mathcal{F}_{T_n}] = \Lambda(T_n) \) a.s. and since \( \{T_n < \infty\} \in \mathcal{F}_{T_n} \) (see Proposition 3.3.13(d)) it follows that

\[ E[\Lambda(\infty); T_n < \infty] = E[\Lambda(T_n); T_n < \infty] \leq n^{-1}, \quad \forall n = 1, 2, \ldots \]  
(5.250)

where the inequality in (5.250) follows since (5.247) ensures that \( \Lambda(T_n) = n^{-1} \) on \( \{T_n < \infty\} \). Combining (5.249) with (5.250) and taking \( n \to \infty \) shows that

\[ E[\Lambda(\infty); T < \infty] = 0. \]  
(5.251)
Now, from (5.243), one has
\[(5.252)\quad \tilde{P}(A) = E[\Lambda(\infty); A], \quad \forall A \in \mathcal{F}_\infty.\]
Taking $A \triangleq \{T < \infty\}$ in (5.252) and using (5.251) shows that $\tilde{P}[T < \infty] = 0$ and hence, by Condition 5.7.1, we have $P[T < \infty] = 0$. Since $[T = +\infty] \subset \Omega_1$, it follows that $P(\Omega_1) = \tilde{P}(\Omega_1) = 1$. \[\blacksquare\]

**Remark 5.7.7** In view of Proposition 5.7.6, we are able to define the continuous adapted process $\{(\Lambda^{-1}(t), \mathcal{F}_t); t \in [0, \infty)\}$ in accordance with Remark 5.5.7. Observe that, if $\{X(t); t \in [0, \infty)\}$ is a continuous $\mathcal{R}$-valued process on $(\Omega, \mathcal{F}, P)$, then the processes $\{\Lambda^{-1}(t)\Lambda(t)X(t); t \in [0, \infty)\}$ and $\{X(t); t \in [0, \infty)\}$ are indistinguishable (but generally not identical, because of the arbitrary definition of $\Lambda^{-1}(t) \triangleq 0$ for all $\omega$ such that $\Lambda(t, \omega) = 0$).

If $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a given adapted process we shall use the notation $\Lambda X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ to indicate that $\{(\Lambda(t)X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a martingale on $(\Omega, \mathcal{F}, P)$. We shall also write $\Lambda^{-1}X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ with an obviously similar interpretation. The following lemma is of basic importance.

**Lemma 5.7.8** Suppose Conditions 5.7.1 and 5.7.4 hold and let $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ be a $\mathcal{R}$-valued right-continuous adapted process on $(\Omega, \mathcal{F}, P)$. Then
(a) $\Lambda X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ if and only if $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$;
(b) $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ if and only if $\Lambda^{-1}X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$;
(c) $\Lambda^{-1} \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$.

Proof: (a) Suppose that $\Lambda X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$, and fix $s, t \in [0, \infty)$ with $s < t$. Then $E|\Lambda(t)X(t)| < \infty$, whence $\tilde{E}|X(t)| < \infty$ (see (5.245) of Remark 5.7.3). Now fix an arbitrary $A \in \mathcal{F}_s$. Then we get
\[
\tilde{E}[X(t)I_A] = E[\Lambda(t)X(t)I_A] = E[\Lambda(s)X(s)I_A] = \tilde{E}[X(s)I_A],
\]
where the first and third equalities follow by (5.246) of Remark 5.7.3, and the second equality is a consequence of $\Lambda X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$. Thus, by the arbitrary choice of $A \in \mathcal{F}_s$, we see that $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$. It follows that $\Lambda X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ implies $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$. The converse implication is similarly proved.
(b) Suppose that $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$, and define the process $\{Y(t); t \in [0, \infty)\}$ by $Y \triangleq \Lambda^{-1}X$. Since the processes $\{X(t); t \in [0, \infty)\}$ and $\{\Lambda(t)Y(t); t \in [0, \infty)\}$ are indistinguishable (see Remark 5.7.7), one sees that $\Lambda Y \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$, whence, in view of (a), it follows that $\Lambda^{-1}X \equiv Y \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$. The converse implication, that $\Lambda^{-1}X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$ implies $X \in \mathcal{M}(\{\mathcal{F}_t\}, \tilde{P})$, is similarly established. (c) Take $X(t, \omega) \triangleq 1$, $\forall (t, \omega) \in [0, \infty) \otimes \Omega$ in (b). \[\blacksquare\]

We next extend Lemma 5.7.8 from martingales to continuous local martingales:

**Lemma 5.7.9** Suppose Conditions 5.7.1 and 5.7.4 hold, and $\{(X(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous, $\mathcal{R}$-valued, adapted process on $(\Omega, \mathcal{F}, P)$. Then
(a) $\Lambda X \in \mathcal{M}^e_{loc}(\{\mathcal{F}_t\}, \tilde{P})$ if and only if $X \in \mathcal{M}^e_{loc}(\{\mathcal{F}_t\}, \tilde{P})$;
(b) $X \in \mathcal{M}^e_{loc}(\{\mathcal{F}_t\}, \tilde{P})$ if and only if $\Lambda^{-1}X \in \mathcal{M}^e_{loc}(\{\mathcal{F}_t\}, \tilde{P})$;
Proof: (a) Suppose $\Lambda X \in M_{\text{loc}}^{\{\mathcal{F}_t\}, P}$, with a localizing sequence $\{T_n, \ n = 1, 2, \ldots\}$ of $\{\mathcal{F}_t\}$-stopping times. Then, for an arbitrary $T_n$, one has that $\{(\Lambda(t \wedge T_n)X(t \wedge T_n), \mathcal{F}_t); \ t \in [0, \infty)\}$ is a continuous martingale on $(\Omega, \mathcal{F}, P)$. By Lemma 5.7.8(b), it then follows that $\{(\Lambda^{-1}(t)\Lambda(t \wedge T_n)X(t \wedge T_n), \mathcal{F}_t); \ t \in [0, \infty)\}$ is a continuous martingale on $(\Omega, \mathcal{F}, \tilde{P})$, whence, stopping this martingale at $T_n$ and using Corollary 4.5.8, we get that $\{(X(t \wedge T_n), \mathcal{F}_t); \ t \in [0, \infty)\}$ is a continuous martingale on $(\Omega, \mathcal{F}, \tilde{P})$. Since Condition 5.7.1 ensures $\tilde{P} \lim_{n \to \infty} T_n = \infty = P \lim_{n \to \infty} T_n = \infty$, we see that $X \in M_{\text{loc}}^{\{\mathcal{F}_t\}, \tilde{P}}$. The converse, namely that $X \in M_{\text{loc}}^{\{\mathcal{F}_t\}, \tilde{P}}$ implies $\Lambda X \in M_{\text{loc}}^{\{\mathcal{F}_t\}, P}$, follows by a similar argument. Next, (b) is established on the basis of (a) in exactly the same way that Lemma 5.7.8(b) is proved.

\textbf{Theorem 5.7.10} Assume Conditions 5.7.1 and 5.7.4. Then $SM_c^{\{\mathcal{F}_t\}, P} = SM_c^{\{\mathcal{F}_t\}, \tilde{P}}$.

Proof: Suppose that $X \in SM_c^{\{\mathcal{F}_t\}, P}$. Without loss of generality we may assume that $X$ is null at the origin, namely $X \in SM_c^{0,0}(\{\mathcal{F}_t\}, P)$. Since $\Lambda \in SM_c^{\{\mathcal{F}_t\}, P}$ (by Condition 5.7.4), one sees from Remark 5.4.20 that $\Lambda X \in SM_c^{0,0}(\{\mathcal{F}_t\}, P)$, hence we can write

\begin{equation}
\Lambda X = M + A,
\end{equation}

for some $M \in SM_c^{0,0}(\{\mathcal{F}_t\}, P)$ and some $A \in FV^{c,0}\{\mathcal{F}_t\}$. By Lemma 5.7.9(b) we have $\Lambda^{-1}M \in SM_c^{0,0}(\{\mathcal{F}_t\}, \tilde{P})$. Now, trivially, $A \in SM_c^{0,0}(\{\mathcal{F}_t\}, \tilde{P})$, and, by Lemma 5.7.8(c), we also have $\Lambda^{-1} \in SM_c^{\{\mathcal{F}_t\}, \tilde{P}}$. Thus, from Remark 5.4.20, it follows that $\Lambda^{-1}A \in SM_c^{0,0}(\{\mathcal{F}_t\}, \tilde{P})$. In view of (5.253) we see that the process $X$ is indistinguishable from the sum of the process $\Lambda^{-1}M$ and $\Lambda^{-1}A$, hence $X \in SM_c^{0,0}(\{\mathcal{F}_t\}, \tilde{P})$. We have shown that $SM_c^{\{\mathcal{F}_t\}, P} \subset SM_c^{\{\mathcal{F}_t\}, \tilde{P}}$. The proof of the opposite set inclusion is similar.

\textbf{Remark 5.7.11} Suppose that Conditions 5.7.1 and 5.7.4 hold, and let $X, Y \in SM_c^{\{\mathcal{F}_t\}, P}$. We know that $[X, Y]^{P}(t)$ is defined in terms of the local martingale parts of $X$ and $Y$ in accordance with Remark 5.4.5. This random variable depends, of course, on the probability measure $P$, and thus should be written $[X, Y]^{P}((t)$. By Theorem 5.7.10, we also have $X, Y \in SM_c^{\{\mathcal{F}_t\}, \tilde{P}}$, hence the random variable $[X, Y]^{\tilde{P}}((t)$ is similarly defined by Remark 5.4.5. We thus have two continuous adapted processes, $\{([X, Y]^{P}(t), \mathcal{F}_t); \ t \in [0, \infty)\}$ and $\{([X, Y]^{\tilde{P}}(t), \mathcal{F}_t); \ t \in [0, \infty)\}$, calculated with respect to the probability measures $P$ and $\tilde{P}$ respectively. The following proposition shows that these quadratic variation processes are actually indistinguishable and hence we may discard the superscripts $P$ and $\tilde{P}$ when writing them down:

\textbf{Proposition 5.7.12} Suppose Conditions 5.7.1 and 5.7.4 hold, and $X, Y \in SM_c^{\{\mathcal{F}_t\}, P}$. Then, with reference to Remark 5.7.11, the processes $\{([X, Y]^{P}(t), \mathcal{F}_t); \ t \in [0, \infty)\}$ and $\{([X, Y]^{\tilde{P}}(t), \mathcal{F}_t); \ t \in [0, \infty)\}$, are indistinguishable.
Proof: Write \( X = X_0 + M + A \) and \( Y = Y_0 + N + B \), for \( M, N \in \mathbf{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P) \) and \( A, B \in \mathbf{FV}^{c,0}\{\mathcal{F}_t\} \), and fix some arbitrary \( t \in [0, \infty) \). By Theorem 5.4.4 we see that the discrete sums

\[
(5.254) \quad \sum_{0 \leq k < \infty} [X(t \wedge \tau^n_{k+1}) - X(t \wedge \tau^n_k)][Y(t \wedge \tau^n_{k+1}) - Y(t \wedge \tau^n_k)]
\]

converge in \( P \)-measure to \( [X, Y]_P(t) \overset{\Delta}{=} [M, N](t) \) as \( n \to \infty \). Since \( \tilde{P} \ll P[\mathcal{F}] \), it follows from Theorem 1.2.29 that the sums in (5.254) also converge in \( \tilde{P} \)-measure to the same limit \( [X, Y]_{\tilde{P}}(t) \). But, since \( X, Y \in \mathbf{SM}^c(\{\mathcal{F}_t\}, \tilde{P}) \), we know that the sums in (5.254) converge in \( \tilde{P} \)-measure to \( [X, Y]_{\tilde{P}}(t) \). It follows from Theorem 1.2.10 that \( [X, Y]_{\tilde{P}}(t) = [X, Y]_{\tilde{P}}(t) \overset{\tilde{P}}{=} \text{a.s. for each } t \in [0, \infty) \). Since \( \tilde{P} \equiv P[\mathcal{F}] \) and \( \{[X, Y]_P(t), t \in [0, \infty]\} \), \( \{[X, Y]_{\tilde{P}}(t), t \in [0, \infty]\} \), are sample-path continuous, the proposition follows.

**Remark 5.7.13** We shall frequently need to associate with \( X \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, P) \) a process \( \tilde{X} \) defined by

\[
(5.255) \quad \tilde{X}(t) \overset{\Delta}{=} X(t) - \int_0^t \Lambda^{-1}(s) \ d[\Lambda, X](s), \quad \forall t \in [0, \infty).
\]

Since \( \{ (\Lambda^{-1}(t), \mathcal{F}_t); t \in [0, \infty) \} \) is a continuous adapted process (by Remark 5.7.7) we see from Remark 5.2.8 that the second term on the right hand side of (5.255) defines a member of \( \mathbf{FV}^{c,0}\{\mathcal{F}_t\} \), from which it follows that \( \tilde{X} \in \mathbf{SM}^c(\{\mathcal{F}_t\}, P) \). The process \( \tilde{X} \) defined by (5.255) is called the **Girsanov transformation** of the continuous local martingale \( X \) by the density process \( \Lambda \) for the pair \( (P, \tilde{P}) \).

We are now ready to establish the first main result of this section:

**Theorem 5.7.14** (Girsanov, Meyer) Suppose Conditions 5.7.1 and 5.7.4 hold. If \( X \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, P) \) and \( \tilde{X} \) is the Girsanov transformation of \( X \) defined by (5.255) then \( \tilde{X} \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, \tilde{P}) \). Moreover, if \( Y \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, P) \) and \( \tilde{Y} \) is the Girsanov transformation of \( Y \), namely

\[
(5.256) \quad \tilde{Y}(t) \overset{\Delta}{=} Y(t) - \int_0^t \Lambda^{-1}(s) \ d[\Lambda, Y](s), \quad \forall t \in [0, \infty),
\]

then we have a.s.:

\[
(5.257) \quad [X, Y](t) = [X, \tilde{Y}](t) = [\tilde{X}, Y](t) = [\tilde{X}, \tilde{Y}](t), \quad \forall t \in [0, \infty),
\]

and, in particular,

\[
(5.258) \quad [X] = [\tilde{X}](t) \quad \forall t \in [0, \infty).
\]

Proof: In view of Lemma 5.7.9(a) it is enough to establish \( \Lambda \tilde{X} \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, P) \) in order to conclude that \( \tilde{X} \in \mathbf{M}^{c}_{loc}(\{\mathcal{F}_t\}, \tilde{P}) \). We have already observed in Remark 5.7.13 that \( \tilde{X} \in \mathbf{SM}^c(\{\mathcal{F}_t\}, P) \), and Condition 5.7.4 obviously ensures that \( \Lambda \in \mathbf{SM}^c(\{\mathcal{F}_t\}, P) \). Thus, from integration-by-parts (Theorem 5.4.19), we have a.s.:

\[
(5.259) \quad \Lambda(t) \tilde{X}(t) = \Lambda(0) \tilde{X}(0) + \int_0^t \tilde{X}(s) \ d\Lambda(s) + \int_0^t \Lambda(s) \ d\tilde{X}(s) + [\Lambda, \tilde{X}](t), \quad \forall t \in [0, \infty).
\]

253
Now put \( A(t) \triangleq \int_0^t \Lambda^{-1}(s) \, d[\Lambda, X](s) \). Since \( \{(\Lambda(t), \mathcal{F}_t); \ t \in [0, \infty)\} \) is continuous and adapted, hence a member of \( \mathbf{B}_{loc}\mathcal{F}_t \), we see from (5.255) and Definition 5.4.10 that:

\[
(5.260) \quad \int_0^t \Lambda(s) \, d\tilde{X}(s) = \int_0^t \Lambda(s) \, dX(s) - \int_0^t \Lambda(s) \, dA(s), \quad \forall \ t \in [0, \infty).
\]

Now, in view of Theorem 5.1.13, (5.255) and Proposition 5.4.6, we get a.s.:

\[
(5.261) \quad \int_0^t \Lambda(s) dA(s) = [\Lambda, X](t) = [\Lambda, \tilde{X}](t), \quad \forall \ t \in [0, \infty).
\]

Combining (5.259), (5.260) and (5.261), we obtain a.s.:

\[
(5.262) \quad \Lambda(t) \tilde{X}(t) = \Lambda(0) \tilde{X}(0) + \int_0^t \tilde{X}(s) \, d\Lambda(s) + \int_0^t \Lambda(s) \, dX(s), \quad \forall \ t \in [0, \infty).
\]

Since \( \Lambda, \tilde{X} \in \mathbf{B}_{loc}\mathcal{F}_t \) and \( X, \Lambda \in \mathbf{M}_{loc}^c(\mathcal{F}_t, P) \), we see that the second and third terms on the right hand side of (5.262) define elements in \( \mathbf{M}_{loc}^{c,0}(\mathcal{F}_t, P) \), whence we conclude \( \Lambda \tilde{X} \in \mathbf{M}_{loc}^{c,0}(\mathcal{F}_t, P) \), as required to establish \( \tilde{X} \in \mathbf{M}_{loc}^c(\mathcal{F}_t, \tilde{P}) \). Now (5.257) follows from the fact that \( X, Y, \tilde{X}, \tilde{Y} \in \mathbf{SM}^c(\mathcal{F}_t, P) \equiv \mathbf{SM}^c(\mathcal{F}_t, \tilde{P}) \), together with Proposition 5.4.6, Remark 5.7.11 and Proposition 5.7.12.

**Remark 5.7.15** Suppose \( X \in \mathbf{SM}^c(\mathcal{F}_t, P) \) and \( \Phi \in \mathbf{B}_{loc}\mathcal{F}_t \). Then the Itô stochastic integral \( \Phi \bullet X \) is defined and seen to be a member of \( \mathbf{SM}^c(\mathcal{F}_t, P) \) (recall Definition 5.4.10). The underlying measure \( P \) is a determining element for the Hilbert space \( \mathbf{M}_{2, b}^c(\mathcal{F}_t, P) \) as well as for the local martingale part of \( X \), and so we should really denote the previous stochastic integral by \( (\Phi \bullet X)^P \) to make clear its dependence on \( P \). Now, when Conditions 5.7.1 and 5.7.4 hold, then we know from Theorem 5.7.10 that \( X \in \mathbf{SM}^c(\mathcal{F}_t, \tilde{P}) \), and hence we can also define an Itô stochastic integral \( (\Phi \bullet X)^{\tilde{P}} \). We can use Theorem 5.7.14 to show that these stochastic integrals are indistinguishable:

**Theorem 5.7.16** Suppose Conditions 5.7.1 and 5.7.4 hold. If \( X \in \mathbf{SM}^c(\mathcal{F}_t, P) \) and \( \Phi \in \mathbf{B}_{loc}\mathcal{F}_t \) then, with reference to Remark 5.7.15, we have a.s.:

\[
(5.263) \quad (\Phi \bullet X)^P(t) = (\Phi \bullet X)^{\tilde{P}}(t), \quad \forall \ t \in [0, \infty).
\]

**Proof:** Suppose first that \( X \in \mathbf{M}_{loc}^{c,0}(\mathcal{F}_t, P) \), and let \( \tilde{X} \) be the Girsanov transformation of \( X \), namely

\[
(5.264) \quad \tilde{X}(t) \triangleq X(t) - \int_0^t \Lambda^{-1}(s) \, d[\Lambda, X](s).
\]

By Theorem 5.7.14, we see that \( X \in \mathbf{SM}^{c,0}(\mathcal{F}_t, \tilde{P}) \) with a local martingale part \( \tilde{X} \in \mathbf{M}_{loc}^{c,0}(\mathcal{F}_t, \tilde{P}) \) and a bounded variation part \( A(t) \triangleq \int_0^t \Lambda^{-1}(s) \, d[\Lambda, X](s) \). Thus, by Definition 5.4.10 and Theorem 5.1.13,

\[
(5.265) \quad (\Phi \bullet X)^{\tilde{P}}(t) = (\Phi \bullet \tilde{X})^{\tilde{P}}(t) + \int_0^t \Phi(s) \Lambda^{-1}(s) \, d[\Lambda, X](s), \quad \forall \ t \in [0, \infty).
\]
Now fix some arbitrary $Z \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$. Since $\tilde{X} \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$, it follows from Definition 5.3.32 that one has a.s.:

$$
(5.266) \quad [(\Phi \cdot \tilde{X})^\tilde{P}, Z](t) = \int_0^t \Phi(s) \ d[\tilde{X}, Z](s), \quad \forall \ t \in [0, \infty).
$$

Let $Y \overset{\Delta}{=} (\Phi \cdot X)^P$, observe that $Y \in M^c_{\text{loc}}(\{F_t\}, P)$, and let $\tilde{Y}$ be the Girsanov transformation of $Y$, namely

$$
(5.267) \quad \tilde{Y}(t) = Y(t) - \int_0^t \Lambda^{-1}(s) \ d[\Lambda, Y](s), \quad \forall \ t \in [0, \infty).
$$

By Theorem 5.7.10 we have $Z \in SM^c(\{F_t\}, P)$, and obviously $\tilde{Y} \in SM^c(\{F_t\}, P)$. Since the second term on the right hand side of (5.267) defines an element of $FV^c_{\text{loc}}(\{F_t\})$, we see from Proposition 5.4.6 along with Proposition 5.4.11 that, a.s.:

$$
(5.268) \quad [\tilde{Y}, Z](t) = [Y, Z](t) = [(\Phi \cdot X)^P, Z](t)
= \int_0^t \Phi(s) d[X, Z](s), \quad \forall \ t \in [0, \infty).
$$

Moreover, from (5.264) and Proposition 5.4.6 we have a.s.:

$$
(5.269) \quad [X, Z](s) = [\tilde{X}, Z](s), \quad \forall \ s \in [0, \infty).
$$

Combining (5.266), (5.268) and (5.269), for each $Z \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$ we have a.s.:

$$
(5.270) \quad [(\Phi \cdot \tilde{X})^\tilde{P}, Z](t) = [\tilde{Y}, Z](t), \quad \forall \ t \in [0, \infty).
$$

Now Theorem 5.7.14 ensures that $\tilde{Y} \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$ (since $Y \overset{\Delta}{=} (\Phi \cdot X)^P \in M^c_{\text{loc}}(\{F_t\}, P)$), while $(\Phi \cdot \tilde{X})^\tilde{P} \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$ (since, as already observed, $\tilde{X} \in M^c_{\text{loc}}(\{F_t\}, \tilde{P})$). In view of these observations, (5.270), and Proposition 4.7.31 we conclude that a.s.:

$$
(5.271) \quad (\Phi \cdot \tilde{X})^\tilde{P} = \tilde{Y}(t), \quad \forall \ t \in [0, \infty).
$$

Now, by Theorem 5.1.13 we easily see that

$$
(5.272) \quad \int_0^t \Lambda^{-1}(s) \ d[\Lambda, Y](s) = \int_0^t \Phi(s) \Lambda^{-1}(s) \ d[\Lambda, X](s), \quad \forall \ t \in [0, \infty).
$$

Combining (5.265), (5.267), (5.271) and (5.272) we see that (5.263) holds when $X \in M^c_{\text{loc}}(\{F_t\}, P)$. The general case, where $X \in SM^c(\{F_t\}, P)$, follows easily from this special case.

Suppose that one has a Wiener process $\{(W(t), F_t); t \in [0, \infty)\}$ and a progressively measurable process $\{(\Phi(t), F_t); t \in [0, \infty)\}$ on probability space $(\Omega, F, P)$, and define the process $\{(\tilde{W}(t), F_t); t \in [0, \infty)\}$ as follows:

$$
(5.273) \quad \tilde{W}(t) \overset{\Delta}{=} W(t) - \int_0^t \Phi(s)ds, \quad \forall \ t \in [0, \infty).
$$

255
There are numerous applications where one must construct a probability measure \( \tilde{P} \) on \((\Omega, \mathcal{F})\) such that \( \{ (\tilde{W}(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a Wiener process on \((\Omega, \mathcal{F}, \tilde{P})\). In the remainder of this section we shall use Theorem 5.7.14 to see how to construct such a probability measure. To this end, we postulate the following conditions:

**Condition 5.7.17** \( \{ \mathcal{F}_t, \ t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\) such that \( \mathcal{F}_0 \) includes all \( P \)-null events of \( \mathcal{F} \). \( \{(W(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a \(d\)-dimensional standard Wiener process on \((\Omega, \mathcal{F}, P)\).

**Condition 5.7.18** \( \{(\Phi(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is an \( \mathbb{R}^d \)-valued progressively measurable process on \((\Omega, \mathcal{F}, P)\) such that

\[
(\Phi \cdot W)(t) \overset{\Delta}{=} \sum_{k=1}^d (\Phi^k \cdot W^k)(t), \quad t \in [0, \infty).
\]

**Remark 5.7.19** Since \( \Phi \cdot W \in \mathcal{M}^c_{\text{loc}}(\mathcal{F}_t, P) \) we see from Proposition 5.5.2(c) that \( \mathcal{E}(\Phi \cdot W) \in \mathcal{M}^c_{\text{loc}}(\mathcal{F}_t, P) \). Furthermore, by Remark 5.5.3, there exists some a.s. unique \( \mathcal{F}_\infty \)-measurable non-negative random variable \( \mathcal{E}(\Phi \cdot W)(\infty) \) such that

\[
\lim_{t \to \infty} \mathcal{E}(\Phi \cdot W)(t) = \mathcal{E}(\Phi \cdot W)(\infty) \quad \text{a.s. and} \quad E[\mathcal{E}(\Phi \cdot W)(\infty)] \leq 1.
\]

Now (recall Definition 5.5.1),

\[
(\Phi \cdot W)(t) \overset{\Delta}{=} \exp \left\{ (\Phi \cdot W)(t) - \frac{1}{2}[\Phi \cdot W](t) \right\}, \quad \forall \ t \in [0, \infty),
\]

and using (5.275), Lemma 4.7.28, Theorem 5.3.34, and Remark 4.7.25, we get

\[
[\Phi \cdot W](t) = \left[ \sum_{j=1}^d \Phi^j \cdot W^j, \sum_{k=1}^d \Phi^k \cdot W^k \right](t)
= \sum_{j,k=1}^d \int_0^t \Phi^j(s)\Phi^k(s) \, d[W^j, W^k](s)
= \sum_{k=1}^d \int_0^t |\Phi^k(s)|^2 \, ds.
\]

Combining (5.277), (5.275), and (5.278), we see that \( \mathcal{E}(\Phi \cdot W)(t) \) can be written as

\[
(5.279) \quad \mathcal{E}(\Phi \cdot W)(t) = \exp \left\{ \sum_{k=1}^d \int_0^t \Phi^k(s) \, dW^k(s) - \frac{1}{2} \sum_{k=1}^d \int_0^t |\Phi^k(s)|^2 \, ds \right\}, \quad \forall \ t \in [0, \infty).
\]
We are going to define a probability measure $\tilde{P}$ on $(\Omega, \mathcal{F})$ in such a way that $\mathcal{E}(\Phi \cdot W)$ is the density process for the pair $(P, \tilde{P})$ (recall Remark 5.7.3). For this purpose it is not enough to just have $\mathcal{E}(\Phi \cdot W) \in \mathcal{M}_{loc}^c(\{\mathcal{F}_t\}, P)$, and we must postulate the following:

**Condition 5.7.20** The progressively measurable process $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ in Condition 5.7.18 is such that $\{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous uniformly integrable martingale with $P[\mathcal{E}(\Phi \cdot W)(\infty) > 0] = 1$.

**Remark 5.7.21** In order to verify Condition 5.7.20 we shall typically use the sufficient condition furnished by Theorem 5.5.10. Indeed, if $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a progressively measurable process such that

$$\mathcal{E} \left[ \exp \left( \frac{1}{2} \sum_{k=1}^d \int_0^\infty |\Phi^k(s)|^2 \, ds \right) \right] < \infty,$$

then we see at once from (5.277), (5.278) and Theorem 5.5.10, that $\{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\}$ is a uniformly integrable martingale. Moreover, it is clear that (5.280) also implies

$$P \left\{ \int_0^\infty \sum_{k=1}^d |\Phi^k(s)|^2 \, ds < \infty \right\} = 1,$$

which, in view of Proposition 5.5.4, (5.277) and (5.278), establishes $P[\mathcal{E}(\Phi \cdot W)(\infty) = 0] = 0$, as required to verify Condition 5.7.20.

**Remark 5.7.22** Condition 5.7.20 and Proposition 5.5.9 ensure that

$$\mathcal{E}(\Phi \cdot W)(\infty) = 1,$$

which, together with the non-negativity of $\mathcal{E}(\Phi \cdot W)(\infty)$, shows that $\tilde{P}$ defined on $(\Omega, \mathcal{F})$ by

$$\tilde{P}(A) \triangleq E[\mathcal{E}(\Phi \cdot W)(\infty); A], \quad \forall A \in \mathcal{F},$$

is a probability measure. It follows at once from (5.282) that $\tilde{P} \ll P[\mathcal{F}]$, and since Condition 5.7.20 also postulates $P[\mathcal{E}(\Phi \cdot W)(\infty) > 0] = 1$, one sees (from Theorem 1.2.24) that $\tilde{P}(A) = 0$ for some $A \in \mathcal{F}$ implies $P(A) = 0$, thus $P \ll \tilde{P}[\mathcal{F}]$. We therefore have $\tilde{P} \equiv P[\mathcal{F}]$.

**Remark 5.7.23** It remains to compute a density process for the pair $(P, \tilde{P})$. Fix some $t \in [0, \infty]$. From (5.282),

$$\tilde{P}(A) = E[\mathcal{E}(\Phi \cdot W)(\infty) | \mathcal{F}_t]; A], \quad \forall A \in \mathcal{F}_t,$$

hence (see Remark 1.2.34), we get

$$\frac{d\tilde{P}}{dP} \bigg|_{\mathcal{F}_t} = E[\mathcal{E}(\Phi \cdot W)(\infty) | \mathcal{F}_t] \quad \text{a.s.}$$

257
Moreover, from Condition 5.7.20 and Theorem 4.4.11(b),

\[(5.284) \quad \mathcal{E}(\Phi \cdot W)(t) = E[\mathcal{E}(\Phi \cdot W)(\infty) \mid \mathcal{F}_t] \quad \text{a.s.}\]

and thus, from (5.283),

\[(5.285) \quad \frac{d\tilde{P}}{dP}\bigg|_{\mathcal{F}_t} = \mathcal{E}(\Phi \cdot W)(t) \quad \text{a.s.} \]

for each \(t \in [0, \infty]\). We conclude the following: when Conditions 5.7.17, 5.7.18, and 5.7.20 hold, and \(\tilde{P}(\cdot)\) is defined by (5.282), then \(\mathcal{E}(\Phi \cdot W)\) given by (5.279) is a density process for the pair \((P, \tilde{P})\). In particular, Condition 5.7.4 is verified.

With the above preliminaries at hand, we can establish the second main result of this section:

**Theorem 5.7.24 (Girsanov)** Suppose that Conditions 5.7.17, 5.7.18 and 5.7.20 hold, and let \(\{\tilde{W}(t); t \in [0, \infty]\}\) be the \(\mathbb{R}^d\)-valued process whose \(k\)-th scalar element is defined by

\[(5.286) \quad \tilde{W}^k(t) \triangleq W^k(t) - \int_0^t \Phi^k(s) \, ds, \quad \forall \ t \in [0, \infty), \]

for each \(k = 1, 2, \ldots, d\). Then \(\{(\tilde{W}(t), \mathcal{F}_t); t \in [0, \infty]\}\) is a \(d\)-dimensional standard Wiener process on \((\Omega, \mathcal{F}, \tilde{P})\) (where \(\tilde{P}\) is the probability measure defined on \((\Omega, \mathcal{F})\) by (5.282)).

Proof: From Remark 5.7.23 we see that the pair \((P, \tilde{P})\) has a density process \(\{\Lambda(t), t \in [0, \infty)\}\) given by

\[(5.287) \quad \Lambda(t) = \mathcal{E}(\Phi \cdot W)(t), \quad \forall \ t \in [0, \infty). \]

For brevity put

\[(5.288) \quad Z(t) \triangleq \Phi \cdot W(t), \quad \forall \ t \in [0, \infty), \]

(see (5.275)). Fix some \(k \in \{1, 2, \ldots, d\}\) and define

\[(5.289) \quad A(t) \triangleq \int_0^t \Lambda^{-1}(s) \, d[\Lambda, W^k](s), \quad \forall \ t \in [0, \infty). \]

From Proposition 5.5.2, (5.277), (5.288), and (5.287), we see that a.s.

\[(5.290) \quad \Lambda(t) = 1 + \int_0^t \Lambda(s) \, dZ(s), \quad \forall \ t \in [0, \infty), \]

thus a.s.

\[(5.291) \quad [\Lambda, W^k](t) = \int_0^t \Lambda(s) \, d[Z, W^k](s), \quad \forall \ t \in [0, \infty). \]

In view of (5.275) clearly

\[Z(t) = \sum_{i=1}^d (\Phi^i \cdot W^i)(t), \]

258
whence, by Lemma 4.7.28, Theorem 5.3.31, and Remark 4.7.25, we find

\begin{align}
(Z, W^k)(t) &= \sum_{i=1}^{d} \Phi_i \cdot [W^i, W^k](t) \text{ a.s.} \\
&= \sum_{i=1}^{d} \int_0^t \Phi_i(s) d[W^i, W^k](s) \text{ a.s.} \\
&= \int_0^t \Phi_k(s) ds \text{ a.s.}
\end{align}

Combining (5.289), (5.291), and (5.292), we see that a.s.

\begin{align}
A(t) = \int_0^t \Phi_k(s) ds, \quad \forall \ t \in [0, \infty).
\end{align}

From (5.289) and (5.293) one sees that the process \( \tilde{W}^k \) defined by (5.286) is the Girsanov transformation of \( W^k \) by the density process \( \Lambda \) for the pair \( (P, \tilde{P}) \) (see Remark 5.7.13). Moreover, we have seen in Remark 5.7.23 that Condition 5.7.4 is verified, and therefore Theorem 5.7.14 ensures that \( \tilde{W}^k \in M^c_{loc}(\{\mathcal{F}_t\}, \tilde{P}) \) and

\begin{align}
[\tilde{W}^j, \tilde{W}^k](t) &= [W^j, W^k](t) \\
&= t\delta_{j,k}, \quad \forall \ t \in [0, \infty),
\end{align}

for each \( j, k = 1, 2, \ldots d \). Now, applying Levy’s Theorem 5.6.2 to the adapted process \( \{(\tilde{W}(t), \mathcal{F}_t); t \in [0, \infty)\} \) on \( (\Omega, \mathcal{F}, \tilde{P}) \) establishes the result. \( \square \)

Theorem 5.7.24 is not particularly useful in the form just stated, because in applications one typically deals with processes \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) for which \( \{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) defined by (5.277) is a continuous martingale on \( (\Omega, \mathcal{F}, P) \) which nevertheless fails to be uniformly integrable. For example, this situation prevails in the simple case where \( d = 1 \), and \( \Phi \equiv 1 \), since Corollary 5.5.11 ensures that \( \mathcal{E}(\Phi \cdot W) \in M^c(\{\mathcal{F}_t\}, P) \), but one sees from Remark 5.5.12 that \( \{\mathcal{E}(\Phi \cdot W)(t); t \in [0, \infty)\} \) fails to be uniformly integrable. To overcome this difficulty we next establish a corollary of Theorem 5.7.24 which, in place of Condition 5.7.20, postulates the following:

**Condition 5.7.25** The adapted process \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) in Condition 5.7.18 is such that \( \{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale on \( (\Omega, \mathcal{F}, P) \).

**Remark 5.7.26** Clearly, if \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is such that

\[ E \left[ \exp \left( \frac{1}{2} \sum_{k=1}^{d} \int_0^t |\Phi^k(s)|^2 ds \right) \right] < \infty, \quad \forall \ t \in [0, \infty), \]

then Condition 5.7.25 holds. This follows from (5.277), (5.278) and Corollary 5.5.11.
Corollary 5.7.27 Suppose Conditions 5.7.17, 5.7.18 and 5.7.25 hold, and let \( \{ \tilde{W}(t); t \in [0, \infty) \} \) be the \( \mathbb{R}^d \)-valued process defined by (5.286), for each \( k = 1, 2, \ldots, d \). If \( T \in [0, \infty) \) is an arbitrary constant and \( \tilde{P}_T \) is the probability measure on \( (\Omega, \mathcal{F}) \) defined by

\[
\tilde{P}_T(A) \overset{\Delta}{=} E[\mathcal{E}(\Phi \cdot W)(T) I_A], \quad \forall \; A \in \mathcal{F},
\]

then \( \{(\tilde{W}(t), \mathcal{F}_t); t \in [0, T]\} \) is a \( d \)-dimensional standard Wiener process on the probability space \( (\Omega, \mathcal{F}, \tilde{P}_T) \) (recall Remark 3.2.15).

Proof: We see at once from Condition 5.7.25 that \( \{(\mathcal{E}(\Phi \cdot W)(t \wedge T), \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable martingale on \( (\Omega, \mathcal{F}, P) \). Since the processes \( \{(\mathcal{E}(\Phi \cdot W)(t \wedge T), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(\mathcal{E}(\Phi[0, T] \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) are indistinguishable, it follows that Condition 5.7.20 holds when \( \Phi \) is replaced by the truncated process \( \Phi[0, T] \), with the probability measure \( \tilde{P} \) in Condition 5.7.20 now being identical to \( \tilde{P}_T \) in (5.294). Thus, we can replace \( \Phi \) by \( \Phi[0, T] \) in Theorem 5.7.24, to conclude that the \( \mathbb{R}^d \)-valued process whose \( k \)-th scalar element is defined by

\[
W^k(t) - \int_0^t \Phi^k[0, T](s) \, ds, \quad \forall \; t \in [0, \infty),
\]

is a \( d \)-dimensional standard Wiener process on \( (\Omega, \mathcal{F}, \tilde{P}_T) \), from which the result follows.

Remark 5.7.28 The hypotheses of Corollary 5.7.27 are weaker than those of Theorem 5.7.24 since now one only postulates Condition 5.7.25 in place of the more stringent Condition 5.7.20. In return, Corollary 5.7.27 delivers a slightly weaker conclusion, namely that \( \tilde{W} \) defined by (5.286) is a Wiener process when restricted to an interval \( [0, T] \) of arbitrary but finite length. This slightly weaker result is quite acceptable for most applications of the Girsanov theorem.

Remark 5.7.29 Corollary 5.7.27 is essentially the form of Girsanov’s theorem as it first appeared in Girsanov ([13], page 287).

Remark 5.7.30 In this chapter we have used Durrett [11], Karatzas and Shreve [19], Revuz and Yor [25], and Rogers and Williams [26] as our primary sources. Attention has been limited throughout to stochastic integration in which the stochastic integrator is a continuous local martingale. One can develop the results of this chapter in the more general context of stochastic integration with respect to local martingales with right-continuous sample-paths. This calls for ideas considerably beyond the level of a simple introduction, and we refrain from studying stochastic integration at this level of generality here. Comprehensive accounts of stochastic integration with respect to right-continuous local martingales may be found in Elliott [12], Protter [24], and Rogers and Williams [26].
5.8 Problems

In each of the following problems \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration in the probability space \((\Omega, \mathcal{F}, P)\), and \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \).

Problem 5.8.1 Suppose that \( X, Y \in SM^c(\{ \mathcal{F}_t \}, P) \), and define the Stratonovich integral of \( X \) with respect to \( Y \) by
\[
\int_0^t X(s) \circ dY(s) \triangleq (X \cdot Y)(t) + \frac{1}{2} [X, Y](t), \quad \forall \, t \in [0, \infty).
\]
If \( f : \mathbb{R} \to \mathbb{R} \) is three-times continuously differentiable, use Itô’s formula to show that
\[
f(X(t)) = f(X(0)) + \int_0^t \frac{\partial f}{\partial x}(X(s)) \circ dX(s), \quad \forall \, t \in [0, \infty).
\]

Problem 5.8.2 Suppose that \( \{(\Theta_{kj}(t), \mathcal{F}_t); t \in [0, \infty)\} \) and \( \{(\Psi_k(t), \mathcal{F}_t); t \in [0, \infty)\} \) are progressively measurable processes on \((\Omega, \mathcal{F}, P), \forall \, k = 1, 2, \ldots d, \forall \, j = 1, 2, \ldots D \), such that
\[
P \left[ \int_0^t |\Theta_{kj}(s)|^2 \, ds < \infty \right] = 1 \quad \text{and} \quad P \left[ \int_0^t |\Psi_k(s)| \, ds < \infty \right] = 1, \quad \forall \, t \in [0, \infty).
\]
Let \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) be an \( \mathbb{R}^d \)-valued vector semimartingale (see Definition 5.4.23) defined by
\[
X_k(t) = X_k(0) + \int_0^t \Psi_k(s) \, ds + \sum_{j=1}^D \int_0^t \Theta_{kj}(s) \, dW^j(s), \quad \forall \, t \in [0, \infty),
\]
\( \forall \, k = 1, 2, \ldots d \), where \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a given \( \mathbb{R}^D \)-valued standard Wiener process and \( X_k(0) \) are given \( \mathcal{F}_0 \)-measurable random variables on \((\Omega, \mathcal{F}, P)\). A vector semimartingale having the particular structure in (5.295), namely the sum of an \( \mathcal{F}_0 \)-measurable random vector, a vector of (random) Lebesgue integrals, and a vector of stochastic integrals with respect to some common Wiener process of appropriate dimension, is called an Itô process, while the vector-valued process \( \{(\Psi(t); t \in [0, \infty)\} \) is called the drift term and the matrix-valued process \( \{(\Theta(t), t \in [0, \infty)\} \) is called the covariance term of the Itô process. Put \( \Gamma \triangleq \Theta\Theta^T \), where \( \Theta(t, \omega) \) is the \( d \) by \( d \) matrix whose \((k, j)\) element is \( \Theta_{kj}(t, \omega) \). Let \( f : \mathbb{R}^d \to \mathbb{R} \) be a given \( C^2 \)-mapping, and define
\[
(Lf)(t, x, \omega) \triangleq \sum_{i=1}^d \Psi_i(t, \omega) \frac{\partial f}{\partial x^i}(x) + \frac{1}{2} \sum_{i,j=1}^d \Gamma_{ij}(t, \omega) \frac{\partial^2 f}{\partial x^i \partial x^j}(x), \quad \forall \, x \in \mathbb{R}^d, \forall \, t \in [0, \infty), \forall \, \omega \in \Omega.
\]
Use Itô’s formula to show that \( f(X(t)) \) is an Itô process of the form:
\[
f(X(t)) = f(X(0)) + \int_0^t (Lf)(s, X(s)) \, ds + \sum_{k=1}^d \sum_{j=1}^D \int_0^t \frac{\partial f}{\partial x^k}(X(s))\Theta_{kj}(s) \, dW^j(s), \quad \forall \, t \in [0, \infty).
\]
This result is also called Itô’s formula in some textbooks.
Problem 5.8.3 Suppose that \((W(t), \mathcal{F}_t); t \in [0, \infty)\) is a standard scalar Wiener process.

(a) Use Itô’s formula to calculate \(E[\cos^2(W(t))]\) as a function of \(t\). Hint: expand \(\cos^2(W(t))\).

(b) Put \(X(t) \triangleq \exp(t/2) \cos(W(t)), \forall t \in [0, \infty)\). Use Itô’s formula to show that \(X \in \mathcal{M}^c_2(\{\mathcal{F}_t\}, P)\) and calculate \(E([X](t))\) as a function of \(t\). Also, show that \(\mathcal{E}(X) \in \mathcal{M}^c(\{\mathcal{F}_t\}, P)\).

Problem 5.8.4 Suppose that \(X \in \mathcal{S}M^{c,0}(\{\mathcal{F}_t\}, P), Y \in \mathcal{S}M^c(\{\mathcal{F}_t\}, P)\), and

\[
Z(t) \triangleq \mathcal{E}(X)(t) \left\{ Y(0) + \int_0^t \mathcal{E}(X)^{-1}(s) \, dY(s) - \int_0^t \mathcal{E}(X)^{-1}(s) \, d[X,Y](s) \right\}, \quad \forall t \in [0, \infty),
\]

where \(\mathcal{E}(X)\) denotes the Itô exponential of semimartingale \(X\). Use the Itô product formula to show that a.s.:

\[
Z(t) = Y(t) + \int_0^t Z(s) \, dX(s), \quad \forall t \in [0, \infty).
\]

Problem 5.8.5 Suppose that \(\{(W_t, \mathcal{F}_t); t \in [0, \infty)\}\) is an \(\mathbb{R}^d\)-valued standard Wiener process on \((\Omega, \mathcal{F}, P)\), and \(\phi: \mathbb{R}^d \to (0, \infty)\) is twice continuously differentiable. Put

\[
X_t \triangleq \phi(W_t) \exp \left[ - \int_0^t \nabla^2 \phi(W_s) \, ds \right],
\]

where \(\nabla^2 \phi(\cdot)\), the Laplacian of \(\phi(\cdot)\), is defined by

\[(5.296) \quad \nabla^2 \phi(x) \triangleq \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 \phi(x)}{\partial x^i \partial x^i}.
\]

Show that

\[
\lim_{t \to \infty} X_t
\]

exists a.s. and is integrable. Hint: Expand \(X_t\) using Itô’s formula.

Problem 5.8.6 Suppose that \(X \in \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P)\), and define the following iterated stochastic integrals:

\[
I_0(t) \triangleq 1 \quad \text{and} \quad I_n(t) \triangleq \int_0^t I_{n-1}(s) \, dX(s), \quad \forall t \in [0, \infty).
\]

For all \(n = 2, 3, \ldots\) show that a.s:

\[
nI_n(t) = I_{n-1}(t)X(t) - I_{n-2}(t)[X](t), \quad \forall t \in [0, \infty).
\]

Problem 5.8.7 (a) Suppose that \(X \in \mathcal{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P)\). Establish the upper-bound

\[
P \left[ \sup_{s \in [0, \infty)} X(s) \geq \alpha, \ [X](\infty) \leq \beta \right] \leq \exp \left( -\frac{\alpha^2}{2\beta} \right), \quad \alpha \in [0, \infty), \ \beta \in (0, \infty).
\]

Hint: For each \(\lambda \in (0, \infty)\) put \(Z^\lambda(t) \triangleq \mathcal{E}(\lambda X)(t)\), and show that

\[
\left\{ \sup_{s \in [0, \infty)} X(s) \geq \alpha, \ [X](\infty) \leq \beta \right\} \subset \left\{ \sup_{s \in [0, \infty)} Z^\lambda(s) \geq \exp \left( \lambda \alpha - \frac{\lambda^2 \beta}{2} \right) \right\},
\]

262
for all $\alpha, \beta \in (0, \infty)$.

(b) Suppose that $X \in M_{loc}^{c,0}(\mathcal{F}_t, P)$ and there is a constant $c \in (0, \infty)$ such that $[X](t) \leq ct$ for all $t \in [0, \infty)$ a.s. Use the result from (a) to establish the following inequality of S.N. Bernstein:

$$P \left[ \sup_{s \in [0,t]} X(s) \geq \alpha t \right] \leq \exp \left( -\frac{\alpha^2 t}{2c} \right), \quad \alpha, t \in [0, \infty).$$

This extremely useful upper-bound includes the result of Problem 4.8.9(b) as a very special case.

**Problem 5.8.8** Suppose that $\Gamma$ and $\Delta$ are open subsets in $\mathbb{R}^d$, with $\Gamma \subset \Delta$ and $\Gamma$ being bounded. Let $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ be an $\mathbb{R}^d$-valued standard Wiener process on $(\Omega, \mathcal{F}, P)$, and, for each $x \in \mathbb{R}^d$, define

$$W^x(t, \omega) \overset{\Delta}{=} x + W(t, \omega), \quad \forall \, t \in [0, \infty), \forall \, \omega \in \Omega,$$

and

$$T^x(\omega) \overset{\Delta}{=} \inf\{t \in [0, \infty): W^x(t, \omega) \not\in \Gamma\}, \quad \forall \, \omega \in \Omega.$$

If $\varphi: \Delta \to \mathbb{R}$ is a twice-continuously differentiable function, for each $x \in \Delta$, put

$$Z^x(t) \overset{\Delta}{=} \varphi(W^x(t \wedge T^x)) - \frac{1}{2} \int_0^{t \wedge T^x} \nabla^2 \varphi(W^x(s))ds, \quad \forall \, t \in [0, \infty),$$

where, exactly as for (5.296),

$$\nabla^2 \varphi(x) \overset{\Delta}{=} \sum_{i=1}^d \frac{\partial^2 \varphi(x)}{\partial x^i \partial x^i}, \quad \forall \, x \in \Delta.$$

(a) Use Itô’s formula to show that $Z^x \in M^{c}(\mathcal{F}_t, P)$, $\forall \, x \in \Gamma$.

(b) If $\varphi(\cdot)$ is such that $\nabla^2 \varphi(x) = 0$, $\forall \, x \in \Gamma$, show that

$$\varphi(x) = E[\varphi(W^x(T^x))], \quad \forall \, x \in \Gamma.$$

Hint: Use the fact that $P\{T^x < \infty\} = 1$, which follows easily from Problem 4.8.10(a).

(c) Suppose that $R \in (0, \infty)$ is a constant, and for each $x \in \mathbb{R}^d$ put

$$T^x \overset{\Delta}{=} \inf\{t \in [0, \infty): \|W^x(t)\| \geq R\}$$

(here $\|\xi\| \overset{\Delta}{=} \left[ \sum_{i=1}^d |\xi^i|^2 \right]^{1/2}$ denotes the usual Pythagorean norm of $\xi \in \mathbb{R}^d$). Show that

$$E[T^x] = \frac{R^2 - \|x\|^2}{d}, \quad \forall \, x \in \mathbb{R}^d, \text{ with } \|x\| < R.$$

Hint: Use the result established in (a) with $\Delta \overset{\Delta}{=} \mathbb{R}^d$, $\Gamma \overset{\Delta}{=} \{x \in \mathbb{R}^d: \|x\| < R\}$, and $\varphi: \Delta \to \mathbb{R}$ given by $\varphi(\xi) \overset{\Delta}{=} \|\xi\|^2$, $\forall \, \xi \in \mathbb{R}^d$. 

263
(d) Consider the single-dimensional case $d = 1$. Suppose that $a_1, a_2 \in \mathbb{R}$ are constants, with $a_1 < a_2$, and for each $x \in \mathbb{R}$ put

$$T_x^1(\omega) \overset{\triangle}{=} \inf\{t \in [0, \infty) : W^x(t, \omega) \leq a_1\}, \quad T_x^2(\omega) \overset{\triangle}{=} \inf\{t \in [0, \infty) : W^x(t, \omega) \geq a_2\}, \quad \forall \omega \in \Omega.$$

Show that

$$P\{T_1^x < T_2^x\} = \frac{a_2 - x}{a_2 - a_1}, \quad \forall x \in (a_1, a_2).$$

Hint: Use the result in (b) with $\Delta = \mathbb{R}$, $\Gamma \overset{\triangle}{=} (a_1, a_2)$, and $\varphi : \Delta \to \mathbb{R}$ given by $\varphi(x) \overset{\triangle}{=} x$, $\forall x \in \Delta$.

(e) Consider the planar case where $d = 2$. Suppose that $r, R \in \mathbb{R}$ are constants, with $0 < r < R$, and for each $x \in \mathbb{R}^2$ put

$$T_x^1(\omega) \overset{\triangle}{=} \inf\{t \in [0, \infty) : \|W^x(t, \omega)\| \leq r\}, \quad T_x^2(\omega) \overset{\triangle}{=} \inf\{t \in [0, \infty) : \|W^x(t, \omega)\| \geq R\}, \quad \forall \omega \in \Omega.$$

Show that

$$P\{T_1^x < T_2^x\} = \frac{\ln(R) - \ln(\|x\|)}{\ln(R) - \ln(r)}, \quad \forall x \in \mathbb{R}^2 \text{ with } r < \|x\| < R.$$

Hint: Use the result in (b) with $\Delta \overset{\triangle}{=} \{x \in \mathbb{R}^d : x \neq 0\}$, $\varphi : \Delta \to \mathbb{R}$ given by $\varphi(x) \overset{\triangle}{=} \ln(\|x\|), \forall x \in \Delta$, and $\Gamma \overset{\triangle}{=} \{x \in \mathbb{R}^d : r < \|x\| < R\}$.

(f) Consider the case where $d \geq 3$. Suppose that $r, R \in \mathbb{R}$ are constants, with $0 < r < R$, and for each $x \in \mathbb{R}^d$ define $T_x^1$ and $T_x^2$ exactly as in (e). Show that

$$P\{T_1^x < T_2^x\} = \frac{\|x\|^{2-d} - R^{2-d}}{r^{2-d} - R^{2-d}}, \quad \forall x \in \mathbb{R}^d \text{ with } r < \|x\| < R.$$

Hint: Use the result in (b) with $\Delta \overset{\triangle}{=} \{x \in \mathbb{R}^d : x \neq 0\}$, $\varphi : \Delta \to \mathbb{R}$ given by $\varphi(x) \overset{\triangle}{=} \|x\|^{2-d}, \forall x \in \Delta$, and $\Gamma \overset{\triangle}{=} \{x \in \mathbb{R}^d : r < \|x\| < R\}$.

**Problem 5.8.9** (a) Suppose that $\{(W(t), \mathcal{F}_t) ; t \in [0, \infty)\}$ is a planar (i.e. $\mathbb{R}^2$-valued) standard Wiener process on $(\Omega, \mathcal{F}, P)$. For some non-zero $x \in \mathbb{R}^2$ put

$$W^x(t, \omega) \overset{\triangle}{=} x + W(t, x), \quad \forall t \in [0, \infty), \quad \forall \omega \in \Omega.$$

Show that

$$(5.297) \quad P\{W^x(t) \neq 0, \forall t \in [0, \infty)\} = 1.$$  

Hint: Use the result from Problem 5.8.8(e) to show that $P\{T = +\infty\} = 1$ for

$$T(\omega) \overset{\triangle}{=} \inf\{t \in [0, \infty) : \|W^x(t, \omega)\| = 0\}.$$
(b) Now suppose that \( \{(W(t), F_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued standard Wiener process on \((\Omega, F, P)\) for some \( d \geq 3 \). For a non-zero \( x \in \mathbb{R}^d \), define \( W^x(t, \omega) \) as in (a), and show that (5.297) still holds. 
Hint: Use the result from (a).

**Problem 5.8.10** Suppose that \( \{(W(t), F_t); t \in [0, \infty)\} \) is a standard \( \mathbb{R}^d \)-valued Wiener process on \((\Omega, F, P)\), fix some non-zero \( x \in \mathbb{R}^d \) for an integer \( d \geq 2 \), and put

\[
W^x(t, \omega) \triangleq x + W(t, \omega), \quad \forall \ t \in [0, \infty), \ \forall \ \omega \in \Omega,
\]

and

\[
R^x(t, \omega) \triangleq \| W^x(t, \omega) \|, \quad \forall \ t \in [0, \infty), \ \forall \ \omega \in \Omega.
\]

Then \( R^x(t) \) gives the Euclidean distance of \( W^x(t) \) from the origin, and \( \{R^x(t), t \in [0, \infty)\} \) is called a **Bessel process**. Establish the following:

(a) \( \{(V(t), F_t); t \in [0, \infty)\} \) is a standard scalar Wiener process on \((\Omega, F, P)\), for

\[
V(t, \omega) \triangleq \sum_{i=1}^{d} \int_{0}^{t} \frac{W^x_i(s)}{R^x(s)} dW^x_i(s), \quad \forall \ t \in [0, \infty).
\]

where \( W^x_i(s) \) denotes the \( i \)-th scalar element of the \( \mathbb{R}^d \)-vector \( W^x(s) \), \( i = 1, \ldots, d \).

(b) The Bessel process \( \{R^x(t), t \in [0, \infty)\} \) satisfies the identity

\[
R^x(t) = \| x \| + \int_{0}^{t} \frac{d-1}{2R^x(s)} ds + V(t), \quad \forall \ t \in [0, \infty).
\]

**Problem 5.8.11** Suppose that \( \{(W(t), F_t); t \in [0, \infty)\} \) is a scalar standard Wiener process and that \( \{(Z(t), F_t); t \in [0, \infty)\} \) is an \( \mathbb{R} \)-valued progressively measurable process on the complete probability space \((\Omega, F, P)\) with \( |Z(t, \omega)| \leq C, \ \forall \ (t, \omega) \in [0, \infty) \otimes \Omega \), for some constant \( C \in [0, \infty) \). Put

\[
Y(t) \triangleq W(t) + \int_{0}^{t} Z(s) \ ds, \quad \forall \ t \in [0, \infty),
\]

\[
\mathcal{Y}_t \triangleq \sigma\{Y(s), s \in [0,t]\} \lor \mathcal{N}, \quad \forall \ t \in [0, \infty),
\]

where \( \mathcal{N} \triangleq \{N \in \mathcal{F} : P(N) = 0\} \). Also define

\[
I(t) \triangleq Y(t) - \int_{0}^{t} \hat{Z}(s) \ ds, \quad \forall \ t \in [0, \infty),
\]

where \( \{\hat{Z}(t); t \in [0, \infty)\} \) denotes an **optional projection** of the process \( \{Z(t); t \in [0, \infty)\} \) onto the filtration \( \sigma\{\mathcal{Y}_t, \ t \in [0, \infty)\} \). (see Problem 4.8.6). Show that \( \{(I(t), \mathcal{Y}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process.

Hint: First use Lemma 4.5.1 to see that \( \{(I_t, \mathcal{Y}_t); t \in [0, \infty)\} \) is a martingale, then apply Theorem 5.6.2.
Problem 5.8.12 Suppose that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a standard scalar Wiener process, and \( T \) is a \( \{\mathcal{F}_t\} \)-stopping time with \( E[e^{T/2}] < \infty \). Show that
\[
E \left[ \exp \left( W(T) - \frac{T}{2} \right) \right] = 1.
\]

Problem 5.8.13 Suppose that \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued progressively measurable process such that
\[
P \left[ \int_0^t |\Phi^k(s)|^2 \, ds < \infty \right] = 1, \quad \forall k = 1, 2, \ldots, d, \quad \forall t \in [0, \infty),
\]
and \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued standard Wiener process, defined on the common probability space \( (\Omega, \mathcal{F}, P) \). For some constants \( T, a \in (0, \infty) \) define
\[
S \triangleq \inf \{t \in [0, \infty) : (\Phi \cdot W)(t) \geq a \} \wedge T.
\]
where
\[
(\Phi \cdot W)(t) \triangleq \sum_{k=1}^d \int_0^t \Phi^k(s) \, dW^k(s), \quad \forall t \in [0, \infty).
\]
Show that \( \tilde{P} \) defined by
\[
\tilde{P}(A) \triangleq E[\mathcal{E}(\Phi \cdot W)(S); A], \quad \forall A \in \mathcal{F},
\]
is a probability measure on \( (\Omega, \mathcal{F}) \), and \( \{(\tilde{W}(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued standard Wiener process on \( (\Omega, \mathcal{F}, \tilde{P}) \), for \( \{\tilde{W}(t); t \in [0, \infty)\} \) defined by
\[
\tilde{W}^k(t) \triangleq W^k(t) - \int_0^{t \wedge S} \Phi^k(s) \, ds, \quad \forall t \in [0, \infty),
\]
for each \( k = 1, 2, \ldots, d \).

Problem 5.8.14 (a) Suppose that \( X \in \mathcal{M}^{c,0}_{loc}(\{\mathcal{F}_t\}, P) \) is such that its quadratic variation process is deterministic, that is for \( P \)-almost all \( \omega \) we have
\[
[X](t, \omega) = q(t), \quad t \in [0, \infty),
\]
for some continuous non-decreasing function \( q : [0, \infty) \to [0, \infty) \) such that \( q(0) = 0 \). Establish that
(i) \( X(t) - X(s) \sim N(0, q(t) - q(s)) \),
(ii) \( X(t) - X(s) \) is independent of \( \mathcal{F}_s \),
for all \( 0 \leq s < t < \infty \).
Hint: generalize the proof of Theorem 5.6.2, but note that \( q(\cdot) \) is only continuous, not smooth (or even absolutely continuous).
(b) Suppose that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process and \( \Phi : [0, \infty) \to [0, \infty) \) is \( \mathcal{B}[0, \infty) \)-measurable such that
\[
\int_0^t |\Phi(\tau)|^2 \, d\tau < \infty, \quad t \in [0, \infty).
\]
Show that
\[ \int_s^t \Phi(\tau) \, dW(\tau) \sim N(0, \int_s^t |\Phi(\tau)|^2 \, d\tau) \] and \( \int_s^t \Phi(\tau) \, dW(\tau) \) is independent of \( \mathcal{F}_s \), for all \( 0 \leq s < t < \infty \).

**Problem 5.8.15** Suppose that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a standard \( d \)-dimensional Wiener process and \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( \mathbb{R}^d \)-valued progressively measurable process on the probability space \((\Omega, \mathcal{F}, P)\). Establish the following:

(a) If, for each \( t \in (0, \infty) \), there is a finite partition \( 0 = t_0 < t_1 < \ldots < t_N = t \) such that
\[
E \left[ \exp \left( \frac{1}{2} \sum_{k=1}^d \int_{t_{n-1}}^{t_n} |\Phi^k(s)|^2 \, ds \right) \right] < \infty, \quad \forall \ n = 1, 2, \ldots, N,
\]
then \( \{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale.

(b) If, for each \( t \in [0, \infty) \), there are constants \( c_1(t), c_2(t) \in (0, \infty) \) such that
\[
E \left[ \exp \left( c_2(t) \sum_{i=1}^d |\Phi^k(s)|^2 \right) \right] \leq c_1(t), \quad \forall \ s \in [0, t],
\]
then \( \{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale.

(c) If \( \Phi(t) \) is Gaussian distributed for each \( t \in [0, \infty) \), and the functions \( t \to E\Phi(t) \) and \( t \to \text{Cov}\{\Phi(t)\} \) are continuous on \([0, \infty)\), then \( \{(\mathcal{E}(\Phi \cdot W)(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a continuous martingale.

**Hint:** If \( U \) is an \( \mathbb{R}^d \)-valued Gaussian distributed random vector with \( EU = 0 \) and \( \text{Cov}\{U\} = I \), then
\[ E[\exp\{\gamma|U|^2\}] < +\infty, \quad \forall \ \gamma \in (-\infty, 1/2), \]
where \(|x|\) denotes the Euclidean norm of a vector \( x \in \mathbb{R}^d \). Use this fact and (b).

**Problem 5.8.16** Suppose that \( \{(X(t), \mathcal{H}_t); t \in [0, \infty)\} \) is a scalar progressively measurable process and \( \{(W(t), \mathcal{W}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process on \((\Omega, \mathcal{F}, P)\), such that \( \mathcal{H}_\infty \) and \( \mathcal{W}_\infty \) are \( P \)-independent and the \( \sigma \)-algebras \( \mathcal{H}_0 \) and \( \mathcal{W}_0 \) include all \( P \)-null events in \( \mathcal{F} \). Put \( \mathcal{F}_t \triangleq \sigma\{\mathcal{H}_\infty, \mathcal{W}_t\}, \ t \in [0, \infty). \)

(a) Show that \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a standard Wiener process;

(b) Suppose that \( h : \mathbb{R} \to \mathbb{R} \) is bounded and continuous, and put
\[
\Lambda(t) \triangleq \mathcal{E}(-h(X) \cdot W)(t), \quad \forall \ t \in [0, \infty),
\]
\[
Y(t) \triangleq W(t) + \int_0^t h(X(s)) \, ds, \quad \forall \ t \in [0, \infty).
\]
Fix an arbitrary $T \in (0, \infty)$ and define mapping $\tilde{P} : \mathcal{F} \to [0, \infty]$ by

$$\tilde{P}(A) \triangleq E[\Lambda(T); A), \quad \forall A \in \mathcal{F}.$$ 

Establish the following:
(i) $\tilde{P}$ defines a probability measure on $(\Omega, \mathcal{F})$ and $P \equiv \tilde{P}[\mathcal{F}]$.
(ii) $\{(Y(t), \mathcal{F}_t); \ t \in [0, T]\}$ is a standard Wiener process on $(\Omega, \mathcal{F}, \tilde{P})$.
(iii) The $\sigma$-algebras $\mathcal{H}_\infty$ and $\sigma\{Y(s), \ s \in [0, T]\}$ are $\tilde{P}$-independent.

**Problem 5.8.17** Suppose that $\Gamma \subset \mathbb{R}^d$ is a bounded open set, $V : \mathbb{R}^d \to \mathbb{R}$ is a continuous mapping, and $\psi : \mathbb{R}^d \to \mathbb{R}$ is a $C^2$-mapping solving the partial differential equation

$$(5.298) \quad \frac{1}{2} \nabla^2 \psi(x) + V(x)\psi(x) = 0, \quad \forall \ x \in \Gamma,$$

where the Laplacian operator $\nabla^2$ is defined by (5.296). If $\{(W(t), \mathcal{F}_t); \ t \in [0, \infty)\}$ is an $\mathbb{R}^d$-valued standard Wiener process on $(\Omega, \mathcal{F}, P)$, put

$$T^x \triangleq \inf\{t \in [0, \infty) : \ x + W(t) \notin \Gamma\}, \quad \forall \ x \in \Gamma.$$ 

Use Itô’s formula to show that, for each $x \in \Gamma$, we have

$$(5.299) \quad \psi(x) = E \left[ \psi(x + W(t \wedge T^x)) \exp \left( \int_0^{t \wedge T^x} V(x + W(s)) \, ds \right) \right], \quad \forall \ t \in [0, \infty).$$

Equation (5.298) is a variant of the Schrodinger equation of mathematical physics, and the relation (5.299) is the first step in establishing the *Feynman-Kac* representation of the solutions of this equation.
Chapter 6

Representation of Martingales as Stochastic Integrals

In the preceding chapter we developed a powerful theory of stochastic calculus in the general setting of integration with respect to continuous local martingales. The most important example of a continuous local martingale is the Wiener process, which has been introduced in Chapter 3, and stochastic integration where the Wiener process is the stochastic integrator occupies an especially important place within the stochastic calculus. For instance, one can use stochastic integration with respect to a Wiener process to generate a host of other continuous local martingales. It is perhaps worth noting that, from an historical point of view, the first definition of stochastic integration was formulated specifically with the Wiener process as the stochastic integrator (see Itô[15]) and this definition was then later generalized by Kunita and Watanabe [21] to include integration with respect to general local martingales. The Wiener process is significant not only because it is the primary example of a continuous local martingale, but also because it is a useful model for many naturally occurring random phenomena, and because the most impressive applications of stochastic integration theory usually involve a subtle interaction between the abstract stochastic calculus of Chapter 5 on the one hand and the special mathematical properties of Wiener processes on the other. Applications where this is most clearly evident are to be found in a variety of fields including nonlinear filtering (Liptser and Shiryaev [22], Kallianpur [18], Wong and Hajek [32]), financial economics (Duffie [8] and Karatzas and Shreve [20]), quantum mechanics (Simon [28]), and many parts of mathematical analysis (Durrett [9]). Central to the role of Wiener processes in stochastic calculus are basic results of Itô and Doob which, under appropriate conditions, allow one to write a given martingale as the stochastic integral of some stochastic integrand with respect to a Wiener process. In the present chapter our goal is to establish these representation results.
6.1 Wiener Filtrations and Stochastic Integrals

Suppose that \( \{W(t); t \in [0, \infty)\} \) is a given \( d \)-dimensional standard Wiener process on a complete probability space \((\Omega, \mathcal{F}, P)\), and put

\[
\mathcal{F}_t \triangleq \sigma\{\mathcal{F}_t^W, \mathbb{Z}P[\mathcal{F}]\}, \quad \forall \ t \in [0, \infty).
\]

We shall always refer to the filtration which arises from a Wiener process in this way as a Wiener filtration. Corollary 3.4.20 ensures that a Wiener filtration is a standard filtration and hence, if \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale on \((\Omega, \mathcal{F}, P)\), then it follows from Corollary 4.2.19 that \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) has a corol modification, that is, there is some corol process \( \{Y(t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) such that \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale and \( P[X(t) = Y(t)] = 1, \forall \ t \in [0, \infty) \).

The fact that the filtration \( \{\mathcal{F}_t, t \in [0, \infty)\} \) arises specifically from a Wiener process in fact allows one to make a much stronger statement, namely: if \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale then there is a continuous modification \( \{Y(t); t \in [0, \infty)\} \) of \( \{X(t); t \in [0, \infty)\} \) such that \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a martingale. Moreover, \( \{Y(t); t \in [0, \infty)\} \) is also given by the stochastic integral of some progressively measurable integrand with respect to the Wiener process \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) from which the filtration arises. Our goal in the present section is to establish this important result. We begin by recalling some useful ideas from complex analysis:

**Definition 6.1.1** Given a positive integer \( n \), we write \( \mathbb{C}^n \) for the set of all \( n \)-tuples \( z = (z_1, z_2, \ldots, z_n) \) of complex numbers \( z_k, k = 1, 2, \ldots, n \). Given some \( a = (a_1, a_2, \ldots, a_n) \in \mathbb{C}^n \) and \( \epsilon > 0 \), put

\[
P(a, \epsilon) \triangleq \{z \triangleq (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n : |z_k - a_k| < \epsilon, \ \forall \ k = 1, 2, \ldots, n\},
\]

where, for \( x \in \mathbb{C} \) with the form \( x = \alpha + i\beta \), we denote \( |x| \triangleq (\alpha^2 + \beta^2)^{1/2} \). The set \( P(a, \epsilon) \) is called the open polydisc of radius \( \epsilon \) and center \( a \). We write \( \mathbb{N}^n \) for the set of all \( n \)-tuples \( \nu = (\nu_1, \nu_2, \ldots, \nu_n) \) of integers \( \nu_k \geq 0, k = 1, 2, \ldots, n \), and put \( |\nu| \triangleq \nu_1 + \nu_2 + \ldots + \nu_n \). A mapping \( F: \mathbb{C}^n \to \mathbb{C} \) is called analytic at a point \( a \in \mathbb{C}^n \) when there exists some \( \epsilon > 0 \) together with a set of complex numbers \( \{c_\nu : \ \nu \in \mathbb{N}^n\} \) (with \( \epsilon \) and \( c_\nu \) generally depending upon \( a \)), such that, for each \( z \in P(a, \epsilon) \), we have

\[
\sum_{0 \leq k < \infty} \sum_{\nu \in \mathbb{N}^n \atop |\nu| = k} |c_\nu(z_1 - a_1)^{\nu_1}(z_2 - a_2)^{\nu_2} \ldots (z_n - a_n)^{\nu_n}| < \infty,
\]

and

\[
F(z) = \sum_{0 \leq k < \infty} \sum_{\nu \in \mathbb{N}^n \atop |\nu| = k} c_\nu(z_1 - a_1)^{\nu_1}(z_2 - a_2)^{\nu_2} \ldots (z_n - a_n)^{\nu_n}.
\]

The mapping \( F: \mathbb{C}^n \to \mathbb{C} \) is called analytic on \( \mathbb{C}^n \) when it is analytic at each \( a \in \mathbb{C}^n \).

Concerning analytic functions, we need only the following simple result:
Theorem 6.1.2 Suppose that $F : \mathbb{C}^n \to \mathbb{C}$ is analytic on $\mathbb{C}^n$, and $F(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0, \forall \alpha_i \in \mathbb{R}, i = 1, 2, \ldots, n$. Then $F(z_1, z_2, \ldots, z_n) = 0, \forall z_i \in \mathbb{C}, i = 1, 2, \ldots, n$.

In short, if an analytic function of $n$-variables is identically zero when its arguments are restricted to the real numbers then its still identically zero when its arguments are allowed to take complex values.

Remark 6.1.3 Suppose that $X$ is an $\mathbb{R}^d$-valued random vector, Gaussian distributed according to $X \sim N(m, Q)$. Then it is easily shown that

$$E[\exp\{(\theta, X)\}] = \exp\{(\theta, m) + \frac{1}{2}(\theta, Q\theta)\}, \theta \in \mathbb{R}^d.$$ 

The mapping $\theta \to E[\exp\{(\theta, X)\}] : \mathbb{R}^d \to \mathbb{R}$ is called the moment generating function of the random vector $X$.

Remark 6.1.4 Suppose that $X_1, X_2, \ldots, X_n$ are Gaussian random variables on a probability space $(\Omega, \mathcal{F}, P)$, $Y \in L^2(\Omega, \mathcal{F}, P)$, and $z_k = \alpha_k + i\beta_k, k = 1, 2, \ldots, n$, are complex numbers. In view of Remark 6.1.3 and the Cauchy-Schwarz inequality,

$$E \left[ \left| \sum_{k=1}^n z_k X_k \right| \right] Y \leq E^{1/2} \left[ \exp \left( \sum_{k=1}^n \alpha_k X_k \right) \right] E^{1/2} [Y^2] < \infty,$$

and similarly

$$E \left[ \left| \sum_{k=1}^n z_k X_k \right| \right] Y < \infty.$$ 

Thus the mapping $F : \mathbb{C}^n \to \mathbb{C}$ given by

$$(6.2) \quad F(z_1, z_2, \ldots, z_n) \triangleq E \left[ \exp \left( \sum_{k=1}^n z_k X_k \right) Y \right], \forall z_k \in \mathbb{C}, \forall k = 1, 2, \ldots, n,$$

is well-defined. It is easy, although tedious, to show that $F$ is an analytic function on $\mathbb{C}^n$.

Proposition 6.1.5 Suppose that $X_1, X_2, \ldots, X_n$ are Gaussian distributed random variables on a probability space $(\Omega, \mathcal{F}, P)$, and $Y \in L^2(\Omega, \mathcal{F}, P)$. If

$$E \left[ \exp \left( \sum_{k=1}^n \alpha_k X_k \right) Y \right] = 0, \forall \alpha_k \in \mathbb{C}, \forall k = 1, 2, \ldots, n,$$

then $E[Y; A] = 0, \forall A \in \sigma\{X_1, X_2, \ldots, X_n\}$.

Proof: Define $F : \mathbb{C}^n \to \mathbb{C}$ in terms of $X_1, X_2, \ldots, X_n$ and $Y$ as in (6.2). By Remark 6.1.4, the mapping $F$ is analytic on $\mathbb{C}^n$, thus Theorem 6.1.2 together with (6.3) ensures that $F$ is identically zero on $\mathbb{C}^n$, hence, in particular,

$$(6.4) \quad E \left[ \exp \left( i \sum_{k=1}^n \beta X_k \right) Y \right] = 0, \forall \beta_k \in \mathbb{C}, \forall k = 1, 2, \ldots, n.$$
Define $Q_+ : \mathcal{F} \to [0, \infty]$ by
\[
Q_+(A) \triangleq E[Y_+; A], \quad \forall \, A \in \mathcal{F},
\]
and define $Q_- : \mathcal{F} \to [0, \infty]$ in the same way in terms of $Y_-$. From Corollary 1.2.13, and the fact that $E[Y^2] < \infty$ ensures $E[Y_+] < \infty$ and $E[Y_-] < \infty$, we see that $Q_\pm$ are bounded measures on $\mathcal{F}$. In view of Theorem 1.2.17, one can write (6.4) as
\[
\int_\Omega \exp\{i(\beta, X)\}dQ_+ = \int_\Omega \exp\{i(\beta, X)\}dQ_-, \quad \forall \, \beta \triangleq (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{R}^n,
\]
where $X \triangleq (X_1, X_2, \ldots, X_n)$. Next, define measures $\nu_\pm : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ by $\nu_\pm(\Gamma) \triangleq Q_\pm(X^{-1}(\Gamma))$, $\forall \, \Gamma \in \mathcal{B}(\mathbb{R}^n)$. In view of Lemma 1.2.19 and (6.6), we have
\[
\int_{\mathbb{R}^n} \exp\{i(\beta, x)\}d\nu_+ = \int_{\mathbb{R}^n} \exp\{i(\beta, x)\}d\nu_-, \quad \forall \, \beta \in \mathbb{R}^n.
\]
Clearly the $\nu_\pm$ are bounded measures, and taking $\beta = 0$ in (6.7) shows that $\nu_+(\mathbb{R}^n) = \nu_-\mathbb{R}^n) < \infty$. Thus, dividing each side of (6.7) by $\nu_+(\mathbb{R}^n)$, without loss of generality we can regard $\nu_\pm$ as probability measures on $\mathcal{B}(\mathbb{R}^n)$, and conclude from (6.7) and Theorem 1.4.7 that $\nu_+(\Gamma) = \nu_-(\Gamma), \forall \, \Gamma \in \mathcal{B}(\mathbb{R}^n)$. This in turn implies that $Q_+(X^{-1}(\Gamma)) = Q_-(X^{-1}(\Gamma)), \forall \, \Gamma \in \mathcal{B}(\mathbb{R}^n)$, which, together with (6.5), establishes the proposition.

**Definition 6.1.6** A mapping $\Phi : [0, \infty) \to \mathbb{R}$ is a **step function** on $[0, \infty)$ when it is of the form
\[
\Phi(t) = \alpha_1 I_{[0]}(t) + \sum_{k=0}^{n-1} \alpha_{k+1} I_{(t_k, t_{k+1}]}(t), \quad \forall \, t \in [0, \infty),
\]
for some positive integer $n$, some $0 = t_0 < t_1 < \ldots < t_n < \infty$, and some $\alpha_k \in \mathbb{R}, \forall \, k = 1, 2, \ldots, n$. We shall write $S[0, \infty)$ for the set of all step functions on $[0, \infty)$.

**Remark 6.1.7** Suppose that $\{W(t); t \in [0, \infty)\}$ is a scalar standard Wiener process on a complete probability space $(\Omega, \mathcal{F}, P)$, and the filtration $\{\mathcal{F}_t, t \in [0, \infty)\}$ is defined by (6.1). If $\Phi \in S[0, \infty)$ then, trivially, we can regard $\Phi$ as a member of $L^2(W, \{\mathcal{F}_t\}, P)$ (see Definition 5.3.39), and, if $\Phi$ has the form in (6.8), then one easily verifies that an Itô stochastic integral of $\{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\}$ with respect to $\{(W(t), \mathcal{F}_t); t \in [0, \infty)\}$ is given by
\[
(\Phi \cdot W)(t, \omega) = \sum_{k=0}^{n-1} \alpha_{k+1}[W(t \wedge t_{k+1}, \omega) - W(t \wedge t_k, \omega)], \quad \forall \, t \in [0, \infty), \, \forall \, \omega \in \Omega,
\]
(c.f. the proof of Proposition 5.3.46). Since $\Phi(t) = 0, \forall \, t > t_n$, we see that $\Phi \cdot W$ and $\mathcal{E}(\Phi \cdot W) - 1$ are members of $M^0_{2, P}(\{\mathcal{F}_t\}, P)$, with
\[
(\Phi \cdot W)(\infty) = \sum_{k=0}^{n-1} \alpha_{k+1}[W(t_{k+1}) - W(t_k)]
\]
and
\begin{equation}
(6.10) \quad \mathcal{E}(\Phi \cdot W)(\infty) = \exp \left\{ (\Phi \cdot W)(\infty) - \frac{1}{2} \sum_{k=0}^{n-1} \alpha_{k+1}^2 (t_{k+1} - t_k) \right\}.
\end{equation}

Thus, \( (\Phi \cdot W)(\infty) \) is a Gaussian distributed zero-mean random variable, and in view of Remark 6.1.3, it follows that \( \mathcal{E}(\Phi \cdot W)(\infty) \in L^2(\Omega, \mathcal{F}_\infty, P) \).

**Lemma 6.1.8** Suppose that \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on a complete probability space \( (\Omega, \mathcal{F}, P) \), and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the standard Wiener filtration defined by (6.1). If \( Y \in L^2(\Omega, \mathcal{F}_\infty, P) \) is such that \( Y \perp \mathcal{E}(\Phi \cdot W)(\infty), \forall \Phi \in \mathcal{S}[0, \infty) \), then \( P[Y = 0] = 1 \).

Proof: Let \( \Phi \in \mathcal{S}[0, \infty) \) have the form given by (6.8) for arbitrary but fixed \( 0 = t_0 < t_1 < \ldots < t_n < \infty \) and \( \alpha_k \in \mathbb{R}, \forall k = 1, 2, \ldots, n \). From (6.9) and (6.10),
\begin{equation}
E[\mathcal{E}(\Phi \cdot W)(\infty)Y] = E \left[ \exp \left\{ \sum_{k=0}^{n-1} \alpha_{k+1}[W(t_{k+1}) - W(t_k)] \right\} Y \right] \exp \left[ \sum_{k=0}^{n-1} \alpha_{k+1}^2 (t_{k+1} - t_k) \right],
\end{equation}
and hence, since \( Y \perp \mathcal{E}(\Phi \cdot W)(\infty) \), we find that
\begin{equation}
E \left[ \exp \left\{ \sum_{k=0}^{n-1} \alpha_{k+1}[W(t_{k+1}) - W(t_k)] \right\} Y \right] = 0, \quad \forall \alpha_k \in \mathbb{R}, \ k = 1, 2, \ldots, n.
\end{equation}
By Proposition 6.1.5, we see that, for each positive integer \( n \), and arbitrary \( 0 = t_0 < t_1 < \ldots < t_n < \infty \), we have
\begin{equation}
(6.12) \quad E[Y_+; A] = 0, \quad \forall A \in \sigma\{W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})\} \equiv \sigma\{W(t_1), W(t_2), \ldots, W(t_n)\}.
\end{equation}
Now let \( \mathcal{D} \) be the collection of all sets \( A \subset \Omega \) with the form \( A = \{W(t_k) \in \Gamma_k, \ k = 0, 1, 2, \ldots, n\} \) for positive integers \( n \), and arbitrary \( 0 = t_0 < t_1 < \ldots < t_n < \infty \), \( \Gamma_k \in \mathcal{B}(\mathbb{R}), \ k = 0, 1, 2, \ldots, n \). Then \( \mathcal{D} \) is a \( \pi \)-class over \( \Omega \), and clearly
\begin{equation}
(6.13) \quad \sigma\{\mathcal{D}\} = \mathcal{F}_\infty^w.
\end{equation}
Now (6.12) ensures that \( E[Y_+; A] = E[Y_-; A], \forall A \in \mathcal{D} \), hence, by Theorem 1.2.4 and (6.13),
\begin{equation}
(6.14) \quad E[Y; A] = 0, \quad \forall A \in \mathcal{F}_\infty^w.
\end{equation}
Next, fix some \( B \in \mathcal{F}_\infty \) and note from Proposition 3.4.22 there exists some \( A \in \mathcal{F}_\infty^w \) such that \( A \triangle B \in \mathcal{Z}^P[\mathcal{F}] \), whence \( E[Y; B] = E[Y; A] \). In view of (6.14) we get \( E[Y; B] = 0, \forall B \in \mathcal{F}_\infty \). Since \( Y \) is \( \mathcal{F}_\infty \)-measurable we see from Theorem 1.2.24 that \( P[Y = 0] = 1 \).
Remark 6.1.9 When \( \{(W(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a scalar standard Wiener process define the inner product on the real vector space \( \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) (recall Definition 5.3.39) by

\[
(\Phi, \Psi)_W \buildrel \Delta \over = E \left[ \int_0^\infty \Phi(s)\Psi(s) \, ds \right].
\]

An argument identical to that used for Proposition 5.3.10 establishes that \( \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) with this inner product is a Hilbert space.

We are now able to establish the principal result of this section:

**Theorem 6.1.10 (Itô)** Suppose that \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on a complete probability space \( (\Omega, \mathcal{F}, P) \) and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the standard Wiener filtration defined by (6.1). For each \( X \in \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \) there is some \( \Phi \in \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) such that

\[
X = EX + (\Phi \cdot W)(\infty) \quad \text{a.s.}
\]

Moreover, \( \Phi \) is unique in the sense that if \( \Psi \in \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) is such that (6.16) holds with \( \Psi \) substituted for \( \Phi \), then

\[
(\lambda \otimes P)\{(t, \omega) \in [0, \infty) \times \Omega : \Phi(t, \omega) \neq \Psi(t, \omega)\} = 0.
\]

Proof: Denote by \( \mathbb{H} \) the collection of all \( X \in \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \) corresponding to which there is some \( \Phi \in \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) such that (6.16) holds. Clearly \( \mathbb{H} \) is a linear subspace of \( \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \). We next show that this linear subspace is closed. If \( X \in \mathbb{H} \) and \( \Phi \in \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \) is such that (6.16) holds then

\[
E[X^2] = (EX)^2 + E[((\Phi \cdot W)(\infty))^2]
\]

(recall from Proposition 5.3.40 that \( E[(\Phi \cdot W)(\infty)] = 0 \)). Suppose that \( \{X_n, n = 1, 2, \ldots\} \) is a sequence in \( \mathbb{H} \) which is \( L^2 \)-convergent to some \( X \in \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \), and let \( \Phi_n \in \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) be such that

\[
X_n = EX_n + (\Phi_n \cdot W)(\infty).
\]

Then \( \{X_n, n = 1, 2, \ldots\} \) is a Cauchy sequence in \( \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \), and consequently \( \{EX_n, n = 1, 2, \ldots\} \) is a Cauchy sequence in \( \mathbb{R} \). Since \( X_n - X_m \in \mathbb{H} \), one then easily sees from the identity (6.18) that \( \{(\Phi_n \cdot W)(\infty), n = 1, 2, \ldots\} \) is a Cauchy sequence in \( \mathbb{L}^2(\Omega, \mathcal{F}_\infty, P) \), whence, by Proposition 5.3.40, it follows that \( \{\Phi_n, n = 1, 2, \ldots\} \) is a Cauchy sequence in \( \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \). Thus, in view of Remark 6.1.9, there is some \( \Phi \in \mathbb{L}^2(W, \{\mathcal{F}_t\}, P) \) such that

\[
\lim_{n \to \infty} E \left[ \int_0^\infty |\Phi_n(s) - \Phi(s)|^2 \, ds \right] = 0,
\]

or, in view of Proposition 5.3.40,

\[
\lim_{n \to \infty} E[|((\Phi_n \cdot W)(\infty) - (\Phi \cdot W)(\infty)|^2] = 0.
\]
Thus, taking \( L^2 \)-limits in (6.19), one obtains
\[
X = \left( \lim_{n \to \infty} E X_n \right) + (\Phi \cdot W)(\infty),
\]
whence \( X \in \mathbf{H} \), as required to establish that \( \mathbf{H} \) is a closed linear subspace of \( L^2(\Omega, \mathcal{F}_\infty, P) \). It remains to show that \( \mathbf{H} = L^2(\Omega, \mathcal{F}_\infty, P) \) in order to get the desired result. Fix an arbitrary \( \Phi \in \mathbf{S}[0, \infty) \). An application of Itô’s formula shows that a.s.:
\[
E(\Phi \cdot W)(t) = 1 + \int_0^t \Phi(s) E(\Phi \cdot W)(s) \, dW(s), \quad \forall \ t \in [0, \infty),
\]
(6.20) (compare with Proposition 5.5.2(b)). From Remark 6.1.7 we observe that \( E(\Phi \cdot W) - 1 \in \mathbf{M}_{2, b}^c(\{\mathcal{F}_t\}, P) \), and, since \( \Phi \in \mathbf{S}[0, \infty) \) one trivially sees that
\[
E \left[ \int_0^t (\Phi(s) E(\Phi \cdot W)(s))^2 \, ds \right] < \infty, \quad \forall \ t \in [0, \infty).
\]
(6.21)
It follows that
\[
E \left[ \int_0^t (\Phi(s) E(\Phi \cdot W)(s))^2 \, ds \right] = E \left[ \left( \int_0^t \Phi(s) E(\Phi \cdot W)(s) \, dW(s) \right)^2 \right]
= E[ (E(\Phi \cdot W)(t) - 1)^2 ] \leq E[(E(\Phi \cdot W)(\infty) - 1)^2] < \infty,
\]
for each \( t \in [0, \infty) \). The first equality in (6.22) is justified by (6.21) and Proposition 5.3.40, the second equality follows from (6.20), and the two inequalities are a consequence of the fact that \( E(\Phi \cdot W) - 1 \in \mathbf{M}_{2, b}^c(\{\mathcal{F}_t\}, P) \) together with Jensen’s inequality. From (6.22) we conclude that
\[
E \left[ \int_0^\infty (\Phi(s) E(\Phi \cdot W)(s))^2 \, ds \right] < \infty,
\]
and hence that \( \Phi E(\Phi \cdot W) \in \mathbf{L}^2(\mathcal{W}, \{\mathcal{F}_t\}, P) \). Thus, since both sides of (6.20) are members of \( \mathbf{M}_{2, b}^c(\{\mathcal{F}_t\}, P) \), we can take \( L^2 \)-limits as \( t \to \infty \) and get
\[
E(\Phi \cdot W)(\infty) = 1 + \int_0^\infty \Phi(s) E(\Phi \cdot W)(s) \, dW(s),
\]
whence it follows that \( E(\Phi \cdot W)(\infty) \in \mathbf{H}, \forall \ \Phi \in \mathbf{S}[0, \infty) \). Thus, if \( Y \in L^2(\Omega, \mathcal{F}_\infty, P) \) is such that \( Y \perp \mathbf{H} \), then \( E[ E(\Phi \cdot W)(\infty) Y] = 0, \forall \ \Phi \in \mathbf{S}[0, \infty) \), and hence, from Lemma 6.1.8, we find \( Y = 0 \). Since \( \mathbf{H} \) is a closed linear subspace of \( L^2(\Omega, \mathcal{F}_\infty, P) \), one sees from Theorem 1.1.1 that \( \mathbf{H} = L^2(\Omega, \mathcal{F}_\infty, P) \). Finally, if \( \Psi \in \mathbf{L}^2(\mathcal{W}, \{\mathcal{F}_t\}, P) \) is such that (6.16) holds with \( \Psi \) in place of \( \Phi \), then Proposition 5.3.40 ensures that
\[
E \left[ \int_0^\infty |\Phi(s) - \Psi(s)|^2 \, ds \right] = 0,
\]
from which we get (6.17).
Corollary 6.1.11 Suppose that \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on a complete probability space \((\Omega, \mathcal{F}, P)\), and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the standard Wiener filtration in \((\Omega, \mathcal{F}, P)\) defined by (6.1). If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^2 \)-bounded martingale then there is some \( \Phi \in L^2(W, \{\mathcal{F}_t\}, P) \) such that
\[
X(t) = EX(0) + \int_0^t \Phi(s) \, dW(s) \quad \text{a.s.}
\]
for each \( t \in [0, \infty] \). In particular \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) has a continuous modification which is defined by the right hand side of (6.23). Moreover, \( \Phi \) is unique in the sense that if \( \Psi \in L^2(W, \{\mathcal{F}_t\}, P) \) is such that (6.23) holds, with \( \Psi \) in place of \( \Phi \), for each \( t \in [0, \infty] \), then (6.17) holds.

Proof: By Corollary 3.4.20 and Corollary 4.2.19 there is some corollary martingale \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) such that \( P[X(t) = Y(t)] = 1, \forall t \in [0, \infty) \). Thus, from Theorem 4.4.12, there exists some \( Y(\infty) \in L^2(\Omega, \mathcal{F}_\infty, P) \) such that
\[
Y(t) = E[Y(\infty) | \mathcal{F}_t] \quad \text{a.s.}
\]
for each \( t \in [0, \infty] \). By Theorem 6.1.10 there exists some \( \Phi \in L^2(W, \{\mathcal{F}_t\}, P) \) such that
\[
Y(\infty) = E[Y(\infty)] + (\Phi \bullet W)(\infty) \quad \text{a.s.}
\]
Since \( (\Phi \bullet W) \in M_{2,0}^c(\{\mathcal{F}_t\}, P) \) clearly
\[
(\Phi \bullet W)(t) = E[(\Phi \bullet W)(\infty) | \mathcal{F}_t] \quad \text{a.s.}
\]
for each \( t \in [0, \infty] \). Combining (6.24), (6.25), (6.26), we have
\[
Y(t) = E[Y(\infty)] + (\Phi \bullet W)(t), \quad \text{a.s.}
\]
for each \( t \in [0, \infty] \). Since \( E[X(0)] = E[Y(0)] = E[Y(\infty)] \) and \( P[X(t) = Y(t)] = 1 \), we obtain (6.23). Finally, uniqueness of \( \Phi \) follows exactly as in Theorem 6.1.10.

We next want to obtain a representation formula similar to that in Corollary 6.1.11, but for local martingales. To this end we must first establish some simple preliminary results:

Proposition 6.1.12 Suppose \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on a complete probability space \((\Omega, \mathcal{F}, P)\) and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the standard Wiener filtration in \((\Omega, \mathcal{F}, P)\) defined by (6.1). If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a uniformly integrable martingale then it has a continuous modification.

Proof: By Corollary 3.4.20 and Corollary 4.2.19 there is some corollary martingale \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty)\} \) on \((\Omega, \mathcal{F}, P)\) such that \( P[X(t) = Y(t)] = 1, \forall t \in [0, \infty) \). Thus, from Theorem 4.4.11 and the obvious uniform integrability of \( \{Y(t); t \in [0, \infty)\} \), there is some \( Y(\infty) \in L^1(\Omega, \mathcal{F}_\infty, P) \) such that
\[
Y(t) = E[Y(\infty) | \mathcal{F}_t] \quad \text{a.s.}
\]
for each $t \in [0, \infty)$, and

$$(6.28) \quad \lim_{t \to \infty} Y(t) = Y(\infty), \quad \text{(a.s.} \& \ L^1 \text{-sense).}$$

By Corollary 2.6.1, for each $n = 1, 2, \ldots$ we have that $E[|Y(\infty)|; |Y(\infty)| > c_n] < 2^{-n}$, for some $c_n \in [0, \infty)$, thus, if we define $Z_n \overset{\Delta}{=} Y(\infty)I\{|Y(\infty)| \leq c_n\}$, then $E[|Z_n - Y(\infty)|] < 2^{-n}, \forall \ n = 1, 2, \ldots$ Now put $Y_n(t) \overset{\Delta}{=} E \left[ Z_n \mid \mathcal{F}_t \right], \forall t \in [0, \infty)$. Since $|Z_n(\omega)| \leq c_n, \forall \omega \in \Omega$, it follows that $E[|Y_n(t)|^2] \leq c_n^2 < \infty, \forall \ t \in [0, \infty)$, thus each $\{(Y_n(t), \mathcal{F}_t); t \in [0, \infty)\}$ is an $L^2$-bounded martingale. In view of Corollary 6.1.11, there is some continuous modification $\{\tilde{Y}_n(t), \mathcal{F}_t); t \in [0, \infty)\}$ of $\{(Y_n(t), \mathcal{F}_t); t \in [0, \infty)\}$. Thus (see Proposition 4.1.12) $\{(|Y(t) - \tilde{Y}_n(t)|, \mathcal{F}_t); t \in [0, \infty)\}$ is a right-continuous submartingale and hence, for each $T \in [0, \infty)$, we have

$$(6.29) \quad P[\sup_{0 \leq t \leq T} |Y(t) - \tilde{Y}_n(t)| > n^{-1}] \leq nE[|Y(T) - Y_n(T)|]$$

$$\quad \leq nE[|Y(\infty) - Z_n|] < n2^{-n}.$$

Here we have used Theorem 4.3.3(i) together with $\tilde{Y}_n(T) = Y_n(T)$ a.s. at the first inequality, and (6.27) along with Jensen’s inequality for conditional expectations at the second. Taking $T \to \infty$ in (6.29) then gives

$$(6.30) \quad P[\sup_{t \in [0, \infty)} |Y(t) - \tilde{Y}_n(t)| > n^{-1}] < n2^{-n}, \forall \ n = 1, 2, \ldots$$

Since $\sum n2^{-n} < \infty$, the Borel-Cantelli Theorem 1.2.5(f) shows that there is some $N \in \mathcal{F}$, with $P(N) = 0$, such that

$$(6.31) \quad \lim_{n \to \infty} \sup_{t \in [0, \infty)} |Y(t, \omega) - \tilde{Y}_n(t, \omega)| = 0, \forall \omega \notin N.$$

Since $t \to \tilde{Y}_n(t, \omega)$ is continuous on $[0, \infty), \forall \ n = 1, 2, \ldots, \forall \omega \in \Omega$, we see from (6.31) that $t \to Y(t, \omega)$ is continuous on $[0, \infty)$ for each $\omega \notin N$. Now define $Y(t, \omega) \overset{\Delta}{=} Y(t, \omega), \forall t \in [0, \infty)$, for each $\omega \notin N$, and $\tilde{Y}(t, \omega) \overset{\Delta}{=} 0, \forall t \in [0, \infty)$, for each $\omega \in N$. Since $N \in \mathcal{F}_0$ we see that $\{\tilde{Y}(t, \mathcal{F}_t); t \in [0, \infty)\}$ is a continuous martingale which is indistinguishable from $\{(Y(t, \mathcal{F}_t); t \in [0, \infty)\}$ and hence a continuous modification of $\{(X(t, \mathcal{F}_t); t \in [0, \infty)\}$. \hfill\blacksquare

**Proposition 6.1.13** Suppose $\{W(t); t \in [0, \infty)\}$ is a scalar standard Wiener process on a complete probability space $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t, t \in [0, \infty)\}$ is the standard Wiener filtration in $(\Omega, \mathcal{F}, P)$ defined by (6.1). If $\{(X(t, \mathcal{F}_t); t \in [0, \infty)\}$ is a local martingale then it has a continuous modification.

Proof: In view of Remark 4.6.2 there is some localizing sequence $\{T_n, n = 1, 2, \ldots\}$ of $\{\mathcal{F}_t\}$-stopping times such that $\{(X(t \land T_n), \mathcal{F}_t); t \in [0, \infty)\}$ is a uniformly integrable martingale with $T_n \leq n, \forall \ n = 1, 2, \ldots$ By Corollary 6.1.12, for each $n = 1, 2, \ldots$ there is some continuous martingale $\{(Y_n(t), \mathcal{F}_t); t \in [0, \infty)\}$ on $(\Omega, \mathcal{F}, P)$ such that

$$(6.32) \quad P[X(t \land T_n) = Y_n(t)] = 1, \forall \ t \in [0, \infty).$$
Next, fix some arbitrary \( n = 1, 2, \ldots \) and \( t \in [0, \infty) \), and observe that \( X(t \wedge T_n) \equiv X(t \wedge T_{n+1}) \) on \( \{ t \leq T_n \} \), whence, from (6.32), we have \( Y_n(t) = Y_{n+1}(t) \) a.s. on \( \{ t \leq T_n \} \), or equivalently,

\begin{equation}
Y_n(t)I\{ t \leq T_n \} = Y_{n+1}(t)I\{ t \leq T_n \} \quad \text{a.s.}
\end{equation}

for each \( t \in [0, \infty) \). Since each side of (6.33) defines a left-continuous process, it follows from Proposition 3.1.11 that these processes are indistinguishable, thus \( P(\Omega_n) = 1, \forall \ n = 1, 2, \ldots \) for

\begin{equation}
\Omega_n \triangleq \{ \omega \in \Omega : \ Y_n(t, \omega) = Y_{n+1}(t, \omega), \ \forall \ t \in [0, T_n(\omega)] \}.
\end{equation}

Clearly then, \( P(\Omega^*) = 1 \) for \( \Omega^* \in \mathcal{F}_\infty \) defined by

\begin{equation}
\Omega^* \triangleq \left[ \bigcap_{1 \leq n < \infty} (\Omega_n) \right] \bigcap \left[ \lim_{n \to \infty} T_n = +\infty \right].
\end{equation}

Next, for each \( \omega \in \Omega \) and \( t \in [0, \infty) \), define

\begin{equation}
Y^*(t, \omega) \triangleq \limsup_{n \to \infty} Y_n(t, \omega),
\end{equation}

and observe from (6.34) and (6.35) that, for each \( \omega \in \Omega^* \), we have

\begin{equation}
Y^*(t, \omega) = Y_n(t, \omega), \quad \forall \ t \in [0, T_n(\omega)],
\end{equation}

for all \( n = 1, 2, \ldots \) Moreover, when \( \omega \in \Omega^* \), then \( \lim_{n \to \infty} T_n(\omega) = +\infty \), thus (6.37) together with continuity on \( [0, \infty] \) of the mapping \( t \to Y_n(t, \omega) \) establishes that \( t \to Y^*(t, \omega) \) is continuous on \( [0, \infty] \).

Since the mapping \( \omega \to Y_n(t, \omega) : \Omega \to \mathbb{R} \) is \( \mathcal{F}_t \)-measurable for each \( n = 1, 2, \ldots \) Proposition 1.2.6 shows that \( \omega \to Y^*(t, \omega) : \Omega \to \mathbb{R} \) is \( \mathcal{F}_t \)-measurable. Finally, define \( \{ Y(t); \ t \in [0, \infty) \} \) as follows: for \( \omega \in \Omega^* \), put \( Y(t, \omega) \triangleq Y^*(t, \omega), \ \forall \ t \in [0, \infty) \), and for \( \omega \notin \Omega^* \) put \( Y(t, \omega) \triangleq 0, \ \forall \ t \in [0, \infty) \). Since \( \mathcal{F}_0 \) includes all \( P \)-null events in \( \mathcal{F} \), we see that \( \{ (Y(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a continuous adapted process. Fixing \( t \in [0, \infty) \), clearly

\begin{equation}
Y(t \wedge T_n) = Y^*(t \wedge T_n) = Y_n(t) \quad \text{a.s.}
\end{equation}

(see (6.37) for the second equality), and combining this with (6.32) gives \( X(t \wedge T_n) = Y(t \wedge T_n) \) a.s. \( \forall \ n = 1, 2, \ldots \) Taking \( n \to \infty \) then gives \( X(t) = Y(t) \) a.s. thus \( \{ Y(t); \ t \in [0, \infty) \} \) is a continuous modification of \( \{ X(t); \ t \in [0, \infty) \} \). Since each \( \{ (Y_n(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a martingale we see from (6.38) that \( \{ (Y(t \wedge T_n), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a martingale, as required to show that \( \{ (Y(t), \mathcal{F}_t); \ t \in [0, \infty) \} \) is a continuous local martingale.

It is now easy to extend the representation formula in Corollary 6.1.11 to the case of local martingales:

**Theorem 6.1.14** Suppose that \( \{ W(t); t \in [0, \infty) \} \) is a scalar standard Wiener process on the complete probability space \( (\Omega, \mathcal{F}, P) \) and \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is the standard Wiener filtration in \( (\Omega, \mathcal{F}, P) \) defined.
by (6.1). If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty]\} \) is a local martingale with \( X(0) \equiv 0 \), then there exists some \( \Phi \in L_{loc}^2(W, \{\mathcal{F}_t\}, P) \) such that
\[
X(t) = \int_0^t \Phi(s)dW(s) \quad \text{a.s.}
\]
for each \( t \in [0, \infty) \). In particular, \( \{(X(t), \mathcal{F}_t); t \in [0, \infty]\} \) has a continuous modification which is defined by the right hand side of (6.39). Moreover, \( \Phi \) is unique in the sense that, if \( \{\{\Psi(t), \mathcal{F}_t\); t \in [0, \infty]\} \in L_{loc}^2(W, \{\mathcal{F}_t\}, P) \) is such that (6.39) holds for each \( t \in [0, \infty) \), but with \( \Psi \) in place of \( \Phi \), then we have (6.17).

Proof: By Proposition 6.1.13 there is some continuous local martingale \( \{(Y(t), \mathcal{F}_t); t \in [0, \infty]\} \) on \( (\Omega, \mathcal{F}, P) \) such that \( P[X(t) = Y(t)] = 1, \forall t \in [0, \infty) \). Proposition 4.6.9 gives a localizing sequence \( \{T_n, n = 1, 2, \ldots\} \) of \( \{\mathcal{F}_t\} \)-stopping times such that each \( \{(Y(t \wedge T_n), \mathcal{F}_t) ; t \in [0, \infty)\} \) is a uniformly bounded and hence \( L^2 \)-bounded martingale. Replacing \( T_n \) by \( T_n \wedge n \) if necessary, we can also assume that \( T_n \leq n \). In view of Corollary 6.1.11, for each \( n = 1, 2, \ldots \) there is some \( \Phi_n \in L^2(W, \{\mathcal{F}_t\}, P) \) such that we have a.s.
\[
Y(t \wedge T_n) = \int_0^t \Phi_n(s) dW(s), \quad \forall t \in [0, \infty).
\]
Since \( Y(t \wedge T_n) \equiv Y((t \wedge T_n) \wedge T_{n+1}) \), it follows from Theorem 5.3.33 and (6.40) that, a.s.
\[
E[\int_0^\infty |\Phi_n(s) - \Phi_{n+1}[0, T_n](s)|^2 ds] = 0,
\]
or, equivalently,
\[
\Phi_n[0, T_n] = \Phi_{n+1}[0, T_n], \quad \lambda \otimes P - \text{a.e. on } [0, \infty) \otimes \Omega,
\]
for each \( n = 1, 2, \ldots \) Now define
\[
\Phi(t, \omega) \triangleq \limsup_{n \to \infty} \Phi_n(t, \omega), \quad \forall t \in [0, \infty), \quad \forall \omega \in \Omega.
\]
Then \( (t, \omega) \to \Phi(t, \omega) : [0, \infty) \otimes \Omega \to \mathbb{R} \) is \( PM \{\mathcal{F}_t\} \)-measurable (see Proposition 3.1.22), and (6.42) together with the fact that \( \lim_{n \to \infty} T_n(\omega) = +\infty, \forall \omega \in \Omega \), shows that \( (\lambda \otimes P)\{\Phi(t, \omega) | \Phi(t, \omega)\} = +\infty \} = 0 \). Thus \( \{(\Phi(t), \mathcal{F}_t) ; t \in [0, \infty)\} \) is a progressively measurable process, and, from (6.42),
\[
\Phi_n[0, T_n] = \Phi[0, T_n], \quad \lambda \otimes P - \text{a.e. on } [0, \infty) \otimes \Omega,
\]
for each \( n = 1, 2, \ldots \) It follows that \( \Phi[0, T_n] \in L^2(W, \{\mathcal{F}_t\}, P) \), or equivalently, \( \Phi \in L^2(W^{T_n}, \{\mathcal{F}_t\}, P) \), \( \forall n = 1, 2, \ldots \) Now \( W^{T_n} \in M_{2,b}^0(\{\mathcal{F}_t\}, P) \) (since \( T_n \leq n \), \( \forall n = 1, 2, \ldots \), thus \( \Phi \in L_{loc}^2(W, \{\mathcal{F}_t\}, P) \) (recall
Proposition 5.3.28). From (6.43) we have a.s.
\begin{equation}
\int_{0}^{t \wedge T_n} \Phi_n(s) \, dW(s) = \int_{0}^{t \wedge T_n} \Phi(s) \, dW(s), \quad \forall \ t \in [0, \infty),
\end{equation}
for each \( n = 1, 2, \ldots \). Fixing some arbitrary \( t \in [0, \infty) \) and combining (6.40) and (6.44), we obtain
\[ Y(t \wedge T_n) = \int_{0}^{t \wedge T_n} \Phi(s) \, dW(s) \quad \text{a.s.} \]
Taking \( n \to \infty \), and recalling that \( X(t) = Y(t) \) a.s., gives (6.39). The proof of uniqueness is straightforward and is left to the reader. \( \square \)

The next result deals with the intermediate case of \( L^2 \)-martingales relative to a Wiener filtration and thus complements the representation formulae in Corollary 6.1.11 and Theorem 6.1.14. The proof is a simple modification of the argument used for Theorem 6.1.14 and is left to the reader:

**Theorem 6.1.15** Suppose that \( \{W(t); t \in [0, \infty)\} \) is a scalar standard Wiener process on the complete probability space \((\Omega, \mathcal{F}, P)\) and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the standard Wiener filtration in \((\Omega, \mathcal{F}, P)\) defined by (6.1). If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^2 \)-martingale, then there exists some progressively measurable process \( \{(\Phi(t), \mathcal{F}_t); t \in [0, \infty)\} \) such that
\begin{equation}
E \left[ \int_{0}^{t} |\Phi(s)|^2 \, ds \right] < \infty \quad \text{and} \quad X(t) = E[X(0)] + \int_{0}^{t} \Phi(s) \, dW(s) \quad \text{a.s.}
\end{equation}
for each \( t \in [0, \infty) \). In particular, \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) has a continuous modification which is defined by the second equation in (6.45). Moreover, \( \Phi \) is unique in the sense that, if \( \{(\Psi(t), \mathcal{F}_t); t \in [0, \infty)\} \) is a progressively measurable process such that (6.45) holds for each \( t \in [0, \infty) \), but with \( \Psi \) in place of \( \Phi \), then we have (6.17).

Finally, we observe that the preceding representation formulae continue to hold in the case where the Wiener process which generates the standard Wiener filtration is multidimensional. The proofs involve only straightforward modification of the previous arguments and are omitted:

**Theorem 6.1.16** Suppose that \( \{W(t); t \in [0, \infty)\} \) is a \( d \)-dimensional Wiener process on a complete probability space \((\Omega, \mathcal{F}, P)\), \( W(t) \triangleq (W^1(t), W^2(t), \ldots, W^d(t)) \), and \( \{\mathcal{F}_t, t \in [0, \infty)\} \) is the Wiener filtration defined by (6.1). Then the following hold:
(a) If \( \{(X_t, \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^2 \)-martingale then there exist progressively measurable processes \( \{(\Phi^k(t), \mathcal{F}_t); t \in [0, \infty)\}, \ k = 1, 2, \ldots, d, \) such that
\begin{equation}
E \left[ \sum_{k=1}^{d} \int_{0}^{t} |\Phi^k(s)|^2 \, ds \right] < \infty \quad \text{and} \quad X(t) = E[X(0)] + \sum_{k=1}^{d} \int_{0}^{t} \Phi^k(s) \, dW^k(s) \quad \text{a.s.}
\end{equation}
for each \( t \in [0, \infty) \). In particular, \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) has a continuous modification which is defined by the second equation in (6.46). If, moreover, \( \{(X(t), \mathcal{F}_t); t \in [0, \infty)\} \) is an \( L^2 \)-bounded martingale, then the integrands \( \Phi^k \) are members of \( L^2(W, \{\mathcal{F}_t\}, P) \) and (6.46) holds for each \( t \in [0, \infty) \).
(b) If \( \{(X(t), \mathcal{F}_t); t \in [0, \infty) \} \) is a local martingale with \( X(0) \equiv 0 \), then there exist some \( \Phi^k \in L^2_{\text{loc}}(W, \{\mathcal{F}_t\}, P), k = 1, 2, \ldots, d \), such that

\[
X(t) = \sum_{k=1}^{d} \int_0^t \Phi(s) dW(s) \quad \text{a.s.}
\]

(6.47)

for each \( t \in [0, \infty) \). In particular, \( \{(X(t), \mathcal{F}_t); t \in [0, \infty) \} \) has a continuous modification which is defined by the right hand side of (6.47).

In cases (a) and (b) the integrands \( \Phi^k \) are unique to within \((\lambda \otimes P)\)-null sets.
Appendix A

Glossary of Notation

A.1 Spaces of Stochastic Processes

We need several spaces of stochastic processes in these notes. These spaces are all defined with reference
to the following set-up: \( \{ \mathcal{F}_t, t \in [0, \infty) \} \) is a given filtration (Definition 3.1.14) in the probability space
\( (\Omega, \mathcal{F}, P) \). In the following list we indicate where each space of processes has been defined in the text:

I. Spaces of Martingales and Semimartingales
1. \( \mathcal{M}(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(a)]
2. \( \mathcal{M}_p(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(b)]
3. \( \mathcal{M}_{p,b}(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(b)]
4. \( \mathcal{M}^c(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(c)]
5. \( \mathcal{M}^c_p(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(c)]
6. \( \mathcal{M}^c_{p,b}(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(c)]
7. \( \mathcal{M}^c_{p,0}(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(d)]
8. \( \mathcal{M}^c_{p,b,0}(\{\mathcal{F}_t\}, P) \) [Remark 4.1.6(d)]
9. \( \mathcal{M}_{loc}(\{\mathcal{F}_t\}, P) \) [Remark 4.6.5]
10. \( \mathcal{M}^c_{loc}(\{\mathcal{F}_t\}, P) \) [Remark 4.6.5]
11. \( \mathcal{M}^c_{loc,0}(\{\mathcal{F}_t\}, P) \) [Remark 4.6.5]
12. \( \mathcal{S}^c(\{\mathcal{F}_t\}, P) \) [Remark 5.4.2]
13. \( \mathcal{S}^c_{loc}(\{\mathcal{F}_t\}, P) \) [Remark 5.4.2]

II. Spaces of Continuous Bounded Variation Processes
1. \( \mathcal{F}^c(\{\mathcal{F}_t\}) \) [Remark 5.2.5]
2. \( \mathcal{F}^c_{loc}(\{\mathcal{F}_t\}) \) [Remark 5.2.5]

III. Spaces of Integrands for Itô Stochastic Integrals
1. $L^2(X, \{\mathcal{F}_t\}, P)$ for $X \in M_{2,b}^c(\mathcal{F}_t, P)$ [Definitions 5.3.8 and 5.3.39]
2. $L^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)$ for $X \in M_{\text{loc}}^c(\mathcal{F}_t, P)$ [Definition 5.3.27]
3. $L^2_{\text{loc}}(X, \{\mathcal{F}_t\}, P)$ for $X \in \mathbf{S}^c(\mathcal{F}_t, P)$ [Definition 5.4.8]
4. $B_{\text{loc}}(\mathcal{F}_t)$ [Remark 5.4.14]
5. $L^2(W, \{\mathcal{F}_t\}, P)$ for a standard Wiener process $W$ [Definition 5.3.39]

A.2 Miscellaneous Symbols

1. $\|f\|_p$ [Eqn. (1.10)]
2. $N_\nu$ [Definition 1.2.28]
3. $\nu \ll \mu[S]$ [Definition 1.2.28]
4. $\nu \equiv \mu[S]$ [Definition 1.2.35]
5. $A \subset B$ a.s. and $A = B$ a.s. [Remark 1.4.17]
6. $Z^\mu[\mathcal{H}]$ [Eqn. (3.37)]
7. $\mathcal{F}_\infty$ [Remark 2.1.3, Remark 3.1.15]
8. $\mathcal{F}_{t-}, \mathcal{F}_{t+}$ [Definition 3.1.26]
9. $\mathcal{F}_T$ (T a stopping time) [Definition 2.1.10, Definition 3.3.11]
10. $Q_+$ [Remark 4.2.8]
11. $X^T$ [Remark 4.5.9]
12. $[X]$ [Remarks 4.7.16 and 4.7.16]
13. $[X, Y]$ [Remark 4.7.24, Theorem 4.7.22]
14. $V[A; s, t], V_+[A; s, t], V_-[A; s, t]$ [Definitions 4.7.7 and 5.1.1]
15. $\tilde{A}(\cdot), \tilde{A}_+(\cdot), \tilde{A}_-(\cdot)$ [Definition 5.1.1, Remark 5.1.2]
16. $\mu_A(\cdot)$ [Theorem 5.1.7, Remark 5.1.8]
17. $\|X\|_M$ [Eqn. (5.75)]
18. $(X, Y)_M$ [Eqn. (5.76)]
19. $\nu_X(\Gamma)$ [Eqn. (5.77)]
20. $\|\Phi\|_X$ [Eqn. (5.81)]
21. $(\Phi, \Psi)_X$ [Eqn. (5.82)]
22. $\Phi[0, T]$ [Remark 5.3.23]
23. $\Phi \cdot X$ [Definitions 5.3.17 and 5.3.35]
24. $\mathcal{G}\{\mathcal{F}_t\}$ [Definition 5.3.43]
25. $\mathcal{E}(X)(t)$ [Definition 5.5.1]
26. $\mathcal{E}(X)(\infty)$ [Remark 5.5.3]
Bibliography


