

Conditional Expectation and L^p -Spaces

In this appendix (A), we briefly review the definition and some basic properties of conditional expectation. We also establish some notation for L^p spaces. For more details we refer the reader to a graduate text in probability such as Chung [9] or D. Williams [38].

Consider a fixed probability space $(\Omega, \mathcal{F}, P)$, where Ω represents a sample space of possible outcomes, $\mathcal{F}$ is a σ -algebra of subsets of Ω representing the events to which probabilities can be assigned, and P is a probability measure on $(\Omega, \mathcal{F})$. Expectation under P will be denoted by $E[\cdot]$. The term a.s. will mean almost surely with respect to the probability measure P . Given a random variable X defined on this space satisfying $E[|X|] < \infty$ and a sub- σ -algebra $\mathcal{G} \subset \mathcal{F}$, the *conditional expectation* of X given $\mathcal{G}$ is denoted by $E[X|\mathcal{G}]$. This represents a member of the equivalence class of $\mathcal{G}$ -measurable random variables satisfying

$$E[E[X|\mathcal{G}]1_A] = E[X1_A] \quad \text{for all } A \in \mathcal{G}. \quad (\text{A.1})$$

Because conditional expectations are determined uniquely only up to almost sure equivalence, (in)equalities between conditional expectations are always interpreted to hold almost surely. If Y is a random variable defined on the probability space and $\sigma(Y)$ denotes the σ -algebra generated by Y , then the conditional expectation $E[X|\sigma(Y)]$ is often simply denoted by $E[X|Y]$. We list several common properties of conditional expectation here. In the

following, X and Y are integrable random variables defined on $(\Omega, \mathcal{F}, P)$, and $\mathcal{G}, \mathcal{H}$ are sub- σ -algebras of $\mathcal{F}$.

- (i) If X is $\mathcal{G}$ -measurable, then $E[X|\mathcal{G}] = X$.
- (ii) If X is independent of $\mathcal{G}$, then $E[X|\mathcal{G}] = E[X]$.
- (iii) If Z is $\mathcal{G}$ -measurable and XZ is integrable (in addition to X), then

$$E[XZ|\mathcal{G}] = ZE[X|\mathcal{G}].$$

- (iv) For $c \in \mathbb{R}$,

$$E[cX + Y|\mathcal{G}] = cE[X|\mathcal{G}] + E[Y|\mathcal{G}].$$

- (v) If $X \leq Y$ a.s., then

$$E[X|\mathcal{G}] \leq E[Y|\mathcal{G}].$$

- (vi) $|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}]$.

- (vii) (Jensen's inequality) If $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is convex and $\phi(X)$ is integrable (in addition to X), then

$$\phi(E[X|\mathcal{G}]) \leq E[\phi(X)|\mathcal{G}].$$

- (viii) (Tower property) If $\mathcal{G} \subset \mathcal{H}$, then

$$E[E[X|\mathcal{H}]|\mathcal{G}] = E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{G}].$$

- (ix) (Monotone convergence) If, for each $n = 1, 2, \dots$, X_n is a random variable such that $0 \leq X_n \leq X$ a.s., and $X_n \uparrow X$ a.s. as $n \rightarrow \infty$ (where X is integrable), then

$$E[X_n|\mathcal{G}] \uparrow E[X|\mathcal{G}] \quad \text{a.s. as } n \rightarrow \infty.$$

- (x) (Dominated convergence) If, for each $n = 1, 2, \dots$, X_n is a random variable such that $|X_n| \leq Y$ a.s., and $X_n \rightarrow X$ a.s. as $n \rightarrow \infty$ (where X, Y are integrable), then

$$E[X_n|\mathcal{G}] \rightarrow E[X|\mathcal{G}] \quad \text{a.s. as } n \rightarrow \infty.$$

The following special case is relevant to the probability models treated in Chapters 2 and 3. When the sample space Ω is a finite set, a sub- σ -algebra $\mathcal{G}$ of $\mathcal{F}$ can be specified by giving a finite partition $\mathcal{P} = \{A_1, \dots, A_n\}$ of Ω that generates $\mathcal{G}$; i.e., A_i , $i = 1, \dots, n$, are disjoint sets in $\mathcal{F}$, $\cup_{i=1}^n A_i = \Omega$ and $\mathcal{G} = \sigma(\mathcal{P})$, the smallest σ -algebra containing $\mathcal{P}$. In this case the random variable $Z = E[X|\mathcal{G}]$ must be constant on each of the sets in the partition $\mathcal{P}$, and its value on any set A_i of positive probability is given by

$$E[X1_{A_i}]/P(A_i).$$

For any σ -algebra $\mathcal{G} \subset \mathcal{F}$ and $p \in [1, \infty)$, $L^p(\Omega, \mathcal{G}, P)$ will denote the space of random variables $X : \Omega \rightarrow \mathbb{R}$ that are $\mathcal{G}$ -measurable; i.e., $X^{-1}(B) \equiv$

$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{G}$ for each Borel set B in $\mathbb{R}$, and that satisfy $E[|X|^p] < \infty$. The norm of $X \in L^p(\Omega, \mathcal{G}, P)$ is given by

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

Furthermore, $L^\infty(\Omega, \mathcal{G}, P)$ denotes the space of $\mathcal{G}$ -measurable random variables $X : \Omega \rightarrow \mathbb{R}$ such that X is bounded P -a.s. The norm on this space is the so-called essential supremum norm given for $X \in L^\infty(\Omega, \mathcal{G}, P)$ by

$$\|X\|_\infty = \inf\{a \geq 0 : P(|X| > a) = 0\}.$$

The following interpretation of $E[X|\mathcal{G}]$ is useful when X is square integrable. If the random variable X is such that $E[X^2] < \infty$, then the conditional expectation $Z = E[X|\mathcal{G}]$ is a version of the orthogonal projection of X onto the L^2 space $L^2(\Omega, \mathcal{G}, P)$; i.e., Z is the $\mathcal{G}$ -measurable, square integrable random variable (unique up to a.s. equivalence) that minimizes the squared distance

$$E[(Z - X)^2].$$

Discrete Time Stochastic Processes

In this appendix (B), we briefly review some concepts and results for the types of discrete time stochastic processes used in this book. For more details and further background we refer the reader to a graduate text in probability such as Chung [9] or D. Williams [38].

Only discrete time stochastic processes indexed by a finite set of times are considered in this appendix, and so, without loss of generality, here we assume that the time index set is the set of consecutive non-negative integers $\{0, 1, \dots, T\}$ for some finite non-negative integer T .

Throughout this appendix (B), we assume that $(\Omega, \mathcal{F}, P)$ is a fixed probability space, where Ω is a sample space representing the set of possible outcomes, $\mathcal{F}$ is a σ -algebra of subsets of Ω representing the events to which we can assign probabilities, and P is a probability measure on $(\Omega, \mathcal{F})$. The expectation with respect to P will be denoted by $E[\cdot]$.

A *filtration* is a family $\{\mathcal{F}_t, t = 0, 1, \dots, T\}$ of sub- σ -algebras of $\mathcal{F}$ indexed by $t = 0, 1, \dots, T$ such that

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \dots \subset \mathcal{F}_T;$$

i.e., the family is increasing with time. If the sample space Ω is a finite set, often the σ -algebra $\mathcal{F}_0$ is trivial, consisting simply of the empty set $\emptyset$ and the whole sample space Ω . Also, since we shall be considering only random variables that are $\mathcal{F}_T$ -measurable, without loss of generality (by redefining $\mathcal{F}$ to equal $\mathcal{F}_T$), we can assume that $\mathcal{F}_T = \mathcal{F}$. We often write $\{\mathcal{F}_t\}$ instead of the more cumbersome $\{\mathcal{F}_t, t = 0, 1, \dots, T\}$. Intuitively, the filtration keeps track of what information is known at each of the times $t = 0, 1, \dots, T$, where

information only increases with time. More precisely, for each $t = 0, 1, \dots, T$, the σ -algebra $\mathcal{F}_t$ tells us which events may be observed by time t . We call the quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ a *filtered probability space*.

For the remainder of this appendix (B) we suppose that in addition to $(\Omega, \mathcal{F}, P)$ being fixed, $\{\mathcal{F}_t\}$ is a fixed filtration, and so $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ is a given filtered probability space. All random variables and processes considered in this appendix are assumed to be defined on this space.

Given a positive integer d , a *d-dimensional (stochastic) process* with time index set $\{0, 1, \dots, T\}$, defined on the given filtered probability space, is a collection $X = \{X_t, t = 0, 1, \dots, T\}$ where each X_t is a d -dimensional random vector; i.e., a function $X_t : \Omega \rightarrow \mathbb{R}^d$ such that $X_t^{-1}(B) \equiv \{\omega \in \Omega : X_t(\omega) \in B\} \in \mathcal{F}$ for each Borel subset B of $\mathbb{R}^d$. The process X is called *adapted* if $X_t^{-1}(B) \in \mathcal{F}_t$ for each Borel set B in $\mathbb{R}^d$ and for each $t = 0, 1, \dots, T$. We often write $X_t \in \mathcal{F}_t$ as shorthand for $X_t^{-1}(B) \in \mathcal{F}_t$ for all Borel sets B in $\mathbb{R}^d$.

Given two d -dimensional processes $Y = \{Y_t, t = 0, 1, \dots, T\}$ and $Z = \{Z_t, t = 0, 1, \dots, T\}$ defined on $(\Omega, \mathcal{F}, P)$, these processes are *modifications of one another* if $P(Y_t = Z_t) = 1$ for each $t = 0, 1, \dots, T$. Since the time index set is finite, this is equivalent to Y and Z being *indistinguishable*; i.e., $P(Y_t = Z_t \text{ for } t = 0, 1, \dots, T) = 1$. We shall regard two such indistinguishable processes as being equal as stochastic processes.

A (discrete) *stopping time* (or optional time) is a function $\tau : \Omega \rightarrow \{0, 1, \dots, T\} \cup \{\infty\}$ such that

$$\{\tau = t\} \in \mathcal{F}_t \quad \text{for } t = 0, 1, \dots, T. \quad (\text{B.1})$$

Note that for such a stopping time τ , we have

$$\{\tau = \infty\} = \Omega \setminus (\cup_{t=0}^T \{\tau = t\}) \in \mathcal{F}_T.$$

For convenience we define $\mathcal{F}_\infty = \mathcal{F}_T$ and then (B.1) also holds with $t = \infty$. A deterministic time in $\{0, 1, \dots, T\} \cup \{\infty\}$ is a stopping time. Furthermore, if τ and σ are two stopping times, then it is straightforward to verify that $\tau \wedge \sigma = \min(\tau, \sigma)$ and $\tau \vee \sigma = \max(\tau, \sigma)$ are also stopping times. There is a σ -algebra associated with any (discrete) stopping time τ , defined by

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau = t\} \in \mathcal{F}_t \text{ for } t = 0, 1, \dots, T\}.$$

Note that, for $A \in \mathcal{F}_\tau$, $A \cap \{\tau = \infty\} \in \mathcal{F}_\infty$.

A collection $M = \{M_t, \mathcal{F}_t, t = 0, 1, \dots, T\}$, where each M_t is a real-valued random variable, is called a *martingale* if the following three properties hold:

- (i) $E[|M_t|] < \infty$ for $t = 0, 1, \dots, T$,
- (ii) M_t is $\mathcal{F}_t$ -measurable for $t = 0, 1, \dots, T$,

$$(iii) \ E[M_t | \mathcal{F}_{t-1}] = M_{t-1} \text{ for } t = 1, \dots, T.$$

In this discrete time setting, condition (iii) can be equivalently replaced by

$$(iii)' \ E[M_t | \mathcal{F}_s] = M_s \text{ for all } s < t \text{ in } \{0, 1, \dots, T\}.$$

We call M a *submartingale* if the “=” in (iii) (or (iii)') is replaced by “ $\geq$ ”, and we call M a *supermartingale* if the “=” in (iii) (or (iii)') is replaced by “ $\leq$ ”. The condition (iii) (or (iii)') is often referred to as the (sub/super)martingale property. Note that M is a submartingale if and only if $-M$ is a supermartingale. If M is a (sub/super)martingale and in addition there is $p \in (1, \infty)$ such that $M_t \in L^p(\Omega, \mathcal{F}, P)$ for each $t = 0, 1, \dots, T$, then we call M an L^p -(sub/super)martingale. In describing (sub/super)martingales, we shall sometimes omit the filtration $\{\mathcal{F}_t\}$ from the notation for M when it is understood.

We adopt the convention that for a d -dimensional process M , the collection $\{M_t, \mathcal{F}_t, t = 0, 1, \dots, T\}$ is called a (sub/super)martingale if and only if each of its one-dimensional components $\{M_t^i, \mathcal{F}_t, t = 0, 1, \dots, T\}$, $i = 1, \dots, d$, is a (sub/super)martingale.

Two basic results from discrete time martingale theory are stated here for ease of reference. We state them for one-dimensional processes.

Theorem B.0.1. (*Doob's L^p -inequality*) For $p \in (1, \infty)$, let $\{M_t, \mathcal{F}_t, t = 0, 1, \dots, T\}$ be a real-valued L^p -martingale. Then, for each $t = 0, 1, \dots, T$, $N_t \equiv \sup\{|M_s|, s = 0, 1, \dots, t\} \in L^p(\Omega, \mathcal{F}_t, P)$ and

$$E[(N_t)^p] \leq q^p E[|M_t|^p], \quad (B.2)$$

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

Theorem B.0.2. (*Doob's stopping theorem*) Let $\{M_t, \mathcal{F}_t, t = 0, 1, \dots, T\}$ be a real-valued (sub)martingale and suppose that τ is a stopping time. Then $\{M_{t \wedge \tau}, \mathcal{F}_{t \wedge \tau}, t = 0, 1, \dots, T\}$ is a (sub)martingale. Furthermore, if σ is another stopping time and $\sigma \leq \tau \leq T$, then M_τ, M_σ are integrable random variables and

$$E[M_\tau] \geq E[M_\sigma], \quad (B.3)$$

where equality holds if M is a martingale.

We shall use the last result above in particular with $\sigma = 0$.

Continuous Time Stochastic Processes

In this appendix (C), we briefly review some concepts and results from the theory of continuous time stochastic processes. Many of these concepts and results are similar to those introduced for discrete time stochastic processes in the preceding appendix. Some notable exceptions are the definitions of stopping time and martingale and the fact that filtrations and stochastic processes in continuous time are frequently assumed to have some regularity (e.g., right continuity) due to the fact that the time index set is now a continuum. In this appendix (C), we restrict ourselves to compact time intervals.

In the following, we consider the compact time interval $[0, T]$ where T is a fixed value in $[0, \infty)$. We denote the Borel σ -algebra on $[0, T]$ by $\mathcal{B}_T$.

A *filtered probability space* is a quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \in [0, T]\}, P)$, where Ω is a set representing the sample space, $\mathcal{F}$ is a σ -algebra of subsets of Ω , P is a probability measure on $(\Omega, \mathcal{F})$, and $\{\mathcal{F}_t, t \in [0, T]\}$ is a filtration; i.e., a family of sub- σ -algebras of $\mathcal{F}$ indexed by $t \in [0, T]$ that is increasing: for each $s, t \in [0, T]$ satisfying $s < t$ we have $\mathcal{F}_s \subset \mathcal{F}_t$. We often write $\{\mathcal{F}_t\}$ instead of the more cumbersome $\{\mathcal{F}_t, t \in [0, T]\}$. Intuitively, we may regard $\mathcal{F}$ as containing all events which might ever be observed or to which we can assign probabilities. The filtration keeps track of what information is known by each of the times $t \in [0, T]$, where information only increases with time. More precisely, for each $t \in [0, T]$, the σ -algebra $\mathcal{F}_t$ tells us which events might be observed by time t . For such a filtered probability space, a set $A \in \mathcal{F}$ whose probability is zero is called a *P -null set*. The expectation with respect to P will be denoted by $E[\cdot]$. Since we shall be

considering only random variables that are $\mathcal{F}_T$ -measurable, without loss of generality (by redefining $\mathcal{F}$ to equal $\mathcal{F}_T$), we can assume that $\mathcal{F}_T = \mathcal{F}$.

In continuous time, some additional regularity assumptions are typically imposed on a filtered probability space. In particular, a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ is said to satisfy the *usual conditions* if the following three conditions hold:

- (i) $(\Omega, \mathcal{F}, P)$ is complete, i.e., if $A \subset B$ where $B \in \mathcal{F}$ and $P(B) = 0$, then $A \in \mathcal{F}$ and $P(A) = 0$;
- (ii) $\mathcal{F}_0$ contains all of the P -null sets;
- (iii) $\{\mathcal{F}_t\}$ is right continuous, i.e., $\mathcal{F}_t = \mathcal{F}_{t+} \equiv \bigcap_{s \in (t, T]} \mathcal{F}_s$, for all $t < T$.

Remark. If we have a filtered probability space that does not satisfy the usual conditions, by a standard procedure of completing the probability space, augmenting the members of the filtration using all of the P -null sets, and replacing $\mathcal{F}_t$ by $\mathcal{F}_{t+}$ for $t < T$, we can ensure that the usual conditions are satisfied (cf. Chung [9], page 29ff.).

For the remainder of this appendix, we suppose that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ is a given filtered probability space satisfying the usual conditions. All random variables and processes considered in this section are assumed to be defined on this space.

Given a positive integer d , a *d-dimensional (stochastic) process* on the given filtered probability space is a collection $X = \{X_t, t \in [0, T]\}$ where each X_t is a d -dimensional random vector; i.e., a function $X_t : \Omega \rightarrow \mathbb{R}^d$ such that $X_t^{-1}(B) \equiv \{\omega \in \Omega : X_t(\omega) \in B\} \in \mathcal{F}$ for each Borel subset B of $\mathbb{R}^d$. (In fact, it is enough to check that this last property holds for all open balls B in $\mathbb{R}^d$.) The process X is called *adapted* if $X_t^{-1}(B) \in \mathcal{F}_t$ for each Borel set B in $\mathbb{R}^d$ and for all $t \in [0, T]$. We often write $X_t \in \mathcal{F}_t$ as shorthand for $X_t^{-1}(B) \in \mathcal{F}_t$ for all Borel sets B in $\mathbb{R}^d$. By setting $X(t, \omega) = X_t(\omega)$ for all $t \in [0, T]$, $\omega \in \Omega$, we may sometimes view X as a function from $[0, T] \times \Omega$ into $\mathbb{R}^d$. The process X is said to be (left/right) continuous if all of its sample paths are (left/right) continuous. (Some authors require only that (left/right) continuous processes have (left/right) continuous paths almost surely. Frequently a process with such a property can be redefined on a P -null set to have (left/right) continuous paths surely. For most practical purposes such a redefined process can be considered to be the same as the original process.)

Two d -dimensional processes $\{Y_t, t \in [0, T]\}$ and $\{Z_t, t \in [0, T]\}$ defined on $(\Omega, \mathcal{F}, P)$ are

- (i) *modifications of one another* if $P(Y_t = Z_t) = 1$ for all $t \in [0, T]$,
- (ii) *indistinguishable* if $P(Y_t = Z_t \text{ for all } t \in [0, T]) = 1$.

(Note that for (ii) it is implicit that the event $\{Y_t = Z_t \text{ for all } t \in [0, T]\}$ is measurable.) Two right continuous processes that are modifications of one another are indistinguishable, and we shall regard them as being equal.

A *stopping time* (or optional time) is a function $\tau : \Omega \rightarrow [0, T] \cup \{\infty\}$ such that

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \in [0, T]. \quad (\text{C.1})$$

Note that for such a stopping time τ , we have $\{\tau = \infty\} = \Omega \setminus \{\tau \leq T\} \in \mathcal{F}_T$, and so we define $\mathcal{F}_\infty = \mathcal{F}_T$. Note that then (C.1) also holds with $t = \infty$. There is a σ -algebra associated with any stopping time τ , defined by

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \in [0, T]\}.$$

It is easily verified that a stopping time τ is measurable as a mapping from $(\Omega, \mathcal{F}_\tau)$ into $([0, T] \cup \{\infty\}, \mathcal{B}_\infty)$ where $\mathcal{B}_\infty$ is the σ -algebra on $[0, T] \cup \{\infty\}$ generated by the Borel sets in $[0, T]$ together with the singleton set $\{\infty\}$. If τ and σ are two stopping times, then it is straightforward to verify that $\tau \wedge \sigma = \min(\tau, \sigma)$ and $\tau \vee \sigma = \max(\tau, \sigma)$ are also stopping times.

Example. If X is a continuous, adapted d -dimensional process, then for any open or closed set $A \subset \mathbb{R}^d$, $\tau_A = \inf\{t \in [0, T] : X_t \in A\}$ is a stopping time. (As usual, the infimum of the empty set is defined to equal ∞ .) For a proof of this that uses the fact that the filtration satisfies the usual conditions, see Chung [9], Section 2.4.

A collection $M = \{M_t, \mathcal{F}_t, t \in [0, T]\}$, where each M_t is a real-valued random variable, is called a *martingale* if the following three properties hold:

- (i) $E[|M_t|] < \infty$ for all $t \in [0, T]$,
- (ii) M_t is $\mathcal{F}_t$ -measurable for each $t \in [0, T]$,
- (iii) $E[M_t | \mathcal{F}_s] = M_s$ for all $s \leq t$ in $[0, T]$.

We call M a *submartingale* if the “=” in (iii) is replaced by “ $\geq$ ”, and we call M a *supermartingale* if the “=” in (iii) is replaced by “ $\leq$ ”. This condition (iii) is referred to as the (sub/super)martingale property whenever M is a (sub/super)martingale. If M is a (sub/super)martingale and in addition there is $p \in (1, \infty)$ such that $M_t \in L^p(\Omega, \mathcal{F}, P)$ for all $t \in [0, T]$, then we call M an L^p -(sub/super)martingale.

It is well known that under the usual conditions that we have assumed to hold for the filtered probability space, every martingale has a modification whose paths are all right continuous with finite left limits; cf. Chung [9], Theorem 3, page 29, and Corollary 1, page 26.

The stochastic integrals that we consider will often define martingales; however, sometimes they will only locally define martingales. This leads us to the notion of a local martingale. A collection $M = \{M_t, \mathcal{F}_t, t \in [0, T]\}$, where each M_t is a real-valued random variable, is called a *local martingale*

if M is an adapted process and there is a sequence of stopping times $\{\tau_k\}_{k=1}^\infty$ such that $\tau_k \leq \tau_{k+1}$ for each k , $\lim_{k \rightarrow \infty} \tau_k = \infty$ P -a.s., and for each k ,

$$M^k = \{M_{t \wedge \tau_k}, \mathcal{F}_t, t \in [0, T]\}$$

is a martingale. We call such a sequence $\{\tau_k\}$ a localizing sequence for M . Here we have used a slightly stronger notion of local martingale than that introduced in [11] and [35]. In the case when the initial value M_0 is a constant, our definition is equivalent to that in [11] and [35]. As that is the only circumstance considered here, we have chosen to give the stronger and simpler definition. Indeed, for stochastic integrals, the initial value is not only constant but zero.

In describing (sub/super/local) martingales, we shall sometimes omit the filtration $\{\mathcal{F}_t\}$ from the notation for M when it is understood. We also adopt the convention that a d -dimensional process $M = \{M_t, t \in [0, T]\}$ is called a (sub/super/local) martingale if and only if each of its one-dimensional components $\{M_t^i, t \in [0, T]\}$, $i = 1, \dots, d$, is a (sub/super/local) martingale. In the case of a local martingale, one can choose a common localizing sequence for all components, since the minimum of finitely many stopping times is again a stopping time. We shall always assume that such a common sequence is being used when we consider multi-dimensional local martingales.

Two basic results from continuous time martingale theory are stated here for ease of reference. We state them for one-dimensional processes.

Theorem C.0.3. (*Doob's L^p -inequality*) For $p \in (1, \infty)$, let $\{M_t, \mathcal{F}_t, t \in [0, T]\}$ be a real-valued right continuous L^p -martingale. Then, for each $t \in [0, T]$, $N_t \equiv \sup_{0 \leq s \leq t} |M_s| \in L^p(\Omega, \mathcal{F}_t, P)$ and

$$E[(N_t)^p] \leq q^p E[|M_t|^p], \quad (\text{C.2})$$

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

Theorem C.0.4. (*Doob's Stopping Theorem*) Let $\{M_t, \mathcal{F}_t, t \in [0, T]\}$ be a real-valued right continuous (sub)martingale and suppose that τ is a stopping time. Then $\{M_{t \wedge \tau}, \mathcal{F}_{t \wedge \tau}, t \in [0, T]\}$ is a right continuous (sub)martingale. Furthermore, if σ is another stopping time and $\sigma \leq \tau \leq T$, then M_τ and M_σ are integrable random variables and

$$E[M_\tau] \geq E[M_\sigma], \quad (\text{C.3})$$

where equality holds if M is a martingale.

We shall use the last result above in particular with $\sigma = 0$.

Brownian Motion and Stochastic Integration

In this appendix (D), we briefly review some definitions and basic results for Brownian motion and stochastic integrals with respect to Brownian motion. For more details and further background we refer the reader to Chung [9] and Chung and Williams [11].

All stochastic processes considered in this appendix are indexed by time lying in the compact time interval $[0, T]$ for a fixed value T in $[0, \infty)$, and all processes are assumed to be defined on a given complete probability space $(\Omega, \mathcal{F}, P)$.

D.1. Brownian Motion

A standard *one-dimensional Brownian motion* (on the time interval $[0, T]$) is a one-dimensional process $W = \{W_t, t \in [0, T]\}$ such that

- (i) $W_0 = 0$ P -a.s.;
- (ii) all of the sample paths of W are continuous;
- (iii) W has independent increments: $W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}$ are independent for any $0 \leq t_0 < t_1 < \dots < t_n < \infty$, $n = 1, 2, \dots$;
- (iv) for $0 \leq s < t < \infty$, $W_t - W_s$ is a normally distributed random variable with mean zero and variance $t - s$.

One can think of standard Brownian motion as a limit in distribution of a sequence of simple symmetric random walks where the time step and walk step size are scaled to zero in a suitable way (see e.g., [27]).

For a positive integer $n \geq 1$, a standard n -dimensional Brownian motion is an n -dimensional process $W = \{W_t, t \in [0, T]\}$ such that the coordinate processes $W^1, W^2, \dots, W^n$ are mutually independent and for each $i \in \{1, \dots, n\}$, $W^i = \{W_t^i, t \in [0, T]\}$ is a standard one-dimensional Brownian motion.

Assuming that W is a standard n -dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, P)$, for each $t \in [0, T]$, let $\mathcal{F}_t^o = \sigma\{W_s : 0 \leq s \leq t\}$, the smallest σ -algebra on Ω with respect to which W_s is measurable for each $0 \leq s \leq t$. For each $t \in [0, T]$, let $\mathcal{F}_t$ denote the smallest σ -algebra containing $\mathcal{F}_t^o$ and all of the P -null sets in $\mathcal{F}$. This is called the augmentation of $\mathcal{F}_t$ by the P -null sets, and $\{\mathcal{F}_t, t \in [0, T]\}$ is called the filtration generated by the Brownian motion W under P . The latter is sometimes referred to as the *standard filtration* associated with the Brownian motion or simply the Brownian filtration. It is well known that this filtration is right continuous; i.e., for each $t \in [0, T)$, $\mathcal{F}_t = \mathcal{F}_{t+} \equiv \cap_{s \in (t, T]} \mathcal{F}_s$ (cf. Chung [9], Section 2.3, Theorem 4). Hence the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfies the usual conditions. It is well known that in this case when the filtration is the standard Brownian filtration, every martingale relative to $\{\mathcal{F}_t\}$ has a continuous modification; cf. Theorem V.3.5 in Revuz and Yor [35]. It is straightforward to check that $\{W_t, \mathcal{F}_t, t \in [0, T]\}$ is a continuous, n -dimensional martingale.

D.2. Stochastic Integrals (with respect to Brownian motion)

In Chapter 3 sums of the form:

$$\sum_{s=1}^t \phi_s \cdot \Delta S_s^*$$

play a key role in arbitrage pricing of European contingent claims in discrete finite market models. For the continuous market models considered in Chapters 4 and 5, we need the continuous time analogue of these discrete time “stochastic integrals” where S^* is replaced by a solution of a stochastic differential equation driven by Brownian motion. This leads us to define stochastic integrals where Brownian motion is the integrator. In general, these stochastic integrals cannot be defined pathwise using Riemann-Stieltjes integrals, since the paths of Brownian motion although continuous are not locally of bounded variation.

Suppose that W is a standard one-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t\}$ is a filtration such that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfies the usual conditions and $\{W_t, \mathcal{F}_t, t \in [0, T]\}$ is a continuous martingale. (One could use the standard filtration associated

with the Brownian motion for $\{\mathcal{F}_t\}$; however, we do not assume that a priori.) Without loss of generality, we assume that $\mathcal{F} = \mathcal{F}_T$.

Let $\mathcal{L}_s$ denote the set of one-dimensional processes $Y = \{Y_t, t \in [0, T]\}$ such that Y can be written in the simple form:

$$Y_t(\omega) = H_0(\omega)1_{\{0\}}(t) + \sum_{i=1}^{n-1} H_i(\omega)1_{(t_i, t_{i+1}]}(t) \quad \text{for all } t \in [0, T], \omega \in \Omega,$$

for some $0 = t_1 < t_2 < \dots < t_n = T < \infty$, $H_i : \Omega \rightarrow \mathbb{R}$ that is bounded and $\mathcal{F}_{t_i}$ -measurable for $i = 0, 1, \dots, n-1$, where $t_0 \equiv 0$. Note that for each fixed ω as a function of t , Y is a step function. Also, the value of Y_0 depends only on the information available at time zero, and for each $i = 1, \dots, n-1$, the value of the random variable Y_t for t in the interval $(t_i, t_{i+1}]$ depends only on the information available up to time t_i . For these reasons Y is called a *simple predictable* function. (The σ -algebra induced on $[0, T] \times \Omega$ by the simple predictable processes is called the *predictable* σ -algebra.)

Define for each $\omega \in \Omega$ and $t \in [0, T]$,

$$\left(\int_0^t Y_s dW_s \right) (\omega) = \sum_{i=1}^{n-1} H_i(\omega) (W_{t \wedge t_{i+1}} - W_{t \wedge t_i})(\omega).$$

This family of stochastic integrals (one integral for each time t) is linear in $Y \in \mathcal{L}_s$ and has the following properties:

- (i) $\left\{ \int_0^t Y_s dW_s, \mathcal{F}_t, t \in [0, T] \right\}$ is a continuous L^2 -martingale,
- (ii) $E \left[\left(\int_0^t Y_s dW_s \right)^2 \right] = E \left[\int_0^t Y_s^2 ds \right]$ for all $t \in [0, T]$,
- (iii) $E \left[\sup_{t \in [0, T]} \left(\int_0^t Y_s dW_s \right)^2 \right] \leq 4E \left[\int_0^T Y_s^2 ds \right]$.

The first two properties follow by simple but tedious verification. Property (ii) is referred to as the L^2 -isometry. The third property is a consequence of (i) and (ii) and Doob's inequality for continuous L^2 -martingales.

Let $\mathcal{L}$ denote the set of all one-dimensional processes $Y = \{Y_t, t \in [0, T]\}$ such that Y is adapted, $Y : [0, T] \times \Omega \rightarrow \mathbb{R}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $Y(t, \omega) = Y_t(\omega)$ for all $t \in [0, T]$ and $\omega \in \Omega$, and

$$E \left[\int_0^T Y_s^2 ds \right] < \infty.$$

Then $\mathcal{L} \subset L^2([0, T] \times \Omega, \mathcal{B}_T \times \mathcal{F}_T, \lambda \times P)$, where λ denotes Lebesgue measure on $[0, T]$. The L^2 -norm of an element $Y \in \mathcal{L}$ is given by

$$\|Y\|_{\mathcal{L}} = \left(E \left[\int_0^T Y_s^2 ds \right] \right)^{\frac{1}{2}}.$$

It is known (cf. Chung and Williams [11], Sections 3.3 and 2.5) that $\mathcal{L}_s$ is dense in $\mathcal{L}$ with respect to the L^2 -norm $\|\cdot\|_{\mathcal{L}}$. This property, the isometry (ii), and its consequence (iii) are the keys to extending the definition of the stochastic integral process $\{\int_0^t Y_s dW_s, t \in [0, T]\}$ to all integrands $Y \in \mathcal{L}$. Briefly, this extension is accomplished as follows. For any $Y \in \mathcal{L}$, there is a sequence $\{Y^{(m)}\}_{m=1}^\infty$ in $\mathcal{L}_s$ such that $\|Y - Y^{(m)}\|_{\mathcal{L}} \rightarrow 0$ as $m \rightarrow \infty$. It follows that such a sequence is Cauchy in $\mathcal{L}$ and hence by the linearity of the stochastic integral and the isometry property (ii) (with $t = T$), the corresponding sequence of stochastic integrals $\{\int_0^T Y_s^{(m)} dW_s\}_{m=1}^\infty$ is Cauchy in $L^2(\Omega, \mathcal{F}_T, P)$. In fact, using (iii) on differences of the stochastic integral processes and a standard Borel-Cantelli argument, one can show that there is a subsequence $\{m_k\}$ of $\{m\}$ and a continuous, adapted process, which we denote by $\{\int_0^t Y_s dW_s, t \in [0, T]\}$, such that P -a.s., as $k \rightarrow \infty$, $\int_0^t Y_s^{(m_k)} dW_s$ converges uniformly for $t \in [0, T]$ to $\int_0^t Y_s dW_s$. Up to indistinguishability, this limit process $\{\int_0^t Y_s dW_s, t \in [0, T]\}$ does not depend on the particular convergent subsequence nor on the original sequence chosen to approximate Y . This process $\{\int_0^t Y_s dW_s, t \in [0, T]\}$ inherits the martingale and isometry properties, (i) and (ii), from the approximating sequence.

A final step in defining stochastic integrals is to localize the class of integrands with the attendant stochastic integrals yielding local martingales starting from zero. For this, let $\mathcal{L}_{loc}$ denote the set of one-dimensional processes $Y = \{Y_t, t \in [0, T]\}$ such that Y is adapted, $Y : [0, T] \times \Omega \rightarrow \mathbb{R}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $Y(t, \omega) = Y_t(\omega)$ for all $t \in [0, T]$ and $\omega \in \Omega$, and

$$\int_0^T Y_s^2 ds < \infty \quad P\text{-a.s.}$$

Given $Y \in \mathcal{L}_{loc}$, for each positive integer m , consider the stopping time

$$\tau_m = \inf \left\{ t \in [0, T] : \int_0^t Y_s^2 ds \geq m \right\}.$$

It is readily verified that for each m , the process $Y^{(m)}$ defined by

$$Y_s^{(m)}(\omega) = 1_{[0, \tau_m(\omega)]}(s) Y_s(\omega) \quad \text{for } s \in [0, T], \omega \in \Omega,$$

is in $\mathcal{L}$. One can show that, as m increases, the stochastic integral processes $\{\int_0^t Y_s^{(m)} dW_s, t \in [0, T]\}$ are consistent (almost surely). It then follows that there is a continuous local martingale that starts from zero (with localizing sequence $\{\tau_m\}$), which we denote by $\{\int_0^t Y_s dW_s, t \in [0, T]\}$, that P -a.s. satisfies

$$\int_0^t Y_s dW_s = \lim_{m \rightarrow \infty} \int_0^t Y_s^{(m)} dW_s \quad \text{for all } t \in [0, T].$$

This limit process is unique up to indistinguishability.

Now suppose that W is an n -dimensional standard Brownian motion W defined on the complete probability space $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t, t \in [0, T]\}$ is a filtration such that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfies the usual conditions and $\{W_t, \mathcal{F}_t, t \in [0, T]\}$ is a continuous, n -dimensional martingale. (Again we could use the standard filtration associated with W for $\{\mathcal{F}_t\}$, but we do not assume that a priori.) Assume without loss of generality that $\mathcal{F} = \mathcal{F}_T$. Define $\mathcal{L}$ and $\mathcal{L}_{loc}$ in an analogous manner to that described above using this filtration. Suppose that $Y = (Y^1, \dots, Y^n)$ is an n -dimensional process such that for each $i = 1, \dots, n$, $Y^i \in \mathcal{L}_{loc}$ for each i . Then each of the one-dimensional stochastic integral processes $\{\int_0^t Y_s^i dW_s^i, t \in [0, T]\}$ can be defined as above, and then we define

$$\int_0^t Y_s \cdot dW_s = \sum_{i=1}^n \int_0^t Y_s^i dW_s^i, \quad t \in [0, T].$$

It can be shown using the independence of the components of W and the isometry for stochastic integrals that if $Y_i \in \mathcal{L}$ for all i , then

$$E \left[\left(\int_0^t Y_s \cdot dW_s \right)^2 \right] = E \left[\int_0^t |Y_s|^2 ds \right], \quad t \in [0, T],$$

where $|Y_s| = \left(\sum_{i=1}^n (Y_s^i)^2 \right)^{\frac{1}{2}}$ for all $s \in [0, T]$.

D.3. Itô Process

Suppose that W is an n -dimensional Brownian motion (for some $n \geq 1$) defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfying the usual conditions such that $\{W_t, \mathcal{F}_t, t \in [0, T]\}$ is a continuous n -dimensional martingale. Assume without loss of generality that $\mathcal{F} = \mathcal{F}_T$.

An *Itô process* driven by the Brownian motion W is a d -dimensional (for some $d \geq 1$), continuous, adapted process $X = \{X_t, t \in [0, T]\}$ satisfying P -a.s.,

$$X_t = X_0 + \int_0^t R_s ds + \int_0^t Z_s \cdot dW_s \quad \text{for all } t \in [0, T], \quad (\text{D.1})$$

where

$$\left(\int_0^t Z_s \cdot dW_s \right)^i = \sum_{j=1}^n \int_0^t Z_s^{ij} dW_s^j \quad \text{for } i = 1, \dots, d, \quad (\text{D.2})$$

and

- (i) $X_0 \in \mathcal{F}_0$;

- (ii) $R = \{R_t, t \in [0, T]\}$ is a d -dimensional, adapted process such that $R : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $R(t, \omega) = R_t(\omega)$ for each $t \in [0, T]$ and $\omega \in \Omega$, and $\int_0^T |R_s| ds < \infty$ P -a.s. where $|R_s| = \left(\sum_{i=1}^d (R_s^i)^2 \right)^{\frac{1}{2}}$ for $s \in [0, T]$;
- (iii) $Z = \{Z_t, t \in [0, T]\}$ is an adapted process taking values in the set of $d \times n$ matrices with real entries (denoted by $\mathbb{R}^{d \times n}$), $Z : [0, T] \times \Omega \rightarrow \mathbb{R}^{d \times n}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $Z(t, \omega) = Z_t(\omega)$ for each $t \in [0, T]$ and $\omega \in \Omega$, and $\int_0^T |Z_s|^2 ds < \infty$ P -a.s. where $|Z_s| = \left(\sum_{i=1}^d \sum_{j=1}^n (Z_s^{ij})^2 \right)^{\frac{1}{2}}$ for $s \in [0, T]$.

Remark. Given the above assumptions on R , P -a.s., the integral $\int_0^t R_s ds$ is well defined pathwise as a continuous function of $t \in [0, T]$ taking values in $\mathbb{R}^d$. For convenience, on an exceptional null set, we define the integrals for all $t \in [0, T]$ to be zero. This defines a continuous, adapted process. The adaptedness can be seen using the fact (cf. Chung and Williams [9], Theorem 3.7, Lemma 3.5(ii)) that there is a d -dimensional predictable process U such that

$$(\lambda \times P)(\{(t, \omega) \in [0, T] \times \Omega : R_t(\omega) \neq U_t(\omega)\}) = 0. \quad (\text{D.3})$$

(A d -dimensional process $U = \{U_t, t \in [0, T]\}$ is *predictable* if $U : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ given by $U(t, \omega) = U_t(\omega)$ for $t \in [0, T]$ and $\omega \in \Omega$, is measurable with the predictable σ -algebra on $[0, T] \times \Omega$ and the Borel σ -algebra on $\mathbb{R}^d$.) Using a suitable definition on an exceptional P -null set, $\{\int_0^t U_s ds, t \in [0, T]\}$ defines a continuous, adapted process. By (D.3), this process is indistinguishable from $\{\int_0^t R_s ds, t \in [0, T]\}$, and so (since the filtration satisfies the usual conditions) the latter is also an adapted process (it was already defined so as to be continuous). The integrals in (D.2) with respect to W are stochastic integrals as defined in Section D.2. By the assumptions on Z , these define continuous local martingales.

Sometimes for brevity we shall write the stochastic differential equation satisfied by an Itô process X in differential form:

$$dX_t = R_t dt + Z_t \cdot dW_t, \quad (\text{D.4})$$

where the rigorous interpretation of this is as (D.1).

For fixed $i, j \in \{1, \dots, d\}$, the *mutual variation process* $\langle X^i, X^j \rangle$ associated with the i^{th} and j^{th} components of the Itô process X is given P -a.s., by

$$\langle X^i, X^j \rangle_t = \int_0^t (ZZ')_s^{ij} ds = \sum_{k=1}^n \int_0^t Z_s^{ik} Z_s^{jk} ds, \quad t \in [0, T], \quad (\text{D.5})$$

where $'$ denotes transpose. When $i = j$, this yields the *quadratic variation process* associated with the component X^i . This process is sometimes also denoted by $[X^i]$, and P -a.s. is given by

$$[X^i]_t = \sum_{k=1}^n \int_0^t (Z_s^{ik})^2 ds, \quad t \in [0, T]. \quad (\text{D.6})$$

D.4. Itô Formula

The following result is known as the *Itô formula*. For a proof of results that imply this, see Chung and Williams [11], Theorem 5.10. Let X be a d -dimensional Itô process as described in Section D.3 above. Suppose that D is a domain (i.e., an open connected set) in $\mathbb{R}^d$ such that

$$P(X_t \in D \text{ for all } t \in [0, T]) = 1.$$

Let $f : [0, T] \times D \rightarrow \mathbb{R}$ be a continuous function such that the partial derivatives $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x_i}, \frac{\partial^2 f}{\partial x_i \partial x_j}$ exist and are continuous on $[0, T] \times D$. Then P -a.s. for all $t \in [0, T]$,

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(s, X_s) R_s^i ds \\ &\quad + \sum_{i=1}^d \sum_{j=1}^n \int_0^t \frac{\partial f}{\partial x_i}(s, X_s) Z_s^{ij} dW_s^j \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s) (ZZ')_s^{ij} ds. \end{aligned}$$

Here the integrals with respect to dW are stochastic integrals and those with respect to ds are defined a.s. as Lebesgue integrals. Using the notation (D.4) and (D.5), it is often convenient to write the above as

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(s, X_s) dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(s, X_s) d\langle X^i, X^j \rangle_s, \end{aligned}$$

or in differential form

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t}(t, X_t) dt + \sum_{i=1}^d \frac{\partial f}{\partial x_i}(t, X_t) dX_t^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t) d\langle X^i, X^j \rangle_t. \end{aligned}$$

In the simple case when $d = 1$ and $f : \mathbb{R} \rightarrow D$ is twice continuously differentiable, the Itô formula reduces to

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s) R_s ds + \int_0^t f'(X_s) Z_s \cdot dW_s \\ &\quad + \frac{1}{2} \int_0^t f''(X_s) |Z_s|^2 ds. \end{aligned}$$

D.5. Girsanov Transformation

The following result is frequently referred to as the Girsanov transformation or Girsanov theorem. The prototype of this formula for one-dimensional Brownian motion was developed by Cameron and Martin [6, 7, 8]. Subsequently, Girsanov [18] generalized their transformation to a multi-dimensional Brownian motion and used the modern notation and terminology of Itô's theory of stochastic integration. For a proof of this result, see Chung and Williams [11], Section 9.4; Karatzas and Shreve [27], Section 3.5; or Revuz and Yor [35], Chapter VIII.

Suppose that W is an n -dimensional Brownian motion (for some $n \geq 1$) defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfying the usual conditions such that $\{W_t, \mathcal{F}_t, t \in [0, T]\}$ is a continuous n -dimensional martingale. Assume without loss of generality that $\mathcal{F} = \mathcal{F}_T$. Let $\kappa = \{\kappa_t, t \in [0, T]\}$ be an n -dimensional, adapted process such that $\kappa : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $\kappa(t, \omega) = \kappa_t(\omega)$ for all $t \in [0, T]$ and $\omega \in \Omega$, and suppose that $\int_0^T |\kappa_t|^2 dt < \infty$ P -a.s. Then the stochastic integral process

$$\int_0^t \kappa_s \cdot dW_s = \sum_{i=1}^n \int_0^t \kappa_s^i dW_s^i, \quad t \in [0, T],$$

is well defined as a continuous local martingale. The integrals $\int_0^t \kappa_s ds$ and $\int_0^t |\kappa_s|^2 ds$ are well defined for all $t \in [0, T]$, P -a.s. As explained in the remark in Section D.3, by defining these integrals to be zero for all $t \in [0, T]$ on an exceptional P -null set, we may assume that they define continuous adapted processes. Define

$$\Lambda_t = \exp \left(\int_0^t \kappa_s \cdot dW_s - \frac{1}{2} \int_0^t |\kappa_s|^2 ds \right), \quad t \in [0, T].$$

Then, one can show using Itô's formula that $\Lambda = \{\Lambda_t, t \in [0, T]\}$ is a continuous local martingale satisfying P -a.s.,

$$\Lambda_t = 1 + \int_0^t \Lambda_s \kappa_s \cdot dW_s, \quad t \in [0, T].$$

Define

$$\tilde{W}_t = W_t - \int_0^t \kappa_s ds, \quad t \in [0, T].$$

For the following theorem, recall that we are assuming that $\mathcal{F} = \mathcal{F}_T$.

Theorem D.5.1. (*Girsanov's theorem*) Suppose that $\{\Lambda_t, t \in [0, T]\}$ is a martingale. Define a new probability measure Q on $(\Omega, \mathcal{F})$ by

$$Q(A) = E^P[1_A \Lambda_T] \quad \text{for all } A \in \mathcal{F}.$$

Then $\tilde{W}$ is a standard n -dimensional Brownian motion under Q .

Remark. In fact, Λ will be a martingale if

$$E^P[\Lambda_T] = 1.$$

A sufficient condition for this is what is known as *Novikov's criterion*:

$$E^P \left[\exp \left(\frac{1}{2} \int_0^T |\kappa_t|^2 dt \right) \right] < \infty.$$

For a proof of this sufficiency, see Karatzas and Shreve [27], Corollary 3.5.13, or Revuz and Yor [35], Proposition VIII.1.15.

D.6. Martingale Representation Theorem

Suppose that $W = \{W_t, t \in [0, T]\}$ defined on the complete probability space $(\Omega, \mathcal{F}, P)$ is a standard n -dimensional Brownian motion (for some $n \geq 1$) and let $\{\mathcal{F}_t, t \in [0, T]\}$ be its standard filtration. Without loss of generality we assume that $\mathcal{F} = \mathcal{F}_T$. The following result is used several times in the chapters on continuous models. It is important for this result that the filtration is the standard one generated by the Brownian motion. For a proof of this theorem, see for example Revuz and Yor [35], Theorem V.3.4.

Theorem D.6.1. (*Martingale Representation Theorem*) Suppose that $\{M_t, \mathcal{F}_t, t \in [0, T]\}$ is a right continuous local martingale. There is an adapted, n -dimensional process $\eta = \{\eta_t, t \in [0, T]\}$ satisfying

- (i) $\eta : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $\eta(t, \omega) = \eta_t(\omega)$ for $t \in [0, T]$ and $\omega \in \Omega$,
- (ii) $\int_0^T |\eta_s|^2 ds < \infty$ P -a.s.,

such that P -a.s.,

$$M_t = M_0 + \int_0^t \eta_s \cdot dW_s \quad \text{for } t \in [0, T].$$

