

Appendix A

Continuous Probability and Stochastic Calculus

In this appendix we briefly review some concepts and results from continuous time probability and stochastic calculus. Here we restrict to finite time intervals and to stochastic integrals with respect to Brownian motion, as that is all that is needed in this book. For more details and further background we refer the reader to Chung [6] and Chung and Williams [7].

In the following, $\mathbb{T}$ will denote a finite closed interval $[0, T]$ for some $T < \infty$. We denote the Borel σ -algebra on $\mathbb{T}$ by $\mathcal{B}_T$. For convenience, we let $\mathbb{T}_\infty = \mathbb{T} \cup \{\infty\}$.

A.1 Some Basic Concepts in Continuous Probability

A *filtered probability space* is a quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, 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 $\{\mathcal{F}_t\}$ is a filtration, i.e., a family of sub- σ -algebras of $\mathcal{F}$ indexed by $t \in \mathbb{T}$ that is increasing: for each $s, t \in \mathbb{T}$ satisfying $s < t$ we have $\mathcal{F}_s \subset \mathcal{F}_t$. For such a 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 $E[\cdot]$.

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 [6], page 29ff.).

For the remainder of this subsection 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 subsection 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 $X = \{X_t : t \in \mathbb{T}\}$ 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 $t \in \mathbb{T}$. We often write $X_t \in \mathcal{F}_t$, or say X_t is adapted, 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 $t \in \mathbb{T}$, $\omega \in \Omega$, we may sometimes view X as a function from $\mathbb{T} \times \Omega$ into $\mathbb{R}^d$. The process X is said to be (left/right) continuous if all of its sample paths are

(left/right) continuous.

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

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

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

A *stopping time (or optional time)* is a function $\tau : \Omega \rightarrow \mathbb{T}_\infty$ such that

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \in \mathbb{T}. \quad (\text{A.1})$$

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 (A.1) also holds with $t = \infty$. There is a σ -algebra associated with any stopping time τ , defined by

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

It is easily verified that a stopping time τ is measurable as a mapping from $(\Omega, \mathcal{F}_\tau)$ into $(\mathbb{T}_\infty, \mathcal{B}_\infty)$ where $\mathcal{B}_\infty$ is the σ -algebra on $\mathbb{T}_\infty$ generated by the Borel sets in $\mathbb{T}$ together with the singleton set $\{\infty\}$.

Example. If X is a continuous, adapted d -dimensional process, then for any open or closed set $A \subset \mathbb{R}^d$, $\tau_A = \inf\{t \in \mathbb{T} : X_t \in A\}$ is a stopping time. (As usual, the infimum of the empty set is defined to equal ∞ .) For a proof of this, see Chung [6], Section 2.4.

A collection $M = \{M_t, \mathcal{F}_t, t \in \mathbb{T}\}$, 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 $t \in \mathbb{T}$,
- (ii) $M_t \in \mathcal{F}_t$ for all $t \in \mathbb{T}$,
- (iii) $E[M_t | \mathcal{F}_s] = M_s$ for all $s \leq t$ in $\mathbb{T}$.

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 $t \in \mathbb{T}$, then we call M an L^p -(sub/super) martingale. The stochastic integrals that we consider will often define martingales starting from zero; however, sometimes they will only locally define martingales starting from zero. Consequently, we introduce the notion of a local martingale starting from zero. (The restriction to a zero starting point allows for a simplification of the definition.) We call M a *local martingale* starting from zero if M is an adapted process, $M_0 = 0$ P -a.s., 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 , $M^k = \{M_{t \wedge \tau_k}, \mathcal{F}_t, t \in \mathbb{T}\}$ is a martingale. We call such a sequence $\{\tau_k\}$ a *localizing sequence* for M . 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 adopt the convention that a d -dimensional process $M = \{M_t, \mathcal{F}_t, t \in \mathbb{T}\}$ is called a (sub/super/local) martingale if and only if each of its one-dimensional components $\{M_t^i, t \in \mathbb{T}\}$, $i = 1, \dots, d$, is a (sub/super/local) martingale.

A.2 Brownian Motion and Stochastic Integration

Throughout this section, all processes are assumed to be defined on a given complete probability space $(\Omega, \mathcal{F}, P)$.

A.2.1 Brownian Motion

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

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

Assuming that W is a standard d -dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, P)$, for each $t \in \mathbb{T}$, 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 $t \in \mathbb{T}$, let $\mathcal{F}_t$ denote the smallest σ -algebra containing $\mathcal{F}_t^o$ and all of the P -null sets in $\mathcal{F}$. Then $\{\mathcal{F}_t : t \in \mathbb{T}\}$ is called the filtration generated by the Brownian motion W under P . It is well known that this filtration is right continuous, i.e., for each $t \in [0, T)$, $\mathcal{F}_t = \mathcal{F}_{t+} \equiv \bigcap_{s \in (t, T]} \mathcal{F}_s$, cf. Chung [6], Section 2.3, Theorem 4. Hence the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfies the usual conditions. Whenever we deal with

Brownian motion on a filtered probability space, we shall always assume that the filtration is chosen in this way and will call it the standard filtration associated with the Brownian motion. It is straightforward to check that $\{W_t, \mathcal{F}_t, t \in \mathbb{T}\}$ is a continuous martingale.

A.2.2 Stochastic Integrals (with respect to Brownian motion)

In Chapter 3 we saw that sums of the form $\sum_{s=1}^t \phi_s \cdot \Delta S_s^*$ played a key role in arbitrage pricing of European contingent claims in discrete finite market models. We now want to consider the continuous time analogue of these discrete “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 of bounded variation.

Let W be a standard one-dimensional Brownian motion defined on the complete probability space $(\Omega, \mathcal{F}, P)$ and let $\{\mathcal{F}_t, t \in \mathbb{T}\}$ denote the standard filtration associated with W .

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

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

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

Define for each $\omega \in \Omega$ and $t \in \mathbb{T}$,

$$\left(\int_0^t Y_s dW_s \right) (\omega) = \sum_{i=0}^{n-1} H_i(\omega) (W_{t \wedge t_{i+1}} - W_{t \wedge t_i}) (\omega).$$

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) $\left\{ \int_0^t Y_s dW_s, \mathcal{F}_t, t \in \mathbb{T} \right\}$ 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 $t \in \mathbb{T}$,
- (iii) $E \left[\sup_{t \in \mathbb{T}} \left(\int_0^t Y_s dW_s \right)^2 \right] \leq 4E \left[\int_0^T Y_s^2 ds \right]$.

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 $Y = \{Y_t, t \in \mathbb{T}\}$ such that Y is adapted, $Y : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $Y(t, \omega) = Y_t(\omega)$ for all $t \in \mathbb{T}$ and $\omega \in \Omega$, and

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

Then $\mathcal{L} \subset L^2(\mathbb{T} \times \Omega, \mathcal{B}_T \times \mathcal{F}_T, \lambda \times P)$, where λ denotes Lebesgue measure on $\mathbb{T}$. 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 [7], 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 $\{\int_0^t Y_s dW_s, t \in \mathbb{T}\}$ to

all integrands $Y \in \mathcal{L}$. Briefly, this extension is accomplished as follows. Firstly, for any $Y \in \mathcal{L}$, there is a sequence $\{Y^n\}_{n=1}^\infty$ in $\mathcal{L}_s$ such that $\|Y - Y^n\|_{\mathcal{L}} \rightarrow 0$ as $n \rightarrow \infty$. 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 $\{\int_0^T Y_s^n dW_s\}_{n=1}^\infty$ 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 $\{n_k\}$ of $\{n\}$ and a continuous, adapted process, which we denote by $\{\int_0^t Y_s dW_s, t \in \mathbb{T}\}$, such that P -a.s., as $k \rightarrow \infty$, $\int_0^t Y_s^{n_k} dW_s$ converges uniformly for $t \in \mathbb{T}$ to $\int_0^t Y_s dW_s$. Up to indistinguishability, this limit process $\{\int_0^t Y_s dW_s, t \in \mathbb{T}\}$ does not depend on the particular convergent subsequence, nor on the original sequence chosen to approximate Y . This process $\{\int_0^t Y_s dW_s, t \in \mathbb{T}\}$ 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 $Y = \{Y_t, t \in \mathbb{T}\}$ such that Y is adapted, $Y : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $Y(t, \omega) = Y_t(\omega)$ for all $t \in \mathbb{T}$ 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 n , consider the stopping time

$$\tau_n = \inf \left\{ t \in \mathbb{T} : \int_0^t Y_s^2 ds \geq n \right\}.$$

It is readily verified that for each n , the process Y^n defined by

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

is in $\mathcal{L}$. One can show that the stochastic integral processes $\{\int_0^t Y_s^n dW_s, t \in \mathbb{T}\}$ are consistent as n increases. It then follows that there is a contin-

uous local martingale that starts from zero (with localizing sequence $\{\tau_n\}$), which we denote by $\{\int_0^t Y_s dW_s, t \in \mathbb{T}\}$, that P -a.s. satisfies

$$\int_0^t Y_s dW_s = \lim_n \int_0^t Y_s^n dW_s \quad \text{for all } t \in \mathbb{T}.$$

This limit process is unique up to indistinguishability.

Now consider a d -dimensional standard Brownian motion W and let $\{\mathcal{F}_t, t \in \mathbb{T}\}$ denote the standard filtration associated with W . Define $\mathcal{L}$ and $\mathcal{L}_{loc}$ in an analogous manner to that described above using this filtration. Suppose that $Y = (Y^1, \dots, Y^d)$ is a d -dimensional process such that for each $i = 1, \dots, d$, $Y^i \in \mathcal{L}_{loc}$ for each i . Then each of the one-dimensional stochastic integral processes $\{\int_0^t Y_s^i dW_s^i, t \in \mathbb{T}\}$ can be defined as above and then we define

$$\int_0^t Y_s \cdot dW_s = \sum_{i=1}^d \int_0^t Y_s^i dW_s^i, \quad t \in \mathbb{T}.$$

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

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

where $|Y| = \left(\sum_{i=1}^d Y_i^2 \right)^{\frac{1}{2}}$.

A.2.3 Itô's Formula

Given a d -dimensional Brownian motion W with its standard filtration $\{\mathcal{F}_t, t \in \mathbb{T}\}$, an *Itô process* is an n -dimensional, continuous, adapted process $X = \{X_t : t \in \mathbb{T}\}$ satisfying P -a.s.,

$$X_t = X_0 + \int_0^t R_s ds + \int_0^t Z_s \cdot dW_s \quad \text{for all } t \in \mathbb{T},$$

where

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

and

- (i) $X_0 \in \mathcal{F}_0$,
- (ii) $R = \{R_t, t \in \mathbb{T}\}$ is an n -dimensional, adapted process such that $R : \mathbb{T} \times \Omega \rightarrow \mathbb{R}^n$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable where $R(t, \omega) = R_t(\omega)$ for each $t \in \mathbb{T}$ and $\omega \in \Omega$, and $\int_0^T |R_s| ds < \infty$ P -a.s.,
- (iii) $Z = \{Z_t, t \in \mathbb{T}\}$ is an adapted process taking values in the set of $n \times d$ matrices with real entries (denoted by $\mathbb{R}^{n \times d}$), $Z : \mathbb{T} \times \Omega \rightarrow \mathbb{R}^{n \times d}$ is $(\mathcal{B}_T \times \mathcal{F}_T)$ -measurable, where $Z(t, \omega) = Z_t(\omega)$ for each $t \in \mathbb{T}$ and $\omega \in \Omega$, and $\int_0^T |Z_s|^2 ds < \infty$ P -a.s. where $|Z| = \left(\sum_{i=1}^n \sum_{j=1}^d Z_{ij}^2 \right)^{\frac{1}{2}}$.

Remark. Given the above assumptions on R , P -a.s., the integral $\int_0^t R_s ds$ is well defined for all $t \in \mathbb{T}$. For convenience, on the exceptional null set, we define these integrals for $t \in \mathbb{T}$ to all be zero. It is known that there is a predictable process U such that $(\lambda \times P)(\{(t, \omega) \in \mathbb{T} \times \Omega : R_t(\omega) \neq U_t(\omega)\}) = 0$ (cf. Chung and Williams [6], Theorem 3.7, Lemma 3.5(ii)). It follows that the continuous, adapted process $\{\int_0^t U_s ds, t \in \mathbb{T}\}$ is indistinguishable from $\{\int_0^t R_s ds, t \in \mathbb{T}\}$, and so (since the filtration satisfies the usual conditions) the latter is also a continuous, adapted process. We always use such a continuous adapted process as the representative for $\{\int_0^t R_s ds, t \in \mathbb{T}\}$.

The following result is known as *Itô's formula*. For a proof of this, see Chung and Williams [7]. Let $f : \mathbb{T} \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x_i}, \frac{\partial^2 f}{\partial x_i \partial x_j}$

exist and are continuous on $\mathbb{T} \times \mathbb{R}^n$. Then P -a.s., for all $t \in \mathbb{T}$,

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

where $'$ denotes transpose. In the simple case when $n = 1$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ is twice continuously differentiable, the above formula reduces to

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) R_s ds + \frac{1}{2} \int_0^t f''(X_s) Z_s^2 ds.$$