

# A Brief Introduction to Stochastic Calculus

---

These notes provide a very brief introduction to stochastic calculus, the branch of mathematics that is most identified with financial engineering and mathematical finance. We will ignore most of the technical details and take an “engineering” approach to the subject. We will only introduce the concepts that are necessary for deriving the Black-Scholes formula later in the course. These concepts include quadratic variation, stochastic integrals and stochastic differential equations. We will of course also introduce Itô’s Lemma, probably the most important result in stochastic calculus.

---

## 1 Martingales, Brownian Motion and Quadratic Variation

We make the following assumptions throughout.

- There is a probability triple  $(\Omega, \mathcal{F}, P)$  where
  - $P$  is the “true” or *physical* probability measure
  - $\Omega$  is the universe of possible outcomes. We use  $\omega \in \Omega$  to represent a generic outcome, typically a sample path of a stochastic process.
  - the set  $\mathcal{F}$  represents the set of possible *events* where an event is a subset of  $\Omega$ .
- There is also a *filtration*,  $\{\mathcal{F}_t\}_{t \geq 0}$ , that models the evolution of information through time. So for example, if it is known by time  $t$  whether or not an event,  $E$ , has occurred, then we have  $E \in \mathcal{F}_t$ . If we are working with a finite horizon,  $[0, T]$ , then we can take  $\mathcal{F} = \mathcal{F}_T$ .
- We also say that a stochastic process,  $X_t$ , is  $\mathcal{F}_t$ -adapted if the value of  $X_t$  is known at time  $t$  when the information represented by  $\mathcal{F}_t$  is known. All the processes we consider will be  $\mathcal{F}_t$ -adapted so we will not bother to state this in the sequel.
- In the continuous-time models that we will study, it will be understood that the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  will be the filtration *generated* by the stochastic processes (usually a Brownian motion,  $W_t$ ) that are specified in the model description.

### 1.1 Martingales and Brownian Motion

**Definition 1** A stochastic process,  $\{W_t : 0 \leq t \leq \infty\}$ , is a standard Brownian motion if

1.  $W_0 = 0$
2. It has continuous sample paths
3. It has independent, stationary increments.
4.  $W_t \sim N(0, t)$ .

**Definition 2** An  $n$ -dimensional process,  $W_t = (W_t^{(1)}, \dots, W_t^{(n)})$ , is a standard  $n$ -dimensional Brownian motion if each  $W_t^{(i)}$  is a standard Brownian motion and the  $W_t^{(i)}$ ’s are independent of each other.

**Definition 3** A stochastic process,  $\{X_t : 0 \leq t \leq \infty\}$ , is a martingale with respect to the filtration,  $\mathcal{F}_t$ , and probability measure,  $P$ , if

1.  $E^P[|X_t|] < \infty$  for all  $t \geq 0$
2.  $E^P[X_{t+s}|\mathcal{F}_t] = X_t$  for all  $t, s \geq 0$ .

**Example 1 (Brownian martingales)**

Let  $W_t$  be a Brownian motion. Then  $W_t$ ,  $W_t^2 - t$  and  $\exp(\theta W_t - \theta^2 t/2)$  are all martingales.

The latter martingale is an example of an *exponential martingale*. Exponential martingales are of particular significance since they are positive and may be used to define new probability measures. ■

**Exercise 1 (Conditional expectations as martingales)** Let  $Z$  be a random variable and set  $X_t := E[Z|\mathcal{F}_t]$ . Show that  $X_t$  is a martingale.

**1.2 Quadratic Variation**

Consider a partition of the time interval,  $[0, T]$  given by

$$0 = t_0 < t_1 < t_2 < \dots < t_n = T.$$

Let  $X_t$  be a Brownian motion and consider the sum of squared changes

$$Q_n(T) := \sum_{i=1}^n [\Delta X_{t_i}]^2 \tag{1}$$

where  $\Delta X_{t_i} := X_{t_i} - X_{t_{i-1}}$ .

**Definition 4 (Quadratic Variation)** The quadratic variation of a stochastic process,  $X_t$ , is equal to the limit of  $Q_n(T)$  as  $\Delta t := \max_i(t_i - t_{i-1}) \rightarrow 0$ .

**Theorem 1** The quadratic variation of a Brownian motion is equal to  $T$  with probability 1.

The functions with which you are normally familiar, e.g. continuous differentiable functions, have quadratic variation equal to zero. Note that any continuous stochastic process or function<sup>1</sup> that has non-zero quadratic variation must have infinite **total variation** where the total variation of a process,  $X_t$ , on  $[0, T]$  is defined as

$$\text{Total Variation} := \lim_{\Delta t \rightarrow 0} \sum_{i=1}^n |X_{t_k} - X_{t_{k-1}}|.$$

This follows by observing that

$$\sum_{i=1}^n (X_{t_k} - X_{t_{k-1}})^2 \leq \left( \sum_{i=1}^n |X_{t_k} - X_{t_{k-1}}| \right) \max_{1 \leq k \leq n} |X_{t_k} - X_{t_{k-1}}|. \tag{2}$$

If we now let  $n \rightarrow \infty$  in (2) then the continuity of  $X_t$  implies the impossibility of the process having finite total variation and non-zero quadratic variation. Theorem 1 therefore implies that the total variation of a Brownian motion is infinite. We have the following important result which proves very useful if we need to price options when there are multiple underlying Brownian motions, as is the case with quanto options for example.

**Theorem 2 (Levy's Theorem)** A continuous martingale is a Brownian motion if and only if its quadratic variation over each interval  $[0, t]$  is equal to  $t$ .

<sup>1</sup>A sample path of a stochastic process can be viewed as a function.

## 2 Stochastic Integrals

We now discuss the concept of a stochastic integral, ignoring the various technical conditions that are required to make our definitions rigorous. In this section, we write  $X_t(\omega)$  instead of the usual  $X_t$  to emphasize that the quantities in question are stochastic.

**Definition 5** A stopping time of the filtration  $\mathcal{F}_t$  is a random time,  $\tau$ , such that the event  $\{\tau \leq t\} \in \mathcal{F}_t$  for all  $t > 0$ .

In non-mathematical terms, we see that a stopping time is a random time whose value is part of the information accumulated by that time.

**Definition 6** We say a process,  $h_t(\omega)$ , is elementary if it is piece-wise constant so that there exists a sequence of stopping times  $0 = t_0 < t_1 < \dots < t_n = T$  and a set of  $\mathcal{F}_{t_i}$ -measurable<sup>2</sup> functions,  $e_i(\omega)$ , such that

$$h_t(\omega) = \sum_i e_i(\omega) I_{[t_i, t_{i+1})}(t)$$

where  $I_{[t_i, t_{i+1})}(t) = 1$  if  $t \in [t_i, t_{i+1})$  and 0 otherwise.

**Definition 7** The stochastic integral of an elementary function,  $h_t(\omega)$ , with respect to a Brownian motion,  $W_t$ , is defined as

$$\int_0^T h_t(\omega) dW_t(\omega) := \sum_{i=0}^{n-1} e_i(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)). \quad (3)$$

Note that our definition of an elementary function assumes that the function,  $h_t(\omega)$ , is evaluated at the **left-hand point** of the interval in which  $t$  falls. This is a key component in the definition of the stochastic integral: without it the results below would no longer hold. Moreover, defining the stochastic integral in this way makes the resulting theory suitable for financial applications. In particular, if we interpret  $h_t(\omega)$  as a trading strategy and the stochastic integral as the gains or losses from this trading strategy, then evaluating  $h_t(\omega)$  at the left-hand point is equivalent to imposing the *non-anticipativity* of the trading strategy, a property that we always wish to impose.

For a more general process,  $X_t(\omega)$ , we have

$$\int_0^T X_t(\omega) dW_t(\omega) := \lim_{n \rightarrow \infty} \int_0^T X_t^{(n)}(\omega) dW_t(\omega)$$

where  $X_t^{(n)}$  is a sequence of elementary processes that converges (in an appropriate manner) to  $X_t$ .

**Example 2** We want to compute  $\int_0^T W_t dW_t$ . Towards this end, let  $0 = t_0^n < t_1^n < t_2^n < \dots < t_n^n = T$  be a partition of  $[0, T]$  and define

$$X_t^n := \sum_{i=0}^{n-1} W_{t_i^n} I_{[t_i^n, t_{i+1}^n)}(t)$$

where  $I_{[t_i^n, t_{i+1}^n)}(t) = 1$  if  $t \in [t_i^n, t_{i+1}^n)$  and is 0 otherwise. Then  $X_t^n$  is an adapted elementary process and, by continuity of Brownian motion, satisfies  $\lim_{n \rightarrow \infty} X_t^n = W_t$  almost surely as  $\max_i |t_{i+1}^n - t_i^n| \rightarrow 0$ . The

<sup>2</sup>Loosely speaking, a function  $f(\omega)$  is  $\mathcal{F}_t$  measurable if its value is known by time  $t$ .

stochastic integral of  $X_t^n$  is given by

$$\begin{aligned}
\int_0^T X_t^n dW_t &= \sum_{i=0}^{n-1} W_{t_i^n} (W_{t_{i+1}^n} - W_{t_i^n}) \\
&= \frac{1}{2} \sum_{i=0}^{n-1} \left( W_{t_{i+1}^n}^2 - W_{t_i^n}^2 - (W_{t_{i+1}^n} - W_{t_i^n})^2 \right) \\
&= \frac{1}{2} W_T^2 - \frac{1}{2} W_0^2 - \frac{1}{2} \sum_{i=0}^{n-1} (W_{t_{i+1}^n} - W_{t_i^n})^2.
\end{aligned} \tag{4}$$

By Theorem 1 the sum on the right-hand-side of (4) converges in probability to  $T$  as  $n \rightarrow \infty$ . And since  $W_0 = 0$  we obtain

$$\int_0^T W_t dW_t = \lim_{n \rightarrow \infty} \int_0^T X_t^n dW_t = \frac{1}{2} W_T^2 - \frac{1}{2} T.$$

Note that we will generally evaluate stochastic integrals using Itô's Lemma (to be discussed later) without having to take limits of elementary processes as we did in Example 2.

**Definition 8** We define the space  $L^2[0, T]$  to be the space of processes,  $X_t(\omega)$ , such that

$$\mathbb{E} \left[ \int_0^T X_t(\omega)^2 dt \right] < \infty.$$

**Theorem 3 (Itô's Isometry)** For any  $X_t(\omega) \in L^2[0, T]$  we have

$$\mathbb{E} \left[ \left( \int_0^T X_t(\omega) dW_t(\omega) \right)^2 \right] = \mathbb{E} \left[ \int_0^T X_t(\omega)^2 dt \right].$$

**Proof: (For the case where  $X_t$  is an elementary process)**

Let  $X_t = \sum_i e_i(\omega) I_{[t_i, t_{i+1})}(t)$  be an elementary process where the  $e_i(\omega)$ 's and  $t_i$ 's are as defined in Definition 6. We therefore have  $\int_0^T X_t(\omega) dW_t(\omega) := \sum_{i=0}^{n-1} e_i(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega))$ . We then have

$$\begin{aligned}
\mathbb{E} \left[ \left( \int_0^T X_t(\omega) dW_t(\omega) \right)^2 \right] &= \mathbb{E} \left[ \left( \sum_{i=0}^{n-1} e_i(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) \right)^2 \right] \\
&= \sum_{i=0}^{n-1} \mathbb{E} \left[ e_i^2(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega))^2 \right] \\
&\quad + 2 \sum_{0 \leq i < j \leq n-1} \mathbb{E} \left[ e_i e_j(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) (W_{t_{j+1}}(\omega) - W_{t_j}(\omega)) \right] \\
&= \sum_{i=0}^{n-1} \mathbb{E} \left[ e_i^2(\omega) \underbrace{\mathbb{E}_{t_i} \left[ (W_{t_{i+1}}(\omega) - W_{t_i}(\omega))^2 \right]}_{= t_{i+1} - t_i} \right] \\
&\quad + 2 \sum_{0 \leq i < j \leq n-1} \mathbb{E} \left[ e_i e_j(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) \underbrace{\mathbb{E}_{t_j} \left[ (W_{t_{j+1}}(\omega) - W_{t_j}(\omega)) \right]}_{=0} \right]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[ \sum_{i=0}^{n-1} e_i^2(\omega)(t_{i+1} - t_i) \right] \\
&= \mathbb{E} \left[ \int_0^T X_t(\omega)^2 dt \right]
\end{aligned}$$

which is what we had to show. ■

**Theorem 4 (Martingale Property of Stochastic Integrals)** *The stochastic integral,  $Y_t := \int_0^t X_s(\omega) dW_s(\omega)$ , is a martingale for any  $X_t(\omega) \in L^2[0, T]$ .*

**Exercise 2** *Check that  $\int_0^t X_s(\omega) dW_t(\omega)$  is indeed a martingale when  $X_t$  is an elementary process. (Hint: Follow the steps we took in our proof of Theorem 3.)*

## 2.1 Stochastic Differential Equations

**Definition 9** *An  $n$ -dimensional Itô process,  $X_t$ , is a process that can be represented as*

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s \quad (5)$$

where  $W$  is an  $m$ -dimensional standard Brownian motion, and  $a$  and  $b$  are  $n$ -dimensional and  $n \times m$ -dimensional  $\mathcal{F}_t$ -adapted<sup>3</sup> processes, respectively<sup>4</sup>.

We often use the notation

$$dX_t = a_t dt + b_t dW_t$$

as shorthand for (5). An  $n$ -dimensional *stochastic differential equation* (SDE) has the form

$$dX_t = a(X_t, t) dt + b(X_t, t) dW_t; \quad X_0 = x \quad (6)$$

where as before,  $W_t$  is an  $m$ -dimensional standard Brownian motion, and  $a$  and  $b$  are  $n$ -dimensional and  $n \times m$ -dimensional adapted processes, respectively. Once again, (6) is shorthand for

$$X_t = x + \int_0^t a(X_s, s) dt + \int_0^t b(X_s, t) dW_s. \quad (7)$$

While we do not discuss the issue here, various conditions exist to guarantee existence and uniqueness of solutions to (7). A useful tool for solving SDE's is Itô's Lemma which we now discuss.

## 3 Itô's Lemma

Itô's Lemma is the most important result in stochastic calculus, the "sine qua non" of the field. We first state and give an outline proof of a basic form of the result.

**Theorem 5 (Itô's Lemma for 1-dimensional Brownian Motion)**

*Let  $W_t$  be a Brownian motion on  $[0, T]$  and suppose  $f(x)$  is a twice continuously differentiable function on  $\mathbf{R}$ . Then for any  $t \leq T$  we have*

$$f(W_t) = f(0) + \frac{1}{2} \int_0^t f''(W_s) ds + \int_0^t f'(W_s) dW_s. \quad (8)$$

<sup>3</sup> $a_t$  and  $b_t$  are  $\mathcal{F}_t$ -'adapted' if  $a_t$  and  $b_t$  are  $\mathcal{F}_t$ -measurable for all  $t$ . We always assume that our processes are  $\mathcal{F}_t$ -adapted.

<sup>4</sup>Additional technical conditions on  $a_t$  and  $b_t$  are also necessary.

**Proof: (Sketch)** Let  $0 = t_0 < t_1 < t_2 < \dots < t_n = t$  be a partition of  $[0, t]$ . Clearly

$$f(W_t) = f(0) + \sum_{i=0}^{n-1} (f(W_{t_{i+1}}) - f(W_{t_i})). \quad (9)$$

Taylor's Theorem implies

$$f(W_{t_{i+1}}) - f(W_{t_i}) = f'(W_{t_i})(W_{t_{i+1}} - W_{t_i}) + \frac{1}{2} f''(\theta_i)(W_{t_{i+1}} - W_{t_i})^2 \quad (10)$$

for some  $\theta_i \in (W_{t_i}, W_{t_{i+1}})$ . Substituting (10) into (9) we obtain

$$f(W_t) = f(0) + \sum_{i=0}^{n-1} f'(W_{t_i})(W_{t_{i+1}} - W_{t_i}) + \frac{1}{2} \sum_{i=0}^{n-1} f''(\theta_i)(W_{t_{i+1}} - W_{t_i})^2. \quad (11)$$

If we let  $\delta := \max_i |t_{i+1} - t_i| \rightarrow 0$  then it can be shown that the terms on the right-hand-side of (11) converge to the corresponding terms on the right-hand-side of (8) as desired. (This should not be surprising as we know the quadratic variation of Brownian motion on  $[0, t]$  is equal to  $t$ .) ■

A more general version of Itô's Lemma can be stated for Itô processes.

### Theorem 6 (Itô's Lemma for 1-dimensional Itô process)

Let  $X_t$  be a 1-dimensional Itô process satisfying the SDE

$$dX_t = \mu_t dt + \sigma_t dW_t.$$

If  $f(t, x) : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$  is a  $C^{1,2}$  function and  $Z_t := f(t, X_t)$  then

$$\begin{aligned} dZ_t &= \frac{\partial f}{\partial t}(t, X_t) dt + \frac{\partial f}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t) (dX_t)^2 \\ &= \left( \frac{\partial f}{\partial t}(t, X_t) + \frac{\partial f}{\partial x}(t, X_t) \mu_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t) \sigma_t^2 \right) dt + \frac{\partial f}{\partial x}(t, X_t) \sigma_t dW_t. \end{aligned}$$

## 3.1 The "Box" Calculus

In the statement of Itô's Lemma, we implicitly assumed that  $(dX_t)^2 = \sigma_t^2 dt$ . The "box calculus" is a series of simple rules for calculating such quantities. In particular, we use the rules

$$\begin{aligned} dt \times dt &= dt \times dW_t = 0 \quad \text{and} \\ dW_t \times dW_t &= dt \end{aligned}$$

when determining quantities such as  $(dX_t)^2$  in the statement of Itô's Lemma above. Note that these rules are consistent with Theorem 1. When we have two correlated Brownian motions,  $W_t^{(1)}$  and  $W_t^{(2)}$ , with correlation coefficient,  $\rho_t$ , then we easily obtain that  $dW_t^{(1)} \times dW_t^{(2)} = \rho_t dt$ . We use the box calculus for computing the quadratic variation of Itô processes.

**Exercise 3** Let  $W_t^{(1)}$  and  $W_t^{(2)}$  be two independent Brownian motions. Use Levy's Theorem to show that

$$W_t := \rho W_t^{(1)} + \sqrt{1 - \rho^2} W_t^{(2)}$$

is also a Brownian motion for a given constant,  $\rho$ .

### 3.2 Some Examples

**Example 3** Suppose a stock price,  $S_t$ , satisfies the SDE

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t.$$

Then we can use the substitution,  $Y_t = \log(S_t)$  and Itô's Lemma applied to the function<sup>5</sup>  $f(x) := \log(x)$  to obtain

$$S_t = S_0 \exp \left( \int_0^t (\mu_s - \sigma_s^2/2) ds + \int_0^t \sigma_s dW_s \right). \quad (12)$$

Note that  $S_t$  does not appear on the right-hand-side of (12) so that we have indeed solved the SDE. When  $\mu_s = \mu$  and  $\sigma_s = \sigma$  are constants we obtain

$$S_t = S_0 \exp \left( (\mu - \sigma^2/2) t + \sigma dW_t \right) \quad (13)$$

so that  $\log(S_t) \sim N \left( (\mu - \sigma^2/2)t, \sigma^2 t \right)$ . ■

**Example 4 (Ornstein-Uhlenbeck Process)**

Let  $S_t$  be a security price and suppose  $X_t = \log(S_t)$  satisfies the SDE

$$dX_t = [-\gamma(X_t - \mu t) + \mu] dt + \sigma dW_t.$$

Then we can apply Itô's Lemma to  $Y_t := \exp(\gamma t)X_t$  to obtain

$$\begin{aligned} dY_t &= \exp(\gamma t) dX_t + X_t d(\exp(\gamma t)) \\ &= \exp(\gamma t) ([-\gamma(X_t - \mu t) + \mu] dt + \sigma dW_t) + X_t \gamma \exp(\gamma t) dt \\ &= \exp(\gamma t) ([\gamma \mu t + \mu] dt + \sigma dW_t) \end{aligned}$$

so that

$$Y_t = Y_0 + \mu \int_0^t e^{\gamma s} (\gamma s + 1) ds + \sigma \int_0^t e^{\gamma s} dW_s \quad (14)$$

or alternatively (after simplifying the Riemann integral in (14))

$$X_t = X_0 e^{-\gamma t} + \mu t + \sigma e^{-\gamma t} \int_0^t e^{\gamma s} dW_s. \quad (15)$$

Once again, note that  $X_t$  does not appear on the right-hand-side of (15) so that we have indeed solved the SDE. We also obtain  $E[X_t] = X_0 e^{-\gamma t} + \mu t$  and

$$\begin{aligned} \text{Var}(X_t) &= \text{Var} \left( \sigma e^{-\gamma t} \int_0^t e^{\gamma s} dW_s \right) = \sigma^2 e^{-2\gamma t} \mathbb{E} \left[ \left( \int_0^t e^{\gamma s} dW_s \right)^2 \right] \\ &= \sigma^2 e^{-2\gamma t} \int_0^t e^{2\gamma s} ds \quad (\text{by Itô's Isometry}) \\ &= \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma t}). \end{aligned}$$

These moments should be compared with the corresponding moments for  $\log(S_t)$  in the previous example. ■

---

<sup>5</sup>Note that  $f(\cdot)$  is not a function of  $t$ .