

## Chapter 3

# Stochastic Integration and Continuous Time Models

### 3.1 Brownian Motion

The single most important continuous time process in the construction of financial models is the Brownian motion process. A Brownian motion is the oldest continuous time model used in finance and goes back to Bachelier(1900) around the turn of the last century. It is also the most common building block for more sophisticated continuous time models called diffusion processes.

The Brownian motion process is a random continuous time process denoted  $W(t)$  or  $W_t$ , defined for  $t \geq 0$  such that  $W(0)$  takes some predetermined value, usually 0, and for each  $0 \leq s < t$ ,  $W(t) - W(s)$  has a normal distribution with mean  $\mu(t-s)$  and variance  $\sigma^2(t-s)$ . The parameters  $\mu$  and  $\sigma$  are the drift and the diffusion parameters of the Brownian motion and in the special case  $\mu = 0, \sigma = 1$ ,  $W(t)$  is often referred to as a standard Brownian motion or a Wiener process. Further properties of the Brownian motion process that are important are:

A Brownian motion process exists such that the sample paths are each continuous functions of  $t$  (with probability one)

The joint distribution of any finite number of increments  $W(t_2) - W(t_1), W(t_4) - W(t_3), \dots, W(t_k) - W(t_{k-1})$  are independent normal random variables provided that  $0 \leq t_1 < t_2 \leq t_3 < t_4 \leq \dots t_{k-1} < t_k$ .

#### Further Properties of the (standard) Brownian Motion Process

The covariance between  $W(t)$  and  $W(s)$ ,  $Cov(W(t), W(s)) = \min(s, t)$ . Brownian motion is an example of a *Gaussian process*, a process for which every finite-dimensional distribution such as  $(W(t_1), W(t_2), \dots, W(t_k))$  is Normal (multivariate or univariate). In fact, Gaussian processes are uniquely determined by their covariance structure. In particular if a Gaussian process has  $E(X_t) = 0$  and  $Cov(X(t), X(s)) =$

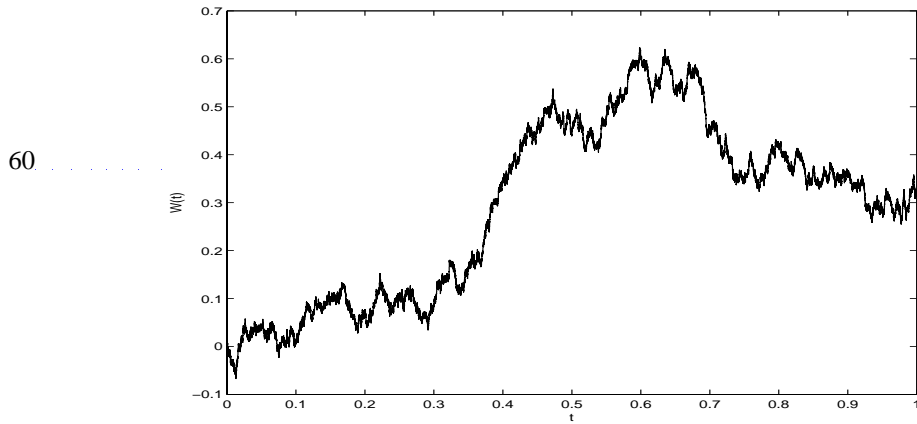


Figure 3.1: A sample Path for the Standard Brownian Motion (Wiener Process)

$\min(s, t)$ , then it has independent increments. In addition it has continuous sample paths and if  $X_0 = 0$ , then it is standard Brownian motion.

Towards the construction of a Brownian motion process, define the triangular function

$$\Delta(t) = \begin{cases} 2t & \text{for } 0 \leq t \leq \frac{1}{2} \\ 2(1-t) & \text{for } \frac{1}{2} \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and similar functions with base of length  $2^{-j}$

$$\begin{aligned} \Delta_{j,k}(t) &= \Delta(2^j t - k) \text{ for } j = 1, 2, \dots \text{ and } k = 0, 1, \dots, 2^j - 1. \\ \Delta_{0,0}(t) &= t, \quad 0 \leq t \leq 1 \end{aligned}$$

**Theorem A38 (Wavelet construction of Brownian motion)**

Suppose the random variables  $Z_{j,k}$  are independent  $N(0, 1)$  random variables. Then series below converges uniformly (almost surely) to a standard Brownian motion process  $B(t)$  on the interval  $[0, 1]$ .

$$B(t) = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} 2^{-j/2-1} Z_{j,k} \Delta_{j,k}(t)$$

The standard Brownian motion process can be extended to the whole interval  $[0, \infty)$  by generating independent Brownian motion processes  $B^{(n)}$  on the interval  $[0, 1]$  and defining  $W(t) = \sum_{j=1}^n B^{(j)}(1) + B^{(n+1)}(t - n)$  whenever  $n \leq t < n + 1$ .

Figure 1.73.1 gives a sample path of the standard Brownian motion. Evidently the path is continuous but if you examine it locally it appears to be just barely continuous, having no higher order smoothness properties. For example derivatives do not appear

to exist because of the rapid fluctuations of the process everywhere. There are various modifications of the Brownian motion process that result in a process with exactly the same distribution.

**Theorem A39**

If  $W(t)$  is a standard Brownian motion process on  $[0, \infty)$ , then so are the processes  $X_t = \frac{1}{\sqrt{a}}W(at)$  and  $Y_t = tW(1/t)$  for any  $a > 0$ .

A standard Brownian motion process is an example of a continuous time martingale, because, for  $s < t$ ,

$$\begin{aligned} E[W(t)|H_s] &= E[W(t) - W(s)|H_s] + E[W(s)|H_s] \\ &= 0 + W(s) \end{aligned}$$

since the increment  $W(t) - W(s)$  is independent of the past and normally distributed with expected value 0. In fact it is a continuous martingale in the sense that sample paths are continuous (with probability one) functions of  $t$ . It is not the only continuous martingale, however. For example it is not difficult to show that both  $X_t = W_t^2 - t$  and  $\exp(\alpha W_t - \alpha^2 t/2)$ , for  $\alpha$  any real number are continuous martingales. Of course neither are Gaussian processes. Their finite dimensional distributions cannot be normal since both processes are restricted to values in the positive reals. We discussed earlier the ruin probabilities for a random walk using martingale theory, and a similar theory can be used to establish the boundary crossing probabilities for a Brownian motion process. The following theorem establishes the relative probability that a Brownian motion hits each of two boundaries, one above the initial value and the other below.

**Theorem A40 (Ruin probabilities for Brownian motion)**

If  $W(t)$  is a standard Brownian motion and the stopping time  $\tau$  is defined by

$$\tau = \inf\{t; W(t) = -b \text{ or } a\}$$

where  $a$  and  $b$  are positive numbers, then  $P(\tau < \infty) = 1$  and

$$P[W_\tau = a] = \frac{b}{a+b}$$

Although this establishes which boundary is hit with what probability, it says nothing about the time at which the boundary is first hit. The distribution of this hitting time (the first passage time distribution) is particularly simple:

**Theorem A41 (Hitting times for a flat boundary)**

If  $W(t)$  is a standard Brownian motion and the stopping time  $\tau$  is defined by

$$\tau_a = \inf\{t; W(t) = a\}$$

where  $a > 0$ , then

**Theorem A42**

1.  $P(\tau_a < \infty) = 1$
2.  $\tau_a$  has a Laplace Transform given by

$$E(e^{-\lambda\tau_a}) = e^{-\sqrt{2\lambda}|a|}.$$

3. The probability density function of  $\tau_a$  is

$$f(t) = at^{-3/2}\phi(at^{-1/2})$$

where  $\phi$  is the standard normal probability density function.

4. The cumulative distribution function of  $\tau_a$  is given by

$$P[\tau_a \leq t] = 2P[W(t) > a] = 2[1 - \Phi(at^{-1/2})] \text{ for } t > 0 \text{ zero otherwise}$$

5.  $E(\tau_a) = \infty$

The last property is surprising. The standard Brownian motion has no general tendency to rise or fall, but because of the fluctuations it is guaranteed to strike a barrier placed at any level  $a > 0$ . However, the time before this barrier is struck can be very long, so long that the expected time is infinite. The following corollary provides an interesting connection between the maximum of a Brownian motion over an interval and its value at the end of the interval.

**Corollary**

If  $W_t^* = \max\{W(s); 0 < s < t\}$  then for  $a \geq 0$ ,

$$P[W_t^* > a] = P[\tau_a \leq t] = 2P[W(t) > a]$$

**Theorem A43 (Time of last return to 0)**

Consider the random time  $L = \sup\{t \leq 1; W(t) = 0\}$ . Then  $L$  has cumulative distribution function

$$P[L \leq s] = \frac{2}{\pi} \arcsin(\sqrt{s}), 0 < s < 1$$

and corresponding probability density function

$$\frac{d}{ds} \frac{2}{\pi} \arcsin(\sqrt{s}) = \frac{1}{\pi\sqrt{s(1-s)}}, 0 < s < 1$$

## 3.2 Continuous Time Martingales

As usual, the value of the stochastic process at time  $t$  may be denoted  $X(t)$  or by  $X_t$  for  $t \in [0, \infty)$  and let  $H_t$  is a sub sigma-field of  $H$  such that  $H_s \subset H_t$  whenever  $s \leq t$ . We call such a sequence a *filtration*.  $X_t$  is said to be *adapted* to the filtration if  $X(t)$  is measurable  $H_t$  for all  $t \in [0, \infty)$ .

Henceforth, we assume that all stochastic processes under consideration are adapted to the filtration  $H_t$ . We also assume that the filtration  $H_t$  is *right continuous*, i.e. that

$$\bigcap_{\epsilon > 0} H_{t+\epsilon} = H_t. \quad (3.1)$$

We can make this assumption without loss of generality because if  $H_t$  is any filtration, then we can make it right continuous by replacing it with

$$H_{t+} = \bigcap_{\epsilon > 0} H_{t+\epsilon}. \quad (3.2)$$

We use the fact that the intersection of sigma fields is a sigma field. Note that any process that was adapted to the original filtration is also adapted to the new filtration  $H_{t+}$ . We also typically assume, by analogy to the definition of the Lebesgue measurable sets, that if  $A$  is any set with  $P(A) = 0$ , then  $A \in H_0$ . These two conditions, that the filtration is right continuous and contains the  $P$ -null sets are referred to as the *standard conditions*. The definition of a martingale is, in continuous time, essentially the same as in discrete time:

### Definition

Let  $X(t)$  be a continuous time stochastic process adapted to a right continuous filtration  $H_t$ , where  $0 \leq t < \infty$ . Then  $X$  is a *martingale* if  $E|X(t)| < \infty$  for all  $t$  and

$$E[X(t)|H_s] = X(s) \quad (3.3)$$

for all  $s < t$ . The process  $X(t)$  is a *submartingale* (respectively a *supermartingale*) if the equality is replaced by  $\geq$  (respectively  $\leq$ ).

### Definition

A random variable  $\tau$  taking values in  $[0, \infty]$  is a *stopping time* for a martingale  $(X_t, H_t)$  if for each  $t \geq 0$ , the event  $[\tau \leq t]$  is in the sigma algebra  $H_t$ .

Stopping a martingale at a sequence of non-decreasing stopping times preserves the martingale property but there are some operations with Brownian motion which preserve the Brownian motion measure:

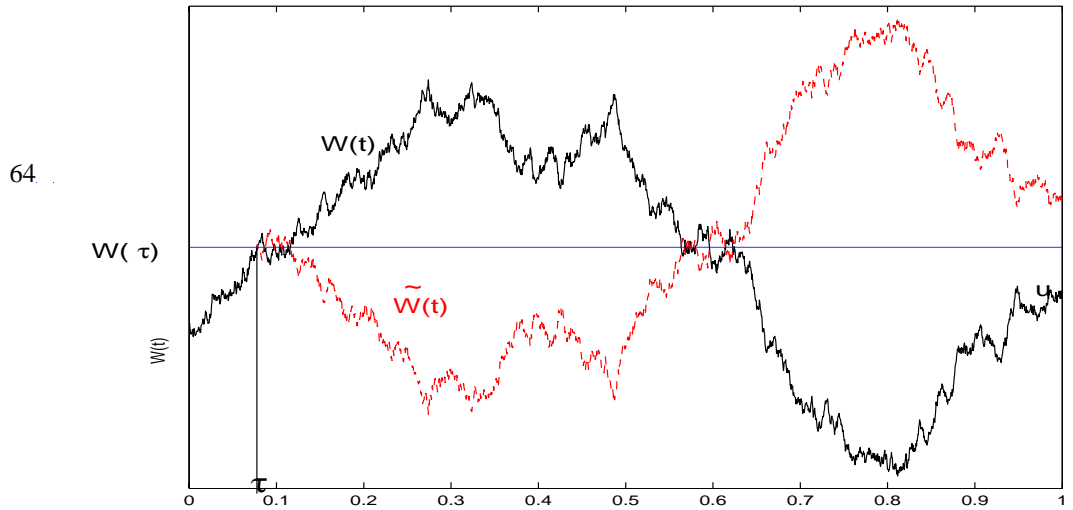


Figure 3.2: The process  $\widetilde{W}(t)$  obtained by reflecting a Brownian motion about  $W(\tau)$ .

**Theorem A44 (Reflection&Strong Markov Property)**

If  $\tau$  is a stopping time with respect to the usual filtration of a standard Brownian motion  $W(t)$ , then the process

$$\widetilde{W}(t) = \begin{cases} W(t) & t < \tau \\ 2W(\tau) - W(t) & t \geq \tau \end{cases}$$

is a standard Brownian motion.

The process  $\widetilde{W}(t)$  is obtained from the Brownian motion process as follows: up to time  $\tau$  the original Brownian motion is left alone, and for  $t > \tau$ , the process  $\widetilde{W}(t)$  is the reflection of  $W(t)$  about a horizontal line drawn at  $y = W(\tau)$ . This is shown in Figure 3.2.

**Theorem A45**

Let  $\{(M_t, H_t), t \geq 0\}$  be a (right-)continuous martingale and assume that the filtration satisfies the standard conditions. If  $\tau$  is a stopping time, then the process

$$X_t = M_{t \wedge \tau}$$

is also a continuous martingale with respect to the same filtration.

Various other results are essentially the same in discrete or continuous time. For example Doob's  $L^p$  inequality

$$\| \sup_{0 \leq t \leq T} M_t \|_p \leq \frac{p}{p-1} \|M_T\|_p, \text{ if } p > 1$$

holds for right-continuous non-negative submartingales and  $p \geq 1$ . Similarly the submartingale convergence theorem holds as stated earlier, but with  $n \rightarrow \infty$  replaced by  $t \rightarrow \infty$ .

### 3.3 Introduction to Stochastic Integrals

The stochastic integral arose from attempts to use the techniques of Riemann-Stieltjes integration for stochastic processes. However, Riemann integration requires that the integrating function have *locally bounded variation* in order that the Riemann-Stieltjes sum converge.

#### Definition (locally bounded variation)

If the process  $A_t$  can be written as the difference of two nondecreasing processes, it is called a *process of locally bounded variation*. A function is said to have locally bounded variation if it can be written as the difference of two non-decreasing processes.

For any function  $G$  of locally bounded variation, random or not, integrals such as  $\int_0^T f dG$  are easy to define because, since we can write  $G = G_1 - G_2$  as the difference between two non-decreasing functions  $G_1, G_2$ , the Riemann-Stieltjes sum

$$\sum_{i=1}^n f(s_i)[G(t_i) - G(t_{i-1})]$$

where  $0 = t_0 < t_1 < t_2 < \dots < t_n = T$  is a partition of  $[0, T]$ , and  $t_{i-1} \leq s_i \leq t_i$  will converge to the same value regardless of where we place  $s_i$  in the interval  $(t_{i-1}, t_i)$  as the mesh size  $\max_i |t_i - t_{i-1}| \rightarrow 0$ .

By contrast, many stochastic processes do not have paths of bounded variation. Consider, for example, a hypothetical integral of the form

$$\int_0^T f dW$$

where  $f$  is a nonrandom function of  $t \in [0, T]$  and  $W$  is a standard Brownian motion. The Riemann-Stieltjes sum for this integral would be

$$\sum_{i=1}^n f(s_i)[W(t_i) - W(t_{i-1})]$$

where again  $0 = t_0 < t_1 < t_2 < \dots < t_n = T$ , and  $t_{i-1} \leq s_i \leq t_i$ . In this case as  $\max_i |t_i - t_{i-1}| \rightarrow 0$  the Riemann-Stieltjes sum will not converge because the

Brownian motion paths are not of bounded variation. When  $f$  has bounded variation, we can circumvent this difficulty by formally defining the integral using integration by parts. Thus if we formally write

$$\int_0^T f dW = f(T)W(T) - f(0)W(0) - \int_0^T W df$$

then the right hand side is well defined and can be used as the definition of the left hand side. Unfortunately this simple interpretation of the stochastic integral does not work for many applications. The integrand  $f$  is often replaced by some function of  $W$  or another stochastic process which does not have bounded variation. There are other difficulties. For example, integration by parts to evaluate the integral

$$\int_0^T W dW$$

leads to  $\int_0^T W dW = W^2(T)/2$  which is not the Ito stochastic integral. Consider for a moment the possible limiting values of the Riemann Stieltjes sums

$$I_\alpha = \sum_{i=1}^n f(s_i) \{W(t_i) - W(t_{i-1})\}. \quad (3.4)$$

where  $s_i = t_{i-1} + \alpha(t_i - t_{i-1})$  for some  $0 \leq \alpha \leq 1$ . If the Riemann integral were well defined, then  $I_1 - I_0 \rightarrow 0$  in probability. However when  $f(s) = W(s)$ , this difference

$$I_1 - I_0 = \sum_{i=1}^n [W(t_i) - W(t_{i-1})]^2$$

and this cannot possibly converge to zero because, in fact, the expected value is

$$E \left( \sum_{i=1}^n [W(t_i) - W(t_{i-1})]^2 \right) = \sum_{i=1}^n (t_i - t_{i-1}) = T.$$

In fact since these increments  $[W(t_i) - W(t_{i-1})]^2$  are independent, we can show by a version of the law of large numbers that

$$\sum_{i=1}^n [W(t_i) - W(t_{i-1})]^2 \rightarrow_p T$$

and more generally  $I_\alpha - I_0 \rightarrow \alpha T$  in probability as the partition grows finer.

In other words, unlike the Riemann-Stieltjes integral, it makes a difference where we place the point  $s_i$  in the interval  $(t_{i-1}, t_i)$  for a stochastic integral. The Ito stochastic integral corresponds to  $\alpha = 0$  and approximates the integral  $\int_0^T W dW$  with partial sums of the form

$$\sum_{i=1}^n W(t_{i-1}) [W(t_i) - W(t_{i-1})]$$



the limit of which is, as the mesh size decreases,  $\frac{1}{2}(W^2(T) - T)$ . If we evaluate the integrand at the right end point of the interval (i.e. taking  $\alpha = 1$ ) we obtain  $\frac{1}{2}(W^2(T) + T)$ . Another natural choice is  $\alpha = 1/2$  (called the Stratonovich integral) and note that this definition gives the answer  $W^2(T)/2$  which is the same result obtained from the usual Riemann integration by parts. Which definition is “correct”? The Stratonovich integral has the advantage that it satisfies most of the traditional rules of deterministic calculus, for example **if the integral below is a Stratonovich integral,**

$$\int_0^T \exp(W_t) dW_t = \exp(W_T) - 1$$

While all definitions of a stochastic integral are useful, the main applications in finance are those in which the values  $f(s_i)$  appearing in (3.4) are the weights on various investments in a portfolio and the increment  $[W(t_i) - W(t_{i-1})]$  represents the changes in price of the components of that portfolio over the next interval of time. Obviously one must commit to ones investments *before* observing the changes in the values of those investments. For this reason the Ito integral ( $\alpha = 0$ ) seems the most natural for these applications.

We now define the class of functions  $f$  to which this integral will apply. We assume that  $H_t$  is a standard Brownian filtration and that the interval  $[0, T]$  is endowed with its Borel sigma field. Let  $\mathcal{H}^2$  be the set of functions  $f(\omega, t)$  on the product space  $\Omega \times [0, T]$  such that

1.  $f$  is measurable with respect to the product sigma field on  $\Omega \times [0, T]$ .
2. For each  $t \in [0, T]$ ,  $f(\cdot, t)$  is measurable  $H_t$ . (in other words the stochastic process  $f(\cdot, t)$  is adapted to  $H_t$ .)
3.  $E[\int_0^T f^2(\omega, t) dt] < \infty$ .

The set of processes  $\mathcal{H}^2$  is the natural domain of the Ito integral. However, before we define the stochastic integral on  $\mathcal{H}^2$  we need to define it in the obvious way on the step functions in  $\mathcal{H}^2$ . Let  $\mathcal{H}_0^2$  be the subset of  $\mathcal{H}^2$  consisting of functions of the form

$$f(\omega, t) = \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}(t_i < t \leq t_{i+1})$$

where the random variables  $a_i$  are measurable with respect to  $H_{t_i}$  and  $0 = t_0 < t_1 < \dots < t_n = T$ . These functions  $f$  are predictable in that their value  $a_i(\omega)$  in the interval  $(t_i, t_{i+1})$  is determined before we reach this interval. A typical step function is graphed in Figure

For such functions, the stochastic integral has only one natural definition:

$$\int_0^T f(\omega, t) dW(t) = \sum_{i=0}^{n-1} a_i(\omega) (W(t_{i+1}) - W(t_i))$$

and note that considered as a function of  $T$ , this forms a continuous time square integrable martingale.

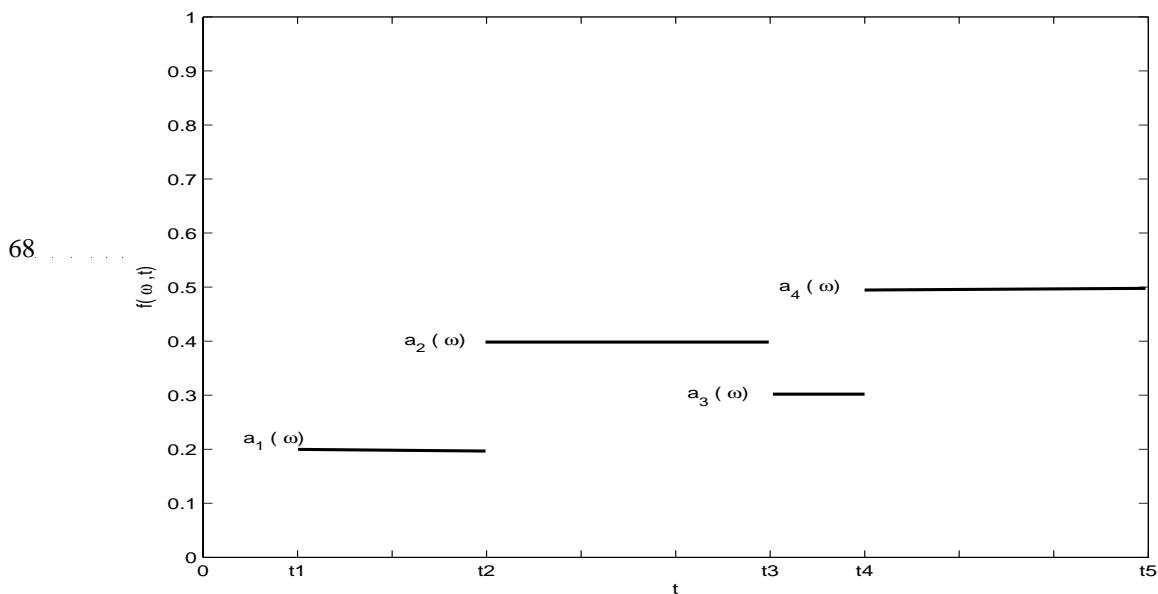


Figure 3.3: A typical step function  $f(\omega, t)$

There is a simple definition of an inner product between two square integrable random variables  $X$  and  $Y$ , namely  $E(XY)$  and we might ask how this inner product behaves when applied to the random variables obtained from stochastic integration like  $X_T(\omega) = \int_0^T f(\omega, t)dW(t)$  and  $Y_T(\omega) = \int_0^T g(\omega, t)dW(t)$ . The answer is simple, in fact, and lies at the heart of Ito's definition of a stochastic integral. For reasons that will become a little clearer later, let us define the predictable covariation process to be the stochastic process described by

$$\langle X, Y \rangle_T(\omega) = \int_0^T f(\omega, t)g(\omega, t)dt$$

**Theorem A46**

For functions  $f$  and  $g$  in  $\mathcal{H}_0^2$ ,

$$E\{\langle X, Y \rangle_T\} = E\{X_T Y_T\}. \tag{3.5}$$

and

$$E(\langle X, X \rangle_T) = E\left\{\int_0^T f^2(\omega, t)dt\right\} = E(X_T^2) \tag{3.6}$$

These identities establish an isometry, a relationship between inner products, at least for two functions in  $\mathcal{H}_0^2$ . The norm on stochastic integrals defined by

$$\left\| \int_0^T f dW \right\|_{L^2(P)}^2 = E\left(\int_0^T f(\omega, t)dW(t)\right)^2$$

agrees with the usual  $L^2$  norm on the space of random functions

$$\|f\|_2^2 = E\left\{\int_0^T f^2(\omega, t)dt\right\}.$$

We use the notation  $\|f\|_2^2 = E\left\{\int_0^T f^2(\omega, t)dt\right\}$ . If we now wish to define a stochastic integral for a general function  $f \in \mathcal{H}^2$ , the method is fairly straightforward. First we approximate any  $f \in \mathcal{H}^2$  using a sequence of step functions  $f_n \in \mathcal{H}_0^2$  such that

$$\|f - f_n\|_2^2 \rightarrow 0$$

To construct the approximating sequence  $f_n$ , we can construct a mesh  $t_i = \frac{i}{2^n}T$  for  $i = 0, 1, \dots, 2^n - 1$  and define

$$f_n(\omega, t) = \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}(t_i < t \leq t_{i+1}) \quad (3.7)$$

with

$$a_i(\omega) = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(\omega, s)ds$$

the average of the function over the previous interval.

The definition of a stochastic integral for any  $f \in \mathcal{H}^2$  is now clear from this approximation. Choose a sequence  $f_n \in \mathcal{H}_0^2$  such that  $\|f - f_n\|_2^2 \rightarrow 0$ . Since the sequence  $f_n$  is Cauchy, the isometry property (3.6) shows that the stochastic integrals  $\int_0^T f_n dW$  also forms a Cauchy sequence in  $L_2(P)$ . Since this space is complete (in the sense that Cauchy sequences converge to a random variable in the space), we can define  $\int_0^T f dW$  to be the limit of the sequence  $\int_0^T f_n dW$  as  $n \rightarrow \infty$ . Of course there is some technical work to be done, for example we need to show that two approximating sequences lead to the same integral and that the Ito isometry (3.5) still holds for functions  $f$  and  $g$  in  $\mathcal{H}^2$ . The details can be found in Steele (2001).

So far we have defined integrals  $\int_0^T f dW$  for a fixed value of  $T$ , but how should we define the stochastic process  $X_t = \int_0^t f dW$  for  $t < T$ ? To do so we define a similar integral but with the function set to 0 for  $s > t$ :

#### **Theorem A47 (Ito integral as a continuous martingale)**

For any  $f$  in  $\mathcal{H}^2$ , there exists a continuous martingale  $X_t$  adapted to the standard Brownian filtration  $H_t$  such that

$$X_t = \int_0^T f(\omega, s) \mathbf{1}(s \leq t) dW(s) \text{ for all } t \leq T.$$

This continuous martingale we will denote by  $\int_0^t f dW$ .

So far we have defined a stochastic integral only for functions  $f$  which are square integral, i.e. which satisfy  $E[\int_0^T f^2(\omega, t)dt] < \infty$  but this condition is too restrictive

for some applications. A larger class of functions to which we can extend the notion of integral is the set of locally square integrable functions,  $\mathcal{L}_{LOC}^2$ . The word “local” in martingale and stochastic integration theory is a bit of a misnomer. A property holds locally if there is a sequence of stopping times  $\nu_n$  each of which is finite but the  $\nu_n \rightarrow \infty$ , and the property holds when restricted to times  $t \leq \nu_n$ .

### Definition

Let  $\mathcal{L}_{LOC}^2$  be the set of functions  $f(\omega, t)$  on the product space  $\Omega \times [0, T]$  such that

1.  $f$  is measurable with respect to the product sigma field on  $\Omega \times [0, T]$
2. For each  $t \in [0, T]$ ,  $f(\cdot, t)$  is measurable  $H_t$  (in other words the stochastic process  $f(\cdot, t)$  is adapted to  $H_t$ ).
3.  $P(\int_0^T f^2(\omega, s)ds < \infty) = 1$

Clearly this space includes  $\mathcal{H}^2$  and arbitrary continuous functions of a Brownian motion. For any function  $f$  in  $\mathcal{L}_{LOC}^2$ , it is possible to define a sequence of stopping times

$$\nu_n = \min(T, \inf\{s; \int_0^s f^2(\omega, t)dt \geq n\})$$

which acts as a *localizing* sequence for  $f$ . Such a sequence has the properties:

1.  $\nu_n$  is a non-decreasing sequence of stopping times
2.  $P[\nu_n = T \text{ for some } n] = 1$
3. The functions  $f_n(\omega, t) = f(\omega, t)1(t \leq \nu_n) \in \mathcal{H}^2$  for each  $n$ .

The purpose of the localizing sequence is essentially to provide approximations of a function  $f$  in  $\mathcal{L}_{LOC}^2$  with functions  $f(\omega, t)1(t \leq \nu_n)$  which are in  $\mathcal{H}^2$  and therefore have a well-defined Ito integral as described above. The integral of  $f$  is obtained by taking the limit as  $n \rightarrow \infty$  of the functions  $f(\omega, t)1(t \leq \nu_n)$

$$\int_0^t f(\omega, s)dW_s = \lim_{n \rightarrow \infty} \int_0^t f(\omega, t)1(t \leq \nu_n)dW_s$$

If  $f$  happens to be a continuous **non-random** function on  $[0, T]$ , the integral  $\int_0^T f(s)dW_s$  is a limit in probability of the Riemann sums,

$$\sum f(s_i)(W_{t_{i+1}} - W_{t_i})$$

for any  $t_i \leq s_i \leq t_{i+1}$ . The integral is the limit of sums of the independent normal zero-mean random variables  $f(s_i)(W_{t_{i+1}} - W_{t_i})$  and is therefore normally distributed. In fact,

$$X_t = \int_0^t f(s)dW_s$$

is a zero mean Gaussian process with  $Cov(X_s, X_t) = \int_0^{\min(s,t)} f^2(u)du$ . Such Gaussian processes are essentially time-changed Brownian motion processes according to the following:

**Theorem A48 (time change to Brownian motion)**

Suppose  $f(s)$  is a continuous non-random function on  $[0, \infty)$  such that

$$\int_0^\infty f^2(s)ds = \infty.$$

Define the function  $t(u) = \int_0^u f^2(s)ds$  and its inverse function  $\tau(t) = \inf\{u; t(u) \geq t\}$ . Then

$$Y(t) = \int_0^{\tau(t)} f(s)dW_s$$

is a standard Brownian motion.

**Definition (local martingale)**

The process  $M(t)$  is a local martingale with respect to the filtration  $H_t$  if there exists a non-decreasing sequence of stopping times  $\tau_k \rightarrow \infty$  a.s. such that the processes

$$M_t^{(k)} = M(\min(t, \tau_k)) - M(0)$$

are martingales with respect to the same filtration.

In general, for  $f \in \mathcal{L}_{LOC}^2$ , stochastic integrals are local martingales or more formally there is a continuous local martingale equal (with probability one) to the stochastic integral  $\int_0^t f(\omega, s)dW_s$  for all  $t$ . We do not usually distinguish among processes that differ on a set of probability zero so we assume that  $\int_0^t f(\omega, s)dW_s$  is a continuous local martingale. There is a famous converse to this result, the martingale representation theorem which asserts that a martingale can be written as a stochastic integral. We assume that  $H_t$  is the standard filtration of a Brownian motion  $W_t$ .

**Theorem A49 (The martingale representation theorem)**

Let  $X_t$  be an  $H_t$  martingale with  $E(X_T^2) < \infty$ . Then there exists  $\phi \in \mathcal{H}^2$  such that

$$X_t = X_0 + \int_0^t \phi(\omega, s)dW_s \text{ for } 0 < t < T$$

and this representation is unique.

### 3.4 Differential notation and Ito's Formula

**Summary 1 (Rules of box Algebra)**

It is common to use the differential notation for stochastic differential equations such as

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

to indicate (this is its only possible meaning) a stochastic process  $X_t$  which is a solution of the equation written in integral form:

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s.$$

We assume that the functions  $\mu$  and  $\sigma$  are such that these two integrals, one a regular Riemann integral and the other a stochastic integral, are well-defined, and we would like conditions on  $\mu, \sigma$  such that existence and uniqueness of a solution is guaranteed. The following result is a standard one in this direction.

**Theorem A50 (existence and uniqueness of solutions of SDE)**

Consider the stochastic DE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad (3.8)$$

with initial condition  $X_0 = x_0$ . Suppose for all  $0 < t < T$ ,

$$|\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K|x - y|^2$$

and

$$|\mu(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2)$$

Then there is a unique (with probability one) continuous adapted solution to (3.8) and it satisfies

$$\sup_{0 < t < T} E(X_t^2) < \infty.$$

It is not difficult to show that some condition is required in the above theorem to ensure that the solution is unique. For example if we consider the purely deterministic equation  $dX_t = 3X_t^{2/3}dt$ , with initial condition  $X(0) = 0$ , it has possible solutions  $X_t = 0, t \leq a$  and  $X_t = (t - a)^3, t > a$  for arbitrary  $a > 0$ . There are at least as many distinct solutions as there are possible values of  $a$ .

Now suppose a process  $X_t$  is a solution of (3.8) and we are interested in a new stochastic process defined as a function of  $X_t$ , say  $Y_t = f(t, X_t)$ . Ito's formula is used to write  $Y_t$  with a stochastic differential equation similar to (3.8). Suppose we attempt this using a Taylor series expansion where we will temporarily regard differentials such as  $dt, dX_t$  as small increments of time and the process respectively (notation such as  $\Delta t, \Delta W$  might have been preferable here). Let the partial derivatives of  $f$  be denoted by

$$f_1(t, x) = \frac{\partial f}{\partial t}, \quad f_2(t, x) = \frac{\partial f}{\partial x}, \quad f_{22}(t, x) = \frac{\partial^2 f}{\partial x^2}, \text{ etc}$$

Then Taylor's series expansion can be written

$$\begin{aligned} dY_t &= f_1(t, X_t)dt + \frac{1}{2}f_{11}(t, X_t)(dt)^2 + \dots \\ &+ f_2(t, X_t)dX_t + \frac{1}{2}f_{22}(t, X_t)(dX_t)^2 + \dots \\ &+ f_{12}(t, X_t)(dt)(dX_t) + \dots \end{aligned} \quad (3.9)$$

and although there are infinitely many terms in this expansion, all but a few turn out to be negligible. The contribution of these terms is largely determined by some simple rules often referred to as the rules of box algebra. In an expansion to terms of order  $dt$ , as  $dt \rightarrow 0$  higher order terms such as  $(dt)^j$  are all negligible for  $j > 0$ . For example  $(dt)^2 = o(dt)$  as  $dt \rightarrow 0$  (intuitively this means that  $(dt)^2$  goes to zero faster than  $dt$  does). Similarly cross terms such as  $(dt)(dW_t)$  are negligible because the increment  $dW_t$  is normally distributed with mean 0 and standard deviation  $(dt)^{1/2}$  and so  $(dt)(dW_t)$  has standard deviation  $(dt)^{3/2} = o(dt)$ . We summarize some of these order arguments with the oversimplified rules below where the symbol “ $\sim$ ” is taken to mean “is order of, as  $dt \rightarrow 0$ ”

$$\begin{aligned}(dt)(dt) &\sim 0 \\ (dt)(dW_t) &\sim 0 \\ (dW_t)(dW_t) &\sim dt\end{aligned}$$

From these we can obtain, for example,

$$\begin{aligned}(dX_t)(dX_t) &= [\mu(t, X_t)dt + \sigma(t, X_t)dW_t][\mu(t, X_t)dt + \sigma(t, X_t)dW_t] \\ &= \mu^2(t, X_t)(dt)^2 + 2\mu(t, X_t)\sigma(t, X_t)(dt)(dW_t) + \sigma^2(t, X_t)(dW_t)(dW_t) \\ &\sim \sigma^2(t, X_t)dt\end{aligned}$$

which indicates the order of the small increments in the process  $X_t$ . If we now use these rules to evaluate (3.9), we obtain

$$\begin{aligned}dY_t &\sim f_1(t, X_t)dt + f_2(t, X_t)dX_t + \frac{1}{2}f_{22}(t, X_t)(dX_t)^2 \\ &\sim f_1(t, X_t)dt + f_2(t, X_t)(\mu(t, X_t)dt + \sigma(t, X_t)dW_t) + \frac{1}{2}f_{22}(t, X_t)\sigma^2(t, X_t)dt\end{aligned}$$

which is the differential expression of Ito’s formula.

#### **Theorem A51 (Ito’s formula).**

Suppose  $X_t$  satisfies  $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$ . Then for any function  $f$  such that  $f_1$  and  $f_{22}$  are continuous, the process  $f(t, X_t)$  satisfies the stochastic differential equation:

$$df(t, X_t) = \{\mu(t, X_t)f_2(t, X_t) + f_1(t, X_t) + \frac{1}{2}f_{22}(t, X_t)\sigma^2(t, X_t)\}dt + f_2(t, X_t)\sigma(t, X_t)dW_t$$

#### **Example (Geometric Brownian Motion)**

Suppose  $X_t$  satisfies

$$dX_t = aX_tdt + \sigma X_t dW_t$$

and  $f(t, X_t) = \ln(X_t)$ . Then substituting in Ito's formula, since  $f_1 = 0$ ,  $f_2 = X_t^{-1}$ ,  $f_{22} = -X_t^{-2}$ ,

$$\begin{aligned} dY_t &= X_t^{-1} a X_t dt - \frac{1}{2} X_t^{-2} \sigma^2 X_t^2 dt + X_t^{-1} \sigma X_t dW_t \\ &= \left(a - \frac{\sigma^2}{2}\right) dt + \sigma dW_t \end{aligned}$$

and so  $Y_t = \ln(X_t)$  is a Brownian motion with drift  $a - \frac{\sigma^2}{2}$  and volatility  $\sigma$ .

### Example (Ornstein-Uhlenbeck process)

Consider the stochastic process defined as

$$X_t = x_0 e^{-\alpha t} + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s$$

for parameters  $\alpha, \sigma > 0$ . Using Ito's lemma,

$$\begin{aligned} dX_t &= (-\alpha) x_0 e^{-\alpha t} + (-\alpha) \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dW_s + \sigma e^{-\alpha t} e^{\alpha t} dW_t \\ &= -\alpha X_t dt + \sigma dW_t. \end{aligned}$$

with the initial condition  $X_0 = x_0$ . This process has Gaussian increments and covariance structure  $\text{cov}(X_s, X_t) = \sigma^2 \int_0^s e^{-\alpha(s+t-u)} ds$ , for  $s < t$  and is called the Ornstein-Uhlenbeck process.

### Example (Brownian Bridge)

Consider the process defined as

$$X_t = (1-t) \int_0^t \frac{1}{1-s} dW_s, \text{ for } 0 < t < 1$$

subject to the initial condition  $X_0 = 0$ . Then

$$\begin{aligned} dX_t &= - \int_0^t \frac{1}{1-s} dW_s + (1-t) \frac{1}{1-t} dW_t \\ &= - \frac{X_t}{1-t} dt + dW_t \end{aligned}$$

This process satisfying  $X_0 = X_1 = 0$  and

$$dX_t = - \frac{X_t}{1-t} dt + dW_t$$

is called the Brownian bridge. It can also be constructed as  $X_t = W_t - tW_1$  and the distribution of the Brownian bridge is identical to the conditional distribution of a standard Brownian motion  $W_t$  given that  $W_0 = 0$ , and  $W_1 = 0$ . The Brownian bridge is a Gaussian process with covariance  $\text{cov}(X_s, X_t) = s(1-t)$  for  $s < t$ .



**Theorem A52 (Ito's formula for two processes)**

If

$$\begin{aligned}dX_t &= a(t, X_t)dt + b(t, X_t)dW_t \\dY_t &= \alpha(t, Y_t)dt + \beta(t, Y_t)dW_t\end{aligned}$$

then

$$\begin{aligned}df(X_t, Y_t) &= f_1(X_t, Y_t)dX_t + f_2(X_t, Y_t)dY_t \\&+ \frac{1}{2}f_{11}(X_t, Y_t)b^2 dt + \frac{1}{2}f_{22}(X_t, Y_t)\beta^2 dt \\&+ f_{12}(X_t, Y_t)b\beta dt\end{aligned}$$

There is an immediate application of this result to obtain the product rule for differentiation of diffusion processes. If we put  $f(x, y) = xy$  above, we obtain

$$d(X_t Y_t) = Y_t dX_t + X_t dY_t + b\beta dt$$

This product rule reduces to the usual with either of  $\beta$  or  $b$  is identically 0.

**3.5 Quadratic Variation**

One way of defining the variation of a process  $X_t$  is to choose a partition  $\pi = \{0 = t_0 \leq t_1 \leq \dots \leq t_n = t\}$  and then define  $Q_\pi(X_t) = \sum_i (X_{t_i} - X_{t_{i-1}})^2$ .

For a diffusion process

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

satisfying standard conditions, as the mesh size  $\max |t_i - t_{i-1}|$  converges to zero, we have  $Q_\pi(X_t) \rightarrow \int_0^t \sigma^2(s, X_s)ds$  in probability. This limit  $\int_0^t \sigma^2(s, X_s)ds$  is the process that we earlier denoted  $\langle X, X \rangle_t$ . For brevity, the redundancy in the notation is usually removed and the process  $\langle X, X \rangle_t$  is denoted  $\langle X \rangle_t$ . For diffusion processes, variation of lower order such as  $\sum_i |X_{t_i} - X_{t_{i-1}}|$  approach infinity and variation of higher order, e.g.  $\sum_i (X_{t_i} - X_{t_{i-1}})^4$  converges to zero as the mesh size converges. We will return to the definition of the predictable covariation process  $\langle X, Y \rangle_t$  in a more general setting shortly.

**The Stochastic Exponential**

Suppose  $X_t$  is a diffusion process and consider a stochastic differential equation

$$dY_t = Y_t dX_t \tag{3.10}$$

with initial condition  $Y_0 = 1$ . If  $X_t$  were an ordinary differentiable function, we could solve this equation by integrating both sides of

$$\frac{dY_t}{Y_t} = dX_t$$

to obtain the exponential function

$$Y_t = c \exp(X_t) \quad (3.11)$$

where  $c$  is a constant of integration. We might try and work backwards from (3.11) to see if this is the correct solution in the general case in which  $X_t$  is a diffusion. Letting  $f(X_t) = \exp(X_t)$  and using Ito's formula,

$$\begin{aligned} df(X_t) &= \left\{ \exp(X_t) + \frac{1}{2} \exp(X_t) \sigma^2(t, X_t) \right\} dt + \exp(X_t) \sigma(t, X_t) dW_t \\ &\neq f(X_t) dX_t \end{aligned}$$

so this solution is not quite right. There is, however, a minor fix of the exponential function which does provide a solution. Suppose we try a solution of the form

$$Y_t = f(t, X_t) = \exp(X_t + h(t))$$

where  $h(t)$  is some differentiable stochastic process. Then again using Ito's lemma, since  $f_1(t, X_t) = Y_t h'(t)$  and  $f_2(t, X_t) = f_{22}(t, X_t) = Y_t$ ,

$$\begin{aligned} dY_t &= f_1(t, X_t) dt + f_2(t, X_t) dX_t - \frac{1}{2} f_{22}(t, X_t) \sigma^2(t, X_t) dt \\ &= Y_t \{ h'(t) + \mu(t, X_t) + \frac{1}{2} \sigma^2(t, X_t) \} dt + Y_t \sigma(t, X_t) dW_t \end{aligned}$$

and if we choose just the right function  $h$  so that  $h'(t) = -\frac{1}{2} \sigma^2(t, X_t)$ , we can get a solution to (3.10). Since  $h(t) = -\frac{1}{2} \int_0^t \sigma^2(s, X_s) ds$  the solution is

$$Y_t = \exp\left(X_t - \frac{1}{2} \int_0^t \sigma^2(s, X_s) ds\right) = \exp\left\{X_t - \frac{1}{2} \langle X \rangle_t\right\}.$$

We may denote this solution  $Y = \mathcal{E}(X)$ . We saw earlier that  $\mathcal{E}(\alpha W)$  is a martingale for  $W$  a standard Brownian motion and  $\alpha$  real. Since the solution to this equation is an exponential in the ordinary calculus, the term "stochastic exponential" seems justified. The "extra" term in the exponent  $\frac{1}{2} \langle X \rangle_t$  is a consequence of the infinite local variation of the process  $X_t$ . One of the most common conditions for  $\mathcal{E}(X)$  to be a martingale is:

*Novikov's condition:* Suppose for  $g \in \mathcal{L}_{LOC}^2$

$$E \exp\left\{\frac{1}{2} \int_0^T g^2(s, X_s) ds\right\} < \infty.$$

Then  $M_t = \mathcal{E}\left(\int_0^t g(\omega, s) dW_s\right)$  is a martingale.

### 3.6 Semimartingales

Suppose  $M_t$  is a continuous martingale adapted to a filtration  $H_t$  and  $A_t$  is a continuous adapted process that is nondecreasing. It is easy to see that the sum  $A_t + M_t$  is a submartingale. But can this argument be reversed? If we are given a submartingale  $X_t$ , is it possible to find a nondecreasing process  $A_t$  and a martingale  $M_t$  such that  $X_t = A_t + M_t$ ? The fundamental result in this direction is the Doob-Meyer decomposition.

**Theorem A53 (Doob-Meyer Decomposition)**

Let  $X$  be a continuous submartingale adapted to a filtration  $H_t$ . Then  $X$  can be uniquely written as  $X_t = A_t + M_t$  where  $M_t$  is a local martingale and  $A_t$  is an adapted nondecreasing process such that  $A_0 = 0$ .

Recall that if  $M_t$  is a square integrable martingale then  $M_t^2$  is a submartingale (this follows from Jensen's inequality). Then according to the Doob-Meyer decomposition, we can decompose  $M_t^2$  into two components, one a martingale and the other a non-decreasing continuous adapted process, which we call the (predictable) *quadratic variation process*  $\langle M \rangle_t$ . In other words,

$$M_t^2 - \langle M \rangle_t$$

is a continuous martingale. We may take this as the more general definition of the process  $\langle M \rangle$ , met earlier for processes obtained as stochastic integrals. For example suppose

$$X_t(\omega) = \int_0^t f(\omega, t) dW(t)$$

where  $f \in \mathcal{H}^2$ . Then with  $\langle X \rangle_t = \int_0^t f^2(\omega, t) dt$ , and  $M_t = X_t^2 - \langle X \rangle_t$ , notice that for  $s < t$

$$\begin{aligned} E[M_t - M_s | H_s] &= E[\{\int_s^t f(\omega, u) dW(u)\}^2 - \int_s^t f^2(\omega, u) du | H_s] \\ &= 0 \end{aligned}$$

by (3.5). This means that our earlier definition of the process  $\langle X \rangle$  coincides with the current one. For two martingales  $X, Y$ , we can define the predictable covariation process  $\langle X, Y \rangle$  by

$$\langle X, Y \rangle_t = \frac{1}{4} \{ \langle X + Y \rangle_t - \langle X - Y \rangle_t \}$$

and once again this agrees for process obtained as stochastic integrals, since if  $X$  and  $Y$  are defined as

$$\begin{aligned} X_t(\omega) &= \int_0^t f(\omega, t) dW(t) \\ Y_t(\omega) &= \int_0^t g(\omega, t) dW(t) \end{aligned}$$

then the predictable covariation is

$$\langle X, Y \rangle_t(\omega) = \int_0^t f(\omega, t) g(\omega, t) dt$$

and this also follows from the Ito isometry.

**Definition (semimartingale)**

A continuous adapted process  $X_t$  is a *semimartingale* if it can be written as the sum  $X_t = A_t + M_t$  of a continuous adapted process  $A_t$  of locally bounded variation, and a continuous local martingale  $M_t$ .

The stochastic integral for square integrable martingales can be extended to the class of semimartingales. Let  $X_t = A_t + M_t$  be a continuous semimartingale. We define

$$\int h(t)dX_t = \int h(t)dA_t + \int h(t)dM_t. \quad (3.12)$$

The first integral on the right hand side of (3.12) is understood to be a Lebesgue-Stieltjes integral while the second is an Ito stochastic integral. There are a number of details that need to be checked with this definition, for example whether when we decompose a semimartingale into the two components, one with bounded variation and one a local martingale in two different ways (this decomposition is not unique), the same integral is obtained.

**3.7 Girsanov's Theorem**

Consider the Brownian motion defined by

$$dX_t = \mu dt + dW_t$$

with  $\mu$  a constant drift parameter and denote by  $E_\mu(\cdot)$  the expectation when the drift is  $\mu$ . Let  $f_\mu(x)$  be the  $N(\mu, T)$  probability density function. Then we can compute expectations under non-zero drift  $\mu$  using a Brownian motion which has drift zero since

$$E_\mu(g(X_T)) = E_0\{g(X_T)M_T(X)\}$$

where

$$M_t(X) = \mathcal{E}(\mu X) = \exp\{\mu X_t - \frac{1}{2}\mu^2 t\}.$$

This is easy to check since the stochastic exponential  $M_T(X)$  happens to be the ratio of the  $N(\mu T, T)$  probability density function to the  $N(0, T)$  density. The implications are many and useful. We can for example calculate moments or simulate under the condition  $\mu = 0$  and apply the results to the case  $\mu \neq 0$ . By a similar calculation, for a bounded Borel measurable function  $g(X_{t_1}, \dots, X_{t_n})$ , where  $0 \leq t_1 \leq \dots \leq t_n$ ,

$$E_\mu\{g(X_{t_1}, \dots, X_{t_n})\} = E_0\{g(X_{t_1}, \dots, X_{t_n})M_{t_n}(X)\}.$$

**Theorem A54 (Girsanov's Theorem for Brownian Motion)**

Consider a Brownian motion with drift  $\mu$  defined by

$$X_t = \mu t + W_t.$$

Then for any bounded measurable function  $g$  defined on the space  $C[0, T]$  of the paths we have

$$E_\mu[g(X)] = E_0[g(X)M_T(X)]$$

where again  $M_T(X)$  is the exponential martingale  $\mathcal{E}(\mu X) = \exp(\mu X_T - \frac{1}{2}\mu^2 T)$ .

Note that if we let  $P_0, P_\mu$  denote the measures on the function space corresponding to drift 0 and  $\mu$  respectively, we can formally write

$$\begin{aligned} E_\mu[g(X)] &= \int g(x) dP_\mu = \int g(x) \frac{dP_\mu}{dP_0} dP_0 \\ &= E_0\left\{g(X) \frac{dP_\mu}{dP_0}\right\} \end{aligned}$$

which means that  $M_T(X)$  plays the role of a likelihood ratio

$$\frac{dP_\mu}{dP_0}$$

for a restriction of the process to the interval  $[0, T]$ . If  $g(X)$  only depended on the process up to time  $t < T$  then, from the martingale property of  $M_t(X)$ ,

$$\begin{aligned} E_\mu[g(X)] &= E_0[g(X)M_T(X)] \\ &= E_0\{E[g(X)M_T(X)|H_t]\} \\ &= E_0\{g(X)M_t(X)\} \end{aligned}$$

which shows that  $M_t(X)$  plays the role of a likelihood ratio for a restriction of the process to the interval  $[0, t]$ .

We can argue for the form of  $M_t(X)$  and show that it “should” be a martingale under  $\mu = 0$  by considering the limit of the ratio of the finite-dimensional probability density functions like

$$\frac{f_\mu(x_{t_1}, \dots, x_{t_n})}{f_0(x_{t_1}, \dots, x_{t_n})}$$

where  $f_\mu$  denotes the joint probability density function of  $X_{t_1}, X_{t_2}, \dots, X_{t_n}$  for  $t_1 < t_2 < \dots < t_n = T$ . These likelihood ratios are discrete-time martingales under  $P_0$ . For a more general diffusion, provided that the diffusion terms are identical, we can still express the Radon-Nikodym derivative as a stochastic exponential.

### Theorem A55 (Girsanov’s Theorem)

Suppose  $P$  is the measure on  $C[0, T]$  induced by  $X_0 = 0$ , and

$$dX_t = \mu(\omega, t)dt + \sigma(\omega, t)dW_t$$

under  $P$ . Assume the standard conditions so that the corresponding stochastic integrals are well-defined. Assume that the function

$$\theta(\omega, t) = \frac{\mu(\omega, t) - \nu(\omega, t)}{\sigma(\omega, t)}$$

is bounded. Then the stochastic exponential

$$M_t = \mathcal{E}\left(-\int_0^t \theta(\omega, s) dW_s\right) = \exp\left\{-\int_0^t \theta(\omega, s) dW_s - \frac{1}{2} \int_0^t \theta^2(\omega, s) ds\right\}$$

is a martingale under  $P$ . Suppose we define a measure  $Q$  on  $C[0, T]$  by

$$\frac{dQ}{dP} = M_T,$$

or, equivalently for measurable subsets  $A$ ,

$$Q(A) = E_P[1(A)M_T].$$

Then under the measure  $Q$ , the process  $W'_t$  defined by

$$W'_t = W_t - \int_0^t \theta(\omega, s) dW_s$$

is a standard Brownian motion and  $X_t$  has the representation

$$dX_t = \nu(\omega, t)dt + \sigma(\omega, t)dW'_t$$