10. Find functions $f_1(t)$ and $f_2(t) = -f_1(t)$ such that a standard Wiener process is between f_1 and f_2 with probability (a) 0.5, (b) 0.95.
11. What is the probability that $-\sqrt{t} < W(t) < \sqrt{t}$?
12. Prove that the transition probability density (11.12) of a Wiener process with drift satisfies the heat equation (11.13).
13. Use Theorem 6.5 to find the characteristic function of $X(s) = x_0 + \mu s + \sigma W(s)$.
14. Let $N = \{N(t)\}$ be a Poisson process with parameter λ. Find the mean and covariance functions of N.
15. Let $M = dN/dt$, where N is as in Exercise 14. What would a sample path of M look like? Use the results of Exercise 7 to ascertain the mean and covariance functions of M.
16. For $s > 0, t > 0$ find the *correlation coefficient* $\rho(s,t)$ (see section 1.3) of $W(s)$ and $W(t)$. Assume $s < t$ and s is fixed. What happens to ρ as $t \to \infty$?

12

Diffusion processes, stochastic differential equations and applications

12.1 DIFFUSION PROCESSES AND THE KOLMOGOROV (OR FOKKER–PLANCK) EQUATIONS

To introduce a wide class of random processes with properties similar to those of a Wiener process with drift, we generalize the constant *drift parameter* μ and *variance parameter* σ of such a process, so that they vary with the value of the process and possibly the time. For a general process X, we have that the *increment* in the small time interval $(t, t+\Delta t]$ is

$$\Delta X(t) = X(t+\Delta t) - X(t).$$

Now the properties of this increment may depend on the time t and the value x of the process at the beginning of the small time interval. We therefore condition on $X(t) = x$ and define the **infinitesimal first moment, or infinitesimal mean,** as

$$\alpha(x,t) = \lim_{\Delta t \to 0} \frac{E[\Delta X(t) \mid X(t) = x]}{\Delta t}. \tag{12.1}$$

Note that because we have taken the expectation, this is not a random quantity. Thus $\alpha(x,t)$ is a deterministic function of x and t. Similarly we define the **infinitesimal second moment,** or, as will be seen in Exercise 1, **infinitesimal variance,**

$$\beta(x,t) = \lim_{\Delta t \to 0} \frac{E[(\Delta X)^2 \mid X(t) = x]}{\Delta t}. \tag{12.2}$$

We assume that the higher order infinitesimal moments are zero, so that, for $n = 3, 4, \ldots,$

$$\lim_{\Delta t \to 0} \frac{E[(\Delta X)^n \mid X(t) = x]}{\Delta t} = 0. \tag{12.3}$$

This indicates that changes in the process in small time intervals will be small, and in fact small enough to make the sample paths continuous for suitably chosen functions α and β. Such a process is called a **diffusion process** and behaves in a fashion similar to a Wiener process – although its paths are continuous, they are with probability one non-differentiable. The **drift** (α) and **diffusion** (β) components depend on the position and the time.

Once the drift and diffusion terms are specified, we are in a position to obtain as much information as we require about the process if we can find its transition probability density function. Fortunately this can always be done because as the following theorem indicates, this function satisfies a partial differential equation which is a general form of the much studied heat equation – the differential equation (11.13) satisfied by the transition probability functions of the Wiener process with drift.

Theorem 12.1 **Let $p(y,t|x,s)$ be the transition probability density function for a diffusion process with first and second infinitesimal moments $\alpha(y,t)$ and $\beta(y,t)$ as defined in equations (12.1) and (12.2) respectively. Then p satisfies the forward Kolmogorov equation**

$$\frac{\partial p}{\partial t} = -\frac{\partial(\alpha p)}{\partial y} + \frac{1}{2}\frac{\partial^2(\beta p)}{\partial y^2}, \tag{12.4}$$

with suitable initial and boundary conditions.

Equation (12.4) is also called a *Fokker-Planck equation*, especially by physical scientists, who sometimes refer to it as a 'Master equation', to emphasize its generality. Proof that this equation follows from the Chapman–Kolmogorov equation (11.14) and the relations (12.1)–(12.3), though not difficult, is rather long and is hence omitted here. Interested readers may refer to, for example, Jaswinski (1970).

The equation (12.4) is called the *forward equation* because the variables x and s which refer to the earlier event are considered to be fixed as the later variables y and t vary. One may also consider p as a function with fixed values of y and t, and allow the earlier variables x and s to vary. This gives rise to the *backward equation* which is often very useful, for problems such as ascertaining times at which a certain value or set of values is first attained.

Theorem 12.2 **Let α and β be the first and second infinitesimal moments of a diffusion process. If the process has a transition probability density function $p(y,t|x,s)$, then this density considered as a function of x and s with y and t fixed, satisfies the backward Kolmogorov equation,**

$$-\frac{\partial p}{\partial s} = \alpha\frac{\partial p}{\partial x} + \frac{1}{2}\beta\frac{\partial^2 p}{\partial x^2}. \tag{12.5}$$

The derivation of the backward equation from the Chapman–Kolmogorov equation is also relatively straightforward but will again not be given here. In addition it will be seen that the transition probability distribution function $P(y,t|x,s)$ also satisfies equation (12.5).

Time-homogeneous processes

In many problems of physical interest, the behaviour of a process depends not on the actual value of the time, but rather the *length* of the time interval since the process was *switched on*. Such a process is called **temporally (or time-) homogeneous** and nearly all diffusion processes which have arisen in applications fall into this category. (Note that some authors refer to such a process as one with stationary transition probabilities.) Clearly the first and second infinitesimal moments of such a process do not depend explicitly on time, so we have $\alpha(x,t)=\alpha(x)$, and $\beta(x,t)=\beta(x)$, being functions only of the state variable. Furthermore we have

$$p(y,t|x,s) = p(y,t-s|x,0),$$

so that we can conveniently drop one of the arguments of the transition density. Thus we can use $p(y,t|x)$ for the density associated with transitions from a state $X(0)=x$. That is,

$$p(y,t|x) = \frac{\partial}{\partial y}\Pr\{X(t) \leqslant y | X(0)=x\}.$$

The forward and backward Kolmogorov equations now take somewhat simpler forms. For the *forward equation* we have

$$\frac{\partial p}{\partial t} = -\frac{\partial(\alpha(y)p)}{\partial y} + \frac{1}{2}\frac{\partial^2(\beta(y)p)}{\partial y^2}, \tag{12.6}$$

and, as is seen in the exercises, the *backward equation* simplifies to

$$\frac{\partial p}{\partial t} = \alpha(x)\frac{\partial p}{\partial x} + \frac{1}{2}\beta(x)\frac{\partial^2 p}{\partial x^2}. \tag{12.7}$$

Boundary conditions

When investigating the properties of a diffusion process by means of the Kolmogorov differential equations, it is necessary to prescribe appropriate boundary conditions in order to solve the latter. To be specific, let us assume that the diffusion process is on the interval (x_1, x_2) and the time at which it commences is $t=0$. Assume from now on also that the process is **time-homogeneous.** For the Kolmogorov equations involving the transition density

Alternatively, the formula (see section 12.2)

$$p^*(y) = \frac{K_2}{\beta(y)\phi(y)},$$

with

$$\phi(y) = \exp\left[\int^y \frac{2(b - ay')}{\sigma^2} \mathrm{d}y'\right].$$

can be employed – see the exercises.

12.5 STOCHASTIC INTEGRALS AND STOCHASTIC DIFFERENTIAL EQUATIONS

We have seen in section 11.4 that a Wiener process with drift can be characterized by the *stochastic differential equation*

$$\mathrm{d}X = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W.$$

The correct interpretation of this equation is in terms of an integral involving a Wiener process – called a **stochastic integral.**

There are a large number of integrals which one may define in connection with random processes. Mathematical complexities arise when integrals involving W are considered because of the irregular properties of the paths of W. This means that the methods of defining integrals given in real-variable calculus courses cannot be used. We will consider stochastic integrals very briefly and somewhat superficially – there are numerous technical accounts – see for example Gihman and Skorohod (1972), Arnold (1974), Lipster and Shiryayev (1977) or Oksendal (1985). Our main purpose is to enable the reader to understand and know how to use a stochastic differential equation of the general form

$$\mathrm{d}X(t) = f(X(t), t)\,\mathrm{d}t + g(X(t), t)\,\mathrm{d}W(t).$$

Equivalently, dropping the reference to t in the random processes, we can write this as

$$\mathrm{d}X = f(X, t)\,\mathrm{d}t + g(X, t)\,\mathrm{d}W, \tag{12.20}$$

where f and g are real-valued functions, W is a standard Wiener process and X is a random process which in cases of interest will be a diffusion process.

However, it must be stated at the outset that (12.20) does **not** always have a unique interpretation. This situation arises for the following reason. Equation (12.20) is interpreted correctly as implying the **stochastic integral equation**

$$X(t) = X(0) + \int_0^t f(X(t'), t')\,\mathrm{d}t' + \int_0^t g(X(t'), t')\,\mathrm{d}W(t'), \tag{12.21}$$

and the process X so defined is called a **solution of the stochastic differential equation (12.20).** Although the first integral here presents no problems, there are many ways of defining the second one,

$$\int_0^t g(X(t'), t')\,\mathrm{d}W(t'),$$

which is called a **stochastic integral.** Furthermore, the different definitions can lead to various solutions, X, with quite different properties.

Despite this apparent ambiguity, there are two useful definitions which are most commonly employed – the **Ito stochastic integral** and the **Stratonovich stochastic integral**; and there is a simple relation between these two.

A note on notation. It is preferable in (12.20) not to 'divide' throughout by dt, because as we have seen, the derivatives of W and hence of X do not exist in the usual sense. However, as long as we keep that in mind, it is possible to display (12.20) as a stochastic differential equation involving *white noise* w, the 'derivative' with respect to time t of W (see section 11.3):

$$\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, t) + g(X, t)w,$$

or perhaps even

$$\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, t) + g(X, t)\dot{W}.$$

Stochastic differential equations written in this form are often called **Langevin equations.**

Let us now make an important observation on the stochastic differential equation (12.20). **If the function g is identically zero, the differential equation is deterministic** and can be written in the usual way

$$\frac{\mathrm{d}X}{\mathrm{d}t} = f(X, t).$$

Assuming the initial value $X(0) = x_0$ is not random then $X(t)$ is non-random for all t and this equation is solved in the usual way.

We expect that the behaviour of solutions of this deterministic equation would be related to those of the stochastic differential equation (12.20), and be close to them when the **noise term** g is small. We would be correct in believing that, in particular, the expected value $E[X(t)]$ of the solution of

(12.20) would not be very far, in most cases, from the solution of the deterministic equation.

This can be illustrated nicely with the Wiener process with linear drift μt and variance parameter σ. From section 11.4, this process has the stochastic differential equation (11.11):

$$\mathrm{d}X = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W.$$

Here

$$f(X,t) = \mu,$$

and

$$g(X,t) = \sigma.$$

If we put $\sigma = 0$ we obtain the deterministic differential equation

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu.$$

The solution of this with initial value x_0 is

$$X(t) = x_0 + \mu t,$$

and this, as seen in section 11.4, is equal to the mean value function of the process satisfying (11.11). The added noise makes the paths of X very irregular, but the mean value is still μt. This was depicted in Fig. 11.4.

Heuristic interpretation

Before proceeding more formally, let us describe roughly how we can understand an equation of the form (12.20). This can perhaps best be accomplished by writing the related difference equation

$$\Delta X = f(X,t)\Delta t + g(X,t)\Delta W. \tag{12.22}$$

Here we may regard the (random) *increment* in X in the time interval $(t, t+\Delta t]$ as having two components. The first component is equal to the value of $f(X;t)$ at the beginning of the time interval multiplied by the length Δt of the time interval. The second component is the value of $g(X,t)$ at the beginning of the time interval, multiplied by the (random) increment $\Delta W = W(t+\Delta t) - W(t)$ that occurs in a standard Wiener process in Δt. As we have seen, ΔW is a Gaussian random variable with mean zero and variance Δt. We have essentially outlined a **method of simulation** of the stochastic differential equation (12.20) – this will be elaborated on below.

It should be realized, however, that even though the functions f and g are functions in the usual deterministic sense, both the components of the increment in X, namely, $f\Delta t$ and $g\Delta W$, are random variables, because $f(X(t),t)$, as well as $g(X(t),t)$ are random variables by virtue of their dependence on the value of the random variable $X(t)$.

We will now proceed to define the Ito integral, then state, with the aid of a proof outline, a change of variable rule called **Ito's formula**. This will be followed by a brief consideration of the Stratonovich integral and the roles of the various integrals in stochastic modelling.

The Ito stochastic integral

We are going to define the Ito stochastic integral

$$I = \int_a^b f(t)\,\mathrm{d}W(t),$$

where $\{f(t), t\in[a,b)\}$ is a suitable random process. (The f here is general and not related to that in (12.20).) Suitable random processes will be said to belong to **class** M. We define this class as containing random processes f which satisfy the following requirements:

(i) f is a *non-anticipative* process. In the present context this means that questions about f up to and including time t can be answered without knowledge of the evolution of the Wiener process W for times beyond t; or one could say that the evolution of f up to and including any particular time is *independent* of future values of W.

(ii) the integral of the square of $f(t)$ over $[a,b)$ is *finite* with probability one; that is,

$$\Pr\left\{\int_a^b [f(t)]^2\,\mathrm{d}t < \infty\right\} = 1.$$

Thus the sample paths of f cannot be often singular or too wildy fluctuating – such paths may occur, but their associated probability is zero. When an event occurs with probability one it is said to be **almost sure** so we could say that the integral of f^2 is almost surely finite. Note that the integral appearing here is, like the Ito integral we are about to define, a random variable, which takes on various values as the various sample paths of f arise.

Definition of Ito stochastic integral for simple random functions

A **simple random function** (or step function), in the present context, is one which is constant on sub-intervals of $[a,b)$. Such a function is represented *schematically* in Fig. 12.2. The constant values on the various sub-intervals are actually *random variables* – fixed ones – rather than numbers as in real-variable calculus. We will insert the underlying sample-space variable ω to make it clear that these are constant random variables. Thus we may put

$$f(t,\omega) = f_j(\omega), \quad t_j \leqslant t < t_{j+1}, \quad j = 0,1,\ldots,n-1,$$

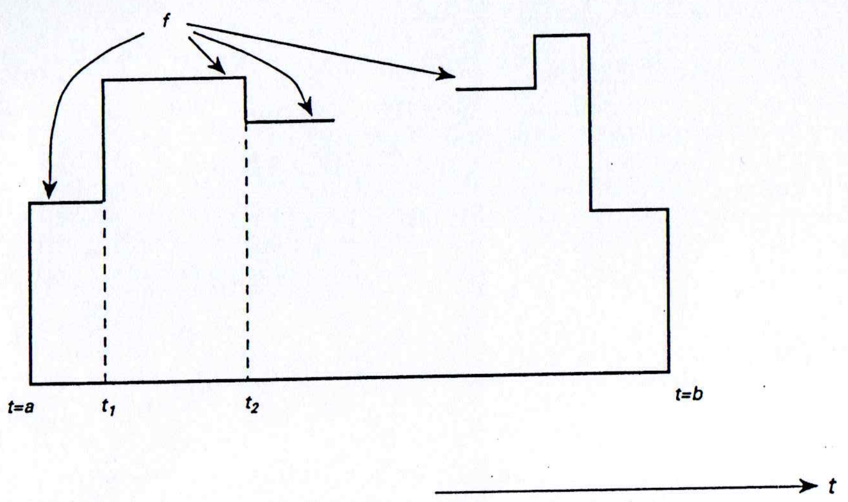

Figure 12.2 A *schematic* representation of a simple function f. In the present context the constant values on sub-intervals are actually random variables.

where $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ is a **partition** of the interval $[a, b)$ and the $\{f_j\}$ are a set of n random variables.

Now for each sub-interval $[t_j, t_{j+1})$ of $[a, b)$ we can define the corresponding increment in a standard Wiener process

$$\Delta W_j = W(t_{j+1}) - W(t_j), \quad j = 0, 1, \ldots, n-1,$$

which defines a further n random variables. We are now ready to define the Ito stochastic integral of a simple random function f with respect to a Wiener process as

$$\int_a^b f(t)\,\mathrm{d}W(t) = \sum_{j=0}^{n-1} f_j \Delta W_j.$$

We note that we could also write this as

$$\int_a^b f(t)\,\mathrm{d}W(t) = \sum_{j=0}^{n-1} f(t_j)[W(t_{j+1}) - W(t_j)] \qquad (12.23)$$

which will be a useful observation when we distinguish between the Ito and Stratonovich integrals.

In order to define the stochastic integral for the general class of random processes we have called M, we state the following lemma without proof – see the references at the beginning of this section. This lemma tells us that for any random process f in M, we can be sure there is a sequence $\{f_n, n = 1, 2, \ldots\}$, also in M, whose members get closer and closer to f as $n \to \infty$, in the following sense.

Lemma **Let $f \in M$ be a random process in the class defined above. Then a sequence of simple functions $f_n(t) \in M$ exists such that as $n \to \infty$,**

$$\int_a^b [f(t) - f_n(t)]^2\,\mathrm{d}t \xrightarrow{p} 0.$$

That is, this sequence of integrals, which measure the distances between f and f_n, converges in probability to zero – recall the definition of this mode of convergence in section 6.6.

Definition of Ito stochastic integral for arbitrary $f \in M$

Since we know how to define the Ito stochastic integral for simple functions, we extend the definition to arbitrary random processes in M by using a sequence of approximating integrals of simple functions. The limit of this sequence is defined as the required integral. Thus we set

$$\int_a^b f(t)\,\mathrm{d}W(t) = \lim_{n \to \infty} \int_a^b f_n(t)\,\mathrm{d}W(t),$$

where the limit is again in the sense of **convergence in probability of a sequence of random variables.** The above lemma guarantees that a suitable sequence $\{f_n(t)\}$ of approximating random functions exists.

We mentioned that other definitions can be given for the Ito stochastic integral according to the different properties which are ascribed to the integrand $f(t)$. Usually these conditions are more restrictive (**stronger**) than the ones we have employed. However, then, and indeed in most cases of practical interest, we have the following results concerning the mean and the variance of the Ito stochastic integral:

(i) Mean

$$E\left(\int_a^b f(t)\,\mathrm{d}W(t)\right) = 0. \qquad (12.24)$$

(ii) Variance

$$E\left(\int_a^b f(t)\,\mathrm{d}W(t)\right)^2 = \int_a^b f^2(t)\,\mathrm{d}t. \qquad (12.25)$$

We will not prove (12.25) but (12.24) will be considered in the exercises. Furthermore, if f and g are both in M then

$$E\left(\int_a^b f(t)\,\mathrm{d}W(t) \times \int_a^b g(t)\,\mathrm{d}W(t)\right) = \int_a^b E[f(t)g(t)]\,\mathrm{d}t. \qquad (12.25')$$

Thus the differential equations (12.4) and (12.5) satisfied by the transition probability density function can be obtained simply from the stochastic differential equation.

In particular, for a **time-homogeneous** process with stochastic equation

$$dX = f(X)\,dt + g(X)\,dW,$$

the forward Kolmogorov equation for the transition density $p(y, t|x)$ is, from Equation (12.6),

$$\frac{\partial p}{\partial t} = -\frac{\partial (f(y)p)}{\partial y} + \frac{1}{2}\frac{\partial^2 (g^2(y)p)}{\partial y^2},$$

and the corresponding backward equation is, from Equation (12.7),

$$\frac{\partial p}{\partial t} = f(x)\frac{\partial p}{\partial x} + \frac{1}{2}g^2(x)\frac{\partial^2 p}{\partial x^2}.$$

If a stochastic differential equation is given with Stratonovich's interpretation, it may first be converted to an Ito equation using Equation (12.30); then Theorem 12.5 will yield the infinitesimal moments of the process.

It might seem that, in modelling a random phenomenon with a stochastic differential equation, an ambiguity arises in the choice of an Ito or Stratonovich interpretation. This, however, is not the case. A diffusion process is specified once its infinitesimal mean and variance are given. Thus, when deriving a model one must be certain that these infinitesimal moments are correct. Once this is done it matters not which stochastic calculus one adopts.

A word of caution is necessary. It is not a good idea to take an existing deterministic differential equation and convert one of its parameters to white noise. The reason is that in most cases there **will be ambiguity** because there is no *a priori* reason why a particular interpretation, Ito or Stratonovich, should be correct. A particular choice would only be defensible if the infinitesimal moments could be ascertained to be the correct ones. This will be illustrated in the exercises. However, in certain cases it has been shown that limits of sequences of discrete stochastic equations have solutions corresponding to the Stratonovich differential (Wong and Zakai, 1965).

12.7 APPLICATIONS

Diffusion approximation to a random walk

Suppose a random walk $X_\varepsilon = \{X_\varepsilon(t), t \geqslant 0\}$ occurs with jumps up or down of magnitude ε. Jumps up form a Poisson process N_1 with intensity λ and jumps down form another Poisson process N_2, independent of N_1, but with the same intensity. One may assume for convenience that the processes start

at zero at time $t = 0$. Note that this random walk, sometimes called a **randomized random walk** (Feller, 1971), differs from the simple random walk of Chapter 7 in that there the jumps (steps) occurred at fixed time intervals whereas now they occur at random times.

We may write the following expression for the value of the process X_ε at time t:

$$X_\varepsilon(t) = \varepsilon N_1(t) - \varepsilon N_2(t).$$

Similarly we may write the following expression for the increment in X_ε in the interval $(t, t + \Delta t]$:

$$\begin{aligned}\Delta X_\varepsilon &= X_\varepsilon(t + \Delta t) - X_\varepsilon(t)\\ &= \varepsilon[N_1(t + \Delta t] - N_1(t)] - \varepsilon[N_2(t + \Delta t) - N_2(t)]\\ &= \varepsilon\Delta N_1 - \varepsilon\Delta N_2.\end{aligned}$$

One can also write the following stochastic differential equation involving Poisson processes:

$$dX_\varepsilon = \varepsilon[dN_1 - dN_2].$$

We will now determine the first two infinitesimal moments of a diffusion process, X, obtained from the X_ε-processes as the jump magnitudes ε go to zero and the jump rates λ go to infinity – but in such a way that these moments remain finite and non-zero in the limit. We may call this a **diffusion approximation** to the original discontinuous random walk – and this approximation will have, as we have seen, continuous sample paths. The situation is illustrated in Fig. 12.3.

Now as we have seen in Equation (9.2), the increments in the Poisson process N_1 are such that:

$$\Delta N_1 = \begin{cases} 1, & \text{with prob. } \lambda\Delta t + o(\Delta t);\\ 0, & \text{with prob. } 1 - \lambda\Delta t + o(\Delta t),\end{cases}$$

and the probabilities of other values are $o(\Delta t)$; similarly for increments ΔN_2 in the process N_2.

Thus, utilizing the fact that the increments are independent of previous values and that the two Poisson processes are independent, we have

$$E[\Delta X_\varepsilon | X_\varepsilon(t) = x] = o(\Delta t) \qquad (12.31)$$

and

$$\begin{aligned}\operatorname{Var}[\Delta X_\varepsilon | X_\varepsilon(t) = x] &= \varepsilon^2[\operatorname{Var}(\Delta N_1) + \operatorname{Var}(\Delta N_2)] \qquad (12.32)\\ &= 2\varepsilon^2\lambda\Delta t + o(\Delta t).\end{aligned}$$

The first two infinitesimal moments, α_ε and β_ε, of X_ε are thus, from (12.31)

When the Ito definition of stochastic integral is used in the interpretation of equation (12.20) we call it an *Ito stochastic differential* or *Ito stochastic differential equation.* The process X is called the solution of the equation. There are many accounts of existence and uniqueness theorems and properties of the solution – see for example the references at the beginning of this section. We will avoid technicalities and simply assume that the functions f and g are *suitable.*

Ito's formula

When one uses Ito's definition of stochastic integral, the usual rules of calculus sometimes must be modified. This is not the case with the Stratonovich definition considered below. In particular, one must be careful when changing variables as the following result, called **Ito's formula**, shows.

We will not give a detailed proof, which can be found for example in Liptser and Shiryayev (1977). We will simply utilize the fact that

$$(\mathrm{d}W^2) \text{ acts like } \mathrm{d}t. \tag{12.26}$$

This statement can be understood in terms of the following probability one relation (see, for example the previous reference) for the so-called *quadratic variation* of W over the interval $[0, t)$:

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} [W(t_{j+1}) - W(t_j)]^2 = t,$$

where $\{0 = t_0 < t_1 < \cdots < t_n = t\}$ is a partition of $[0, t)$. This suggests we can write

$$\int_0^t (\mathrm{d}W)^2 = t.$$

and the latter can be written as (12.26). The significance of this result for our purposes is that whereas second order differentials such as $(\mathrm{d}t)^2$ can usually be ignored, $(\mathrm{d}W)^2$ is of the same order as $\mathrm{d}t$ and hence cannot be neglected.

We are now ready to discuss Ito's formula.

Notation. In the following we will use subscripts to denote partial differentiation with respect to a variable. Thus, for example, if $h = h(x, t)$ then

$$h_{xt} = \frac{\partial^2 h}{\partial t \partial x}.$$

Theorem 12.4 **Let the random process X have the Ito stochastic differential**

$$\mathrm{d}X = f(X, t)\,\mathrm{d}t + g(X, t)\,\mathrm{d}W, \tag{12.27}$$

for suitable functions f and g. Suppose we change variables by putting

$$Y = h(X, t).$$

Then Y satisfies the Ito stochastic differential equation

$$\mathrm{d}Y = \left[f h_x + \frac{g^2}{2} h_{xx} + h_t \right] \mathrm{d}t + h_x g\,\mathrm{d}W, \tag{12.28}$$

where h_x and h_t are the first partial derivatives of the function h with respect to its first and second variables, respectively, and h_{xx} is the second partial derivative of h with respect to its first variable. Equation (12.28) is called Ito's formula.

Proof outline It we use the first few terms in a Taylor's expansion for h we get

$$\begin{aligned}
\mathrm{d}Y &= h_x(X,t)\,\mathrm{d}X + h_t(X,t)\,\mathrm{d}t \\
&\quad + \tfrac{1}{2} h_{xx}(X,t)(\mathrm{d}X)^2 + h_{xt}(X,t)\,\mathrm{d}X\,\mathrm{d}t + \tfrac{1}{2} h_{tt}(\mathrm{d}t)^2 + \cdots \\
&= h_x(X,t)[f\,\mathrm{d}t + g\,\mathrm{d}W] + h_t(X,t)\,\mathrm{d}t \\
&\quad + \tfrac{1}{2} h_{xx}(X,t)[f^2(\mathrm{d}t)^2 + 2fg\,\mathrm{d}t\,\mathrm{d}W + g^2(\mathrm{d}W)^2] \\
&\quad + h_{xt}(X,t)[f\,\mathrm{d}t + g\,\mathrm{d}W]\,\mathrm{d}t + \tfrac{1}{2} h_{tt}(X,t)(\mathrm{d}t)^2 + \cdots
\end{aligned}$$

Retaining only those terms of order $\mathrm{d}t$ or $\mathrm{d}W$, we obtain (12.28). There is the extra term involving h_{xx} which is absent in standard calculus.

Example

Suppose X is a standard Wiener process and so satisfies the stochastic differential equation

$$\mathrm{d}X = \mathrm{d}W.$$

What Ito stochastic equation does the new variable

$$Y = X^2$$

satisfy?

Solution

Relating this to the standard form (12.27), we find $f(X, t) = 0$ and $g(X, t) = 1$. Now $h(X, t) = X^2$, so we have

$$h_x(X,t) = 2X$$
$$h_{xx}(X,t) = 2$$
$$h_t(X,t) = 0.$$

Using (12.28) we see

$$dY = [2X.0 + \tfrac{1}{2}.2 + 0]\,dt + 2X\,dW$$

or

$$dY = 1\,dt + 2\sqrt{Y}\,dW,$$

which is the required stochastic differential equation for Y. The extra term is 1 dt. Other examples will appear as exercises.

Stratonovich's stochastic integral

Consider equation (12.23) where the Ito integral was defined for simple functions. The Stratonovich integral differs in that instead of employing the value of f at the beginning t_j of the sub-interval $[t_j, t_{j+1})$, the value at the mid-point $(t_j + t_{j+1})/2$ is used. The integral is thus said to be *symmetrized.* This difference is sufficient to alter sometimes the properties of the resulting random variable significantly.

With the Stratonovich definition, the change of variable formula (12.28) becomes simply

$$dY = [fh_x + h_t]\,dt + h_x g\,dW, \qquad (12.29)$$

and usual calculus rules (product rule, quotient rule, etc.) can be used.

However, in most cases encountered in modelling real-world phenomena, one may switch back and forth from the Ito and Stratonovich schemes by using the following simple result. If X satisfies the *Stratonovich stochastic differential equation*

$$dX = f(X,t)\,dt + g(X,t)\,dW,$$

then the corresponding Ito equation is

$$dX = [f(X,t) + \tfrac{1}{2}g(X,t)g_x(X,t)]\,dt + g(X,t)\,dW. \qquad (12.30)$$

Note that if g contains no explicit X-dependence, then g_x is zero and there is no difference between the two definitions. In particular, **for all equations of the form**

$$dX = f(X,t)\,dt + g(t)\,dW,$$

it makes absolutely no difference whether the interpretation is through an Ito or a Stratonovich stochastic integral.

Example

If X satisfies the Stratonovich equation

$$dX = X\,dt + X^2\,dW,$$

what is the equivalent Ito equation?

Solution

Here $f = X$ and $g = X^2$. Hence $g_x = 2X$ and we find

$$dX = [X + \tfrac{1}{2}X^2.2X]\,dt + X^2\,dW$$

or

$$dX = [X + X^3]\,dt + X^2\,dW.$$

is the corresponding Ito equation.

12.6 MODELLING WITH STOCHASTIC DIFFERENTIAL EQUATIONS

We have seen in the last section that a wide class of random processes may be defined directly by writing down a stochastic differential equation of the form

$$dX = f(X,t)\,dt + g(X,t)\,dW,$$

where we call $f(X,t)\,dt$ the **drift term** and $g(X,t)\,dW$ the **noise term.** In most cases of interest, the solutions of such equations are *diffusion processes* which were defined by analytical methods in section 12.1. In that section, diffusion processes were defined in terms of their *first infinitesimal moment* – given as $\alpha(x,t)$ in Equation (12.1) – and *infinitesimal variance* – given as $\beta(x,t)$ in Equation (12.2). Such processes have sample paths which are **continuous** with probability one.

It would be very convenient to have a relationship between these two representations so that given a stochastic differential equation for X one could immediately ascertain its infinitesimal mean and variance, and *vice versa*: – given the infinitesimal moments, write down the corresponding stochastic differential equation. It is indeed possible to accomplish these switches very easily as the following result shows.

Theorem 12.5 Let X be a diffusion process with Ito stochastic differential equation

$$dX = f(X,t)\,dt + g(X,t)\,dW,$$

where f and g are suitable functions. Then the infinitesimal mean of X is given by

$$\alpha(x,t) = f(x,t),$$

and the infinitesimal variance of X is given by

$$\beta(x,t) = g^2(x,t).$$

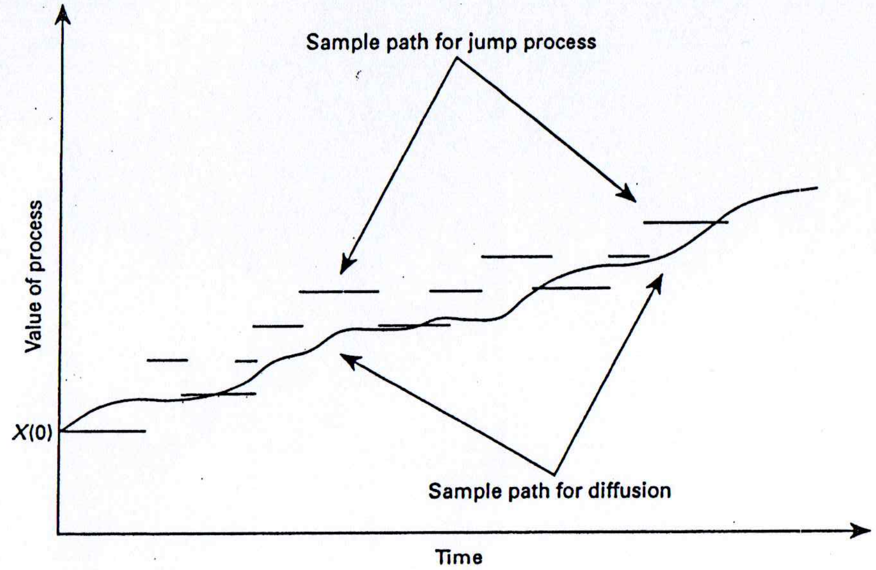

Figure 12.3 Showing how the discontinuous random walk is approximated by a diffusion process with continuous sample paths.

and (12.32) respectively, and the definitions (12.1) and (12.2),

$$\alpha_\varepsilon(x,t) = \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0,$$

and

$$\beta_\varepsilon(x,t) = \lim_{\Delta t \to 0} \frac{2\varepsilon^2\lambda\Delta t}{\Delta t} + \frac{o(\Delta t)}{\Delta t}$$
$$= 2\varepsilon^2\lambda.$$

Let us now make ε vanish and λ grow to infinity but insist that the infinitesimal variance remains finite and nonzero. This can be easily achieved if, for example, we require

$$\lambda = \frac{1}{2\varepsilon^2}. \quad (12.33)$$

In this case, the first and second infinitesimal moments of the limiting process X are

$$\alpha(x,t) = 0,$$

and

$$\beta(x,t) = 1.$$

With our choice of the relation (12.33) and letting $\varepsilon \to 0$, it will be seen in the exercises that all the higher order infinitesimal moments of X vanish as required by (12.3) – although some are zero regardless of the value of ε.

We can now use Theorem 12.5. Since $\beta(x,t) = 1$ we have $g^2(x,t) = 1$, so, choosing the positive root,

$$g(x,t) = 1.$$

Hence the diffusion process approximating the random walk has the stochastic differential equation

$$\mathrm{d}X = \mathrm{d}W.$$

That is, X is a standard Wiener process. One may also illustrate this **convergence to a Wiener process** using characteristic functions – see the exercises.

A mathematical model for the activity of a nerve cell (neuron)

In Fig. 12.4 we depict a *nerve cell* or *neuron* – a component of the nervous system. The one depicted is a pyramidal cell of the part of the human brain called the cerebral cortex. Such cells are specialized to receive signals from other cells at the junctions called *synapses.* If no signals are received, the electrical potential difference V_M across the target cell membrane tends to a fixed value called the *resting potential,* which we designate by V_R. Usually V_R is about 70 millivolts (the inside being negative) or about $\frac{1}{20}$th of the voltage of the familiar AA battery.

Let us denote the difference between V_M and V_R by V:

$$V = V_M - V_R,$$

– see the left part of Fig. 12.5. V is called the *depolarization.*

As explained at the end of Chapter 7, the incoming signals may *excite* the target cell: this makes V_M less negative, in which case V is increased. We assume that each such incoming signal increases V by an amount a_E and that such signals occur as a Poisson process N_E with intensity λ_E. Alternatively an incoming signal may *inhibit* the target cell, thereby decreasing V (which can go negative), by an amount a_I. Let such inhibitory effects be generated by a Poisson process N_I with intensity λ_I. These events result in V's executing a random walk as depicted on the right in Fig. 12.5.

The state of the neuron at time t can thus be characterized by

$$V(t) = a_E N_E(t) - a_I N_I(t),$$

and we assume $V(0) = 0$, so the cell is initially at resting level. The reader can show that the mean and variance of the change in V during $(t, t+\Delta t]$ are

$$E[\Delta V] = a_E\lambda_E\Delta t - a_I\lambda_I\Delta t + o(\Delta t),$$

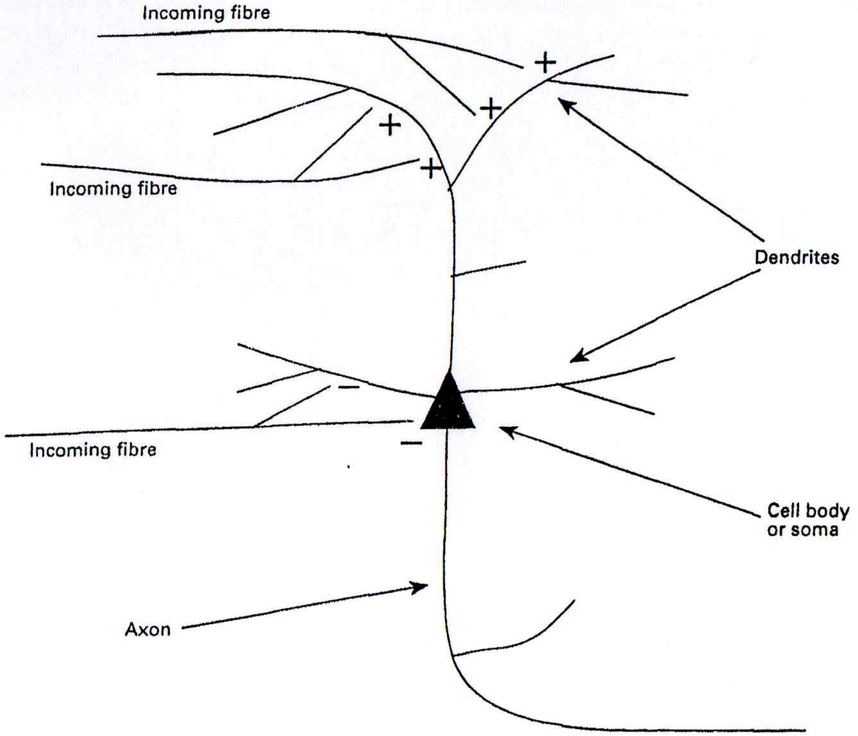

Figure 12.4 Here we depict a class of neuron called a pyramidal cell, showing the main components of axon, dendrites and soma. The neuron is shown connected into a network through synapses at which other cells transmit signals to the target cell. Here a (+) indicates an excitatory junction whereas a (−) denotes an inhibitory one.

and

$$\mathrm{Var}[\Delta V] = a_E^2\lambda_E\Delta t + a_I^2\lambda_I\Delta t + o(\Delta t).$$

Dividing by Δt and taking the limit $\Delta t \to 0$, we obtain the first and second infinitesimal moments. We may construct a diffusion approximation $V^*(t)$ to V with the same infinitesimal mean and variance, viz.,

$$\alpha(x,t) = a_E\lambda_E - a_I\lambda_I,$$
$$\beta(x,t) = a_E^2\lambda_E + a_I^2\lambda_I.$$

Using Theorem 12.5, we find the following stochastic differential for V^*:

$$\mathrm{d}V^* = (a_E\lambda_E - a_I\lambda_I)\,\mathrm{d}t + \sqrt{a_E^2\lambda_E + a_I^2\lambda_I}\,\mathrm{d}W.$$

This will be recognized as the stochastic equation for a *Wiener process with drift*.

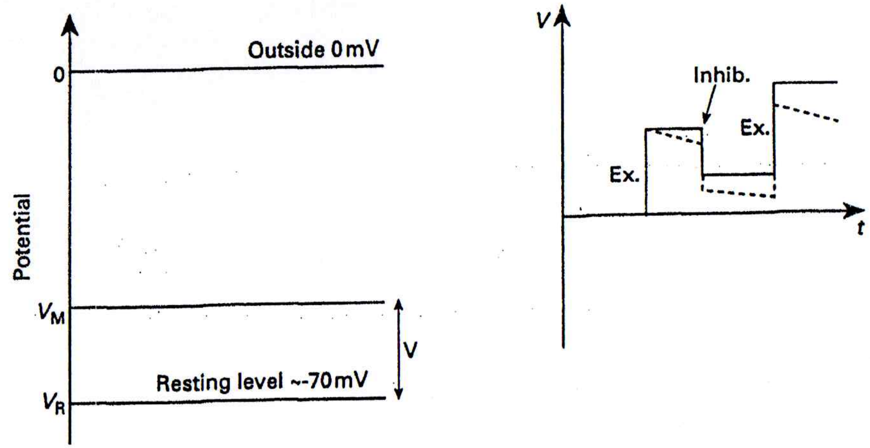

Figure 12.5 On the left we show the relation between V, V_R and V_M. On the right are shown the variations which occur in V as signals come in from other cells. Two excitatory inputs (Ex.) and one inhibitory input (Inhib.) are depicted. The solid line is a trajectory for the process $a_EN_E(t) - a_IN_I(t)$ used in the text to model V. The dashed line represents the more physiologically realistic trajectory involving decay to resting level ($V = 0$) between input events.

The mean and variance, as well as the transition density of the state of a nerve cell can be estimated using this approximation. In addition, the time taken for V or V^* to reach a **threshold value,** θ, say, can be determined as a **first passage time.** When V reaches this threshold value, which is usually about 10 millivolts (i.e., 0.01 volts) from resting level, our target cell itself sends out a signal along its axon (see Fig. 12.3) to excite or inhibit other neurons. The whole collection of neurons is sometimes called a **neural network.** Details of such calculations can be found in Tuckwell (1988) where it is also seen that the *Ornstein–Uhlenbeck process* (see section 12.4) is a somewhat more realistic mathematical model for a neuron than the Wiener process with drift. In the absence of inputs the OUP model correctly predicts that the electrical potential across the neuron's membrane will return exponentially to resting level.

A model for population growth in a random environment

As we have seen in section 9.1, when a population grows in an unrestrained way, the population size $x(t)$ at time t may be approximated by the solution of the deterministic differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = rx, \qquad (12.34)$$

where the assumed constant growth rate r is the difference between the birth and death rates. This is the *Malthusian law* with exponential solutions

$$x(t) = x(0)\, e^{rt},$$

$x(0)$ being the initial population size.

There are often factors, particularly environmental ones, which make the birth and death rates chop and change from generation to generation. For example, in very cold or very dry seasons we might expect a higher death rate in populations of many species, so r would drop below some long-term average value. The same might happen in very hot and very wet seasons. On the other hand, if climatic conditions lead to an abundance of food, r might increase above its average value, but this would be complicated if an organism's predators also benefited.

This suggests that a more accurate mathematical model of population growth would result if the growth rate r and hence the population size x are random processes. The simplest (but not necessarily accurate) assumption is that the growth rate $r(t)$ is the sum of its constant mean value $\bar{r}$ and a fluctuating white noise,

$$r(t) = \bar{r} + \sigma w(t),$$

where σ is a variance parameter and $w(t)$ is standard white noise, defined in section 11.3. Then we may write the model for the random population size $X(t)$,

$$\frac{dX}{dt} = (\bar{r} + \sigma w(t))X(t),$$

or, more satisfactorily as the stochastic differential equation,

$$dX = \bar{r}X\, dt + \sigma X\, dW. \qquad (12.35)$$

We let the initial value be

$$X(0) = x_0.$$

As mentioned in the last section, the approach we have used has the apparent drawback of necessitating a choice of stochastic integral. Fortunately, however, in this case the derivation of a diffusion approximation to Malthusian growth has been carried out (Tuckwell and Walsh, 1983a) from first principles. This approximation satisfies Equation (12.35) interpreted as a Stratonovich equation. Note that this does not imply that the Stratonovich integral is better or more correct than the Ito integral.

If we proceed with Equation (12.35) as a Stratonovich equation we can use the usual rules of calculus. In particular, we find that the change of variable

$$Y = \ln(X)$$

leads to

$$\frac{dY}{dt} = \frac{dY}{dX}\frac{dX}{dt} = \frac{1}{X}\frac{dX}{dt}.$$

Thus

$$\frac{dY}{dt} = \bar{r} + \sigma\frac{dW}{dt}$$

or, equivalently,

$$dY = \bar{r}\, dt + \sigma\, dW.$$

Thus the transformed process Y is simply a *Wiener process with drift*. This is a particular case of a general transformation method (Tuckwell, 1973; 1974) in which the equation

$$dX = f(X)[\mu\, dt + \sigma\, dW]$$

is transformed to a Wiener process with drift using the transformation

$$Y = \int^X \frac{1}{f(X')}\, dX'$$

– see the exercises.

Note that in our present application the possible values of X are in $(0, \infty)$ whereas the possible values of Y are in $(-\infty, \infty)$. In addition, if $X(0) = x_0$ and $Y(0) = y_0$, then $y_0 = \ln(x_0)$.

The transition density of a Wiener process with drift was given by Equation (12.15). If, in conjunction with that result, we use Equation (1.6) for the transformation of a density under a monotonic change of variable, we are immediately able to find the **transition probability density of the population size** X as

$$p(x, t\,|\,x_0) = \frac{1}{x\sqrt{2\pi\sigma^2 t}} \exp\left(\frac{-[\ln(x/x_0) - \bar{r}t]^2}{2\sigma^2 t}\right).$$

Now, we can see that the **probability** P_E **that the population eventually becomes extinct** can be estimated from

$$P_E = \lim_{t\to\infty} \Pr\{X(t) < \varepsilon\,|\,X(0) = x_0\}$$

where $\varepsilon > 0$ is arbitrarily small. Note that X can never reach exactly zero in this model – for we know that Y can never reach $-\infty$. However, an extremely small population size implies extinction in a continuous model.

It will be seen in the exercises that

$$P_E = \begin{cases} 0, & \text{if } \bar{r} > 0; \\ \frac{1}{2}, & \text{if } \bar{r} = 0; \\ 1, & \text{if } \bar{r} < 0. \end{cases}$$

Thus, if the mean growth rate is negative, the population will become extinct with certainty, regardless of either the initial population size or how great the fluctuations (σ) are in the growth rate. In this model, if a zero population growth policy is sustained on average, there is a 50% chance of long-term survival.

In addition, diffusion processes and stochastic differential equations have been usefully employed to model gene frequencies (Kimura, 1964; Tuckwell, 1976a; Watterson, 1979) – providing continuous approximations which are usually easier to work with than the Markov chain models we considered in Chapter 8, where additional references can be found. With these tools, important questions in the theory of evolution can be addressed.

Applications in financial modelling

As exemplified by Figure 11.2b, many quantities of financial or economic interest undergo random fluctuations. Needless to say, if the random component was absent, stock prices, for example, would be predictable so that there would be no risk involved in buying or selling shares. This would imply little possibility of profit and stock markets might cease to exist *per se.*

It is clear that a sound mathematical model for fluctuations in financial entities such as stock and commodity prices, exchange rates, etc. would be very useful as one could then make quantitative estimates for the probabilities of future values, expected profits, waiting times to reach certain levels, etc..

As early as 1900, the French mathematician Bachelier proposed the modelling of stock market fluctuations using a Wiener process. It was subsequently recognized that because the prices of shares could not be negative, a modification to the simple Wiener process was necessary. One solution proposed was the adoption of so-called *geometric Brownian motion* or the *geometric Wiener process* to represent certain financial entities. Thus $Y = \ln(X)$ should be a Wiener process with drift – so that $X = e^Y$ might represent a stock price. We then have $Y \in (-\infty, \infty)$ but $X \in (0, \infty)$.

Another assumption can be made that the expectation of a stock price grows in an exponential fashion (Samuelson, 1965). Thus if $X(t)$ is a stock price at time t, then

$$E[X(t) | X(0) = x_0] = x_0 e^{\mu t}, \tag{12.36}$$

where μ represents a constant growth rate.

To satisfy these requirements we need only put

$$\mathrm{d}X = [\mu\,\mathrm{d}t + \sigma\,\mathrm{d}W]X, \tag{12.37}$$

and interpret this as in Ito stochastic differential equation. That this gives the correct relation (12.36) will be verified in the exercises. The process X is called geometric Brownian motion and was analysed in the context of general growth models, and in particular tumour growth, by Smith and Tuckwell (1974).

In the economics and finance literature, Equation (12.37) is often written

$$\frac{\mathrm{d}X}{X} = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W.$$

This is an essential component of the Black–Scholes model for option prices (Black and Scholes, 1973) which is often used by financial analysts. This model, which provides a starting point for more elaborate models (Aase, 1988), leads to a formula called the *Black–Scholes option formula* for an option price in terms of variables such as term to maturity and risk-free interest rate. Wiener process models have been employed in the analysis of exchange rates also (Werner, 1993).

Practical considerations – simulation, numerical methods and parameter estimation

When employing stochastic differential equations as mathematical models of empirical phenomena, it is often worthwhile, and in many cases necessary, to perform computer simulations. The computer simulation of diffusion processes is not difficult (Tuckwell and Walsh, 1983b) and can be performed as follows.

Let us suppose that we wish to approximate solutions of the Ito equation

$$\mathrm{d}X = f(X,t)\,\mathrm{d}t + g(X,t)\,\mathrm{d}W.$$

We choose a time step Δt and call our simulation X^* with

$$X^*(k\Delta t) \simeq X(k\Delta t)$$

where $k = 0, 1, 2, \ldots$. We are given (or draw from a distribution) the value $X^*(0) = X(0)$. Then we put, for $k = 1, 2, \ldots$

$$\begin{aligned} X^*(k\Delta t) = X^*((k-1)\Delta t) &+ f(X^*(k-1)\Delta t, (k-1)\Delta t))\Delta t \\ &+ g(X^*((k-1)\Delta t), (k-1)\Delta t)\sqrt{\Delta t}N_k, \end{aligned} \tag{12.38}$$

where $\{N_k, k = 1, 2, \ldots\}$ is a sequence of i.i.d. standard normal random variables. The latter can be generated by the methods outlined in Chapter 5 – or simply by using a library random number generator. One performs a large number of runs or trials for the process and can easily determine estimates of quantities such as moments, distributions, and first passage or first exit times, etc. .

An alternative to (12.38) is to employ a sequence of independent Bernoulli random variables B_k (see page 2) instead of the N_k with

$$\Pr\{B_k = 1\} = 1/2 = \Pr\{B_k = -1\}, k = 1, 2, \ldots .$$

Then since $E[B_k] = 0$ and $\mathrm{Var}[B_k] = 1$, one may simply replace N_k in (12.38) by B_k. Since most computers have in their libraries a generator of uniformly distributed random variables (see Chapter 5), U_k, say, the sequence of Bernoulli variables can be easily obtained by putting $B_k = 1$ if $U_k < \frac{1}{2}$ or $B_k = -1$ if $U_k \geqslant \frac{1}{2}$. It has been found that when Bernoulli rather than normal random variables are employed, the simulations may be performed about three times faster (Tuckwell and Lansky, 1994). A detailed treatment of theoretical considerations on the numerical solutions of stochastic differential equations can be found in Kloeden and Platen (1992).

As mentioned earlier, another approach to the study of diffusion processes is to use the analytical method. This will usually involve solving a Kolmogorov or similar equation. Fortunately, even though the stochastic differential equation may be nonlinear, the Kolmogorov equation is always linear. It is then simply a matter of solving the latter using finite-difference approximations either by explicit or implicit methods (Ames, 1977). The analytical approach to first passage and first exit time problems is outlined in Siegert (1951), Tuckwell (1976b; 1981) and Tuckwell and Wan (1984).

Finally, assuming that one has a satisfactory mathematical representation, in the form of a stochastic differential equation, for an empirical process that one wishes to model, it is often desirable to estimate values of the various parameters which occur in the model using observed sample paths. Various methods are available, including maximum likelihood (Feigin, 1976) and quasi-likelihood (Heyde, 1993).

REFERENCES

Aase, K.K. (1988). Contingent claims valuation when the security price is a combination of an Ito process and a random point process. *Stoch. Proc. Appl.*, **28**, 185–220.

Ames, W.F. (1977). *Numerical Methods for Partial Differential Equations*. Nelson, London.

Arnold, L. (1974). *Stochastic Differential Equations*. Wiley, New York.

Bachelier, L. (1900). Theorie de la speculation. *Ann. Sci. Ecole. Norm. Sup.*, **3**, 21–86. English translation in Cootner, P. H. (ed.) (1964), *The Random Character of Stock Market Prices*, pp. 17–78. MIT Press, Cambridge.

Berkey, D.D. (1984). *Calculus*. Holt-Saunders, Philadelphia.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *J. Pol. Econ.*, **81**, 637–54.

Chandrasekhar, S. (1943). Dynamical friction. *Astrophys. J.*, **97**, 263–73.

Feigin, P.D. (1976). Maximum likelihood estimation for continuous-time stochastic processes. *Adv. Appl. Prob.*, **8**, 712–36.

Feller, W. (1952). The parabolic differential equations and the associated semi-groups of transformations. *Ann. Math.*, **55**, 468–519.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications II*. Wiley, New York.

Gihman, I.I. and Skorohod, A.V. (1972). *Stochastic Differential Equations*. Springer-Verlag, Berlin.

Gradshteyn, I.S. and Ryzhik, I.M. (1980). *Tables of Integrals Series and Products*. Academic Press, New York.

Heyde, C.C. (1993). Quasi-likelihood and general theory of inference for stochastic processes. *Lecture Notes. 7th International Summer School on Probability Theory and Mathematical Statistics* (eds A. Obretenov and V.T. Stefanov), pp. 122–52. Science Culture Technology Publishing, Singapore.

Hida, T. (1980). *Brownian Motion*. Springer-Verlag, Berlin.

Jaswinski, A.H. (1970). *Stochastic Process and Filtering Theory*. Academic Press, New York.

Kimura, M. (1964). Diffusion models in population genetics. *J. Appl. Prob.*, **1**, 177–232.

Kloeden, P. and Platen, E. (1992). *The Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, Berlin.

Liptser, R.S. and Shiryayev, A.N. (1977). *Statistics of Random Processes I. General Theory*. Springer-Verlag, New York.

McNeil, D.R. and Schach, S. (1973). Central limit analogues for Markov population processes. *J. Roy. Stat. Soc. B.*, **35**, 1–23.

Noble, B. and Daniel, J.W. (1977). *Applied Linear Algebra*. Prentice-Hall, Englewood Cliffs.

Oksendal, B. (1985). *Stochastic Differential Equations*. Springer-Verlag, Berlin.

Ricciardi, L.M. (1977). *Diffusion Processes and Related Topics in Biology*. Springer-Verlag, Berlin.

Samuelson, P.A. (1965). Rational theory of warrant pricing. *Indust. Manag. Rev.*, **6**, 13–31.

Shields, P.C. (1973). *Elementary Linear Algebra*. Worth, New York.

Siegert, A.J.F. (1951) On the first passage time probability problem. *Phys. Rev.* **81**, 617–23.

Smith, C.E. and Tuckwell, H.C. (1974). Some stochastic growth processes. In *Mathematical Problems in Biology Proceedings 1973*, pp. 211–25. Springer-Verlag, Berlin.

Stratonovich, R.L. (1963). *Topics in the Theory of Random Noise*. Gordon and Breach, New York.

Tuckwell, H.C. (1973). Diffusion Processes in Biology. *SM Thesis*. University of Chicago.

Tuckwell, H.C. (1974). A study of some diffusion models of population growth. *Theor. Pop. Biol.*, **5**, 345–57.

Tuckwell, H.C. (1976a). The effects of random selection on gene frequency. *Math. Biosci.*, **30**, 113–28.

Tuckwell, H.C. (1976b). On the first exit-time problem for temporally homogeneous Markov processes. *J. Appl. Prob.*, **13**, 39–48.

Tuckwell, H.C. (1981). Poisson processes in biology. In *Stochastic Nonlinear Systems*. pp. 162–71, Springer-Verlag, Berlin.

Tuckwell, H.C. (1987). Diffusion approximations to channel noise. *J. Theor. Biol.*, **127**, 427–38.