13

Diffusion processes

Summary. An elementary description of the Wiener process (Brownian motion) is presented, and used to motivate an account of diffusion processes based on the instantaneous mean and variance. This leads to the forward and backward equations for diffusions. First-passage probabilities of the Wiener process are explored using the reflection principle. Interpretations of absorbing and reflecting barriers for diffusions are presented. There is a brief account of excursions, and of the Brownian bridge. The Itô calculus is summarized, and used to construct a class of diffusions which are martingales. The theory of financial mathematics based on the Wiener process is described, including option pricing and the Black–Scholes formula. Finally, there is a discussion of the links between diffusions, harmonic functions, and potential theory.

13.1 Introduction

Random processes come in many types. For example, they may run in discrete time or continuous time, and their state spaces may also be discrete or continuous. In the main, we have so far considered processes which are *discrete* either in time or space; our purpose in this chapter is to approach the theory of processes indexed by continuous time and taking values in the real line $\mathbb{R}$. Many important examples belong to this category: meteorological data, communication systems with noise, molecular motion, and so on. In other important cases, such random processes provide useful approximations to the physical process in question: processes in population genetics or population evolution, for example.

The archetypal diffusion process is the Wiener process W of Example (9.6.13), a Gaussian process with stationary independent increments. Think about W as a description of the motion of a particle moving randomly but continuously about $\mathbb{R}$. There are various ways of *defining* the Wiener process, and each such definition has two components. First of all, we require a *distributional* property, such as that the finite-dimensional distributions are Gaussian, and so on. The second component, not explored in Chapter 9, is that the sample paths of the process $\{W(t;\omega) : t \geq 0\}$, thought of as random functions on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, are almost surely continuous. This assumption is important and natural, and of particular relevance when studying first passage times of the process.

Similar properties are required of a diffusion process, and we reserve the term 'diffusion' for a process $\{X(t) : t \geq 0\}$ having the strong Markov property and whose sample paths are almost surely continuous.

13.2 Brownian motion

Suppose we observe a container of water. The water may appear to be motionless, but this is an illusion. If we are able to approach the container so closely as to be able to distinguish individual molecules then we may perceive that each molecule enjoys a motion which is unceasing and without any apparent order. The disorder of this movement arises from the frequent occasions at which the molecule is repulsed by other molecules which are nearby at the time. A revolutionary microscope design enabled the Dutch scientist A. van Leeuwenhoek (1632–1723) to observe the apparently random motion of micro-organisms dubbed 'animalcules', but this motion was *biological* in cause. Credit for noticing that all sufficiently tiny particles enjoy a random movement of *physical* origin is usually given to the botanist R. Brown (1773–1858). Brown studied in 1827 the motion of tiny particles suspended in water, and he lent his name to the type of erratic movement thus observed. It was a major thrust of mathematics in the 20th century to model such phenomena, and this has led to the mathematical object termed the 'Wiener process', an informal motivation for which is presented in this section.

Brownian motion takes place in continuous time and continuous space. Our first attempt to model it might proceed by approximating to it by a discrete process such as a random walk. At any epoch of time the position of an observed particle is constrained to move about the points $\{(a\delta, b\delta, c\delta) : a, b, c = 0, \pm 1, \pm 2, \dots\}$ of a three-dimensional 'cubic' lattice in which the distance between neighbouring points is δ; the quantity δ is a fixed positive number which is very small. Suppose further that the particle performs a symmetric random walk on this lattice (see Problem (6.13.9) for the case $\delta = 1$) so that its position $\mathbf{S}_n$ after n jumps satisfies

$$\mathbb{P}(\mathbf{S}_{n+1} = \mathbf{S}_n + \delta\boldsymbol{\epsilon}) = \tfrac{1}{6} \quad \text{if} \quad \boldsymbol{\epsilon} = (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1).$$

Let us concentrate on the x coordinate of the particle, and write $\mathbf{S}_n = (S_n^1, S_n^2, S_n^3)$. Then

$$S_n^1 - S_0^1 = \sum_{i=1}^{n} X_i$$

as in Section 3.9, where $\{X_i\}$ is an independent identically distributed sequence with

$$\mathbb{P}(X_i = k\delta) = \begin{cases} \frac{1}{6} & \text{if } k = -1, \\ \frac{1}{6} & \text{if } k = +1, \\ \frac{2}{3} & \text{if } k = 0. \end{cases}$$

We are interested in the displacement $S_n^1 - S_0^1$ when n is large; the central limit theorem (5.10.4) tells us that the distribution of this displacement is approximately $N(0, \frac{1}{3}n\delta^2)$. Now suppose that the jumps of the random walk take place at time epochs $\tau, 2\tau, 3\tau, \dots$ where $\tau > 0$; τ is the time between jumps and is very small, implying that a very large number of jumps occur in any 'reasonable' time interval. Observe the particle after some time t (> 0)

has elapsed. By this time it has experienced $n = \lfloor t/\tau \rfloor$ jumps, and so its x coordinate $S^1(t)$ is such that $S^1(t) - S^1(0)$ is approximately $N(0, \frac{1}{3}t\delta^2/\tau)$. At this stage in the analysis we let the inter-point distance δ and the inter-jump time τ approach zero; in so doing we hope that the discrete random walk may approach some limit whose properties have something in common with the observed features of Brownian motion. We let $\delta \downarrow 0$ and $\tau \downarrow 0$ in such a way that $\frac{1}{3}\delta^2/\tau$ remains constant, since the variance of the distribution of $S^1(t) - S^1(0)$ fails to settle down to a non-trivial limit otherwise. Set

$$\text{(1)} \qquad \tfrac{1}{3}\delta^2/\tau = \sigma^2$$

where σ^2 is a positive constant, and pass to the limit to obtain that the distribution of $S^1(t) - S^1(0)$ approaches $N(0, \sigma^2 t)$. We can apply the same argument to the y coordinate and to the z coordinate of the particle to deduce that the particle's position $\mathbf{S}(t) = (S^1(t), S^2(t), S^3(t))$ at time t is such that the asymptotic distribution of the coordinates of the displacement $\mathbf{S}(t) - \mathbf{S}(0)$ is multivariate normal whenever $\delta, \tau \downarrow 0$, and (1) holds; furthermore, it is not too hard to see that $S^1(t)$, $S^2(t)$, and $S^3(t)$ are independent of each other.

We may guess from the asymptotic properties of this random walk that an adequate model for Brownian motion will involve a process $\mathbf{X} = \{\mathbf{X}(t) : t \geq 0\}$ taking values in $\mathbb{R}^3$ with a coordinate representation $\mathbf{X}(t) = (X^1(t), X^2(t), X^3(t))$ such that:

(a) $\mathbf{X}(0) = (0, 0, 0)$, say,

(b) X^1, X^2, and X^3 are independent and identically distributed processes,

(c) $X^1(s+t) - X^1(s)$ is $N(0, \sigma^2 t)$ for any $s, t \geq 0$,

(d) X^1 has *independent increments* in that $X^1(v) - X^1(u)$ and $X^1(t) - X^1(s)$ are independent whenever $u \leq v \leq s \leq t$.

We have not yet shown the existence of such a process $\mathbf{X}$; the foregoing argument only indicates certain plausible distributional properties without showing that they are attainable. However, properties (c) and (d) are not new to us and remind us of the Wiener process of Example (9.6.13); we deduce that such a process $\mathbf{X}$ indeed exists, and is given by $\mathbf{X}(t) = (W^1(t), W^2(t), W^3(t))$ where W^1, W^2, and W^3 are independent Wiener processes.

This conclusion is gratifying in that it demonstrates the existence of a random process which seems to enjoy at least some of the features of Brownian motion. A more detailed and technical analysis indicates some weak points of the Wiener model. This is beyond the scope of this text, and we able only to skim the surface of the main difficulty. For each ω in the sample space Ω, $\{\mathbf{X}(t; \omega) : t \geq 0\}$ is a sample path of the process along which the particle may move. It can be shown that, in some sense to be discussed in the next section,

(a) the sample paths are continuous functions of t,

(b) almost all sample paths are nowhere differentiable functions of t.

Property (a) is physically necessary, but (b) is a property which *cannot* be shared by the physical phenomenon which we are modelling, since mechanical considerations, such as Newton's laws, imply that only particles with zero mass can move along routes which are nowhere differentiable. As a model for the local movement (over a short time interval) of particles, the Wiener process is poor; over longer periods of time the properties of the Wiener process are indeed very similar to experimental results.

A popular improved model for the local behaviour of Brownian paths is the so-called Ornstein–Uhlenbeck process. We close this section with a short account of this. Roughly, it is founded on the assumption that the velocity of the particle (rather than its position) undergoes a random walk; the ensuing motion is damped by the frictional resistance of the fluid. The result is a 'velocity process' with continuous sample paths; their integrals represent the sample paths of the particle itself. Think of the motion in one dimension as before, and write V_n for the velocity of the particle after the nth jump. At the next jump the change $V_{n+1} - V_n$ in the velocity is assumed to have two contributions: the frictional resistance to motion, and some random fluctuation owing to collisions with other particles. We shall assume that the former damping effect is directly proportional to V_n, so that $V_{n+1} = V_n + X_{n+1}$; this is the so-called *Langevin equation.* We require that:

$$\mathbb{E}(X_{n+1} \mid V_n) = -\beta V_n \qquad : \text{frictional effect,}$$

$$\operatorname{var}(X_{n+1} \mid V_n) = \sigma^2 \qquad : \text{collision effect,}$$

where β and σ^2 are constants. The sequence $\{V_n\}$ is no longer a random walk on some regular grid of points, but it can be shown that the distributions converge as before, after suitable passage to the limit. Furthermore, there exists a process $V = \{V(t) : t \geq 0\}$ with the corresponding distributional properties, and whose sample paths turn out to be almost surely continuous. These sample paths do not represent possible routes of the particle, but rather describe the development of its velocity as time passes. The possible paths of the particle through the space which it inhabits are found by integrating the sample paths of V with respect to time. The resulting paths are almost surely continuously differentiable functions of time.

13.3 Diffusion processes

We say that a particle is 'diffusing' about a space $\mathbb{R}^n$ whenever it experiences erratic and disordered motion through the space; for example, we may speak of radioactive particles diffusing through the atmosphere, or even of a rumour diffusing through a population. For the moment, we restrict our attention to one-dimensional diffusions, for which the position of the observed particle at any time is a point on the real line; similar arguments will hold for higher dimensions. Our first diffusion model is the Wiener process.

(1) Definition. A Wiener process $W = \{W(t) : t \geq 0\}$, starting from $W(0) = w$, say, is a real-valued Gaussian process such that:

(a) W has independent increments (see Lemma (9.6.16)),

(b) $W(s+t) - W(s)$ is distributed as $N(0, \sigma^2 t)$ for all $s, t \geq 0$ where σ^2 is a positive constant,

(c) the sample paths of W are continuous.

Clearly (1a) and (1b) specify the finite-dimensional distributions (fdds) of a Wiener process W, and the argument of Theorem (9.6.1) shows there exists a Gaussian process with these fdds. In agreement with Example (9.6.13), the autocovariance function of W is given by

$$\begin{aligned} c(s,t) &= \mathbb{E}\big([W(s) - W(0)][W(t) - W(0)]\big) \\ &= \mathbb{E}\big([W(s) - W(0)]^2 + [W(s) - W(0)][W(t) - W(s)]\big) \\ &= \sigma^2 s + 0 \qquad \text{if } \ 0 \leq s \leq t, \end{aligned}$$

which is to say that

(2) $$c(s,t) = \sigma^2 \min\{s,t\} \quad \text{for all } s, t \geq 0.$$

The process W is called a *standard* Wiener process if $\sigma^2 = 1$ and $W(0) = 0$. If W is non-standard, then $W_1(t) = (W(t) - W(0))/\sigma$ is standard. The process W is said to have 'stationary' independent increments since the distribution of $W(s+t) - W(s)$ depends on t alone. A simple application of Theorem (9.6.7) shows that W is a Markov process.

The Wiener process W can be used to model the apparently random displacement of Brownian motion in any chosen direction. For this reason, W is sometimes called 'Brownian motion', a term which *we* reserve to describe the motivating physical phenomenon.

Does the Wiener process exist? That is to say, does there exist a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Gaussian process W thereon, satisfying (1a, b, c)? The answer to this non-trivial question is of course in the affirmative, and we defer to the end of this section an explicit construction of such a process. The difficulty lies not in satisfying the distributional properties (1a, b) but in showing that this may achieved with *continuous* sample paths.

Roughly speaking, there are two types of statement to be made about diffusion processes in general, and the Wiener process in particular. The first deals with sample path properties, and the second with distributional properties.

Figure 13.1 is a diagram of a typical sample path. Certain distributional properties of continuity are immediate. For example, W is 'continuous in mean square' in that

$$\mathbb{E}\big([W(s+t) - W(s)]^2\big) \to 0 \quad \text{as} \quad t \to 0;$$

this follows easily from equation (2).

Let us turn our attention to the distributions of a standard Wiener process W. Suppose we are given that $W(s) = x$, say, where $s \geq 0$ and $x \in \mathbb{R}$. Conditional on this, $W(t)$ is distributed as $N(x, t-s)$ for $t \geq s$, which is to say that the conditional distribution function

$$F(t, y \mid s, x) = \mathbb{P}\big(W(t) \leq y \bigm| W(s) = x\big)$$

has density function

(3) $$f(t, y \mid s, x) = \frac{\partial}{\partial y} F(t, y \mid s, x)$$

which is given by

(4) $$f(t, y \mid s, x) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right), \quad -\infty < y < \infty.$$

This is a function of four variables, but just grit your teeth. It is easy to check that f is the solution of the following differential equations.

(5) ***Forward diffusion equation:*** $$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}.$$

(6) ***Backward diffusion equation:*** $$\frac{\partial f}{\partial s} = -\frac{1}{2}\frac{\partial^2 f}{\partial x^2}.$$

We ought to specify the boundary conditions for these equations, but we avoid this at the moment. Subject to certain conditions, (4) is the unique density function which solves (5) or

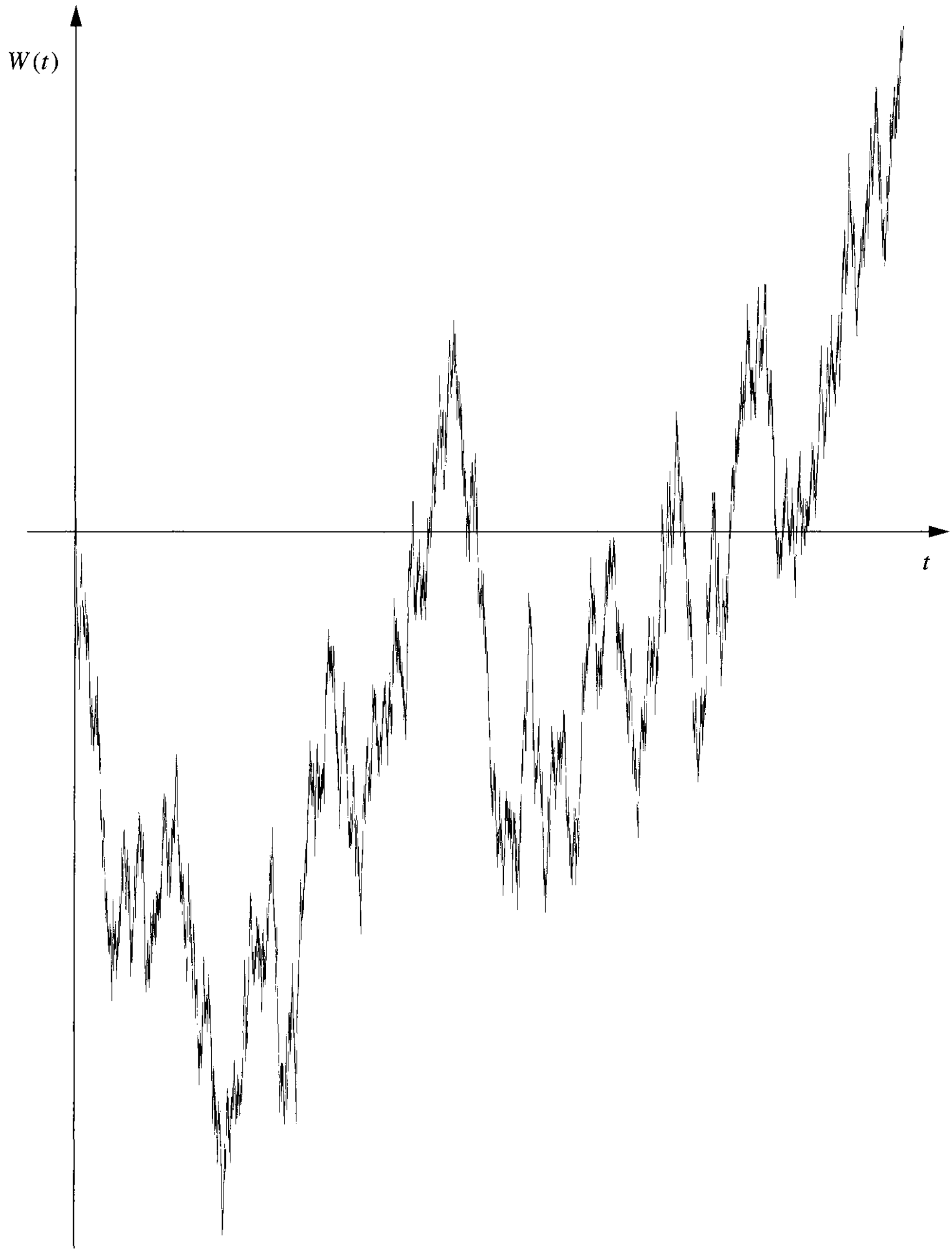

Figure 13.1. A typical realization of a Wiener process W. This is a scale drawing of a sample path of W over the time interval $[0, 1]$. Note that the path is continuous but very spiky. This picture indicates the general features of the path only; the dense black portions indicate superimposed fluctuations which are too fine for this method of description. Any magnification of part of the path would reveal fluctuations of order comparable to those of the original path. This picture was drawn with the aid of a computer, using nearly 90,000 steps of a symmetric random walk and the scaling method of Section 13.2.

(6). There is a good reason why (5) and (6) are called the *forward* and *backward* equations. Remember that W is a Markov process, and use arguments similar to those of Sections 6.8 and 6.9. Equation (5) is obtained by conditioning $W(t+h)$ on the value of $W(t)$ and letting $h \downarrow 0$; (6) is obtained by conditioning $W(t)$ on the value of $W(s+h)$ and letting $h \downarrow 0$. You are treading in Einstein's footprints as you perform these calculations. The derivatives in (5) and (6) have coefficients which do not depend on x, y, s, t; this reflects the fact that the Wiener process is homogeneous in space and time, in that:

(a) the increment $W(t) - W(s)$ is independent of $W(s)$ for all $t \geq s$,

(b) the increments are stationary in time.

Next we turn our attention to diffusion processes which *lack* this homogeneity.

The Wiener process is a Markov process, and the Markov property provides a method for deriving the forward and backward equations. There are other Markov diffusion processes to which this method may be applied in order to obtain similar forward and backward equations; the coefficients in these equations will *not* generally be constant. The existence of such processes can be demonstrated rigorously, but here we explore their distributions only. Let $D = \{D(t) : t \geq 0\}$ denote a diffusion process. In addition to requiring that D has (almost surely) continuous sample paths, we need to impose some conditions on the transitions of D in order to derive its diffusion equations; these conditions take the form of specifying the mean and variance of increments $D(t+h) - D(t)$ of the process over small time intervals $(t, t+h)$. Suppose that there exist functions $a(t,x)$, $b(t,x)$ such that:

$$\begin{aligned}
\mathbb{P}\big(|D(t+h) - D(t)| > \epsilon \,\big|\, D(t) = x\big) &= \mathrm{o}(h) \quad \text{for all } \epsilon > 0, \\
\mathbb{E}\big(D(t+h) - D(t) \,\big|\, D(t) = x\big) &= a(t,x)h + \mathrm{o}(h), \\
\mathbb{E}\big([D(t+h) - D(t)]^2 \,\big|\, D(t) = x\big) &= b(t,x)h + \mathrm{o}(h).
\end{aligned}$$

The functions a and b are called the 'instantaneous mean' (or 'drift') and 'instantaneous variance' of D respectively. Subject to certain other technical conditions (see Feller 1971, pp. 332–335), if $s \leq t$ then the conditional density function of $D(t)$ given $D(s) = x$,

$$f(t, y \mid s, x) = \frac{\partial}{\partial y}\mathbb{P}\big(D(t) \leq y \,\big|\, D(s) = x\big),$$

satisfies the following partial differential equations.

(7) *Forward equation*: $$\frac{\partial f}{\partial t} = -\frac{\partial f}{\partial y}[a(t,y)f] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[b(t,y)f].$$

(8) *Backward equation*: $$\frac{\partial f}{\partial s} = -a(s,x)\frac{\partial f}{\partial x} - \frac{1}{2}b(s,x)\frac{\partial^2 f}{\partial x^2}.$$

It is a noteworthy fact that the density function f is specified as soon as the instantaneous mean a and variance b are known; we need no further information about the distribution of a typical increment. This is very convenient for many applications, since a and b are often specified in a natural manner by the physical description of the process.

(9) Example. The Wiener process. If increments of any given length have zero means and constant variances then

$$a(t,x) = 0, \quad b(t,x) = \sigma^2,$$

for some $\sigma^2 > 0$. Equations (7) and (8) are of the form of (5) and (6) with the inclusion of a factor σ^2. ●

(10) Example. The Wiener process with drift. Suppose a particle undergoes a type of one-dimensional Brownian motion, in which it experiences a drift at constant rate in some particular direction. That is to say,

$$a(t,x) = m, \quad b(t,x) = \sigma^2,$$

for some drift rate m and constant σ^2. The forward diffusion equation becomes

$$\frac{\partial f}{\partial t} = -m\frac{\partial f}{\partial y} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial y^2}$$

and it follows that the corresponding diffusion process D is such that $D(t) = \sigma W(t) + mt$ where W is a standard Wiener process. ●

(11) The Ornstein–Uhlenbeck process. Recall the discussion of this process at the end of Section 13.2. It experiences a drift towards the origin of magnitude proportional to its dislacement. That is to say,

$$a(t,x) = -\beta x, \quad b(t,x) = \sigma^2,$$

and the forward equation is

$$\frac{\partial f}{\partial t} = \beta\frac{\partial}{\partial y}(yf) + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial y^2}.$$

See Problem (13.8.4) for one solution of this equation. ●

(12) Example. Diffusion approximation to the branching process. Diffusion models are sometimes useful as continuous approximations to discrete processes. In Section 13.2 we saw that the Wiener process approximates to the random walk under certain circumstances; here is another example of such an approximation. Let $\{Z_n\}$ be the size of the nth generation of a branching process, with $Z_0 = 1$ and such that $\mathbb{E}(Z_1) = \mu$ and $\mathrm{var}(Z_1) = \sigma^2$. A typical increment $Z_{n+1} - Z_n$ has mean and variance given by

$$\mathbb{E}(Z_{n+1} - Z_n \mid Z_n = x) = (\mu - 1)x,$$
$$\mathrm{var}(Z_{n+1} - Z_n \mid Z_n = x) = \sigma^2 x;$$

these are directly proportional to the size of Z_n. Now, suppose that the time intervals between successive generations become shorter and shorter, but that the means and variances of the increments retain this proportionality; of course, we need to abandon the condition that the process be integer-valued. This suggests a diffusion model as an approximation to the branching process, with instantaneous mean and variance given by

$$a(t,x) = ax, \quad b(t,x) = bx,$$

and the forward equation of such a process is

$$\frac{\partial f}{\partial t} = -a\frac{\partial}{\partial y}(yf) + \frac{1}{2}b\frac{\partial^2}{\partial y^2}(yf). \tag{13}$$

Subject to appropriate boundary conditions, this equation has a unique solution; this may be found by taking Laplace transforms of (13) in order to find the moment generating function of the value of the diffusion process at time t. ●

(14) Example. A branching diffusion process. The next example is a modification of the process of (6.12.15) which modelled the distribution in space of the members of a branching process. Read the first paragraph of (6.12.15) again before proceeding with this example. It is often the case that the members of a population move around the space which they inhabit during their lifetimes. With this in mind we introduce a modification into the process of (6.12.15). Suppose a typical individual is born at time s and at position x. We suppose that this individual moves about $\mathbb{R}$ until its lifetime T is finished, at which point it dies and divides, leaving its offspring at the position at which it dies. We suppose further that it moves according to a standard Wiener process W, so that it is at position $x + W(t)$ at time $s + t$ whenever $0 \le t \le T$. We assume that each individual moves independently of the positions of all the other individuals. We retain the notation of (6.12.15) whenever it is suitable, writing N for the number of offspring of the initial individual, W for the process describing its motion, and T for its lifetime. This individual dies at the point $W(T)$.

We no longer seek complete information about the distribution of the individuals around the space, but restrict ourselves to a less demanding task. It is natural to wonder about the rate at which members of the population move away from the place of birth of the founding member. Let $M(t)$ denote the position of the individual who is furthest right from the origin at time t. That is,

$$M(t) = \sup\{x : Z_1(t, x) > 0\}$$

where $Z_1(t, x)$ is the number of living individuals at time t who are positioned at points in the interval $[x, \infty)$. We shall study the distribution function of $M(t)$

$$F(t, x) = \mathbb{P}\big(M(t) \le x\big),$$

and we proceed roughly as before, noting that

$$F(t, x) = \int_0^\infty \mathbb{P}\big(M(t) \le x \,\big|\, T = s\big) f_T(s)\,ds \tag{15}$$

where f_T is the density function of T. However,

$$\mathbb{P}\big(M(t) \le x \,\big|\, T = s\big) = \mathbb{P}\big(W(t) \le x\big) \quad \text{if} \quad s > t,$$

whilst, if $s \le t$, use of conditional probabilities gives

$$\begin{aligned}&\mathbb{P}\big(M(t) \le x \,\big|\, T = s\big)\\ &\qquad = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \mathbb{P}\big(M(t) \le x \,\big|\, T = s,\ N = n,\ W(s) = w\big)\mathbb{P}(N = n) f_{W(s)}(w)\,dw\end{aligned}$$

where $f_{W(s)}$ is the density function of $W(s)$. However, if $s \le t$, then

$$\mathbb{P}\big(M(t) \le x \,\big|\, T = s,\ N = n,\ W(s) = w\big) = \big[\mathbb{P}\big(M(t-s) \le x - w\big)\big]^n,$$

and so (15) becomes

(16)
$$F(t,x) = \int_{s=0}^{t}\int_{w=-\infty}^{\infty} G_N[F(t-s, x-w)] f_{W(s)}(w) f_T(s)\, dw\, ds + \mathbb{P}\big(W(t) \le x\big)\int_t^{\infty} f_T(s)\, ds.$$

We consider here only the Markovian case when T is exponentially distributed, so that

$$f_T(s) = \mu e^{-\mu s} \quad \text{for} \quad s \ge 0.$$

Multiply throughout (16) by $e^{\mu t}$, substitute $t - s = u$ and $x - w = v$ within the integral, and differentiate with respect to t to obtain

$$\begin{aligned} e^{\mu t}\left(\mu F + \frac{\partial F}{\partial t}\right) = \mu \int_{-\infty}^{\infty} G_N(F(t,v)) f_{W(0)}(x-v) e^{\mu t}\, dv \\ + \mu \int_{u=0}^{t}\int_{v=-\infty}^{\infty} G_N(F(u,v))\left(\frac{\partial}{\partial t} f_{W(t-u)}(x-v)\right) e^{\mu u}\, dv\, du \\ + \frac{\partial}{\partial t}\mathbb{P}\big(W(t) \le x\big). \end{aligned}$$

Now differentiate the same equation twice with respect to x, remembering that $f_{W(s)}(w)$ satisfies the diffusion equations and that $\delta(v) = f_{W(0)}(x - v)$ needs to be interpreted as the Dirac δ function at the point $v = x$ to find that

(17)
$$\mu F + \frac{\partial F}{\partial t} = \mu G_N(F) + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}.$$

Many eminent mathematicians have studied this equation; for example, Kolmogorov and Fisher were concerned with it in connection with the distribution of gene frequencies. It is difficult to extract precise information from (17). One approach is to look for solutions of the form $F(t, x) = \psi(x - ct)$ for some constant c to obtain the following second-order ordinary differential equation for ψ:

(18)
$$\psi'' + 2c\psi' + 2\mu H(\psi) = 0$$

where $H(\psi) = G_N(\psi) - \psi$. Solutions to (18) yield information about the asymptotic distribution of the so-called 'advancing wave' of the members of the process. ●

Finally in this section, we show that Wiener processes exist. The difficulty is the requirement that sample paths be continuous. Certainly there exist Gaussian processes with independent normally distributed increments as required in (1a, b), but there is no reason in general why such a process should have continuous sample paths. We shall show next that one may construct such a Gaussian process with this extra property of continuity.

Let us restrict ourselves for the moment to the time interval $[0, 1]$, and suppose that X is a Gaussian process on $[0, 1]$ with independent increments, such that $X(0) = 0$, and $X(s+t) - X(s)$ is $N(0, t)$ for $s, t \geq 0$. We shall concentrate on a certain countable subset Q of $[0, 1]$, namely the set of 'dyadic rationals', being the set of points of the form $m2^{-n}$ for some $n \geq 1$ and $0 \leq m \leq 2^n$. For each $n \geq 1$, we define the process $X_n(t)$ by $X_n(t) = X(t)$ if $t = m2^{-n}$ for some integer m, and by linear interpolation otherwise; that is to say,

$$X_n(t) = X(m2^{-n}) + 2^n(t - m2^{-n})\big[X((m+1)2^{-n}) - X(m2^{-n})\big]$$

if $m2^{-n} < t < (m+1)2^{-n}$. Thus X_n is a piecewise-linear and continuous function comprising 2^n line segments. Think of X_{n+1} as being obtained from X_n by repositioning the centres of these line segments by amounts which are independent and normally distributed. It is clear that

(19) $$X_n(t) \to X(t) \quad \text{for} \quad t \in Q,$$

since, if $t \in Q$, then $X_n(t) = X(t)$ for all large n. The first step is to show that the convergence in (19) is (almost surely) uniform on Q, since this will imply that the limit function X is (almost surely) continuous on Q. Now

(20) $$X_n(t) = \sum_{j=1}^{n} Z_j(t)$$

where $Z_j(t) = X_j(t) - X_{j-1}(t)$ and $X_0(t) = 0$. This series representation for X_n converges uniformly on Q if

(21) $$\sum_{j=1}^{\infty} \sup_{t \in Q} |Z_j(t)| < \infty.$$

We note that $Z_j(t) = 0$ for values of t having the form $m2^{-j}$ where m is even. It may be seen by drawing a diagram that

$$\sup_{t \in Q} |Z_j(t)| = \max\big\{|Z_j(m2^{-j})| : m = 1, 3, \dots, 2^j - 1\big\}$$

and therefore

(22) $$\mathbb{P}\left(\sup_{t \in Q} |Z_j(t)| > x\right) \leq \sum_{m \text{ odd}} \mathbb{P}\big(|Z_j(m2^{-j})| > x\big).$$

Now

$$\begin{aligned} Z_j(2^{-j}) &= X(2^{-j}) - \tfrac{1}{2}\big[X(0) + X(2^{-j+1})\big] \\ &= \tfrac{1}{2}\big[X(2^{-j}) - X(0)\big] - \tfrac{1}{2}\big[X(2^{-j+1}) - X(2^{-j})\big], \end{aligned}$$

and therefore $\mathbb{E}Z_j(2^{-j}) = 0$ and, using the independence of increments, $\text{var}(Z_j(2^{-j})) = 2^{-j-1}$; a similar calculation is valid for $Z_j(m2^{-j})$ for $m = 1, 3, \dots, 2^j - 1$. It follows by the bound in Exercise (4.4.8) on the tail of the normal distribution that, for all such m,

$$\mathbb{P}\big(|Z_j(m2^{-j})| > x\big) \leq \frac{1}{x2^{j/2}} e^{-x^2 2^j}, \quad x > 0.$$

Setting $x = c\sqrt{j2^{-j}\log 2}$, we obtain from (22) that

$$\mathbb{P}\left(\sup_{t\in Q}|Z_j(t)| > x\right) \le 2^{j-1}\frac{2^{-c^2 j}}{c\sqrt{j\log 2}}.$$

Choosing $c > 1$, the last term is summable in j, implying by the Borel–Cantelli lemma (7.3.10a) that

$$\sup_{t\in Q}|Z_j(t)| > c\sqrt{\frac{j\log 2}{2^j}}$$

for only finitely many values of j (almost surely). Hence

$$\sum_j \sup_{t\in Q}|Z_j(t)| < \infty \quad \text{almost surely,}$$

and the argument prior to (21) yields that X is (almost surely) continuous on Q.

We have proved that X has (almost surely) continuous sample paths on the set of dyadic rationals; a similar argument is valid for other countable dense subsets of $[0, 1]$. It is quite another thing for X to be continuous on the entire interval $[0, 1]$, and actually this need not be the case. We can, however, extend X by continuity from the dyadic rationals to the whole of $[0, 1]$: for $t \in [0, 1]$, define

$$Y(t) = \lim_{\substack{s\to t\\ s\in Q}} X(s),$$

the limit being taken as s approaches t through the dyadic rationals. Such a limit exists almost surely for all t since X is almost surely continuous on Q. It is not difficult to check that the extended process Y is indeed a Gaussian process with covariance function $\text{cov}(Y(s), Y(t)) = \min\{s, t\}$, and, most important, the sample paths of Y are (almost surely) continuous.

Finally we remove the 'almost surely' from the last conclusion. Let Ω' be the subset of the sample space Ω containing all ω for which the corresponding path of Y is continuous on $\mathbb{R}$. We now restrict ourselves to the smaller sample space Ω', with its induced σ-field and probability measure. Since $\mathbb{P}(\Omega') = 1$, this change is invisible in all calculations of probabilities. Conditions (1a) and (1b) remain valid in the restricted space.

This completes the proof of the existence of a Wiener process on $[0, 1]$. A similar argument can be made to work on the time interval $[0, \infty)$, but it is easier either: (a) to patch together continuous Wiener processes on $[n, n+1]$ for $n = 0, 1, \dots$, or (b) to use the result of Problem (9.7.18c).

Exercises for Section 13.3

1. Let $X = \{X(t) : t \ge 0\}$ be a simple birth–death process with parameters $\lambda_n = n\lambda$ and $\mu_n = n\mu$. Suggest a diffusion approximation to X.

2. Bartlett's equation. Let D be a diffusion with instantaneous mean and variance $a(t, x)$ and $b(t, x)$, and let $M(t, \theta) = \mathbb{E}(e^{\theta D(t)})$, the moment generating function of $D(t)$. Use the forward diffusion equation to derive *Bartlett's equation*:

$$\frac{\partial M}{\partial t} = \theta a\left(t, \frac{\partial}{\partial\theta}\right)M + \frac{1}{2}\theta^2 b\left(t, \frac{\partial}{\partial\theta}\right)M$$

where we interpret

$$g\left(t, \frac{\partial}{\partial \theta}\right) M = \sum_n \gamma_n(t) \frac{\partial^n M}{\partial \theta^n}$$

if $g(t, x) = \sum_{n=0}^{\infty} \gamma_n(t) x^n$.

3. Write down Bartlett's equation in the case of the Wiener process D having drift m and instantaneous variance 1, and solve it subject to the boundary condition $D(0) = 0$.

4. Write down Bartlett's equation in the case of an Ornstein–Uhlenbeck process D having instantaneous mean $a(t, x) = -x$ and variance $b(t, x) = 1$, and solve it subject to the boundary condition $D(0) = 0$.

5. Bessel process. If $W_1(t)$, $W_2(t)$, $W_3(t)$ are independent Wiener processes, then $R(t)$ defined as $R^2 = W_1^2 + W_2^2 + W_3^2$ is the three-dimensional *Bessel process*. Show that R is a Markov process. Is this result true in a general number n of dimensions?

6. Show that the transition density for the Bessel process defined in Exercise (5) is

$$\begin{aligned} f(t, y \mid s, x) &= \frac{\partial}{\partial y} \mathbb{P}\big(R(t) \le y \,\big|\, R(s) = x\big) \\ &= \frac{y/x}{\sqrt{2\pi(t-s)}} \left\{ \exp\left(-\frac{(y-x)^2}{2(t-s)}\right) - \exp\left(-\frac{(y+x)^2}{2(t-s)}\right) \right\}. \end{aligned}$$

7. If W is a Wiener process and the function $g : \mathbb{R} \to \mathbb{R}$ is continuous and strictly monotone, show that $g(W)$ is a continuous Markov process.

8. Let W be a Wiener process. Which of the following define martingales?
(a) $e^{\sigma W(t)}$, (b) $cW(t/c^2)$, (c) $tW(t) - \int_0^t W(s)\,ds$.

9. Exponential martingale, geometric Brownian motion. Let W be a standard Wiener process and define $S(t) = e^{at + bW(t)}$. Show that:
(a) S is a Markov process,
(b) S is a martingale (with respect to the filtration generated by W) if and only if $a + \frac{1}{2}b^2 = 0$, and in this case $\mathbb{E}(S(t)) = 1$.

10. Find the transition density for the Markov process of Exercise (9a).

13.4 First passage times

We have often been interested in the time which elapses before a Markov chain visits a specified state for the first time, and we continue this chapter with an account of some of the corresponding problems for a diffusion process.

Consider first a standard Wiener process W. The process W_1 given by

(1) $$W_1(t) = W(t+T) - W(T), \quad t \ge 0,$$

is a standard Wiener process for any fixed value of T and, conditional on $W(T)$, W_1 is independent of $\{W(s) : s < T\}$; the Poisson process enjoys a similar property, which in Section 6.8 we called the 'weak Markov property'. It is a very important and useful fact that this holds even when T is a random variable, so long as T is a stopping time for W. We encountered stopping times in the context of continuous-time martingales in Section 12.7.

(2) Definition. Let $\mathcal{F}_t$ be the smallest σ-field with respect to which $W(s)$ is measurable for each $s \le t$. The random variable T is called a **stopping time** for W if $\{T \le t\} \in \mathcal{F}_t$ for all t.

We say that W has the 'strong Markov property' in that this independence holds for all stopping times T. Why not try to prove this? Here, we make use of the strong Markov property for certain particular stopping times T.

(3) Definition. The **first passage time** $T(x)$ to the point $x \in \mathbb{R}$ is given by

$$T(x) = \inf\{t : W(t) = x\}.$$

The continuity of sample paths is essential in order that this definition make sense: a Wiener process cannot jump over the value x, but must pass through it. The proof of the following lemma is omitted.

(4) Lemma. *The random variable $T(x)$ is a stopping time for W.*

(5) Theorem. *The random variable $T(x)$ has density function*

$$f_{T(x)}(t) = \frac{|x|}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right), \quad t \ge 0.$$

Clearly $T(x)$ and $T(-x)$ are identically distributed. For the case when $x = 1$ we encountered this density function and its moment generating function in Problems (5.12.18) and (5.12.19); it is easy to deduce that $T(x)$ has the same distribution as Z^{-2} where Z is $N(0, x^{-2})$. In advance of giving the proof of Theorem (5), here is a result about the size of the maximum of a Wiener process.

(6) Theorem. *The random variable $M(t) = \max\{W(s) : 0 \le s \le t\}$ has the same distribution as $|W(t)|$. Thus $M(t)$ has density function*

$$f_{M(t)}(m) = \sqrt{\frac{2}{\pi t}} \exp\left(-\frac{m^2}{2t}\right), \quad m \ge 0.$$

You should draw your own diagrams to illustrate the translations and reflections used in the proofs of this section.

Proof of (6). Suppose $m > 0$, and observe that

$$(7) \qquad T(m) \le t \quad \text{if and only if} \quad M(t) \ge m.$$

Then

$$\mathbb{P}\big(M(t) \ge m\big) = \mathbb{P}\big(M(t) \ge m,\ W(t) - m \ge 0\big) + \mathbb{P}\big(M(t) \ge m,\ W(t) - m < 0\big).$$

However, by (7),

$$\begin{aligned}\mathbb{P}\big(M(t) \ge m,\ W(t) - m < 0\big) &= \mathbb{P}\big(W(t) - W(T(m)) < 0 \mid T(m) \le t\big)\mathbb{P}\big(T(m) \le t\big)\\ &= \mathbb{P}\big(W(t) - W(T(m)) \ge 0 \mid T(m) \le t\big)\mathbb{P}\big(T(m) \le t\big)\\ &= \mathbb{P}\big(M(t) \ge m,\ W(t) - m \ge 0\big)\end{aligned}$$

since $W(t) - W(T(m))$ is symmetric whenever $t \geq T(m)$ by the strong Markov property; we have used sample path continuity here, and more specifically that $\mathbb{P}(W(T(m)) = m) = 1$. Thus

$$\mathbb{P}\big(M(t) \geq m\big) = 2\mathbb{P}\big(M(t) \geq m,\ W(t) \geq m\big) = 2\mathbb{P}\big(W(t) \geq m\big)$$

since $W(t) \leq M(t)$. Hence $\mathbb{P}(M(t) \geq m) = \mathbb{P}(|W(t)| \geq m)$ and the theorem is proved on noting that $|W(t)|$ is the absolute value of an $N(0, t)$ variable. ■

Proof of (5). This follows immediately from (7), since if $x > 0$ then

$$\begin{aligned}\mathbb{P}\big(T(x) \leq t\big) &= \mathbb{P}\big(M(t) \geq x\big) = \mathbb{P}\big(|W(t)| \geq x\big)\\ &= \sqrt{\frac{2}{\pi t}} \int_x^\infty \exp\left(-\frac{m^2}{2t}\right) dm\\ &= \int_0^t \frac{|x|}{\sqrt{2\pi y^3}} \exp\left(-\frac{x^2}{2y}\right) dy\end{aligned}$$

by the substitution $y = x^2 t/m^2$. ■

We are now in a position to derive some famous results about the times at which W returns to its starting point, the origin. We say that 'W has a zero at time t' if $W(t) = 0$, and we write $\cos^{-1}$ for the inverse trigonometric function, sometimes written arc cos.

(8) Theorem. *Suppose $0 \leq t_0 < t_1$. The probability that a standard Wiener process W has a zero in the time interval (t_0, t_1), is $(2/\pi)\cos^{-1}\sqrt{t_0/t_1}$.*

Proof. Let $E(u, v)$ denote the event

$$E(u, v) = \{W(t) = 0 \text{ for some } t \in (u, v)\}.$$

Condition on $W(t_0)$ to obtain

$$\begin{aligned}\mathbb{P}(E(t_0, t_1)) &= \int_{-\infty}^{\infty} \mathbb{P}\big(E(t_0, t_1) \mid W(t_0) = w\big) f_0(w)\, dw\\ &= 2\int_{-\infty}^{0} \mathbb{P}\big(E(t_0, t_1) \mid W(t_0) = w\big) f_0(w)\, dw\end{aligned}$$

by the symmetry of W, where f_0 is the density function of $W(t_0)$. However, if $a > 0$,

$$\mathbb{P}\big(E(t_0, t_1) \mid W(t_0) = -a\big) = \mathbb{P}\big(T(a) < t_1 - t_0 \mid W(0) = 0\big)$$

by the homogeneity of W in time and space. Use (5) to obtain that

$$\begin{aligned}\mathbb{P}(E(t_0, t_1)) &= 2\int_{a=0}^{\infty}\int_{t=0}^{t_1-t_0} f_{T(a)}(t) f_0(-a)\, dt\, da\\ &= \frac{1}{\pi\sqrt{t_0}} \int_{t=0}^{t_1-t_0} t^{-\frac{3}{2}} \int_{a=0}^{\infty} a \exp\left[-\frac{1}{2}a^2\left(\frac{t+t_0}{t t_0}\right)\right] da\, dt\\ &= \frac{\sqrt{t_0}}{\pi}\int_0^{t_1-t_0} \frac{dt}{(t+t_0)\sqrt{t}}\\ &= \frac{2}{\pi}\tan^{-1}\sqrt{\frac{t_1}{t_0} - 1} \qquad \text{by the substitution } t = t_0 s^2\\ &= \frac{2}{\pi}\cos^{-1}\sqrt{t_0/t_1} \qquad \text{as required.}\end{aligned}$$

■

The result of (8) indicates some remarkable properties of the sample paths of W. Set $t_0 = 0$ to obtain

$$\mathbb{P}\big(\text{there exists a zero in } (0, t) \,\big|\, W(0) = 0\big) = 1 \quad \text{for all } t > 0,$$

and it follows that

$$T(0) = \inf\{t > 0 : W(t) = 0\}$$

satisfies $T(0) = 0$ almost surely. A deeper analysis shows that, with probability 1, W has infinitely many zeros in any non-empty time interval $[0, t]$; it is no wonder that W has non-differentiable sample paths! The set $Z = \{t : W(t) = 0\}$ of zeros of W is rather a large set; in fact it turns out that Z has Hausdorff dimension $\frac{1}{2}$ (see Mandelbrot (1983) for an introduction to fractional dimensionality).

The proofs of Theorems (5), (6), and (8) have relied heavily upon certain symmetries of the Wiener process; these are similar to the symmetries of the random walk of Section 3.10. Other diffusions may not have these symmetries, and we may need other techniques for investigating their first passage times. We illustrate this point by a glance at the Wiener process with drift. Let $D = \{D(t) : t \geq 0\}$ be a diffusion process with instantaneous mean and variance given by

$$a(t, x) = m, \qquad b(t, x) = 1,$$

where m is a constant. It is easy to check that, if $D(0) = 0$, then $D(t)$ is distributed as $N(mt, t)$. It is not so easy to find the distributions of the sizes of the maxima of D, and we take this opportunity to display the usefulness of martingales and optional stopping.

(9) Theorem. *Let* $U(t) = e^{-2mD(t)}$. *Then* $U = \{U(t) : t \geq 0\}$ *is a martingale.*

Our only experience to date of continuous-time martingales is contained in Section 12.7.

Proof. The process D is Markovian, and so U is a Markov process also. To check that the continuous-martingale condition holds, it suffices to show that

(10)
$$\mathbb{E}\big(U(t+s) \,\big|\, U(t)\big) = U(t) \quad \text{for all } s, t \geq 0.$$

However,

(11)
$$\begin{aligned}
\mathbb{E}\big(U(t+s) \,\big|\, U(t) = e^{-2md}\big) &= \mathbb{E}(e^{-2mD(t+s)} \mid D(t) = d) \\
&= \mathbb{E}\big(\exp\{-2m[D(t+s) - D(t)] - 2md\} \,\big|\, D(t) = d\big) \\
&= e^{-2md}\mathbb{E}\big(\exp\{-2m[D(t+s) - D(t)]\}\big) \\
&= e^{-2md}\mathbb{E}(e^{-2mD(s)})
\end{aligned}$$

because D is Markovian with stationary independent increments. Now, $\mathbb{E}(e^{-2mD(s)}) = M(-2m)$ where M is the moment generating function of an $N(ms, s)$ variable; this function M is given in Example (5.8.5) as $M(u) = e^{msu + \frac{1}{2}su^2}$. Thus $\mathbb{E}(e^{-2mD(s)}) = 1$ and so (10) follows from (11). ■

We can use this martingale to find the distribution of first passage times, just as we did in Example (12.5.6) for the random walk. Let $x, y > 0$ and define

$$T(x, -y) = \inf\big\{t : \text{either } D(t) = x \text{ or } D(t) = -y\big\}$$

to be the first passage time of D to the set $\{x, -y\}$. It is easily shown that $T(x, -y)$ is a stopping time which is almost surely finite.

(12) Theorem. $\mathbb{E}(U[T(x, -y)]) = 1$ *for all* $x, y > 0$.

Proof. This is just an application of a version of the optional stopping theorem (12.7.12). The process U is a martingale and $T(x, -y)$ is a stopping time. Therefore

$$\mathbb{E}(U[T(x, -y)]) = \mathbb{E}(U(0)) = 1. \qquad \blacksquare$$

(13) Corollary. *If* $m < 0$ *and* $x > 0$, *the probability that* D *ever visits the point* x *is*

$$\mathbb{P}(D(t) = x \text{ for some } t) = e^{2mx}.$$

Proof. By Theorem (12),

$$1 = e^{-2mx}\mathbb{P}(D[T(x, -y)] = x) + e^{2my}\{1 - \mathbb{P}(D[T(x, -y)] = x)\}.$$

Let $y \to \infty$ to obtain

$$\mathbb{P}(D[T(x, -y)] = x) \to e^{2mx}$$

so long as $m < 0$. Now complete the proof yourself. $\blacksquare$

The condition of Corollary (13), that the drift be negative, is natural; it is clear that if $m > 0$ then D almost surely visits all points on the positive part of the real axis. The result of (13) tells us about the size of the maximum of D also, since if $x > 0$,

$$\left\{\max_{t \geq 0} D(t) \geq x\right\} = \{D(t) = x \text{ for some } t\},$$

and the distribution of $M = \max\{D(t) : t \geq 0\}$ is easily deduced.

(14) Corollary. *If* $m < 0$ *then* M *is exponentially distributed with parameter* $-2m$.

Exercises for Section 13.4

1. Let W be a standard Wiener process and let $X(t) = \exp\{i\theta W(t) + \frac{1}{2}\theta^2 t\}$ where $i = \sqrt{-1}$. Show that X is a martingale with respect to the filtration given by $\mathcal{F}_t = \sigma(\{W(u) : u \leq t\})$.

2. Let T be the (random) time at which a standard Wiener process W hits the 'barrier' in space–time given by $y = at+b$ where $a < 0$, $b \geq 0$; that is, $T = \inf\{t : W(t) = at+b\}$. Use the result of Exercise (1) to show that the moment generating function of T is given by $\mathbb{E}(e^{\psi T}) = \exp\{-b(\sqrt{a^2 - 2\psi} + a)\}$ for $\psi < \frac{1}{2}a^2$. You may assume that the conditions of the optional stopping theorem are satisfied.

3. Let W be a standard Wiener process, and let T be the time of the last zero of W prior to time t. Show that $\mathbb{P}(T \leq u) = (2/\pi)\sin^{-1}\sqrt{u/t}$, $0 \leq u \leq t$.

13.5 Barriers

Diffusing particles are rarely allowed to roam freely, but are often restricted to a given part of space; for example, Brown's pollen particles were suspended in fluid which was confined to a container. What may happen when a particle hits a barrier? As with random walks, two simple types of barrier are the *absorbing* and the *reflecting*, although there are various other types of some complexity.

We begin with the case of the Wiener process. Let $w > 0$, let W be a standard Wiener process, and consider the shifted process $w + W(t)$ which starts at w. The Wiener process W^{a} *absorbed* at 0 is defined to be the process given by

(1)
$$W^{\mathrm{a}}(t) = \begin{cases} w + W(t) & \text{if } t < T, \\ 0 & \text{if } t \geq T, \end{cases}$$

where $T = \inf\{t : w + W(t) = 0\}$ is the hitting time of the position 0. The Wiener process W^{r} *reflected* at 0 is defined as the process $W^{\mathrm{r}}(t) = |w + W(t)|$.

Viewing the diffusion equations (13.3.7)–(13.3.8) as forward and backward equations, it is clear that W^{a} and W^{r} satisfy these equations so long as they are away from the barrier. That is to say, W^{a} and W^{r} are diffusion processes. In order to find their transition density functions, we might solve the diffusion equations subject to suitable boundary conditions. For the special case of the Wiener process, however, it is simpler to argue as follows.

(2) Theorem. *Let $f(t, y)$ denote the density function of the random variable $W(t)$, and let W^{a} and W^{r} be given as above.*

(a) *The density function of $W^{\mathrm{a}}(t)$ is*

$$f^{\mathrm{a}}(t, y) = f(t, y - w) - f(t, y + w), \quad y > 0.$$

(b) *The density function of $W^{\mathrm{r}}(t)$ is*

$$f^{\mathrm{r}}(t, y) = f(t, y - w) + f(t, y + w), \quad y > 0.$$

The function $f(t, y)$ is the $N(0, t)$ density function,

(3)
$$f(t, y) = \frac{1}{\sqrt{2\pi t}} \exp(-\tfrac{1}{2}y^2/t).$$

Proof. Let I be a subinterval of $(0, \infty)$, and let $I^{\mathrm{r}} = \{x \in \mathbb{R} : -x \in I\}$ be the reflection of I in the point 0. Then

$$\begin{aligned} \mathbb{P}\big(W^{\mathrm{a}}(t) \in I\big) &= \mathbb{P}\big(\{w + W(t) \in I\} \cap \{T > t\}\big) \\ &= \mathbb{P}\big(w + W(t) \in I\big) - \mathbb{P}\big(\{w + W(t) \in I\} \cap \{T \leq t\}\big) \\ &= \mathbb{P}\big(w + W(t) \in I\big) - \mathbb{P}\big(w + W(t) \in I^{\mathrm{r}}\big) \end{aligned}$$

using the reflection principle and the strong Markov property. The result follows.

The result of part (b) is immediate from the fact that $W^{\mathrm{r}}(t) = |w + W(t)|$. ■

We turn now to the absorption and reflection of a *general* diffusion process. Let $D = \{D(t) : t \geq 0\}$ be a diffusion process; we write a and b for the instantaneous mean and variance functions of D, and shall suppose that $b(t, x) > 0$ for all x (≥ 0) and t. We make a further assumption, that D is *regular* in that

(4) $$\mathbb{P}\big(D(t) = y \text{ for some } t \,\big|\, D(0) = x\big) = 1 \quad \text{for all } x, y \geq 0.$$

Suppose that the process starts from $D(0) = d$ say, where $d > 0$. Placing an absorbing barrier at 0 amounts to killing D when it first hits 0. The resulting process D^{a} is given by

$$D^{\mathrm{a}}(t) = \begin{cases} D(t) & \text{if } T > t, \\ 0 & \text{if } T \leq t, \end{cases}$$

where $T = \inf\{t : D(t) = 0\}$; this formulation requires D to have continuous sample paths.

Viewing the diffusion equations (13.3.7)–(13.3.8) as forward and backward equations, it is clear that they are satisfied away from the barrier. The presence of the absorbing barrier affects the solution to the diffusion equations through the boundary conditions.

Denote by $f^{\mathrm{a}}(t, y)$ the density function of $D^{\mathrm{a}}(t)$; we might write $f^{\mathrm{a}}(t, y) = f^{\mathrm{a}}(t, y \mid 0, d)$ to emphasize the value of $D^{\mathrm{a}}(0)$. The boundary condition appropriate to an absorbing barrier at 0 is

(5) $$f^{\mathrm{a}}(t, 0) = 0 \quad \text{for all } t.$$

It is not completely obvious that (5) is the correct condition, but the following rough argument may be made rigorous. The idea is that, if the particle is near to the absorbing barrier, then small local fluctuations, arising from the non-zero instantaneous variance, will carry it to the absorbing barrier extremely quickly. Therefore the chance of it being near to the barrier but unabsorbed is extremely small.

A slightly more rigorous justification for (5) is as follows. Suppose that (5) does not hold, which is to say that there exist $\epsilon, \eta > 0$ and $0 < u < v$ such that

(6) $$f^{\mathrm{a}}(t, y) > \eta \quad \text{for} \quad 0 < y \leq \epsilon,\ u \leq t \leq v.$$

There is probability at least $\eta\, dx$ that $0 < D^{\mathrm{a}}(t) \leq dx$ whenever $u \leq t \leq v$ and $0 < dx \leq \epsilon$. Hence the probability of absorption in the time interval $(t, t + dt)$ is at least

(7) $$\eta\, dx \mathbb{P}\big(D^{\mathrm{a}}(t + dt) - D^{\mathrm{a}}(t) < -dx \,\big|\, 0 < D^{\mathrm{a}}(t) \leq dx\big).$$

The instantaneous variance satisfies $b(t, x) \geq \beta$ for $0 < x \leq \epsilon$, $u \leq t \leq v$, for some $\beta > 0$, implying that $D^{\mathrm{a}}(t + dt) - D^{\mathrm{a}}(t)$ has variance at least $\beta\, dt$, under the condition that $0 < D^{\mathrm{a}}(t) \leq dx$. Therefore,

$$\mathbb{P}\big(D^{\mathrm{a}}(t + dt) - D^{\mathrm{a}}(t) < -\gamma\sqrt{dt} \,\big|\, 0 < D^{\mathrm{a}}(t) \leq dx\big) \geq \delta$$

for some $\gamma, \delta > 0$. Substituting $dx = \gamma\sqrt{dt}$ in (7), we obtain $\mathbb{P}(t < T < t + dt) \geq (\eta\gamma\delta)\sqrt{dt}$, implying by integration that $\mathbb{P}(u < T < v) = \infty$, which is clearly impossible. Hence (5) holds.

(8) Example. Wiener process with drift. Suppose that $a(t,x) = m$ and $b(t,x) = 1$ for all t and x. Put an absorbing barrier at 0 and suppose $D(0) = d > 0$. We wish to find a solution $g(t, y)$ to the forward equation

$$\frac{\partial g}{\partial t} = -m\frac{\partial g}{\partial y} + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}, \quad y > 0, \tag{9}$$

subject to the boundary conditions

$$g(t,0) = 0, \qquad t \geq 0, \tag{10}$$

$$g(0,y) = \delta_d(y), \qquad y \geq 0, \tag{11}$$

where δ_d is the Dirac δ function centred at d. We know from Example (13.3.10), and in any case it is easy to check from first principles, that the function

$$g(t, y \mid x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x-mt)^2}{2t}\right) \tag{12}$$

satisfies (9), for all possible 'sources' x. Our target is to find a linear combination of such functions $g(\cdot,\cdot \mid x)$ which satisfies (10) and (11). It turns out that

$$f^{\mathrm{a}}(t,y) = g(t, y \mid d) - e^{-2md} g(t, y \mid -d), \quad y > 0, \tag{13}$$

is such a function; assuming the solution is unique (which it is), this is therefore the density function of $D^{\mathrm{a}}(t)$. We may think of it as a mixture of the function $g(\cdot,\cdot \mid d)$ with source d together with a corresponding function from the 'image source' $-d$, being the reflection of d in the barrier at 0.

It is a small step to deduce the density function of the time T until the absorption of the particle. At time t, either the process has been absorbed, or its position has density function given by (13). Hence

$$\mathbb{P}(T \leq t) = 1 - \int_0^\infty f^{\mathrm{a}}(t,y)\,dy = 1 - \Phi\left(\frac{mt+d}{\sqrt{t}}\right) + e^{-2md}\Phi\left(\frac{mt-d}{\sqrt{t}}\right)$$

by (12) and (13), where Φ is the $N(0,1)$ distribution function. Differentiate with respect to t to obtain

$$f_T(t) = \frac{d}{\sqrt{2\pi t^3}} \exp\left(-\frac{(d+mt)^2}{2t}\right), \quad t > 0. \tag{14}$$

It is easily seen that

$$\mathbb{P}(\text{absorption takes place}) = \mathbb{P}(T < \infty) = \begin{cases} 1 & \text{if } m \leq 0, \\ e^{-2md} & \text{if } m > 0. \end{cases}$$

●

Turning to the matter of a reflecting barrier, suppose once again that D is a regular diffusion process with instantaneous mean a and variance b, starting from $D(0) = d > 0$. A reflecting barrier at the origin has the effect of disallowing infinitesimal negative jumps at the origin and replacing them by positive jumps. A formal definition requires careful treatment of the sample paths, and this is omitted here. Think instead about a reflecting barrier as giving rise to an appropriate boundary condition for the diffusion equations. Let us denote the reflected process by D^{r}, and let $f^{\mathrm{r}}(t, y)$ be its density function at time t. The reflected process lives on $[0, \infty)$, and therefore

$$\int_0^\infty f^{\mathrm{r}}(t, y)\,dy = 1 \quad \text{for all } t.$$

Differentiating with respect to t and using the forward diffusion equation, we obtain at the expense of mathematical rigour that

$$\begin{aligned} 0 &= \frac{\partial}{\partial t}\int_0^\infty f^{\mathrm{r}}(t, y)\,dy \\ &= \int_0^\infty \frac{\partial f^{\mathrm{r}}}{\partial t}\,dy = \int_0^\infty \left(-\frac{\partial}{\partial y}(af^{\mathrm{r}}) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(bf^{\mathrm{r}})\right)dy \\ &= \left[-af^{\mathrm{r}} + \frac{1}{2}\frac{\partial}{\partial y}(bf^{\mathrm{r}})\right]_0^\infty = \left(af^{\mathrm{r}} - \frac{1}{2}\frac{\partial}{\partial y}(bf^{\mathrm{r}})\right)\bigg|_{y=0}. \end{aligned}$$

This indicates that the density function $f^{\mathrm{r}}(t, y)$ of $D^{\mathrm{r}}(t)$ is obtained by solving the forward diffusion equation

$$\frac{\partial g}{\partial t} = -\frac{\partial}{\partial y}(ag) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(bg)$$

subject to the boundary condition

(15) $$\left(ag - \frac{1}{2}\frac{\partial}{\partial y}(bg)\right)\bigg|_{y=0} = 0 \quad \text{for} \quad t \geq 0,$$

as well as the initial condition

(16) $$g(0, y) = \delta_d(y) \quad \text{for} \quad y \geq 0.$$

(17) Example. Wiener process with drift. Once again suppose that $a(t, x) = m$ and $b(t, x) = 1$ for all x, t. This time we seek a linear combination of the functions g given in (12) which satisfies equations (15) and (16). It turns out that the answer contains an image at $-d$ together with a continuous line of images over the range $(-\infty, -d)$. That is to say, the solution has the form

$$f^{\mathrm{r}}(t, y) = g(t, y \mid d) + Ag(t, y \mid -d) + \int_{-\infty}^{-d} B(x)g(t, y \mid x)\,dx$$

for certain A and $B(x)$. Substituting this into equation (15), one obtains after some work that

(18) $$A = e^{-2md}, \quad B(x) = -2me^{2mx}.$$ ●

Exercise for Section 13.5

1. Let D be a standard Wiener process with drift m starting from $D(0) = d > 0$, and suppose that there is a reflecting barrier at the origin. Show that the density function $f^{\mathrm{r}}(t, y)$ of $D(t)$ satisfies $f^{\mathrm{r}}(t, y) \to 0$ as $t \to \infty$ if $m \geq 0$, whereas $f^{\mathrm{r}}(t, y) \to 2|m|e^{-2|m|y}$ for $y > 0$, as $t \to \infty$ if $m < 0$.

13.6 Excursions and the Brownian bridge

This section is devoted to properties of the Wiener process conditioned on certain special events. We begin with a question concerning the set of zeros of the process. Let $W = \{W(t) : t \geq 0\}$ be a Wiener process with $W(0) = w$, say, and with variance-parameter $\sigma^2 = 1$. What is the probability that W has no zeros in the time interval $(0, v]$ given that it has none in the smaller interval $(0, u]$? The question is not too interesting if $w \neq 0$, since in this case the probability in question is just the ratio

$$\frac{\mathbb{P}(\text{no zeros in } (0, v] \mid W(0) = w)}{\mathbb{P}(\text{no zeros in } (0, u] \mid W(0) = w)} \tag{1}$$

each term of which is easily calculated from the distribution of maxima (13.4.6). The difficulty arises when $w = 0$, since both numerator and denominator in (1) equal 0. In this case, it may be seen that the required probability is the limit of (1) as $w \to 0$. We have that this limit equals $\lim_{w\to 0}\{g_w(v)/g_w(u)\}$ where $g_w(x)$ is the probability that a Wiener process starting at w fails to reach 0 by time x. Using symmetry and Theorem (13.4.6),

$$g_w(x) = \sqrt{\frac{2}{\pi x}} \int_0^{|w|} \exp(-\tfrac{1}{2}m^2/x)\, dm,$$

whence $g_w(v)/g_w(u) \to \sqrt{u/v}$ as $w \to 0$, which we write as

$$\mathbb{P}\big(W \neq 0 \text{ on } (0, v] \,\big|\, W \neq 0 \text{ on } (0, u],\ W(0) = 0\big) = \sqrt{u/v}, \quad 0 < u \leq v. \tag{2}$$

A similar argument results in

$$\mathbb{P}\big(W > 0 \text{ on } (0, v] \,\big|\, W > 0 \text{ on } (0, u],\ W(0) = 0\big) = \sqrt{u/v}, \quad 0 < u \leq v, \tag{3}$$

by the symmetry of the Wiener process.

An 'excursion' of W is a trip taken by W away from 0. That is to say, if $W(u) = W(v) = 0$ and $W(t) \neq 0$ for $u < t < v$, then the trajectory of W during the time interval $[u, v]$ is called an *excursion* of the process; excursions are *positive* if $W > 0$ throughout (u, v), and *negative* otherwise. For any time $t > 0$, let $t - Z(t)$ be the time of the last zero prior to t, which is to say that $Z(t) = \sup\{s : W(t - s) = 0\}$; we suppose that $W(0) = 0$. At time t, some excursion is in progress whose current duration is $Z(t)$.

(4) Theorem. *Let $Y(t) = \sqrt{Z(t)}\operatorname{sign}\{W(t)\}$, and $\mathcal{F}_t = \sigma(\{Y(u) : 0 \leq u \leq t\})$. Then $(Y, \mathcal{F})$ is a martingale, called the* excursions martingale.

Proof. Clearly $Z(t) \le t$, so that $\mathbb{E}|Y(t)| \le \sqrt{t}$. It suffices to prove that

$$\textbf{(5)} \qquad \mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s\big) = Y(s) \quad \text{for} \quad s < t.$$

Suppose $s < t$, and let A be the event that $W(u) = 0$ for some $u \in [s, t]$. With a slight abuse of notation,

$$\mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s\big) = \mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s, A\big)\mathbb{P}(A \mid \mathcal{F}_s) + \mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s, A^{\mathrm{c}}\big)\mathbb{P}(A^{\mathrm{c}} \mid \mathcal{F}_s).$$

Now,

$$\textbf{(6)} \qquad \mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s, A\big) = 0$$

since, on the event A, the random variable $Y(t)$ is symmetric. On the other hand,

$$\textbf{(7)} \qquad \mathbb{E}\big(Y(t) \bigm| \mathcal{F}_s, A^{\mathrm{c}}\big) = \sqrt{t - s + Z(s)}\,\operatorname{sign}\{W(s)\}$$

since, given $\mathcal{F}_s$ and A^{c}, the current duration of the excursion at time t is $(t - s) + Z(s)$, and $\operatorname{sign}\{W(t)\} = \operatorname{sign}\{W(s)\}$. Furthermore $\mathbb{P}(A^{\mathrm{c}} \mid \mathcal{F}_s)$ equals the probability that W has strictly the same sign on $(s - Z(s), t]$ given the corresponding event on $(s - Z(s), s]$, which gives

$$\mathbb{P}(A^{\mathrm{c}} \mid \mathcal{F}_s) = \sqrt{\frac{Z(s)}{t - s + Z(s)}} \quad \text{by (3).}$$

Combining this with equations (6) and (7), we obtain $\mathbb{E}(Y(t) \mid \mathcal{F}_s) = Y(s)$ as required. ■

(8) Corollary. *The probability that the standard Wiener process W has a positive excursion of total duration at least a before it has a negative excursion of total duration at least b is $\sqrt{b}/(\sqrt{a} + \sqrt{b})$.*

Proof. Let $T = \inf\{t : Y(t) \ge \sqrt{a} \text{ or } Y(t) \le -\sqrt{b}\}$, the time which elapses before W records a positive excursion of duration at least a or a negative excursion of duration at least b. It may be shown that the optional stopping theorem for continuous-time martingales is applicable, and hence $\mathbb{E}(Y(T)) = \mathbb{E}(Y(0)) = 0$. However,

$$\mathbb{E}(Y(T)) = \pi\sqrt{a} - (1 - \pi)\sqrt{b}$$

where π is the required probability. ■

We turn next to the Brownian bridge. Think about a sample path of W on the time interval $[0, 1]$ as the shape of a random string with its left end tied to the origin. What does it look like if you tie down its right end also? That is to say, what sort of process is $\{W(t) : 0 \le t \le 1\}$ conditioned on the event that $W(1) = 0$? This new process is called the 'tied-down Wiener process' or the 'Brownian bridge'. There are various ways of studying it, the most obvious of which is perhaps to calculate the fdds of W conditional on the event $\{W(1) \in (-\eta, \eta)\}$, and then take the limit as $\eta \downarrow 0$. This is easily done, and leads to the next theorem.

(9) Theorem. *Let $B = \{B(t) : 0 \le t \le 1\}$ be a process with continuous sample paths and the same fdds as $\{W(t) : 0 \le t \le 1\}$ conditioned on $W(0) = W(1) = 0$. The process B is a diffusion process with instantaneous mean a and variance b given by*

$$\textbf{(10)} \qquad a(t, x) = -\frac{x}{1 - t}, \quad b(t, x) = 1, \quad x \in \mathbb{R},\ 0 \le t \le 1.$$

Note that the Brownian bridge has the same instantaneous variance as W, but its instantaneous mean increases in magnitude as $t \to 1$ and has the effect of guiding the process to its finishing point $B(1) = 0$.

Proof. We make use of an elementary calculation involving conditional density functions. Let W be a standard Wiener process, and suppose that $0 \le u \le v$. It is left as an *exercise* to prove that, conditional on the event $\{W(v) = y\}$, the distribution of $W(u)$ is normal with mean yu/v and variance $u(v-u)/v$. In particular,

$$\mathbb{E}\big(W(u) \,\big|\, W(0) = 0,\ W(v) = y\big) = \frac{yu}{v}, \tag{11}$$

$$\mathbb{E}\big(W(u)^2 \,\big|\, W(0) = 0,\ W(v) = y\big) = \left(\frac{yu}{v}\right)^2 + \frac{u(v-u)}{v}. \tag{12}$$

Returning to the Brownian bridge B, after a little reflection one sees that it is Gaussian and Markov, since W has these properties. Furthermore the instantaneous mean is given by

$$\mathbb{E}\big(B(t+h) - B(t) \,\big|\, B(t) = x\big) = -\frac{xh}{1-t}$$

by (11) with $y = -x$, $u = h$, $v = 1-t$; similarly the instantaneous variance is given by the following consequence of (12):

$$\mathbb{E}\big(|B(t+h) - B(t)|^2 \,\big|\, B(t) = x\big) = h + \mathrm{o}(h). \qquad \blacksquare$$

An elementary calculation based on equations (11) and (12) shows that

$$\operatorname{cov}\big(B(s), B(t)\big) = \min\{s,t\} - st, \quad 0 \le s,t \le 1. \tag{13}$$

Exercises for Section 13.6

1. Let W be a standard Wiener process. Show that the conditional density function of $W(t)$, given that $W(u) > 0$ for $0 < u < t$, is $g(x) = (x/t)e^{-x^2/(2t)}$, $x > 0$.

2. Show that the autocovariance function of the Brownian bridge is $c(s,t) = \min\{s,t\} - st$, $0 \le s,t \le 1$.

3. Let W be a standard Wiener process, and let $\widehat{W}(t) = W(t) - tW(1)$. Show that $\{\widehat{W}(t) : 0 \le t \le 1\}$ is a Brownian bridge.

4. If W is a Wiener process with $W(0) = 0$, show that $\widetilde{W}(t) = (1-t)W(t/(1-t))$ for $0 \le t < 1$, $\widetilde{W}(1) = 0$, defines a Brownian bridge.

5. Let $0 < s < t < 1$. Show that the probability that the Brownian bridge has no zeros in the interval (s,t) is $(2/\pi)\cos^{-1}\sqrt{(t-s)/[t(1-s)]}$.

13.7 Stochastic calculus

We have so far considered a diffusion process† $D = \{D_t : t \geq 0\}$ as a Markov process with continuous sample paths, having some given 'instantaneous mean' $\mu(t, x)$ and 'instantaneous variance' $\sigma^2(t, x)$. The most fundamental diffusion process is the standard Wiener process $W = \{W_t : t \geq 0\}$, with instantaneous mean 0 and variance 1. We have seen in Section 13.3 how to use this characterization of W in order to construct more general diffusions. With this use of the word 'instantaneous', it may seem natural, after a quick look at Section 13.3, to relate increments of D and W in the infinitesimal form

(1) $$dD_t = \mu(t, D_t)\,dt + \sigma(t, D_t)\,dW_t,$$

or equivalently its integrated form

(2) $$D_t - D_0 = \int_0^t \mu(s, D_s)\,ds + \int_0^t \sigma(s, D_s)\,dW_s.$$

The last integral has the form $\int_0^t \psi(s)\,dW_s$ where ψ is a random process. Whereas we saw in Section 9.4 how to construct such an integral for deterministic functions ψ, this more general case poses new problems, not least since a Wiener process is not differentiable. This section contains a general discussion of the stochastic integral, the steps necessary to establish it rigorously being deferred to Section 13.8.

For an example of the infinitesimal form (1) as a modelling tool, suppose that X_t is the price of some stock, bond, or commodity at time t. How may we represent the change dX_t over a small time interval $(t, t + dt)$? It may be a matter of observation that changes in the price X_t are proportional to the price, and otherwise appear to be as random in sign and magnitude as are the displacements of a molecule. It would be plausible to write $dX_t = bX_t\,dW_t$, or $X_t - X_0 = \int_0^t bX_s\,dW_s$, for some constant b. Such a process X is called a *geometric Wiener process*, or *geometric Brownian motion*; see Example (13.9.9) and Section 13.10.

We have already constructed certain representations of diffusion processes in terms of W. For example we have from Problem (13.12.1) that $tW_{1/t}$ and $\alpha W_{t/\alpha^2}$ are Wiener processes. Similarly, the process $D_t = \mu t + \sigma W_t$ is a Wiener process with drift. In addition, Ornstein–Uhlenbeck processes arise in a multiplicity of ways, for example as the processes U_i given by:

$$U_1(t) = e^{-\beta t} W(e^{2\beta t} - 1), \quad U_2(t) = e^{-\beta t} W(e^{2\beta t}), \quad U_3(t) = W(t) - \beta \int_0^t e^{-\beta(t-s)} W(s)\,ds.$$

(See Problem (13.12.3) and Exercises (13.7.4) and (13.7.5).) Expressions of this form enable us to deduce sample path properties of the process in question from those of the underlying Wiener process. For example, since W has continuous sample paths, so do the U_i.

It is illuminating to start with such an expression and to derive a differential form such as equation (1). Let X be a process which is a function of a standard Wiener process W, that is, $X_t = f(W_t)$ for some given f. Experience of the usual Newton–Leibniz chain rule would suggest that $dX_t = f'(W_t)\,dW_t$ but this turns out to be incorrect in this context. If f is sufficiently smooth, a formal application of Taylor's theorem gives

$$X_{t+\delta t} - X_t = f'(W_t)(\delta W_t) + \tfrac{1}{2} f''(W_t)(\delta W_t)^2 + \cdots$$

†For notational convenience, we shall write X_t or $X(t)$ interchangeably in the next few sections.

where $\delta W_t = W_{t+\delta t} - W_t$. In the usual derivation of the chain rule, one uses the fact that the second term on the right side is $o(\delta t)$. However, $(\delta W_t)^2$ has mean δt, and something new is needed. It turns out that δt is indeed an acceptable approximation for $(\delta W_t)^2$, and that the subsequent terms in the Taylor expansion are insignificant in the limit as $\delta t \to 0$. One is therefore led to the formula

$$dX_t = f'(W_t)\,dW_t + \tfrac{1}{2}f''(W_t)\,dt. \tag{3}$$

Note the extra term over that suggested by the usual chain rule. Equation (3) may be written in its integrated form

$$X_t - X_0 = \int_0^t f'(W_s)\,dW_s + \int_0^t \tfrac{1}{2}f''(W_s)\,ds.$$

Sense can be made of this only when we have a proper definition of the stochastic integral $\int_0^t f'(W_s)\,dW_s$. Equation (3) is a special case of what is known as Itô's formula, to which we return in Section 13.9.

Let us next work with a concrete example in the other direction, asking for an non-rigorous interpretation of the stochastic integral $\int_0^t W_s\,dW_s$. By analogy with the usual integral, we take $t = n\delta$ where δ is small and positive, and we partition the interval $(0, t]$ into the intervals $(j\delta, (j+1)\delta], 0 \le j < n$. Following the usual prescription, we take some $\theta_j \in [j\delta, (j+1)\delta]$ and form the sum $I_n = \sum_{j=0}^{n-1} W_{\theta_j}(W_{(j+1)\delta} - W_{j\delta})$.

In the context of the usual Riemann integral, the values $W_{j\delta}$, W_{θ_j}, and $W_{(j+1)\delta}$ would be sufficiently close to one another for I_n to have a limit as $n \to \infty$ which is independent of the choice of the θ_j. The Wiener process W has sample paths with unbounded variation, and therein lies the difference.

Suppose that we take $\theta_j = j\delta$ for each j. It is easy to check that

$$2I_n = \sum_{j=0}^{n-1}(W_{(j+1)\delta}^2 - W_{j\delta}^2) - \sum_{j=0}^{n-1}(W_{(j+1)\delta} - W_{j\delta})^2 = W_t^2 - W_0^2 - Z_n$$

where $Z_n = \sum_{j=0}^{n-1}(W_{(j+1)\delta} - W_{j\delta})^2$. It is the case that

$$\mathbb{E}\big((Z_n - t)^2\big) \to 0 \quad \text{as } n \to \infty, \tag{4}$$

which is to say that $Z_n \to t$ in mean square (see Exercise (13.7.2)). It follows that $I_n \to \frac{1}{2}(W_t^2 - t)$ in mean square as $n \to \infty$, and we are led to the interpretation

$$\int_0^t W_s\,dW = \tfrac{1}{2}(W_t^2 - t). \tag{5}$$

This proposal is verified in Example (13.9.7).

The calculation above is an example of what is called an *Itô integral*. The choice of the θ_j was central to the argument which leads to (5), and other choices lead to different answers. In Exercise (13.7.3) is considered the case when $\theta_j = (j + 1)\delta$, and this leads to the value $\frac{1}{2}(W_t^2 + t)$ for the integral. When θ_j is the midpoint of the interval $[j\delta, (j+1)\delta]$, the answer is the more familiar W_t^2, and the corresponding integral is termed the *Stratonovich integral*.

Exercises for Section 13.7

1. **Doob's L_2 inequality.** Let W be a standard Wiener process, and show that

$$\mathbb{E}\Big(\max_{0\le s\le t}|W_s|^2\Big)\le 4\mathbb{E}(W_t^2).$$

2. Let W be a standard Wiener process. Fix $t>0$, $n\ge 1$, and let $\delta=t/n$. Show that $Z_n=\sum_{j=0}^{n-1}(W_{(j+1)\delta}-W_{j\delta})^2$ satisfies $Z_n\to t$ in mean square as $n\to\infty$.

3. Let W be a standard Wiener process. Fix $t>0$, $n\ge 1$, and let $\delta=t/n$. Let $V_j=W_{j\delta}$ and $\Delta_j=V_{j+1}-V_j$. Evaluate the limits of the following as $n\to\infty$:
(a) $I_1(n)=\sum_j V_j\Delta_j$,
(b) $I_2(n)=\sum_j V_{j+1}\Delta_j$,
(c) $I_3(n)=\sum_j \frac{1}{2}(V_{j+1}+V_j)\Delta_j$,
(d) $I_4(n)=\sum_j W_{(j+\frac{1}{2})\delta}\Delta_j$.

4. Let W be a standard Wiener process. Show that $U(t)=e^{-\beta t}W(e^{2\beta t})$ defines a stationary Ornstein–Uhlenbeck process.

5. Let W be a standard Wiener process. Show that $U_t=W_t-\beta\int_0^t e^{-\beta(t-s)}W_s\,ds$ defines an Ornstein–Uhlenbeck process.

13.8 The Itô integral

Our target in this section is to present a definition of the integral $\int_0^\infty \psi_s\,dW_s$, where ψ is a random process satisfying conditions to be stated. Some of the details will be omitted from the account which follows.

Integrals of the form $\int_0^\infty \phi(s)\,dS_s$ were explored in Section 9.4 for deterministic functions ϕ, subject to the following assumptions on the process S:
(a) $\mathbb{E}(|S_t|^2)<\infty$ for all t,
(b) $\mathbb{E}(|S_{t+h}-S_t|^2)\to 0$ as $h\downarrow 0$, for all t,
(c) S has orthogonal increments.
It was required that ϕ satisfy $\int_0^\infty |\phi(s)|^2\,dG(s)<\infty$ where $G(t)=\mathbb{E}(|S_t-S_0|^2)$.

It is a simple exercise to check that conditions (a)–(c) are satisfied by the standard Wiener process W, and that $G(t)=t$ in this case. We turn next to conditions to be satisfied by the integrand ψ.

Let $W=\{W_t:t\ge 0\}$ be a standard Wiener process on the complete probability space $(\Omega,\mathcal{F},\mathbb{P})$. Let $\mathcal{F}_t$ be the smallest sub-σ-field of $\mathcal{F}$ with respect to which the variables W_s, $0\le s\le t$, are measurable and which contains the null events $\mathcal{N}=\{A\in\mathcal{F}:\mathbb{P}(A)=0\}$. We write $\boldsymbol{\mathcal{F}}=\{\mathcal{F}_t:t\ge 0\}$ for the consequent filtration.

A random process ψ is said to be *measurable* if, when viewed as a function $\psi_t(\omega)$ of both t and the elementary event $\omega\in\Omega$, it is measurable with respect to the product σ-field $\mathcal{B}\otimes\mathcal{F}$; here, $\mathcal{B}$ denotes the Borel σ-field of subsets of $[0,\infty)$. The measurable process ψ is said to be *adapted* to the filtration $\boldsymbol{\mathcal{F}}$ if ψ_t is $\mathcal{F}_t$-measurable for all t. It will emerge that adapted processes may be integrated against the Wiener process so long as they satisfy the integral condition

(1)
$$\mathbb{E}\left(\int_0^\infty |\psi_t|^2\, dt\right) < \infty,$$

and we denote by $\mathcal{A}$ the set of all adapted processes satisfying (1). It may be shown that $\mathcal{A}$ is a Hilbert space (and is thus Cauchy complete) with the norm†

(2)
$$\|\psi\| = \sqrt{\mathbb{E}\left(\int_0^\infty |\psi_t|^2\, dt\right)}.$$

We shall see that $\int_0^\infty \psi_s\, dW_s$ may be defined‡ for $\psi \in \mathcal{A}$.

We follow the scheme laid out in Section 9.4. The integral $\int_0^\infty \psi_s\, dW_s$ is first defined for a random step function $\psi_t = \sum_j C_j I_{(a_j, a_{j+1}]}(t)$ where the a_j are constants and the C_j are random variables with finite second moments which are $\mathcal{F}_{a_j}$-measurable. One then passes to limits of such step functions, finally checking that any process satisfying (1) may be expressed as such a limit. Here are some details.

Let $0 = a_0 < a_1 < \cdots < a_n = t$, and let $C_0, C_1, \ldots, C_{n-1}$ be random variables with finite second moments and such that each C_j is $\mathcal{F}_{a_j}$-measurable. Define the random variable ϕ_t by

$$\phi_t = \sum_{j=0}^{n-1} C_j I_{(a_j, a_{j+1}]}(t) = \begin{cases} 0 & \text{if } t \le 0 \text{ or } t > a_n, \\ C_j & \text{if } a_j < t \le a_{j+1}. \end{cases}$$

We call the function ϕ a 'predictable step function'. The stochastic integral $I(\phi)$ of ϕ with respect to W is evidently to be given by

(3)
$$I(\phi) = \sum_{j=0}^{n-1} C_j (W_{a_{j+1}} - W_{a_j}).$$

It is easily seen that $I(\alpha\phi^1 + \beta\phi^2) = \alpha I(\phi^1) + \beta I(\phi^2)$ for two predictable step functions ϕ^1, ϕ^2 and $\alpha, \beta \in \mathbb{R}$.

The following 'isometry' asserts the equality of the norm $\|\phi\|$ and the L_2 norm of the integral of ϕ. As before, we write $\|U\|_2 = \sqrt{\mathbb{E}|U^2|}$ where U is a random variable.

(4) Lemma. *If ϕ is a predictable step function, $\|I(\phi)\|_2 = \|\phi\|$.*

Proof. Evidently,

(5)
$$\begin{aligned} \mathbb{E}(|I(\phi)|^2) &= \mathbb{E}\left(\sum_{j=0}^{n-1} C_j (W_{a_{j+1}} - W_{a_j}) \sum_{k=0}^{n-1} C_k (W_{a_{k+1}} - W_{a_k})\right) \\ &= \mathbb{E}\left(\sum_{j=0}^{n-1} C_j^2 (W_{a_{j+1}} - W_{a_j})^2 + 2 \sum_{0 \le j < k \le n-1} C_j C_k (W_{a_{j+1}} - W_{a_j})(W_{a_{k+1}} - W_{a_k})\right). \end{aligned}$$

†Actually $\|\cdot\|$ is not a norm, since $\|\psi\| = 0$ does not imply that $\psi = 0$. It is however a norm on the set of equivalence classes obtained from the equivalence relation given by $\psi \sim \phi$ if $\mathbb{P}(\psi = \phi) = 1$.

‡Integrals over bounded intervals are defined similarly, by multiplying the integrand by the indicator function of the interval in question. That ψ be adapted is not really the 'correct' condition. In a more general theory of stochastic integration, the process W is replaced by a so-called semimartingale, and the integrand ψ by a locally bounded predictable process.

Using the fact that C_j is $\mathcal{F}_{a_j}$-measurable,

$$\mathbb{E}\big(C_j^2(W_{a_{j+1}} - W_{a_j})^2\big) = \mathbb{E}\big[\mathbb{E}\big(C_j^2(W_{a_{j+1}} - W_{a_j})^2 \mid \mathcal{F}_{a_j}\big)\big] = \mathbb{E}(C_j^2)(a_{j+1} - a_j).$$

Similarly, by conditioning on $\mathcal{F}_{a_k}$, we find that the mean of the final term in (5) equals 0. Therefore,

$$\mathbb{E}\big(|I(\phi)|^2\big) = \sum_j \mathbb{E}(C_j^2)(a_{j+1} - a_j) = \mathbb{E}\left(\int_0^\infty |\phi(t)|^2\,dt\right) = \|\phi\|^2. \qquad \blacksquare$$

Next we consider limits of sequences of predictable step functions. Let $\psi \in \mathcal{A}$. It may be shown that there exists a sequence $\boldsymbol{\phi} = \{\phi^{(n)}\}$ of predictable step functions such that $\|\phi^{(n)} - \psi\| \to 0$ as $n \to \infty$. We prove this under the assumption that ψ has continuous sample paths, although this continuity condition is not necessary.

(6) Theorem. *Let $\psi \in \mathcal{A}$ be a process with continuous sample paths. There exists a sequence $\boldsymbol{\phi} = \{\phi^{(n)}\}$ of predictable step functions such that $\|\phi^{(n)} - \psi\| \to 0$ as $n \to \infty$.*

Proof. Define the predictable step function

$$\phi_t^{(n)} = \begin{cases} n\displaystyle\int_{(j-1)/n}^{j/n} \psi_s\,ds & \text{for } \dfrac{j}{n} < t \le \dfrac{j+1}{n},\ 1 \le j < n^2, \\ 0 & \text{otherwise.} \end{cases}$$

By a standard use of the Cauchy–Schwarz inequality,

$$\int_{j/n}^{(j+1)/n} |\phi_t^{(n)}|^2\,dt = n\left|\int_{(j-1)/n}^{j/n} \psi_s\,ds\right|^2 \le \int_{(j-1)/n}^{j/n} |\psi_s|^2\,ds \quad \text{for } j \ge 1.$$

Hence

$$\text{(7)} \qquad \int_T^\infty |\phi_s^{(n)}|^2\,ds \le \int_{T-(2/n)}^\infty |\psi_s|^2\,ds \quad \text{for } T \ge 0.$$

Now,

$$\text{(8)} \qquad \int_0^\infty |\phi_s^{(n)} - \psi_s|^2\,ds = \int_0^T |\phi_s^{(n)} - \psi_s|^2\,ds + \int_T^\infty |\phi_s^{(n)} - \psi_s|^2\,ds.$$

Using the continuity of the sample paths of ψ, $|\phi_s^{(n)} - \psi_s| \to 0$ as $n \to \infty$, uniformly on the interval $[0, T]$, whence the penultimate term in (8) tends to 0 as $n \to \infty$. Since $|x + y|^2 \le 2(|x|^2 + |y|^2)$ for $x, y \in \mathbb{R}$, the last term in (8) is by (7) no greater than $4\int_{T-(2/n)}^\infty |\psi_s|^2\,ds$. We let $n \to \infty$ and then $T \to \infty$ in (8). Since $\psi \in \mathcal{A}$, it is the case that $\int_0^\infty |\psi_s|^2\,ds < \infty$ almost surely, and therefore

$$\int_0^\infty |\phi_s^{(n)} - \psi_s|^2\,ds \to 0 \qquad \text{almost surely, as } n \to \infty.$$

By the same argument used to bound the last term in (8),

$$0 \le \int_0^\infty |\phi_s^{(n)} - \psi_s|^2 \, ds \le 4 \int_0^\infty |\psi_s|^2 \, ds,$$

and it follows by the dominated convergence theorem that $\|\phi^{(n)} - \psi\| \to 0$ as $n \to \infty$. ■

Let $\psi \in \mathcal{A}$ and let $\boldsymbol{\phi} = \{\phi^{(n)}\}$ be a sequence of predictable step functions converging in $\mathcal{A}$ to ψ. Since $\phi^{(m)} - \phi^{(n)}$ is itself a predictable step function, we have that

$$\begin{aligned}
\|I(\phi^{(m)}) - I(\phi^{(n)})\|_2 &= \|I(\phi^{(m)} - \phi^{(n)})\|_2 \\
&= \|\phi^{(m)} - \phi^{(n)}\| && \text{by Lemma (4)} \\
&\le \|\phi^{(m)} - \psi\| + \|\phi^{(n)} - \psi\| && \text{by the triangle inequality} \\
&\to 0 && \text{as } m, n \to \infty.
\end{aligned}$$

Therefore the sequence $I(\phi^{(n)})$ is mean-square Cauchy convergent, and hence converges in mean square to some limit random variable denoted $I(\boldsymbol{\phi})$. It is not difficult to show as follows that $\mathbb{P}(I(\boldsymbol{\phi}) = I(\boldsymbol{\rho})) = 1$ for any other sequence $\boldsymbol{\rho}$ of predictable step functions converging in $\mathcal{A}$ to ψ. We have by the triangle inequality that

$$\|I(\boldsymbol{\phi}) - I(\boldsymbol{\rho})\|_2 \le \|I(\boldsymbol{\phi}) - I(\phi^{(n)})\|_2 + \|I(\phi^{(n)}) - I(\rho^{(n)})\|_2 + \|I(\rho^{(n)}) - I(\boldsymbol{\rho})\|_2.$$

The first and third terms on the right side tend to 0 as $n \to \infty$. By Lemma (4) and the linearity of the integral operator on predictable step functions, the second term satisfies

$$\begin{aligned}
\|I(\phi^{(n)}) - I(\rho^{(n)})\|_2 = \|I(\phi^{(n)} - \rho^{(n)})\|_2 &= \|\phi^{(n)} - \rho^{(n)}\| \\
&\le \|\phi^{(n)} - \psi\| + \|\rho^{(n)} - \psi\|
\end{aligned}$$

which tends to zero as $n \to \infty$. Therefore $\|I(\boldsymbol{\phi}) - I(\boldsymbol{\rho})\|_2 = 0$, implying as claimed that $\mathbb{P}(I(\boldsymbol{\phi}) = I(\boldsymbol{\rho})) = 1$.

The (almost surely) unique such quantity $I(\boldsymbol{\phi})$ is denoted by $I(\psi)$, which we call the *Itô integral* of the process ψ. It is usual to denote $I(\psi)$ by $\int_0^\infty \psi_s \, dW_s$, and we adopt this notation forthwith. We define $\int_0^t \psi_s \, dW_s$ to be $\int_0^\infty \psi_s I_{(0,t]}(s) \, dW_s$.

With the (Itô) stochastic integral defined, we may now agree to write

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

as a shorthand form of

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

A continuous process X defined by (9), by which we mean satisfying (10), is called an *Itô process*, or a diffusion process, with infinitesimal mean and variance $\mu(t, x)$ and $\sigma(t, x)^2$. The proof that there exists such a process is beyond our scope. Thus we may define diffusions via stochastic integrals, and it may be shown conversely that all diffusions previously considered in this book may be written as appropriate stochastic integrals.

It is an important property of the above stochastic integrals that they define martingales†. Once again, we prove this under the assumption that ψ has continuous sample paths.

(11) Theorem. *Let $\psi \in \mathcal{A}$ be a process with continuous sample paths. The process $J_t = \int_0^t \psi_u \, dW_u$ is a martingale with respect to the filtration $\mathcal{F}$.*

Proof. Let $0 < s < t$ and, for $n \geq 1$, let $a_0, a_1, \dots, a_n$ be such that $0 = a_0 < a_1 < \cdots < a_m = s < a_{m+1} < \cdots < a_n = t$ for some m. We define the predictable step function

$$\psi_u^{(n)} = \sum_{j=0}^{n-1} \psi_{a_j} I_{(a_j, a_{j+1}]}(u),$$

with integral

$$J_v^{(n)} = \int_0^v \psi_u^{(n)} \, dW_u = \sum_{j=0}^{n-1} \psi_{a_j} (W_{a_{j+1} \wedge v} - W_{a_j \wedge v}), \qquad v \geq 0,$$

where $x \wedge y = \min\{x, y\}$. Now,

$$\mathbb{E}(J_t^{(n)} \mid \mathcal{F}_s) = \sum_{j=0}^{n-1} \mathbb{E}\big(\psi_{a_j}(W_{a_{j+1}} - W_{a_j}) \,\big|\, \mathcal{F}_s\big)$$

where

$$\mathbb{E}\big(\psi_{a_j}(W_{a_{j+1}} - W_{a_j}) \,\big|\, \mathcal{F}_s\big) = \psi_{a_j}(W_{a_{j+1}} - W_{a_j}) \qquad \text{if } j < m,$$

and

$$\mathbb{E}\big(\psi_{a_j}(W_{a_{j+1}} - W_{a_j}) \,\big|\, \mathcal{F}_s\big) = \mathbb{E}\big[\mathbb{E}\big(\psi_{a_j}(W_{a_{j+1}} - W_{a_j}) \,\big|\, \mathcal{F}_{a_j}\big) \,\big|\, \mathcal{F}_s\big] = 0 \qquad \text{if } j \geq m,$$

since $W_{a_{j+1}} - W_{a_j}$ is independent of $\mathcal{F}_{a_j}$ and has zero mean. Therefore,

$$\mathbb{E}(J_t^{(n)} \mid \mathcal{F}_s) = J_s^{(n)}. \tag{12}$$

We now let $n \to \infty$ and assume that $\max_j |a_{j+1} - a_j| \to 0$. As shown in the proof of Theorem (6),

$$\|\psi^{(n)} I_{(0,s]} - \psi I_{(0,s]}\| \to 0 \quad \text{and} \quad \|\psi^{(n)} I_{(0,t]} - \psi I_{(0,t]}\| \to 0,$$

whence $J_s^{(n)} \to \int_0^s \psi_u \, dW_u$ and $J_t^{(n)} \to \int_0^t \psi_u \, dW_u$ in mean square. We let $n \to \infty$ in (12), and use the result of Exercise (13.8.5), to find that $\mathbb{E}(J_t \mid \mathcal{F}_s) = J_s$ almost surely. It follows as claimed that J is a martingale. ■

There is a remarkable converse to Theorem (11) of which we omit the proof.

(13) Theorem. *Let M be a martingale with respect to the filtration $\mathcal{F}$. There exists an adapted random process ψ such that*

$$M_t = M_0 + \int_0^t \psi_u \, dW_u, \qquad t \geq 0.$$

---

†In the absence of condition (1), we may obtain what is called a 'local martingale', but this is beyond the scope of this book.

Exercises for Section 13.8

In the absence of any contrary indication, W denotes a standard Wiener process, and $\mathcal{F}_t$ is the smallest σ-field containing all null events with respect to which every member of $\{W_u : 0 \le u \le t\}$ is measurable.

1. (a) Verify directly that $\int_0^t s\,dW_s = tW_t - \int_0^t W_s\,ds$.

(b) Verify directly that $\int_0^t W_s^2\,dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s\,ds$.

(c) Show that $\mathbb{E}\left(\left[\int_0^t W_s\,dW_s\right]^2\right) = \int_0^t \mathbb{E}(W_s^2)\,ds$.

2. Let $X_t = \int_0^t W_s\,ds$. Show that X is a Gaussian process, and find its autocovariance and autocorrelation function.

3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and suppose that $X_n \xrightarrow{\text{m.s.}} X$ as $n \to \infty$. If $\mathcal{G} \subseteq \mathcal{F}$, show that $\mathbb{E}(X_n \mid \mathcal{G}) \xrightarrow{\text{m.s.}} \mathbb{E}(X \mid \mathcal{G})$.

4. Let ψ_1 and ψ_2 be predictable step functions, and show that

$$\mathbb{E}\{I(\psi_1)I(\psi_2)\} = \mathbb{E}\left(\int_0^\infty \psi_1(t)\psi_2(t)\,dt\right),$$

whenever both sides exist.

5. Assuming that *Gaussian white noise* $G_t = dW_t/dt$ exists in sufficiently many senses to appear as an integrand, show by integrating the stochastic differential equation $dX_t = -\beta X_t\,dt + dW_t$ that

$$X_t = W_t - \beta\int_0^t e^{-\beta(t-s)}W_s\,ds,$$

if $X_0 = 0$.

6. Let ψ be an adapted process with $\|\psi\| < \infty$. Show that $\|I(\psi)\|_2 = \|\psi\|$.

13.9 Itô's formula

The 'stochastic differential equation', or 'SDE',

$$dX = \mu(t, X)\,dt + \sigma(t, X)\,dW \tag{1}$$

is a shorthand for the now well-defined integral equation

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

solutions to which are called *Itô processes*, or *diffusions*. Under rather weak conditions on μ, σ, and X_0, it may be shown that the SDE (1) has a unique solution which is a Markov process with continuous sample paths. The proof of this is beyond our scope.

We turn to a central question. If X satisfies the SDE (1) and $Y_t = f(t, X_t)$ for some given $f : [0, \infty) \times \mathbb{R} \to \mathbb{R}$, what is the infinitesimal formula for the process Y?

(2) Theorem. Itô's formula. *If* $dX = \mu(t, X)\,dt + \sigma(t, X)\,dW$ *and* $Y_t = f(t, X_t)$, *where* f *is twice continuously differentiable on* $[0, \infty) \times \mathbb{R}$, *then* Y *is also an Itô process, given by*†

$$\text{(3)}\quad dY = \left[f_x(t, X)\mu(t, X) + f_t(t, X) + \tfrac{1}{2} f_{xx}(t, X)\sigma^2(t, X)\right] dt + f_x(t, X)\sigma(t, X)\,dW.$$

This formula may be extended to cover multivariate diffusions. We do not prove Itô's formula at the level of generality of (2), instead specializing to the following special case when X is the standard Wiener process. The differentiability assumption on the function f may be weakened.

(4) Theorem. Itô's simple formula. *Let* $f(s, w)$ *be thrice continuously differentiable on* $[0, \infty) \times \mathbb{R}$, *and let* W *be a standard Wiener process. The process* $Y_t = f_t(t, W_t)$ *is an Itô process with*

$$dY = \left[f_t(t, W) + \tfrac{1}{2} f_{ww}(t, W)\right] dt + f_w(t, W)\,dW.$$

Sketch proof. Let $n \geq 1$, $\delta = t/n$, and write $\Delta_j = W_{(j+1)\delta} - W_{j\delta}$. The idea is to express $f(t, W_t)$ as the sum

$$\text{(5)}\qquad f(t, W_t) - f(0, W_0) = \sum_{j=0}^{n-1} \left[f\big((j+1)\delta, W_{(j+1)\delta}\big) - f(j\delta, W_{(j+1)\delta})\right] + \sum_{j=0}^{n-1} \left[f(j\delta, W_{(j+1)\delta}) - f(j\delta, W_{j\delta})\right]$$

and to use Taylor's theorem to study its behaviour as $n \to \infty$. We leave out the majority of details necessary to achieve this, presenting instead the briefest summary.

By (5) and Taylor's theorem, there exist random variables $\theta_j \in [j\delta, (j+1)\delta]$ and $\Omega_j \in [W_{j\delta}, W_{(j+1)\delta}]$ such that

$$\text{(6)}\quad f(t, W_t) - f(0, W_0) = \sum_{j=0}^{n-1} f_t(\theta_j, W_{(j+1)\delta})\delta + \sum_{j=0}^{n-1} f_w(j\delta, W_{j\delta})\Delta_j + \tfrac{1}{2} \sum_{j=0}^{n-1} f_{ww}(j\delta, W_{j\delta})\Delta_j^2 + \tfrac{1}{6} \sum_{j=0}^{n-1} f_{www}(j\delta, \Omega_j)\Delta_j^3.$$

We consider these terms one by one, as $n \to \infty$.

(i) It is a consequence of the continuity properties of f and W that

$$\sum_{j=0}^{n-1} f_t(\theta_j, W_{(j+1)\delta})\delta \xrightarrow{\text{a.s.}} \int_0^t f_t(s, W_s)\,ds.$$

(ii) Using the differentiability of f, one may see that $\sum_{j=0}^{n-1} f_w(j\delta, W_{j\delta})\Delta_j$ converges in mean square as $n \to \infty$ to the Itô integral $\int_0^t f_w(s, W_s)\,dW_s$.

†Here $f_t(t, X)$ and $f_x(t, X)$ denote the derivatives of f with respect to its first and second arguments respectively, and evaluated at (t, X_t).

(iii) We have that $\mathbb{E}(\Delta_j^2) = \delta$, and Δ_j^2 and Δ_k^2 are independent for $j \neq k$. It follows after some algebra that

$$\sum_{j=0}^{n-1} f_{ww}(j\delta, W_{j\delta})\Delta_j^2 - \sum_{j=0}^{n-1} f_{ww}(j\delta, W_{j\delta})\delta \xrightarrow{\text{m.s.}} 0.$$

This implies the convergence of the third sum in (6) to the integral $\frac{1}{2}\int_0^t f_{ww}(s, W_s)\,ds$.

(iv) It may be shown after some work that the fourth term in (6) converges in mean square to zero as $n \to \infty$, and the required result follows by combining (i)–(iv). ■

(7) Example. (a) Let $dX = \mu(t, X)\,dt + \sigma(t, X)\,dW$ and let $Y = X^2$. By Theorem (2),

$$dY = (2\mu X + \sigma^2)\,dt + 2\sigma X\,dW = \sigma(t, X)^2\,dt + 2X\,dX.$$

(b) Let $Y_t = W_t^2$. Applying part (a) with $\mu = 0$ and $\sigma = 1$ (or alternatively using Itô's simple formula (4)), we find that $dY = dt + 2W\,dW$. By integration,

$$\int_0^t W_s\,dW_s = \tfrac{1}{2}(Y_t - Y_0 - t) = \tfrac{1}{2}(W_t^2 - t)$$

in agreement with formula (13.7.5). ●

(8) Example. Product rule. Suppose that

$$dX = \mu_1(t, X)\,dt + \sigma_1(t, X)\,dW, \quad dY = \mu_2(t, Y)\,dt + \sigma_2(t, Y)\,dW,$$

in the notation of Itô's formula (2). We have by Example (7) that

$$\begin{aligned} d(X^2) &= \sigma_1(t, X)^2\,dt + 2X\,dX, \\ d(Y^2) &= \sigma_2(t, Y)^2\,dt + 2Y\,dY, \\ d((X+Y)^2) &= \big(\sigma_1(t, X) + \sigma_2(t, Y)\big)^2\,dt + 2(X+Y)(dX + dY). \end{aligned}$$

Using the representation $XY = \frac{1}{2}\{(X+Y)^2 - X^2 - Y^2\}$, we deduce the product rule

$$d(XY) = X\,dY + Y\,dX + \sigma_1(t, X)\sigma_2(t, Y)\,dt.$$

Note the extra term over the usual Newton–Leibniz rule for differentiating a product. ●

(9) Example. Geometric Brownian motion. Let $Y_t = \exp(\mu t + \sigma W_t)$ for constants μ, σ. Itô's simple formula (4) yields

$$dY = (\mu + \tfrac{1}{2}\sigma^2)Y\,dt + \sigma Y\,dW,$$

so that Y is a diffusion with instantaneous mean $a(t, y) = (\mu + \frac{1}{2}\sigma^2)y$ and instantaneous variance $b(t, y) = \sigma^2 y^2$. As indicated in Example (12.7.10), the process Y is a martingale if and only if $\mu + \frac{1}{2}\sigma^2 = 0$. ●

Exercises for Section 13.9

In the absence of any contrary indication, W denotes a standard Wiener process, and $\mathcal{F}_t$ is the smallest σ-field containing all null events with respect to which every member of $\{W_u : 0 \le u \le t\}$ is measurable.

1. Let X and Y be independent standard Wiener processes. Show that, with $R_t^2 = X_t^2 + Y_t^2$,

$$Z_t = \int_0^t \frac{X_s}{R_s}\, dX_s + \int_0^t \frac{Y_s}{R_s}\, dY_s$$

is a Wiener process. [Hint: Use Theorem (13.8.13).] Hence show that R^2 satisfies

$$R_t^2 = 2\int_0^t R_s\, dW_s + 2t.$$

Generalize this conclusion to n dimensions.

2. Write down the SDE obtained via Itô's formula for the process $Y_t = W_t^4$, and deduce that $\mathbb{E}(W_t^4) = 3t^2$.

3. Show that $Y_t = tW_t$ is an Itô process, and write down the corresponding SDE.

4. **Wiener process on a circle.** Let $Y_t = e^{iW_t}$. Show that $Y = X_1 + iX_2$ is a process on the unit circle satisfying

$$dX_1 = -\tfrac{1}{2}X_1\, dt - X_2\, dW, \quad dX_2 = -\tfrac{1}{2}X_2\, dt + X_1\, dW.$$

5. Find the SDEs satisfied by the processes:
(a) $X_t = W_t/(1+t)$,
(b) $X_t = \sin W_t$,
(c) [Wiener process on an ellipse] $X_t = a\cos W_t$, $Y_t = b\sin W_t$, where $ab \ne 0$.

13.10 Option pricing

It was essentially the Wiener process which Bachelier proposed in 1900 as a model for the evolution of stock prices. Interest in the applications of diffusions and martingales to stock prices has grown astonishingly since the fundamental work of Black, Scholes, and Merton in the early 1970s. The theory of mathematical finance is now well developed, and is one of the most striking modern applications of probability theory. We present here one simple model and application, namely the Black–Scholes solution for the pricing of a European call option. Numerous extensions of this result are possible, and the reader is referred to one of the many books on mathematical finance for further details (see Appendix II).

The Black–Scholes model concerns an economy which comprises two assets, a 'bond' (or 'money market account') whose value grows at a continuously compounded constant interest rate r, and a 'stock' whose price per unit is a stochastic process $S = \{S_t : t \ge 0\}$ indexed by time t. It is assumed that any quantity, *positive or negative* real valued, of either asset may be purchased at any time†. Writing M_t for the cost of one unit of the bond at time t, and

†No taxes or commissions are payable, and the possession of stock brings no dividends. The purchase of negative quantities of bond or stock is called 'short selling' and can lead to a 'short position'.

normalizing so that $M_0 = 1$, we have that

(1) $$dM_t = rM_t\,dt \quad \text{or} \quad M_t = e^{rt}.$$

A basic assumption of the Black–Scholes model is that S satisfies the stochastic differential equation

(2) $$dS_t = S_t(\mu\,dt + \sigma\,dW_t) \quad \text{with solution} \quad S_t = \exp\big((\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t\big),$$

where W is a standard Wiener process and we have normalized by setting $S_0 = 1$. That is to say, S is a geometric Brownian motion (13.9.9) with parameters μ, σ; in this context, σ is usually called the *volatility* of the price process.

The market permits individuals to buy so-called 'forward options' on the stock, such products being termed 'derivatives'. One of the most important derivatives, the 'European call option', permits the buyer to purchase one unit of the stock at some given future time and at some predetermined price. More precisely, the option gives the holder the right to buy one unit of stock at time T, called the 'exercise date', for the price K, called the 'strike price'; the holder is not *required* to exercise this right. The fundamental question is to determine the 'correct price' of this option at some time t satisfying $t \le T$. The following elucidation of market forces leads to an interpretation of the notion of 'correct price', and utilizes some beautiful mathematics†.

We have at time T that:

(a) if $S_T > K$, a holder of the option can buy one unit of the stock for K and sell immediately for S_T, making an immediate profit of $S_T - K$,

(b) if $S_T \le K$, it would be preferable to buy K/S_T (≥ 1) units of the stock on the open market than to exercise the option.

It follows that the value ϕ_T of the option at time T is given by $\phi_T = \max\{S_T - K, 0\} = (S_T - K)^+$. The discounted value of ϕ_T at an earlier time t is $e^{-r(T-t)}(S_T - K)^+$, since an investment at time t of this sum in the bond will be valued at ϕ_T at the later time T. One might naively suppose that the value of the option at an earlier time is given by its expectation; for example, the value at time 0 might be $\phi_0 = \mathbb{E}(e^{-rT}(S_T - K)^+)$. The financial market does not operate in this way, and *this answer is wrong*. It turns out in general that, in a market where options are thus priced according to the mean of their discounted value, the buyer of the option can devise a strategy for making a certain profit. Such an opportunity to make a risk-free profit is called an *arbitrage opportunity* and it may be assumed in practice that no such opportunity exists. In order to define the notion of arbitrage more properly‡, we discuss next the concept of 'portfolio'.

Let $\mathcal{F}_t$ be the σ-field generated by the random variables $\{S_u : 0 \le u \le t\}$. A *portfolio* is a pair $\alpha = \{\alpha_t : t \ge 0\}$, $\beta = \{\beta_t : t \ge 0\}$ of stochastic processes which are adapted to the filtration $\boldsymbol{\mathcal{F}} = \{\mathcal{F}_t : t \ge 0\}$. We interpret the pair (α, β) as a time-dependent portfolio comprising α_t units of stock and β_t units of the bond at time t. The value at time t of the portfolio (α, β) is given by the *value function*

(3) $$V_t(\alpha, \beta) = \alpha_t S_t + \beta_t M_t,$$

†It would in practice be a mistake to adhere over rigidly to strategies based on the mathematical facts presented in this section and elsewhere. Such results are well known across the market and their use can be disadvantageous, as some have found out to their cost.

‡Such a concept was mentioned for a discrete system in Exercise (6.6.3).

and the portfolio is called *self-financing* if

$$dV_t(\alpha, \beta) = \alpha_t \, dS_t + \beta_t \, dM_t, \tag{4}$$

which is to say that changes in value may be attributed to changes in the market only and not to the injection or withdrawal of funds. Condition (4) is a consequence of the modelling assumption implicit in (2) that S is an Itô integral. It is is explained slightly more fully via the following discretization of time. Suppose that $\epsilon > 0$, and that time is divided into the intervals $I_n = [n\epsilon, (n+1)\epsilon)$. We assume that prices remain constant within each interval I_n. We exit interval I_{n-1} having some portfolio $(\alpha_{(n-1)\epsilon}, \beta_{(n-1)\epsilon})$. At the start of I_n this portfolio has value $v_n = \alpha_{(n-1)\epsilon} S_{n\epsilon} + \beta_{(n-1)\epsilon} M_{n\epsilon}$. The self-financing of the portfolio implies that the value at the end of I_n equals v_n, which is to say that

$$\alpha_{(n-1)\epsilon} S_{n\epsilon} + \beta_{(n-1)\epsilon} M_{n\epsilon} = \alpha_{n\epsilon} S_{n\epsilon} + \beta_{n\epsilon} M_{n\epsilon}. \tag{5}$$

Now,

$$\begin{aligned} v_{n+1} - v_n &= \alpha_{n\epsilon} S_{(n+1)\epsilon} + \beta_{n\epsilon} M_{(n+1)\epsilon} - \alpha_{(n-1)\epsilon} S_{n\epsilon} - \beta_{(n-1)\epsilon} M_{n\epsilon} \\ &= \alpha_{n\epsilon}(S_{(n+1)\epsilon} - S_{n\epsilon}) + \beta_{n\epsilon}(M_{(n+1)\epsilon} - M_{n\epsilon}) \end{aligned} \tag{6}$$

by (5). Condition (4) is motivated by passing to the limit $\epsilon \downarrow 0$.

We say that a self-financing portfolio (α, β) *replicates* the given European call option if its value $V_T(\alpha, \beta)$ at time T satisfies $V_T(\alpha, \beta) = (S_T - K)^+$ almost surely.

We now utilize the assumption that the market contains no arbitrage opportunities. Let $t < T$, and suppose that two options are available at a given time t. Option I costs c_1 per unit and yields a (strictly positive) value ϕ at time T; Option II costs c_2 per unit and yields the same value ϕ at time T. We may assume without loss of generality that $c_1 \geq c_2$. Consider the following strategy: at time t, buy $-c_2$ units of Option I and c_1 units of Option II. The total cost is $(-c_2)c_1 + c_1c_2 = 0$, and the value at time T is $(-c_2)\phi + c_1\phi = (c_1 - c_2)\phi$. If $c_1 > c_2$, there exists a strategy which yields a risk-free profit, in contradiction of the assumption of no arbitrage. Therefore $c_1 = c_2$.

Assume now that there exists a self-financing portfolio (α, β) which replicates the European call option. At time t $(< T)$ we may either invest in the option, or we may buy into the portfolio. Since their returns at time T are equal, they must by the argument above have equal cost at time t. That is to say, in the absence of arbitrage, the 'correct value' of the European call option at time t is $V_t(\alpha, \beta)$. In order to price the option, it remains to show that such a portfolio exists, and to find its value function.

First we calculate the value function of such a portfolio, and later we shall return to the question of its existence. Assume that (α, β) is a self-financing portfolio which replicates the European call option. It would be convenient if its discounted value function $e^{-rt}V_t$ were a martingale, since it would follow that $e^{-rt}V_t = \mathbb{E}(e^{-rT}V_T \mid \mathcal{F}_t)$ where $V_T = (S_T - K)^+$. This is not generally the case, but the following clever argument may be exploited. Although $e^{-rt}V_t$ is not a martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, it turns out that there exists an alternative probability measure $\mathbb{Q}$ on the measurable pair $(\Omega, \mathcal{F})$ such that $e^{-rt}V_t$ is indeed a martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. The usual proof of this statement makes use of a result known as the Cameron–Martin–Girsanov formula which is beyond the scope of this book. In the case of the Black–Scholes model, one may argue directly via the following 'change of measure' formula.

(7) Theorem. *Let $B = \{B_t : 0 \le t \le T\}$ be a Wiener process with drift 0 and instantaneous variance σ^2 on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\nu \in \mathbb{R}$. Define the random variable*

$$\Lambda = \exp\left\{\frac{\nu}{\sigma^2}B_T - \frac{\nu^2}{2\sigma^2}T\right\},$$

and the measure $\mathbb{Q}$ by $\mathbb{Q}(A) = \mathbb{E}(\Lambda I_A)$. Then $\mathbb{Q}$ is a probability measure and, regarded as a process on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, B is a Wiener process with drift ν and instantaneous variance σ^2.

Proof. That $\mathbb{Q}$ is a probability measure is a consequence of the fact that

$$\mathbb{Q}(\Omega) = \mathbb{E}(\Lambda) = e^{-\nu^2 T/(2\sigma^2)}\mathbb{E}(e^{(\nu/\sigma^2)B_T}) = e^{-\nu^2 T/(2\sigma^2)}e^{\frac{1}{2}(\nu/\sigma^2)^2\sigma^2 T} = 1.$$

The distribution of B under $\mathbb{Q}$ is specified by its finite-dimensional distributions (we recall the discussion of Section 8.6). Let $0 = t_0 < t_1 < \cdots < t_n = T$ and $x_0, x_1, \dots, x_n = x \in \mathbb{R}$. The notation used in the following is informal but convenient. The process B has independent normal increments under the measure $\mathbb{P}$. Writing $\{B_{t_i} \in dx_i\}$ for the event that $x_i < B_{t_i} \le x_i + dx_i$, we have that

$$\begin{aligned}
&\mathbb{Q}\big(B_{t_1} \in dx_1,\ B_{t_2} \in dx_2, \dots,\ B_{t_n} \in dx_n\big) \\
&= \mathbb{E}\left(\exp\left(\frac{\nu}{\sigma^2}B_T - \frac{\nu^2}{2\sigma^2}T\right) I_{\{B_{t_1}\in dx_1\}\cap\cdots\cap\{B_{t_n}\in dx_n\}}\right) \\
&= \exp\left(\frac{\nu}{\sigma^2}x - \frac{\nu^2}{2\sigma^2}T\right)\prod_{i=1}^n\left\{\frac{1}{\sqrt{2\pi\sigma^2(t_i - t_{i-1})}}\exp\left(-\frac{(x_i - x_{i-1})^2}{2\sigma^2(t_i - t_{i-1})}\right)dx_i\right\} \\
&= \prod_{i=1}^n\left\{\frac{1}{\sqrt{2\pi\sigma^2(t_i - t_{i-1})}}\exp\left(-\frac{(x_i - x_{i-1} - \nu(t_i - t_{i-1}))^2}{2\sigma^2(t_i - t_{i-1})}\right)dx_i\right\}.
\end{aligned}$$

It follows that, under $\mathbb{Q}$, the sequence $B_{t_1}, B_{t_2}, \dots, B_{t_n}$ is distributed in the manner of a Wiener process with drift ν and instantaneous variance σ^2. Since this holds for all sequences $t_1, t_2, \dots, t_n$ and since B has continuous sample paths, the claim of the theorem follows. ■

With W the usual standard Wiener process, and $\nu \in \mathbb{R}$, there exists by Theorem (7) a probability measure $\mathbb{Q}_\nu$ under which σW is a Wiener process with drift ν and instantaneous variance σ^2. Therefore, under $\mathbb{Q}_\nu$, the process $\widetilde{W}$ given by $\sigma\widetilde{W}_t = -\nu t + \sigma W_y$ is a standard Wiener process. By equation (2) and the final observation of Example (13.9.9), under $\mathbb{Q}_\nu$ the process

$$e^{-rt}S_t = \exp\big((\mu - \tfrac{1}{2}\sigma^2 - r)t + \sigma W_t\big) = \exp\big((\mu - \tfrac{1}{2}\sigma^2 - r + \nu)t + \sigma\widetilde{W}_t\big)$$

is a diffusion with instantaneous mean and variance $a(t, x) = (\mu - r + \nu)x$ and $b(t, x) = \sigma^2 x^2$. By Example (12.7.10), it is a martingale under $\mathbb{Q}_\nu$ if $\mu - r + \nu = 0$, and we set $\nu = r - \mu$ accordingly, and write $\mathbb{Q} = \mathbb{Q}_\nu$. The fact that there exists a measure $\mathbb{Q}$ under which $e^{-rt}S_t$ is a martingale is pivotal for the solution to this problem and its generalizations

It is a consequence that, under $\mathbb{Q}$, $e^{-rt}V_t$ constitutes a martingale. This may be seen as follows. By the product rule of Example (13.9.8),

$$
\begin{aligned}
\textbf{(8)} \qquad d(e^{-rt}V_t) &= e^{-rt}\,dV_t - re^{-rt}V_t\,dt \\
&= e^{-rt}\alpha_t\,dS_t - re^{-rt}\alpha_t S_t\,dt + e^{-rt}\beta_t(dM_t - rM_t) && \text{by (4) and (3)} \\
&= \alpha_t e^{-rt}S_t\big((\mu - r)\,dt + \sigma\,dW_t\big) && \text{by (1) and (2)} \\
&= \alpha_t e^{-rt}S_t(-\nu\,dt + \sigma\,dW_t),
\end{aligned}
$$

where $\nu = r - \mu$ as above. Under $\mathbb{Q}$, σW is a Wiener process with drift ν and instantaneous variance σ^2, whence $\sigma\widetilde{W} = -\nu t + \sigma W$ is a Wiener process with drift 0 and instantanteous variance σ^2. By (8),

$$
e^{-rt}V_t = V_0 + \int_0^t \alpha_u e^{-ru}S_u\sigma\,d\widetilde{W}_u
$$

which, by Theorem (13.8.11), defines a martingale under $\mathbb{Q}$. Now V_t equals the value of the European call option at time t and, by the martingale property,

$$
\textbf{(9)} \qquad V_t = e^{rt}(e^{-rt}V_t) = e^{rt}\mathbb{E}_{\mathbb{Q}}(e^{-rT}V_T \mid \mathcal{F}_t)
$$

where $\mathbb{E}_{\mathbb{Q}}$ denotes expectation with respect to $\mathbb{Q}$. The right side of (9) may be computed via the result of Exercise (13.10.1), leading to the following useful form of the value† of the option.

(10) Theorem. Black–Scholes formula. *Let $t < T$. The value at time t of the European call option is*

$$
\textbf{(11)} \qquad S_t\Phi(d_1(t, S_t)) - Ke^{-r(T-t)}\Phi(d_2(t, S_t))
$$

where Φ is the $N(0,1)$ distribution function and

$$
\textbf{(12)} \qquad d_1(t,x) = \frac{\log(x/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2(t,x) = d_1(t,x) - \sigma\sqrt{T-t}.
$$

Note that the Black–Scholes formula depends on the price process through r and σ^2 and not through the value of μ. A similar formula may be derived for any adapted contingent claim having finite second moment.

The discussion prior to the theorem does not constitute a full proof, since it was based on the assumption that there exists a self-financing strategy which replicates the European call option. In order to prove the Black–Scholes formula, we shall show the existence of a self-financing replicating strategy with value function (11). This portfolio may be identified from (11), since (11) is the value function of the portfolio (α, β) given by

$$
\textbf{(13)} \qquad \alpha_t = \Phi(d_1(t, S_t)), \quad \beta_t = -Ke^{-rT}\Phi(d_2(t, S_t)).
$$

†The value given in Theorem (10) is sometimes called the 'no arbitrage value' or the 'risk-neutral value' of the option.

Let $\xi(t, x)$, $\psi(t, x)$ be smooth functions of the real variables t, x, and consider the portfolio denoted (ξ, ψ) which at time t holds $\xi(t, S_t)$ units in stock and $\psi(t, S_t)$ units in the bond. This portfolio has value function $W_t(\xi, \psi) = w(t, S_t)$ where

$$w(t, x) = \xi(t, x)x + \psi(t, x)e^{rt}. \tag{14}$$

(15) Theorem. *Let ξ, ψ be such that the function w given by* (14) *is twice continuously differentiable. The portfolio (ξ, ψ) is self-financing if and only if:*

$$x\xi_x + e^{rt}\psi_x = 0, \tag{16}$$

$$\tfrac{1}{2}\sigma^2 x^2 \xi_x + x\xi_t + e^{rt}\psi_t = 0, \tag{17}$$

where f_x, f_t denote derivatives with respect to x, t.

Proof. We apply Itô's formula (13.9.2) to the function w of equation (14) to find via (2) that

$$dw(t, S_t) = w_x(t, S_t)\,dS_t + \left\{w_t(t, S_t) + \tfrac{1}{2}\sigma^2 S_t^2 w_{xx}(t, S_t)\right\}dt$$

whereas, by (4) and (1), (ξ, ψ) is self-financing if and only if

$$dw(t, S_t) = \xi(t, S_t)\,dS_t + r\psi(t, S_t)e^{rt}\,dt. \tag{18}$$

Equating coefficients of the infinitesimals, we deduce that (ξ, ψ) is self-financing if and only if $\xi = w_x$ and $r\psi e^{rt} = w_t + \frac{1}{2}\sigma^2 x^2 w_{xx}$, which is to say that:

$$\xi = \xi + \xi_x x + \psi_x e^{rt}, \tag{19}$$

$$r\psi e^{rt} = \xi_t x + \psi_t e^{rt} + r\psi e^{rt} + \tfrac{1}{2}\sigma^2 x^2(\xi_{xx}x + 2\xi_x + \psi_{xx}e^{rt}). \tag{20}$$

Differentiating (19) with respect to x yields

$$0 = \xi_{xx}x + \xi_x + \psi_{xx}e^{rt}, \tag{21}$$

which may be inserted into (20) to give as required that

$$0 = \xi_t x + \psi_t e^{rt} + \tfrac{1}{2}\sigma^2 x^2 \xi_x.$$ ∎

Theorem (15) leads to the following characterization of value functions of self-financing portfolios.

(22) Corollary. Black–Scholes equation. *Suppose that $w(t, x)$ is twice continuously differentiable. Then $w(t, S_t)$ is the value function of a self-financing portfolio if and only if*

$$\tfrac{1}{2}\sigma^2 x^2 w_{xx} + rxw_x + w_t - rw = 0. \tag{23}$$

The Black–Scholes equation provides a means for finding self-financing portfolios which replicate general contingent claims. One 'simply' solves equation (23) subject to the boundary condition imposed by the particular claim in question. In the case of the European call option, the appropriate boundary condition is $w(T, x) = (x - K)^+$. It is not always easy to find the solution, but there is a general method known as the 'Feynman–Kac formula', not discussed further here, which allows a representation of the solution in terms of a diffusion process. When the solution exists, the claim is said to be 'hedgeable', and the self-financing portfolio which replicates it is called the 'hedge'.

Proof. Assume that w satisfies (23), and set

$$\xi = w_x, \quad \psi = e^{-rt}(w - xw_x).$$

It is easily checked that the portfolio (ξ, ψ) has value function $w(t, S_t)$ and, via (23), that the pair ξ, ψ satisfy equations (16) and (17).

Conversely, if $w(t, S_t)$ is the value of a self-financing portfolio then $w(t, x) = \xi(t, x)x + \psi(t, x)e^{rt}$ for some pair ξ, ψ satisfying equations (16) and (17). We compute w_x and compare with (16) to find that $\xi = w_x$. Setting $\psi = e^{-rt}(w - xw_x)$, we substitute into (17) to deduce that equation (23) holds. ■

Proof of Theorem (10). Finally we return to the proof of the Black–Scholes formula, showing first that the portfolio (α, β), given in equations (13), is self-financing. Set

$$\alpha(t, x) = \Phi(d_1(t, x)), \quad \beta(t, x) = -Ke^{-rT}\Phi(d_2(t, x)),$$

where d_1 and d_2 are given in (12). We note from (12) that

(24) $$d_2^2 = d_1^2 - 2\log(x/K) - 2r(T - t),$$

and it is straightforward to deduce by substitution that the pair α, β satisfy equations (16) and (17). Therefore, the portfolio (α, β) is self-financing. By construction, it has value function $V_t(\alpha, \beta)$ given in (11).

We may take the limit in (11) as $t \uparrow T$. Since

$$d_i(t, S_t) \to \begin{cases} -\infty & \text{if } S_T < K, \\ \infty & \text{if } S_T > K, \end{cases}$$

for $i = 1, 2$, we deduce that $V_T(\alpha, \beta) = (S_T - K)^+$ whenever $S_T \neq K$. Now $\mathbb{P}(S_T = K) = 0$, and therefore $V_T(\alpha, \beta) = (S_T - K)^+$ almost surely. It follows as required that the portfolio (α, β) replicates the European call option. ■

Exercises for Section 13.10

In the absence of any contrary indication, W denotes a standard Wiener process, and $\mathcal{F}_t$ is the smallest σ-field containing all null events with respect to which every member of $\{W_u : 0 \le u \le t\}$ is measurable. The process $S_t = \exp((\mu - \frac{1}{2}\sigma^2) + \sigma W_t)$ is a geometric Brownian motion, and $r \ge 0$ is the interest rate.

1. (a) Let Z have the $N(\gamma, \tau^2)$ distribution. Show that

$$\mathbb{E}\big((ae^Z - K)^+\big) = ae^{\gamma + \frac{1}{2}\tau^2}\Phi\left(\frac{\log(a/K) + \gamma}{\tau} + \tau\right) - K\Phi\left(\frac{\log(a/K) + \gamma}{\tau}\right)$$

where Φ is the $N(0, 1)$ distribution function.

(b) Let $\mathbb{Q}$ be a probability measure under which σW is a Wiener process with drift $r - \mu$ and instantaneous variance σ^2. Show for $0 \le t \le T$ that

$$\mathbb{E}_{\mathbb{Q}}\big((S_T - K)^+ \mid \mathcal{F}_t\big) = S_t e^{r(T-t)} \Phi(d_1(t, S_t)) - K \Phi(d_2(t, S_t))$$

where

$$d_1(t, x) = \frac{\log(x/K) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}, \quad d_2(t, x) = d_1(t, x) - \sigma\sqrt{T - t}.$$

2. Consider a portfolio which, at time t, holds $\xi(t, S)$ units of stock and $\psi(t, S)$ units of bond, and assume these quantities depend only on the values of S_u for $0 \le u \le t$. Find the function ψ such that the portfolio is self-financing in the three cases:
(a) $\xi(t, S) = 1$ for all t, S,
(b) $\xi(t, S) = S_t$,
(c) $\xi(t, S) = \int_0^t S_v \, dv$.

3. Suppose the stock price S_t is itself a Wiener process and the interest rate r equals 0, so that a unit of bond has unit value for all time. In the notation of Exercise (2), which of the following define self-financing portfolios?
(a) $\xi(t, S) = \psi(t, S) = 1$ for all t, S,
(b) $\xi(t, S) = 2S_t$, $\psi(t, S) = -S_t^2 - t$,
(c) $\xi(t, S) = -t$, $\psi(t, S) = \int_0^t S_s \, ds$,
(d) $\xi(t, S) = \int_0^t S_s \, ds$, $\psi(t, S) = -\int_0^t S_s^2 \, ds$.

4. An 'American call option' differs from a European call option in that it may be exercised by the buyer *at any time up to the expiry date*. Show that the value of the American call option is the same as that of the corresponding European call option, and that there is no advantage to the holder of such an option to exercise it strictly before its expiry date.

5. Show that the Black–Scholes value at time 0 of the European call option is an increasing function of the initial stock price, the exercise date, the interest rate, and the volatility, and is a decreasing function of the strike price.

13.11 Passage probabilities and potentials

In this final section, we study in a superficial way a remarkable connection between probability theory and classical analysis, namely the relationship between the sample paths of a Wiener process and the Newtonian theory of gravitation.

We begin by recalling some fundamental facts from the theory of scalar potentials. Let us assume that matter is distributed about regions of $\mathbb{R}^d$. According to the laws of Newtonian attraction, this matter gives rise to a function $\phi : \mathbb{R}^d \to \mathbb{R}$ which assigns a *potential* $\phi(\mathbf{x})$ to each point $\mathbf{x} = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$. In regions of space which are empty of matter, the potential function ϕ satisfies

(1) *Laplace's equation*: $\nabla^2 \phi = 0$,

where the *Laplacian* $\nabla^2 \phi$ is given by

(2) $$\nabla^2 \phi = \sum_{i=1}^{d} \frac{\partial^2 \phi}{\partial x_i^2}.$$

It is an important application of Green's theorem that solutions to Laplace's equation are also solutions to a type of integral equation. We make this specific as follows. Let $\mathbf{x}$ lie in the interior of a region R of space which is empty of matter, and consider a ball B contained in R with radius a and centre $\mathbf{x}$. The potential $\phi(\mathbf{x})$ at the centre of B is the average of the potential over the surface Σ of B. That is to say, $\phi(\mathbf{x})$ may be expressed as the surface integral

(3) $$\phi(\mathbf{x}) = \int_{y \in \Sigma} \frac{\phi(\mathbf{y})}{4\pi a^2}\, dS.$$

Furthermore, ϕ satisfies (3) for all such balls if and only if ϕ is a solution to Laplace's equation (1) in the appropriate region.

We turn now to probabilities. Let $\mathbf{W}(t) = (W_1(t), W_2(t), \dots, W_d(t))$ be a d-dimensional Wiener process describing the position of a particle which diffuses around $\mathbb{R}^d$, so that the W_i are independent one-dimensional Wiener processes. We assume that $\mathbf{W}(0) = \mathbf{w}$ and that the W_i have variance parameter σ^2. The vector $\mathbf{W}(t)$ contains d random variables with joint density function

(4) $$f_{\mathbf{W}(t)}(\mathbf{x}) = \frac{1}{(2\pi\sigma^2 t)^{d/2}} \exp\left(-\frac{1}{2\sigma^2 t}|\mathbf{x} - \mathbf{w}|^2\right), \quad \mathbf{x} \in \mathbb{R}^d.$$

Let H, J be disjoint subsets of $\mathbb{R}^d$ which are 'nice' in some manner which we will not make specific. Suppose the particle starts from $\mathbf{W}(0) = \mathbf{w}$, and let us ask for the probability that it visits some point of H before it visits any point of J. A particular case of this question might arise as follows. Suppose that $\mathbf{w}$ is a point in the interior of some closed bounded connected domain D of $\mathbb{R}^d$, and suppose that the surface ∂D which bounds D is fairly smooth (if D is a ball then ∂D is the bounding spherical surface, for example). Sooner or later the particle will enter ∂D for the first time. If $\partial D = H \cup J$ for some disjoint sets H and J, then we may ask for the probability that the particle enters ∂D at a point in H rather than at a point in J (as an example, take D to be the ball of radius 1 and centre $\mathbf{w}$, and let H be a hemisphere of D).

In the above example, the process was bound (almost surely) to enter $H \cup J$ at some time. This is not true for general regions H, J. For example, the hitting time of a point in $\mathbb{R}^2$ is almost surely infinite, and we shall see that the hitting time of a sphere in $\mathbb{R}^3$ is infinite with strictly positive probability if the process starts outside the sphere. In order to include all eventualities, we introduce the hitting time $T_A = \inf\{t : \mathbf{W}(t) \in A\}$ of the subset A of $\mathbb{R}^d$, with the usual convention that the infimum of the empty set is ∞. We write $\mathbb{P}_\mathbf{w}$ for the probability measure which governs the Wiener process $\mathbf{W}$ when it starts from $\mathbf{W}(0) = \mathbf{w}$.

(5) Theorem. *Let H and J be disjoint 'nice' subsets† of $\mathbb{R}^d$ such that $H \cup J$ is closed, and let $p(\mathbf{w}) = \mathbb{P}_\mathbf{w}(T_H < T_J)$. The function p satisfies Laplace's equation, $\nabla^2 p(\mathbf{w}) = 0$, at all points $\mathbf{w} \notin H \cup J$, with the boundary conditions*

$$p(\mathbf{w}) = \begin{cases} 1 & \text{if } \mathbf{w} \in H, \\ 0 & \text{if } \mathbf{w} \in J. \end{cases}$$

†We do not explain the condition that H and J be 'nice', but note that sets with smooth surfaces, such as balls or Platonic solids, are nice.

Proof. Let $\mathbf{w} \notin H \cup J$. Since $H \cup J$ is assumed closed, there exists a ball B contained in $\mathbb{R}^d \setminus (H \cup J)$ with centre $\mathbf{w}$. Let a be the radius of B and Σ its surface. Let $T = \inf\{t : \mathbf{W}(t) \in \Sigma\}$ be the first passage time of $\mathbf{W}$ to the set Σ. The random variable T is a stopping time for $\mathbf{W}$, and it is not difficult to see as follows that $\mathbb{P}_{\mathbf{w}}(T < \infty) = 1$. Let $A_i = \{|\mathbf{W}(i) - \mathbf{W}(i-1)| \leq 2a\}$ and note that $\mathbb{P}_{\mathbf{w}}(A_1) < 1$, whence

$$\begin{aligned} \mathbb{P}_{\mathbf{w}}(T > n) &\leq \mathbb{P}_{\mathbf{w}}(A_1 \cap A_2 \cap \cdots \cap A_n) \\ &= \mathbb{P}_{\mathbf{w}}(A_1)^n \qquad \text{by independence} \\ &\to 0 \qquad\qquad\quad \text{as } n \to \infty. \end{aligned}$$

We now condition on the hitting point $W(T)$. By the strong Markov property, given $W(T)$, the path of the process after time T is a Wiener process with the new starting point $\mathbf{W}(T)$. It follows that the (conditional) probability that $\mathbf{W}$ visits H before it visits J is $p(\mathbf{W}(T))$, and we are led to the following formula:

$$\text{(6)} \qquad p(\mathbf{w}) = \int_{\mathbf{y} \in \Sigma} \mathbb{P}_{\mathbf{w}}\big(T_H < T_J \,\big|\, \mathbf{W}(T) = \mathbf{y}\big) f_{\mathbf{w}}(\mathbf{y})\, dS$$

where $f_{\mathbf{w}}$ is the conditional density function of $\mathbf{W}(T)$ given $\mathbf{W}(0) = \mathbf{w}$. Using the spherical symmetry of the density function in (4), we have that $W(T)$ is uniformly distributed on Σ, which is to say that

$$f_{\mathbf{w}}(\mathbf{y}) = \frac{1}{4\pi a^2} \quad \text{for all } \mathbf{y} \in \Sigma,$$

and equation (6) becomes

$$\text{(7)} \qquad p(\mathbf{w}) = \int_{\mathbf{y} \in \Sigma} \frac{p(\mathbf{y})}{4\pi a^2}\, dS.$$

This integral equation holds for any ball B with centre $\mathbf{w}$ whose contents do not overlap $H \cup J$, and we recognize it as the characteristic property (3) of solutions to Laplace's equation (1). Thus p satisfies Laplace's equation. The boundary conditions are derived easily. ■

Theorem (5) provides us with an elegant technique for finding the probabilities that $\mathbf{W}$ visits certain subsets of $\mathbb{R}^d$. The principles of the method are simple, although some of the ensuing calculations may be lengthy since the difficulty of finding explicit solutions to Laplace's equation depends on the boundary conditions (see Example (14) and Problem (13.12.12), for instance).

(8) Example. Take $d = 2$, and start a two-dimensional Wiener process $\mathbf{W}$ at a point $\mathbf{W}(0) = \mathbf{w} \in \mathbb{R}^2$. Let H be a circle with radius ϵ (> 0) and centre at the origin, such that $\mathbf{w}$ does not lie within the inside of H. What is the probability that $\mathbf{W}$ ever visits H?

Solution. We shall need two boundary conditions in order to find the appropriate solution to Laplace's equation. The first arises from the case when $\mathbf{w} \in H$. To find the second, we introduce a second circle J, with radius R and centre at the origin, and suppose that R is much larger than ϵ. We shall solve Laplace's equation in polar coordinates,

$$\text{(9)} \qquad \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 p}{\partial \theta^2} = 0,$$

in the region $\epsilon \le r \le R$, and use the boundary conditions

(10)
$$p(\mathbf{w}) = \begin{cases} 1 & \text{if } \mathbf{w} \in H, \\ 0 & \text{if } \mathbf{w} \in J. \end{cases}$$

Solutions to equation (9) having circular symmetry take the form

$$p(\mathbf{w}) = A \log r + B \quad \text{if} \quad \mathbf{w} = (r, \theta),$$

where A and B are arbitrary constants. We use the boundary conditions to obtain the solution

$$p_R(\mathbf{w}) = \frac{\log(r/R)}{\log(\epsilon/R)}, \qquad \epsilon \le r \le R,$$

and we deduce by Theorem (5) that $\mathbb{P}_{\mathbf{w}}(T_H < T_J) = p_R(\mathbf{w})$.

In the limit as $R \to \infty$, we have that $T_J \to \infty$ almost surely, whence

$$p_R(\mathbf{w}) \to \mathbb{P}_{\mathbf{w}}(T_H < \infty) = 1.$$

We conclude that $\mathbf{W}$ almost surely visits any ϵ-neighbourhood of the origin regardless of its starting point. Such a process is called *persistent* (or *recurrent*) since its sample paths pass arbitrarily closely to every point in the plane with probability 1. ●

(11) Example. We consider next the same question as Example (8) but in three dimensions. Let H be the sphere with radius ϵ and centre at the origin of $\mathbb{R}^3$. We start a three-dimensional Wiener process $\mathbf{W}$ from some point $\mathbf{W}(0) = \mathbf{w}$ which does not lie within H. What is the probability that $\mathbf{W}$ visits H?

Solution. As before, let J be a sphere with radius R and centre at the origin, where R is much larger than ϵ. We seek a solution to Laplace's equation in spherical polar coordinates

(12)
$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial p}{\partial\theta}\right) + \frac{1}{\sin^2\phi}\frac{\partial^2 p}{\partial\phi^2} = 0,$$

subject to the boundary conditions (10). Solutions to equation (12) with spherical symmetry have the form

(13)
$$p(\mathbf{w}) = \frac{A}{r} + B \quad \text{if} \quad \mathbf{w} = (r, \theta, \phi).$$

We use the boundary conditions (10) to obtain the solution

$$p_R(\mathbf{w}) = \frac{r^{-1} - R^{-1}}{\epsilon^{-1} - R^{-1}}.$$

Let $R \to \infty$ to obtain by Theorem (5) that

$$p_R(\mathbf{w}) \to \mathbb{P}(T_H < \infty) = \frac{\epsilon}{r}, \qquad r > \epsilon.$$

That is to say, $\mathbf{W}$ ultimately visits H with probability ϵ/r. It is perhaps striking that the answer is *directly* proportional to ϵ.

We have shown that the three-dimensional Wiener process is *not* persistent, since its sample paths do not pass through every ϵ-neighbourhood with probability 1. This mimics the behaviour of symmetric random walks; recall from Problems (5.12.6) and (6.15.9) that the two-dimensional symmetric random walk is persistent whilst the three-dimensional walk is transient. ●

(14) Example. Let Σ be the surface of the unit sphere in $\mathbb{R}^3$ with centre at the origin, and let

$$H = \left\{(r, \theta, \phi) : r = 1,\ 0 \le \theta \le \tfrac{1}{2}\pi\right\}$$

be the upper hemisphere of Σ. Start a three-dimensional Wiener process $\mathbf{W}$ from a point $\mathbf{W}(0) = \mathbf{w}$ which lies in the *inside* of Σ. What is the probability that $\mathbf{W}$ visits H before it visits $J = \Sigma \setminus H$, the lower hemisphere of Σ?

Solution. The function $p(\mathbf{w}) = \mathbb{P}_{\mathbf{w}}(T_H < T_J)$ satisfies Laplace's equation (12) subject to the boundary conditions (10). Solutions to (12) which are independent of ϕ are also solutions to the simpler equation

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial p}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial p}{\partial\theta}\right) = 0.$$

We abandon the calculation at this point, leaving it to the reader to complete. Some knowledge of Legendre polynomials and the method of separation of variables may prove useful. ●

We may think of a Wiener process as a continuous space–time version of a symmetric random walk, and it is not surprising that Wiener processes and random walks have many properties in common. In particular, potential theory is of central importance to the theory of random walks. We terminate this chapter with a brief but electrifying demonstration of this.

Let $G = (V, E)$ be a finite connected graph with vertex set V and edge set E. For simplicity we assume that G has neither loops nor multiple edges. A particle performs a random walk about the vertex set V. If it is at vertex v at time n, then it moves at time $n+1$ to one of the neighbours of v, each such neighbour being chosen with equal probability, and independently of the previous trajectory of the particle. We write X_n for the position of the particle at time n, and $\mathbb{P}_w$ for the probability measure governing the X_n when $X_0 = w$.

For $A \subseteq V$, we define the passage time $T_A = \inf\{n : X_n \in A\}$. Let H and J be disjoint non-empty sets of vertices. We see by conditioning on X_1 that the function

(15) $$p(w) = \mathbb{P}_w(T_H < T_J)$$

satisfies the difference equation

$$p(w) = \sum_{x \in V} \mathbb{P}_w(X_1 = x)p(x) \qquad \text{for } w \notin H \cup J.$$

This expresses $p(w)$ as the average of the p-values of the neighbours of w:

(16) $$p(w) = \frac{1}{d(w)} \sum_{x : x \sim w} p(x) \qquad \text{for } w \notin H \cup J,$$

where $d(w)$ is the degree of the vertex w and we write $x \sim w$ to mean that x and w are neighbours. Equation (16) is the discrete analogue of the integral equation (7). The boundary conditions are given as before by (10).

Equations (16) have an interesting interpretation in terms of electrical network theory. We may think of G as an electrical network in which each edge is a resistor with resistance 1 ohm. We connect a battery into the network in such a way that the points in H are raised to the potential 1 volt and the points in J are joined to earth. It is physically clear that this potential difference induces a potential $\phi(w)$ at each vertex w, together with a current along each wire. These potentials and currents satisfy a well-known collection of equations called Kirchhoff's laws and Ohm's law, and it is an easy consequence of these laws (*exercise*) that ϕ is the unique solution to equations (16) subject to (10). It follows that

(17) $$\phi(w) = p(w) \quad \text{for all} \quad w \in V.$$

This equality between first passage probabilities and electrical potentials is the discrete analogue of Theorem (5).

As a beautiful application of this relationship, we shall show that random walk on an infinite connected graph is persistent if and only if the graph has *infinite* resistance when viewed as an electrical network.

Let $G = (V, E)$ be an infinite connected graph with countably many vertices and finite vertex degrees, and let 0 denote a chosen vertex of G. We may turn G into an (infinite) electrical network by replacing each edge by a unit resistor. For $u, v \in V$, let $d(u, v)$ be the number of edges in the shortest path joining u and v, and define $\Delta_n = \{v \in V : d(0, v) = n\}$. Let R_n be the electrical resistance between 0 and the set Δ_n. That is to say, $1/R_n$ is the current which flows in the circuit obtained by setting 0 to earth and applying a unit potential to the vertices in Δ_n. It is a standard fact from potential theory that $R_n \le R_{n+1}$, and we define the *resistance* of G to be the limit $R(G) = \lim_{n\to\infty} R_n$.

(18) Theorem. *A random walk on the graph G is persistent if and only if $R(G) = \infty$.*

Proof. Since G is connected with finite vertex degrees, a random walk on G is an irreducible Markov chain on the countable state space V. It suffices therefore to show that the vertex 0 is a persistent state of the chain. We write $\mathbb{P}_x$ for the law of the random walk started from $X_0 = x$.

Let ϕ_n be the potential function in the electrical network obtained from G by earthing 0 and applying unit potential to all vertices in Δ_n. Note that $0 \le \phi_n(x) \le 1$ for all vertices x (this is an application of what is termed the *maximum principle*). We have from the above discussion that $\phi_n(x) = \mathbb{P}_x(T_{\Delta_n} < T_0)$, where T_A denotes the first hitting time of the set A. Now T_{Δ_n} is at least the minimum distance from x to Δ_n, which is at least $n - d(0, x)$, and therefore $\mathbb{P}_x(T_{\Delta_n} \to \infty \text{ as } n \to \infty) = 1$ for all x. It follows that

(19) $$\phi_n(x) \to \mathbb{P}_x(T_0 = \infty) \quad \text{as } n \to \infty.$$

Applying Ohm's law to the edges incident with 0, we have that the total current flowing out of 0 equals

$$\sum_{x:x\sim 0} \phi_n(x) = \frac{1}{R_n}.$$

We let $n \to \infty$ and use equation (19) to find that

(20)
$$\sum_{x:x\sim 0} \mathbb{P}_x(T_0 = \infty) = \frac{1}{R(G)}$$

where $1/\infty$ is interpreted as 0.

We have by conditioning on X_1 that

$$\begin{aligned}\mathbb{P}_0(X_n = 0 \text{ for some } n \geq 1) &= \frac{1}{d(0)} \sum_{x:x\sim 0} \mathbb{P}_x(T_0 < \infty) \\ &= 1 - \frac{1}{d(0)} \sum_{x:x\sim 0} \mathbb{P}_x(T_0 = \infty) = 1 - \frac{1}{d(0)R(G)}.\end{aligned}$$

The claim follows. ■

(21) Theorem. Persistence of two-dimensional random walk. *Symmetric random walk on the two-dimensional square lattice $\mathbb{Z}^2$ is persistent.*

Proof. It suffices by Theorem (18) to prove that $R(\mathbb{Z}^2) = \infty$. We construct a lower bound for R_n in the following way. For each $r \leq n$, short out all the points in Δ_r (draw your own diagram), and use the parallel and series resistance laws to find that

$$R_n \geq \frac{1}{4} + \frac{1}{12} + \cdots + \frac{1}{8n-4}.$$

This implies that $R_n \to \infty$ as $n \to \infty$, and the result is shown. ■

(22) Theorem. Transience of three-dimensional random walk. *Symmetric random walk on the three-dimensional cubic lattice $\mathbb{Z}^3$ is transient.*

Proof. It is a non-trivial and interesting *exercise* to prove that $R(\mathbb{Z}^3) < \infty$. See the solution of Problem (6.15.9) for another method of proof. ■

Exercises for Section 13.11

1. Let G be the closed sphere with radius ϵ and centre at the origin of $\mathbb{R}^d$ where $d \geq 3$. Let $\mathbf{W}$ be a d-dimensional Wiener process starting from $\mathbf{W}(0) = \mathbf{w} \notin G$. Show that the probability that $\mathbf{W}$ visits G is $(\epsilon/r)^{d-2}$, where $r = |\mathbf{w}|$.

2. Let G be an infinite connected graph with finite vertex degrees. Let Δ_n be the set of vertices x which are distance n from 0 (that is, the shortest path from x to 0 contains n edges), and let N_n be the total number of edges joining pairs x, y of vertices with $x \in \Delta_n$, $y \in \Delta_{n+1}$. Show that a random walk on G is persistent if $\sum_i N_i^{-1} = \infty$.

3. Let G be a connected graph with finite vertex degrees, and let H be a connected subgraph of G. Show that a random walk on H is persistent if a random walk on G is persistent, but that the converse is not generally true.

13.12 Problems

1. Let W be a standard Wiener process, that is, a process with independent increments and continuous sample paths such that $W(s+t) - W(s)$ is $N(0, t)$ for $t > 0$. Let α be a positive constant. Show that:
(a) $\alpha W(t/\alpha^2)$ is a standard Wiener process,
(b) $W(t+\alpha) - W(\alpha)$ is a standard Wiener process,
(c) the process V, given by $V(t) = tW(1/t)$ for $t > 0$, $V(0) = 0$, is a standard Wiener process.

2. Let $X = \{X(t) : t \geq 0\}$ be a Gaussian process with continuous sample paths, zero means, and autocovariance function $c(s,t) = u(s)v(t)$ for $s \leq t$ where u and v are continuous functions. Suppose that the ratio $r(t) = u(t)/v(t)$ is continuous and strictly increasing with inverse function r^{-1}. Show that $W(t) = X(r^{-1}(t))/v(r^{-1}(t))$ is a standard Wiener process on a suitable interval of time.

If $c(s,t) = s(1-t)$ for $s \leq t < 1$, express X in terms of W.

3. Let $\beta > 0$, and show that $U(t) = e^{-\beta t} W(e^{2\beta t} - 1)$ is an Ornstein–Uhlenbeck process if W is a standard Wiener process.

4. Let $V = \{V(t) : t \geq 0\}$ be an Ornstein–Uhlenbeck process with instantaneous mean $a(t,x) = -\beta x$ where $\beta > 0$, with instantaneous variance $b(t,x) = \sigma^2$, and with $U(0) = u$. Show that $V(t)$ is $N(ue^{-\beta t}, \sigma^2(1 - e^{-2\beta t})/(2\beta))$. Deduce that $V(t)$ is asymptotically $N(0, \frac{1}{2}\sigma^2/\beta)$ as $t \to \infty$, and show that V is strongly stationary if $V(0)$ is $N(0, \frac{1}{2}\sigma^2/\beta)$.

Show that such a process is the *only* stationary Gaussian Markov process with continuous autocovariance function, and find its spectral density function.

5. Let $D = \{D(t) : t \geq 0\}$ be a diffusion process with instantaneous mean $a(t,x) = \alpha x$ and instantaneous variance $b(t,x) = \beta x$ where α and β are positive constants. Let $D(0) = d$. Show that the moment generating function of $D(t)$ is

$$M(t,\theta) = \exp\left\{ \frac{2\alpha d\theta e^{\alpha t}}{\beta\theta(1 - e^{\alpha t}) + 2\alpha} \right\}.$$

Find the mean and variance of $D(t)$, and show that $\mathbb{P}(D(t) = 0) \to e^{-2d\alpha/\beta}$ as $t \to \infty$.

6. Let D be an Ornstein–Uhlenbeck process with $D(0) = 0$, and place reflecting barriers at $-c$ and d where $c, d > 0$. Find the limiting distribution of D as $t \to \infty$.

7. Let $X_0, X_1, \ldots$ be independent $N(0,1)$ variables, and show that

$$W(t) = \frac{t}{\sqrt{\pi}} X_0 + \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} \frac{\sin(kt)}{k} X_k$$

defines a standard Wiener process on $[0, \pi]$.

8. Let W be a standard Wiener process with $W(0) = 0$. Place absorbing barriers at $-b$ and b, where $b > 0$, and let W^{a} be W absorbed at these barriers. Show that $W^{\mathrm{a}}(t)$ has density function

$$f^{\mathrm{a}}(y,t) = \frac{1}{\sqrt{2\pi t}} \sum_{k=-\infty}^{\infty} (-1)^k \exp\left\{ -\frac{(y - 2kb)^2}{2t} \right\}, \qquad -b < y < b,$$

which may also be expressed as

$$f^{\mathrm{a}}(y,t) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n t} \sin\left(\frac{n\pi(y+b)}{2b} \right), \qquad -b < y < b,$$

where $a_n = b^{-1}\sin(\frac{1}{2}n\pi)$ and $\lambda_n = n^2\pi^2/(8b^2)$.

Hence calculate $\mathbb{P}(\sup_{0\le s\le t} |W(s)| > b)$ for the unrestricted process W.

9. Let D be a Wiener process with drift m, and suppose that $D(0) = 0$. Place absorbing barriers at the points $x = -a$ and $x = b$ where a and b are positive real numbers. Show that the probability p_a that the process is absorbed at $-a$ is given by

$$p_a = \frac{e^{2mb} - 1}{e^{2m(a+b)} - 1}.$$

10. Let W be a standard Wiener process and let $F(u, v)$ be the event that W has no zero in the interval (u, v).
(a) If $ab > 0$, show that $\mathbb{P}\big(F(0, t) \mid W(0) = a,\ W(t) = b\big) = 1 - e^{-2ab/t}$.
(b) If $W(0) = 0$ and $0 < t_0 \le t_1 \le t_2$, show that

$$\mathbb{P}\big(F(t_0, t_2) \mid F(t_0, t_1)\big) = \frac{\sin^{-1}\sqrt{t_0/t_2}}{\sin^{-1}\sqrt{t_0/t_1}}.$$

(c) Deduce that, if $W(0) = 0$ and $0 < t_1 \le t_2$, then $\mathbb{P}(F(0, t_2) \mid F(0, t_1)) = \sqrt{t_1/t_2}$.

11. Let W be a standard Wiener process. Show that

$$\mathbb{P}\left(\sup_{0\le s\le t} |W(s)| \ge w\right) \le 2\mathbb{P}(|W(t)| \ge w) \le \frac{2t}{w^2} \qquad \text{for } w > 0.$$

Set $t = 2^n$ and $w = 2^{2n/3}$ and use the Borel–Cantelli lemma to show that $t^{-1}W(t) \to 0$ a.s. as $t \to \infty$.

12. Let $\mathbf{W}$ be a two-dimensional Wiener process with $\mathbf{W}(0) = \mathbf{w}$, and let F be the unit circle. What is the probability that $\mathbf{W}$ visits the upper semicircle G of F before it visits the lower semicircle H?

13. Let W_1 and W_2 be independent standard Wiener processes; the pair $\mathbf{W}(t) = (W_1(t), W_2(t))$ represents the position of a particle which is experiencing Brownian motion in the plane. Let l be some straight line in $\mathbb{R}^2$, and let P be the point on l which is closest to the origin O. Draw a diagram. Show that
(a) the particle visits l, with probability one,
(b) if the particle hits l for the first time at the point R, then the distance PR (measured as positive or negative as appropriate) has the Cauchy density function $f(x) = d/\{\pi(d^2+x^2)\}$, $-\infty < x < \infty$, where d is the distance OP,
(c) the angle $\widehat{\text{POR}}$ is uniformly distributed on $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$.

14. Let $\phi(x + iy) = u(x, y) + iv(x, y)$ be an analytic function on the complex plane with real part $u(x, y)$ and imaginary part $v(x, y)$, and assume that

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = 1.$$

Let (W_1, W_2) be the planar Wiener process of Problem (13) above. Show that the pair $u(W_1, W_2)$, $v(W_1, W_2)$ is also a planar Wiener process.

15. Let $M(t) = \max_{0\le s\le t} W(s)$, where W is a standard Wiener process. Show that $M(t) - W(t)$ has the same distribution as $M(t)$.

16. Let W be a standard Wiener process, $u \in \mathbb{R}$, and let $Z = \{t : W(t) = u\}$. Show that Z is a null set (i.e., has Lebesgue measure zero) with probability one.

17. Let $M(t) = \max_{0\le s\le t} W(s)$, where W is a standard Wiener process. Show that $M(t)$ is attained at exactly one point in $[0, t]$, with probability one.

18. Sparre Andersen theorem. Let $s_0 = 0$ and $s_m = \sum_{j=1}^{m} x_j$, where $(x_j : 1 \le j \le n)$ is a given sequence of real numbers. Of the $n!$ permutations of $(x_j : 1 \le j \le n)$, let A_r be the number of permutations in which exactly r values of $(s_m : 0 \le m \le n)$ are strictly positive, and let B_r be the number of permutations in which the maximum of $(s_m : 0 \le m \le n)$ first occurs at the rth place. Show that $A_r = B_r$ for $0 \le r \le n$. [Hint: Use induction on n.]

19. Arc sine laws. For the standard Wiener process W, let A be the amount of time u during the time interval $[0, t]$ for which $W(u) > 0$; let L be the time of the last visit to the origin before t; and let R be the time when W attains its maximum in $[0, t]$. Show that A, L, and R have the same distribution function $F(x) = (2/\pi)\sin^{-1}\sqrt{x/t}$ for $0 \le x \le t$. [Hint: Use the results of Problems (13.12.15)–(13.12.18).]

20. Let W be a standard Wiener process, and let U_x be the amount of time spent below the level x (≥ 0) during the time interval $(0, 1)$, that is, $U_x = \int_0^1 I_{\{W(t)<x\}}\,dt$. Show that U_x has density function

$$f_{U_x}(u) = \frac{1}{\pi\sqrt{u(1-u)}}\exp\left(-\frac{x^2}{2u}\right), \qquad 0 < u < 1.$$

Show also that

$$V_x = \begin{cases} \sup\{t \le 1 : W_t = x\} & \text{if this set is non-empty,} \\ 1 & \text{otherwise,} \end{cases}$$

has the same distribution as U_x.

21. Let $\operatorname{sign}(x) = 1$ if $x > 0$ and $\operatorname{sign}(x) = -1$ otherwise. Show that $V_t = \int_0^t \operatorname{sign}(W_s)\,dW_s$ defines a standard Wiener process if W is itself such a process.

22. After the level of an industrial process has been set at its desired value, it wanders in a random fashion. To counteract this the process is periodically reset to this desired value, at times $0, T, 2T, \dots$. If W_t is the deviation from the desired level, t units of time after a reset, then $\{W_t : 0 \le t < T\}$ can be modelled by a standard Wiener process. The behaviour of the process after a reset is independent of its behaviour before the reset. While W_t is outside the range $(-a, a)$ the output from the process is unsatisfactory and a cost is incurred at rate C per unit time. The cost of each reset is R. Show that the period T which minimises the long-run average cost per unit time is T^*, where

$$R = C\int_0^{T^*} \frac{a}{\sqrt{(2\pi t)}}\exp\left(-\frac{a^2}{2t}\right)dt.$$

23. An economy is governed by the Black–Scholes model in which the stock price behaves as a geometric Brownian motion with volatility σ, and there is a constant interest rate r. An investor likes to have a constant proportion γ ($\in (0, 1)$) of the current value of her self-financing portfolio in stock and the remainder in the bond. Show that the value function of her portfolio has the form $V_t = f(t)S_t^{\gamma}$ where $f(t) = c\exp\{(1-\gamma)(\frac{1}{2}\gamma\sigma^2 + r)t\}$ for some constant c depending on her initial wealth.

24. Let $u(t, x)$ be twice continuously differentiable in x and once in t, for $x \in \mathbb{R}$ and $t \in [0, T]$. Let W be the standard Wiener process. Show that u is a solution of the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$$

if and only if the process $U_t = u(T - t, W_t)$, $0 \le t \le T$, has zero drift.