

## Chapter 2

# Diffusion in Cells: Random Walks and Brownian Motion

When one first encounters the concept of diffusion, it is usually within the context of a conservation law describing the flux of many particles moving from regions of high concentration to regions of low concentration at a rate that depends on the local concentration gradient (Fick's law). However, there are some limitations of the standard macroscopic derivation of the diffusion equation. First, it does not take into account microscopic features of the environment within which the particles diffuse. This is crucial when considering diffusive processes within a cell, since the interior of the cell is highly heterogeneous (see Box 1A). The same applies to surface diffusion within the plasma membrane. Second, with the use of advanced imaging techniques such as SPT (Sect. 1.2), it is possible to observe the movement of individual molecules, which is highly stochastic, whereas the classical diffusion equation describes the collective motion of many particles and is deterministic.

In this chapter, we consider two different microscopic theories of diffusion: random walks and overdamped Brownian motion. Both approaches will be used to model diffusion within the complex cellular environment in Chap. 7. We begin by considering a discrete random walk on a 1D lattice, which is a simple example of a discrete Markov process (Sect. 2.1). The probability distribution specifying the likelihood that the walker is at a particular lattice site after  $n$  time steps evolves according to a master equation. We show how the master equation can be solved using discrete Fourier and Laplace transforms, which in probability theory are known as characteristic functions and generating functions, respectively. The resulting solution is given by a binomial distribution, which reduces to a Gaussian distribution in an appropriate continuum limit; the latter is the fundamental solution of the diffusion equation. Background material on Laplace and Fourier transforms, and their discrete analogs, is also provided. Random walk models and various generalizations will later be used to model a variety of cellular processes, including molecular motors, polymerization of cytoskeletal filaments (Chap. 4), and anomalous diffusion (Chap. 7).

We then consider an alternative microscopic theory of diffusion based on an overdamped Brownian particle moving in a fluid-like environment (such as the cytoplasm of a cell), which is modeled in terms of a Langevin equation or SDE

(Sect. 2.2). The latter describes the motion of the particle subject to a combination of external forces and a fluctuating force that is due to collisions with molecules in the surrounding fluid; the fluctuating force is idealized as a Wiener process. Solutions of the Langevin equation represent random sample paths or trajectories of the particle. We show how the probability density on the space of sample paths evolves according to a Fokker–Planck (FP) equation, which is a generalization of the diffusion equation applicable at the single-particle level. Other topics include the distinction between additive and multiplicative noise, Ito vs. Stratonovich interpretations of continuous stochastic processes, power spectra, and correlations. Note that continuous stochastic processes and the FP equation will appear in many chapters of the book. For in addition to describing diffusive-like motion of microscopic particles in solution, it also frequently appears in diffusion approximations of discrete Markov processes, where the continuous variable now represents the fraction of open ion channels (Chap. 3), say, or the concentration of a gene product (Chap. 6). A large-dimensional FP equation will be used to describe stochastic reaction–diffusion systems in Chap. 9 and applied to self-organizing phenomena such as cell polarization.

In Sect. 2.3 we introduce one of the most important characteristics of a diffusion process, namely, the FPT to reach a given target or boundary. This is then used to calculate the Smoluchowski reaction rate formula for diffusion-limited reactions, under the assumption that when reacting molecules come within a certain distance of each other they react immediately (Sect. 2.4). In Sect. 2.5 we tackle the general problem of diffusion in bounded domains (boundary value problems). Here we introduce some basic methods in the analysis of linear PDEs, including separation of variables and transform methods, eigenfunction expansions, and Green’s functions. Finally, in Sect. 2.6 we give an informal introduction to stochastic calculus and numerical methods for simulating continuous stochastic processes.

## 2.1 Discrete-Time Random Walk

Consider a particle that hops at discrete times between neighboring sites on a one-dimensional (1D) lattice with unit spacing [289, 651] (see Fig. 2.1). At each step, the random walker moves a unit distance to the right with probability  $p$  or to the left with probability  $q = 1 - p$ . Let  $P_N(r)$  denote the probability that the particle is at site  $r$  at the  $N$ th time step. The evolution of the probability distribution is described by the discrete-time master equation

$$P_N(r) = pP_{N-1}(r-1) + qP_{N-1}(r+1), \quad r \in \mathbb{Z}, \quad N \geq 1. \quad (2.1.1)$$

If  $q = p = 1/2$ , then the random walk is symmetric or unbiased, whereas for  $p > q$  ( $p < q$ ) it is biased to the right (left). We will analyze this equation using transform methods, since these can be generalized to more complex random walk models such as continuous-time random walks (see Sect. 7.1.3). An introduction to continuous

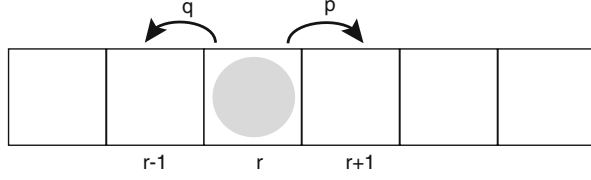


Fig. 2.1: A random walk on a 1D lattice

and discrete transform methods can be found in Box 2A. The first step is to introduce the *characteristic function* (discrete Fourier transform) for fixed  $N$ :

$$G_N(k) = \sum_{r=-\infty}^{\infty} e^{ikr} P_N(r), \quad k \in [-\pi, \pi]. \quad (2.1.2)$$

The characteristic function generates moments of the random displacement variable  $r$  according to

$$\left( -i \frac{d}{dk} \right)^m G_N(k) \Big|_{k=0} = \sum_{r=-\infty}^{\infty} r^m P_N(r) = \langle r^m \rangle, \quad (2.1.3)$$

where  $\langle r^m \rangle$  is the  $m$ th order moment of  $r$ . Multiplying both sides of the master equation by  $e^{ikr}$  and summing over  $r$  gives

$$G_N(k) = (pe^{ik} + qe^{-ik})G_{N-1}(k).$$

Assuming that the particle starts at the origin,  $P_0(r) = \delta_{r,0}$  and  $G_0(k) = 1$ , we have

$$G_N(k) = u(k)^N \quad u(k) = pe^{ik} + qe^{-ik}.$$

Here  $u(k)$  is the discrete Fourier transform of the single-step hopping probability. Finally, taking the inverse Fourier transform,

$$\begin{aligned} P_N(r) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikr} u(k)^N dk \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikr} \sum_{m=0}^N \binom{N}{m} p^m q^{N-m} e^{-ik(N-2m)} dk \\ &= \frac{N!}{\left(\frac{N+r}{2}\right)! \left(\frac{N-r}{2}\right)!} p^{(N+r)/2} q^{(N-r)/2} \end{aligned} \quad (2.1.4)$$

when  $N+r$  is an even integer and zero otherwise. We have used the result (see Box 2A)

$$\int_{-\pi}^{\pi} e^{-ik(N-2m+r)} \frac{dk}{2\pi} = \delta_{N+r,2m}.$$

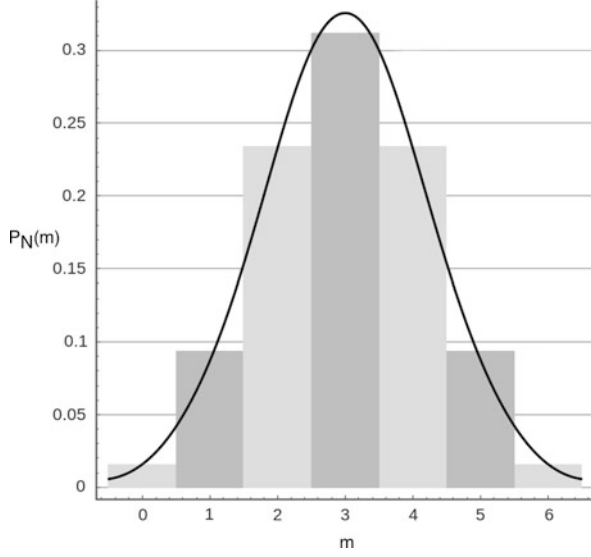


Fig. 2.2: Binomial distribution for  $N = 6$  and  $p = q = 1/2$ . Also shown is a Gaussian fit of the binomial distribution

The distribution (2.1.4) is known as the *binomial distribution* (see Fig. 2.2). In the unbiased case  $p = q = 1/2$ , it gives the probability of a total of  $r$  heads in tossing a fair coin  $N$  times and is known as the *Bernoulli distribution*.

Evaluating  $\log P_N(r)$  for large  $N$  using Stirling's approximation

$$\log N! \approx N \log N - N + \frac{1}{2} \log(2\pi N), \quad (2.1.5)$$

and assuming  $p, q \approx 1/2$ , one finds that (see Ex. 2.1 for the unbiased case)

$$P_N(r) \sim \frac{1}{\sqrt{2\pi N}} e^{-[r-N(p-q)]^2/2N}. \quad (2.1.6)$$

Indeed, the Gaussian form of  $P_N(r)$  in the long-time limit arises universally whenever the mean and variance of the displacement  $\Delta r = r - r'$  in a single step are finite, that is,

$$\langle \Delta r \rangle = \sum_{\Delta r} \Delta r p(\Delta r) < \infty, \quad \langle \Delta r^2 \rangle = \sum_{\Delta r} (\Delta r)^2 p(\Delta r) < \infty,$$

where  $p(\Delta r)$  is the probability of a step of length  $\Delta r$ . In the standard 1D random walk,  $\Delta r = \pm 1$  and  $p(1) = p, p(-1) = q$ . One way to see this is to note that  $u(k)$  has the small- $k$  series expansion

$$\begin{aligned}
u(k) &= \sum_{\Delta r} e^{ik\Delta r} p(\Delta r) \\
&= 1 + ik\langle\Delta r\rangle - \frac{1}{2}k^2\langle\Delta r^2\rangle + \dots \\
&\sim e^{ik\langle\Delta r\rangle - \frac{1}{2}k^2[\langle\Delta r^2\rangle - \langle\Delta r\rangle^2]}.
\end{aligned}$$

Substituting this approximation into the first line of equation (2.1.4) using the fact that the integral is dominated by the behavior in the region around  $k = 0$  when  $N$  is large, the resulting Gaussian integral yields the approximation

$$P_N(r) \sim \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-(r-N\langle\Delta r\rangle)^2/2N\sigma^2}, \quad (2.1.7)$$

with  $\sigma^2 = \langle\Delta r^2\rangle - \langle\Delta r\rangle^2$ . This result is a consequence of the central limit theorem [242] (see also Sect. 1.3).

Another useful quantity when analyzing random walks is the *generating function* (discrete Laplace transform or one-sided  $z$ -transform):

$$\Gamma(r, z) = \sum_{N=0}^{\infty} z^N P_N(r). \quad (2.1.8)$$

It is often simpler to evaluate the generating function in Fourier space,

$$\hat{\Gamma}(k, z) \equiv \sum_{r=-\infty}^{\infty} e^{ikr} \Gamma(r, z) = \sum_{N=0}^{\infty} z^N G_N(k),$$

assuming that we can reverse the order of summations. Since  $G_N(k) = u(k)^N$ , we can sum the resulting geometric series to obtain the result

$$\hat{\Gamma}(k, z) = \frac{1}{1 - zu(k)}.$$

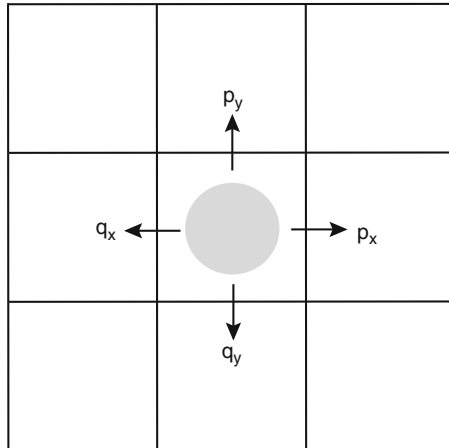


Fig. 2.3: A random walk on a 2D square lattice with  $p_x + q_x + p_y + q_y = 1$

The generating function is thus given by the inverse Fourier transform

$$\Gamma(r, z) = \int_{-\pi}^{\pi} \frac{e^{-ikr}}{1 - zu(k)} \frac{dk}{2\pi}.$$

It can be shown (see Ex. 2.2) that for  $r = 0$  and  $p = q = 1/2$  (unbiased random walk),

$$\Gamma(0, z) = (1 - z^2)^{-1/2}.$$

One immediate consequence of this result is that an unbiased 1D random walk is *recurrent*, which means that the walker is certain to return to the origin; a random walk is said to be *transient* if the probability of returning to the origin is less than one. Recurrence follows from the observation that  $\Gamma(0, 1) = \sum_{N=0}^{\infty} P_N(0)$  is the mean number of times that the walker visits the origin, and

$$\lim_{z \rightarrow 1^-} \Gamma(0, z) = \infty$$

for the 1D random walk. Interestingly, although the 1D random walk is recurrent, the mean time to return to the origin for the first time is infinite. This result can also be established using transform methods and generating functions (see Ex. 2.11). An unbiased random walk in 2D is also recurrent, but in 3D it is transient. An example of a 2D random walk is illustrated in Fig. 2.3. Finally, note that discrete random walks have also been used to describe the coiling of flexible polymers [53, 295] (see Sect. 4.5).

### Box 2A. Transform methods.

Throughout this book we will make extensive use of transform methods, in particular, Laplace and Fourier integral transforms and their discrete analogs. Here we provide a basic introduction to such methods (see also [395]).

**Laplace transforms.** Let  $u(t)$  be a piecewise continuous function that is of exponential order, that is,

$$u(t) \leq ce^{at}, \text{ as } t \rightarrow \infty,$$

for constants  $a, c > 0$ . The *Laplace transform* of  $u$  is defined by

$$\mathcal{L}u(s) \equiv \tilde{u}(s) = \int_0^{\infty} u(t)e^{-st} ds, \quad (2.1.9)$$

and one often writes  $\mathcal{L}u = \tilde{u}$ . The Laplace transform operator  $\mathcal{L}$  is linear, since

$$\mathcal{L}(c_1 u_1 + c_2 u_2) = c_1 \mathcal{L}u_1 + c_2 \mathcal{L}u_2$$

for constants  $c_1, c_2$ . One of the important features of the Laplace transform (and the Fourier transform) is that it converts differential operations

in the time domain into multiplication operations in the transform domain. For example, setting  $u' = du/dt$  etc.,

$$\mathcal{L}u'(s) = s\tilde{u}(s) - u(0) \quad (2.1.10a)$$

$$\mathcal{L}u''(s) = s^2\tilde{u}(s) - su(0) - u'(0), \quad (2.1.10b)$$

which can be proved using integration by parts. It follows that Laplace transforming an ordinary differential equation for  $u(t)$  yields an algebraic equation for  $\tilde{u}(s)$ . The most difficult step, once one has solved the algebraic equation, is to find the inverse Laplace transform to recover  $u(t)$ . The general formula for the inverse transform requires knowledge of contour integration and takes the form

$$u(t) = \mathcal{L}^{-1}\tilde{u}(t) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \tilde{u}(s)e^{st} ds. \quad (2.1.11)$$

The complex contour integral is taken over the infinite vertical line (the Bromwich path) in the complex plane from  $a - i\infty$  to  $a + i\infty$ . The real number  $a$  is chosen so that the Bromwich path lies to the right of any singularities (poles, branch points and cuts, essential points) of the function  $\tilde{u}(s)$ . The evaluation of the contour integral is often difficult. However, many of the Laplace transforms encountered in this book can be found in Table 2.1. One additional useful property of Laplace transforms is expressed by the *convolution theorem*.

**Theorem 2.1.** *Let  $u$  and  $v$  be piecewise continuous for  $t \geq 0$  and of exponential order. Then*

$$\mathcal{L}(u * v)(s) = \tilde{u}(s)\tilde{v}(s), \quad (2.1.12)$$

where

$$u * v(t) \equiv \int_0^t u(t-y)v(y)dy \quad (2.1.13)$$

is the convolution of  $u$  and  $v$ . It immediately follows that  $\mathcal{L}^{-1}(\tilde{u}\tilde{v}) = u * v$ .

In the case of a discrete-time linear process, we can use a discrete version of the Laplace transform (also known as a one-sided  $z$ -transform)

$$\tilde{u}(z) = \sum_{n=0}^{\infty} z^n u_n. \quad (2.1.14)$$

Applying this to the first-order difference equation  $u_n = au_{n-1}$  for  $n \geq 1$  yields

$$\tilde{u}(z) = az\tilde{u}(z) + u_0 \implies \tilde{u}(z) = \frac{u_0}{1-az} = u_0 \sum_{n=0}^{\infty} a^n z^n.$$

The series converges provided that  $|az| < 1$ , in which case we immediately see that  $u_n = a^n u_0$ . More generally, the inverse  $z$ -transform is given by the complex integral around a closed contour  $C$  around the origin in the  $z$ -plane that does not contain any singularities of  $\tilde{u}(z)$ :

$$u_n = \oint_C \frac{\tilde{u}(z)}{z^{n+1}} \frac{dz}{2\pi i}. \quad (2.1.15)$$

However, one often avoids using contour integration by simply Taylor expanding the  $z$ -transform in powers of  $z$  and reading off the coefficient of  $z^n$ , as in the above example.

**Fourier transforms.** The *Fourier transform* of a function of one variable  $u(x)$ ,  $x \in \mathbb{R}$ , is defined by the equation

$$\mathcal{F}u(k) \equiv \hat{u}(k) = \int_{-\infty}^{\infty} u(x) e^{ikx} dx. \quad (2.1.16)$$

The corresponding inverse Fourier transform is

$$\mathcal{F}^{-1}\hat{u}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}(k) e^{-ikx} dk. \quad (2.1.17)$$

An important issue is to determine the set of functions for which the Fourier transform (and its inverse) is well defined. For example, if  $u$  is integrable on  $\mathbb{R}$  so that  $\int_{-\infty}^{\infty} |u(x)| dx < \infty$ , then

$$|\hat{u}(k)| = \left| \int_{-\infty}^{\infty} u(x) e^{ikx} dx \right| \leq \int_{-\infty}^{\infty} |u(x)| dx < \infty,$$

and  $\hat{u}$  exists. However, the latter may itself not be integrable. Therefore, in the application of Fourier transforms, it is common to restrict  $u$  to a much smaller class of functions such as the space of square-integrable functions denoted by  $L^2(\mathbb{R})$ . A few important properties of the Fourier transform are as follow. First, it converts derivatives into algebraic expressions, that is,

$$\mathcal{F}u^{(n)}(k) = (-ik)^n \hat{u}(k), \quad (2.1.18)$$

where  $u^{(n)}$  denotes the  $n$ th derivative of  $u$ , and assuming that  $u$ , and its derivatives are continuous and integrable. There also exists a convolution theorem.

**Theorem 2.2.** *If  $u$  and  $v$  are in  $L^2(\mathbb{R})$ , then  $u * v \in L^2(\mathbb{R})$  and*

$$\mathcal{F}(u * v)(k) = \hat{u}(k) \hat{v}(k), \quad (2.1.19)$$

where

$$(u * v)(x) \equiv \int_{-\infty}^{\infty} u(x-y) v(y) dy. \quad (2.1.20)$$



*Proof.* The theorem is established by interchanging the order of integration:

$$\begin{aligned}
 \mathcal{F}(u * v)(k) &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} u(x-y)v(y)dy \right) e^{ikx} dx \\
 &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} u(x-y)v(y)e^{ikx} dx \right) dy \\
 &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} u(r)v(y)e^{ikr}e^{iky} dr \right) dy \\
 &= \int_{-\infty}^{\infty} u(r)e^{ikr} dr \int_{-\infty}^{\infty} v(y)e^{iky} dy = \hat{u}(k)\hat{v}(k).
 \end{aligned}$$

Yet another useful property is Parseval's theorem

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{u}(k)|^2 dk. \quad (2.1.21)$$

Just as one can define a discrete Laplace transform for discrete-time processes, one can also introduce a discrete Fourier transform of spatial processes such as a random walk, which are defined on a discrete lattice. Therefore, suppose that  $u$  is a function on the space of integers  $\mathbf{Z}$ . The discrete Fourier transform of  $u$  is defined according to

$$(\mathcal{F}u)(k) \equiv \hat{u}(k) = \sum_{r=-\infty}^{\infty} u(r)e^{ikr}, \quad (2.1.22)$$

where  $k$  is now restricted to the finite domain  $(-\pi, \pi)$ . The intuition behind this is that for  $|k| > \pi$ , the spatial oscillations  $\cos(kr)$  and  $\sin(kr)$  probe the function on spatial scales smaller than a unit lattice spacing where there is no information and are thus redundant. The inverse transform is

$$u(r) = \int_{-\pi}^{\pi} \hat{u}(k)e^{-ikr} \frac{dk}{2\pi}. \quad (2.1.23)$$

This is straightforward to prove using the identities

$$\int_{-\pi}^{\pi} e^{ik(r-s)} \frac{dk}{2\pi} = \frac{1}{2\pi i(r-s)} \left[ e^{i\pi(r-s)} - e^{-i\pi(r-s)} \right] = 0 \text{ for } r \neq s,$$

and  $\int_{-\pi}^{\pi} dk/2\pi = 1$ . That is, substituting for  $\hat{u}(k)$  in the inverse transform and revering the order of integration and summation,

$$\begin{aligned}
\int_{-\pi}^{\pi} \hat{u}(k) e^{-ikr} \frac{dk}{2\pi} &= \int_{-\pi}^{\pi} \left( \sum_{s=-\infty}^{\infty} u(s) e^{iks} \right) e^{-ikr} \frac{dk}{2\pi} \\
&= \sum_{s=-\infty}^{\infty} u(s) \int_{-\pi}^{\pi} e^{ik(s-r)} \frac{dk}{2\pi} = \sum_{s=-\infty}^{\infty} u(s) \delta_{s,r} = u(r).
\end{aligned}$$

Note that the discrete Fourier transform should be distinguished from a Fourier series, which is an expansion of a periodic function of  $x$  in terms of a countable set of Fourier components. In other words, in a Fourier series  $k$  is unbounded but takes discrete values. Finally, consider a higher-dimensional square lattice with points  $\ell = n_1 \mathbf{i} + n_2 \mathbf{j}$ . The corresponding discrete Fourier transform (for  $d = 2$ ) is

$$(\mathcal{F}u)(\mathbf{k}) \equiv \hat{u}(\mathbf{k}) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} u(\ell) e^{i\mathbf{k} \cdot \ell}, \quad (2.1.24)$$

with  $\mathbf{k}$  the dual vector

$$\mathbf{k} = k_1 \mathbf{i} + k_2 \mathbf{j}, \quad k_1, k_2 \in (-\pi, \pi).$$

We will consider more general planar lattices (rhombic, hexagonal) and discrete Fourier transforms in Sect. 9.1.

$u(t)$	$\tilde{u}(s)$	$u(t)$	$\tilde{u}(s)$
1	$s^{-1}, \quad s > 0$	$f(t)e^{-at}$	$\tilde{f}(s+a)$
$e^{at}$	$\frac{1}{s-a}, \quad s > a$	$\delta(t-a)$	$\exp(-as)$
$t^n$	$\frac{n!}{s^{n+1}}, \quad s > 0$	$H(t-a)f(t-a)$	$\tilde{f}(s)e^{-as}$
$\sin(at), \cos(at)$	$\frac{a}{s^2+a^2}, \frac{s}{s^2+a^2} \quad s > 0$	$\operatorname{erf}(\sqrt{t})$	$s^{-1}(1+s)^{-1/2}, \quad s > 0$
$\sinh(at), \cosh(at)$	$\frac{a}{s^2-a^2}, \frac{s}{s^2-a^2} \quad s >  a $	$t^{-1/2} \exp(-a^2/4t)$	$\sqrt{\pi/s} \exp(-a\sqrt{s}), \quad s > 0$
$e^{at} \sin(bt)$	$\frac{b}{(s-a)^2+b^2} \quad s > a$	$1 - \operatorname{erf}(a/2\sqrt{t})$	$s^{-1} \exp(-a\sqrt{s}), \quad s > 0$
$e^{at} \cos(bt)$	$\frac{s-a}{(s-a)^2+b^2}, \quad s > a$	$\frac{a}{2t^{3/2}} \exp(-a^2/4t)$	$\sqrt{\pi} \exp(-a\sqrt{s}), \quad s > 0$

Table 2.1: Some common Laplace transforms

### 2.1.1 Continuum Limit of a Random Walk

Having analyzed the discrete random walk, it is now possible to take an appropriate continuum limit to obtain a diffusion equation in continuous space and time. First, introduce infinitesimal step lengths  $\delta x$  and  $\delta t$  for space and time and set  $P_N(r) = \rho(x, t)\delta x$  with  $x = r\delta x, t = N\delta t$ . Substituting into the master equation (2.1.1) gives the following equation for the probability density  $\rho(x, t)$ :

$$\begin{aligned}\rho(x, t) &= p\rho(x - \delta x, t - \delta t) + q\rho(x + \delta x, t - \delta t) \\ &\approx (p + q) \left[ \rho(x, t) - \frac{\partial \rho}{\partial t} \delta t \right] - (p - q) \frac{\partial \rho}{\partial x} \delta x + \frac{(p + q)}{2} \frac{\partial^2 \rho}{\partial x^2} \delta x^2,\end{aligned}$$

where  $\rho$  has been Taylor expanded to first order in  $\delta t$  and to second order in  $\delta x$ . Note that  $p + q = 1$ . Dividing through by  $\delta t$  and taking the continuum limit  $\delta x, \delta t \rightarrow 0$  such that the quantities  $V, D$  are finite, where

$$V = \lim_{\delta x, \delta t \rightarrow 0} (p - q) \frac{\delta x}{\delta t}, \quad D = \lim_{\delta x, \delta t \rightarrow 0} \frac{\delta x^2}{2\delta t},$$

yields the advection–diffusion equation with constant drift  $V$  and diffusivity  $D$ :

$$\frac{\partial \rho(x, t)}{\partial t} = -V \frac{\partial [\rho(x, t)]}{\partial x} + D \frac{\partial^2 \rho(x, t)}{\partial x^2}. \quad (2.1.25)$$

Note that  $p = 0.5 + \kappa\delta x$  and  $q = 0.5 - \kappa\delta x$  with  $\kappa = O(1)$ . For the moment, we will focus on the case of zero drift ( $V = 0$ ), for which Eq. (2.1.25) reduces to the standard diffusion equation.

Although we have derived the diffusion equation from an unbiased random walk, it is more typically interpreted in terms of an evolution equation for a conserved quantity such as particle number rather than a probability density for a single random walker. In order to link these two interpretations, consider  $N$  noninteracting, identical diffusing particles and let  $u(x, t) = N\rho(x, t)$ . For sufficiently large  $N$ , we can treat  $u(x, t)dx$  as the deterministic number of particles in the infinitesimal interval  $[x, x + dx]$  at time  $t$ , with  $u(x, t)$  evolving according to the diffusion equation written in the conservation form

$$\frac{\partial u}{\partial t} = -\frac{\partial J}{\partial x}, \quad J(x, t) = -D \frac{\partial u}{\partial x}, \quad (2.1.26)$$

where  $J(x, t)$  is the Fickian flux of particles. Integrating the diffusion equation (2.1.26) over the interval  $[x, x + dx]$  and reversing the order of integration and differentiation show that

$$\frac{d}{dt} \int_x^{x+dx} u(y, t) dy = J(x, t) - J(x + dx, t),$$

which is an expression of particle conservation. That is, the rate of change of the number of particles in  $[x, x + dx]$  is equal to the net flux crossing the endpoints of the interval. Consider the initial value problem

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2}, x \in \mathbb{R}, t > 0; \quad u(x, 0) = f(x), x \in \mathbb{R},$$

where  $f(x)$  specifies the initial data. For simplicity, we assume that  $u, f \in L^2(\mathbb{R})$ , that is, they are square-integrable. Taking Fourier transforms of the equation with respect to  $x$  gives

$$\frac{\partial \hat{u}(k, t)}{\partial t} = -k^2 D \hat{u}(k, t),$$

which is an ODE in  $t$  with  $k$  treated as a parameter. Its solution is

$$\hat{u}(k, t) = c(k) e^{-k^2 D t},$$

with the coefficient  $c(k)$  determined by the initial data. That is, Fourier transforming the initial condition implies  $\hat{u}(k, 0) = \hat{f}(k)$  and, hence,

$$\hat{u}(k, t) = \hat{f}(k) e^{-k^2 D t}.$$

Applying the convolution Theorem 2.2, we have

$$u(x, t) = \int_{-\infty}^{\infty} K(x - y, t) f(y) dy,$$

where  $K(x, t)$  is the inverse Fourier transform of  $e^{-k^2 D t}$ :

$$K(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} e^{-k^2 D t} dk = \frac{1}{\sqrt{4\pi D t}} e^{-x^2/4Dt}.$$

We thus obtain the result

$$u(x, t) = \frac{1}{\sqrt{4\pi D t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4Dt} f(y) dy. \quad (2.1.27)$$

Note the above solution still holds if we relax the requirement  $f, u \in L^2(\mathbb{R})$ . In particular, if we take the initial condition  $f(x) = \delta(x)$ , where  $\delta(x)$  is the Dirac delta function (see Box 2B), then we obtain the so-called fundamental solution

$$u(x, t) = \frac{1}{\sqrt{4\pi D t}} e^{-x^2/4Dt}. \quad (2.1.28)$$

(Strictly speaking,  $u(x, t)$  is a weak solution of the underlying diffusion equation [554].) Also observe that the fundamental solution corresponds to the continuum limit of the Gaussian distribution (2.1.7) for an unbiased random walk.

**Box 2B. The Dirac delta function**

A heuristic definition of the Dirac delta function would be that it is a “function” with the following properties:

$$\delta(0) = \infty, \quad \delta(x) = 0 \text{ for all } x \neq 0, \quad \int_{\mathbb{R}} \delta(x) dx = 1.$$

However, this definition is not compatible with the classical concept of a function. A rigorous definition of the Dirac delta function requires the theory of generalized functions or distributions [554]. However, an operational definition of the Dirac delta function can be constructed in terms of the limit of a sequence of Heaviside functions. Let  $H(x) = 1$  if  $x \geq 0$  and  $H(x) = 0$  if  $x < 0$ . It follows from this definition that

$$I_{\varepsilon}(x) \equiv \frac{H(x + \varepsilon) - H(x - \varepsilon)}{2\varepsilon} = \begin{cases} \frac{1}{2\varepsilon} & \text{if } -\varepsilon \leq x \leq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

It can be seen that  $I_{\varepsilon}(x)$  has the following properties, (see Fig. 2.4):

- (i) For all  $\varepsilon > 0$ ,

$$\int_{\mathbb{R}} I_{\varepsilon}(x) dx = \frac{1}{2\varepsilon} \times 2\varepsilon = 1.$$

- (ii)

$$\lim_{\varepsilon \rightarrow 0} I_{\varepsilon}(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{if } x = 0 \end{cases}$$

- (iii) If  $\varphi(x)$  is a smooth function that vanishes outside a bounded interval (a test function), then

$$\int_{\mathbb{R}} I_{\varepsilon}(x) \varphi(x) dx = \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \varphi(x) dx \xrightarrow{\varepsilon \rightarrow 0} \varphi(0).$$

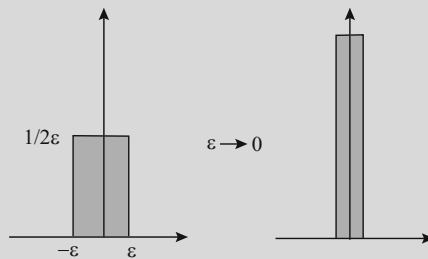


Fig. 2.4: Approximation of Dirac delta function

The third property suggests that we can define the Dirac delta function in terms of how it acts on test functions. Thus the Dirac delta function is defined as a distribution with the following properties:

$$\int_{\mathbb{R}} \delta(x) dx = 1, \quad \int_{\mathbb{R}} \delta(x) \varphi(x) dx = \varphi(0).$$

One can also introduce a shifted Dirac delta function  $\delta_y(x) \equiv \delta(x - y)$ ,

$$\int_{\mathbb{R}} \delta_y(x) dx = 1, \quad \int_{\mathbb{R}} \delta_y(x) \varphi(x) dx = \int_{\mathbb{R}} \delta(x - y) \varphi(x) dx = \varphi(y).$$

The Heaviside construction also suggests that we can formally write  $H'(x) = \delta(x)$ , although again this only really makes sense in terms of test functions:

$$\int_{\mathbb{R}} H'(x) \varphi(x) dx = [H(x) \varphi(x)]_0^{\infty} - \int_{\mathbb{R}} H(x) \varphi'(x) dx = - \int_0^{\infty} \varphi'(x) dx = \varphi(0).$$

We have used integration by parts and the fact that  $\varphi(x) = 0$  at  $x = \infty$ . Finally, note that alternative representations of the Dirac delta function include the Fourier integral,

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} dk, \quad (2.1.29)$$

and the  $t \rightarrow 0$  limit of the fundamental solution (2.1.28),

$$\delta(x) = \lim_{t \rightarrow 0} \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}. \quad (2.1.30)$$

## 2.2 Continuous Stochastic Processes and the Fokker–Planck (FP) Equation

So far we have considered one approach to modeling diffusive processes, which is based on the continuum limit of a random walk; the resulting diffusion equation can be interpreted at the macroscopic level in terms of an equation for particle conservation. In this section we consider an alternative approach to modeling diffusion based on a microscopic particle moving in a fluid, such as the aqueous environment found within the interior of a cell (the cytoplasm or cytosol). The motion of the particle is modeled in terms of a continuous stochastic process evolving according to a Langevin equation or SDE. The probability density of this stochastic process satisfies a generalization of the diffusion equation known as the Fokker–Planck (FP) equation.

### 2.2.1 Derivation of the FP Equation from a Langevin Equation

Consider a microscopic particle such as a macromolecule moving within the cytoplasm of a cell that it is subject to some external force of size  $F$ . Collisions with fluid molecules have two distinct effects. First, they induce an apparent diffusive or Brownian motion of the particle, and second they generate an effective frictional force that opposes motion induced by the external force. In the case of microscopic particles, water acts as a highly viscous medium (low Reynolds number) so that any particle quickly approaches terminal velocity and inertial effects can be ignored (see also Box 5B). The effects of all collisions on the motion of the particle can then be represented in terms of the Langevin equation or SDE [204]

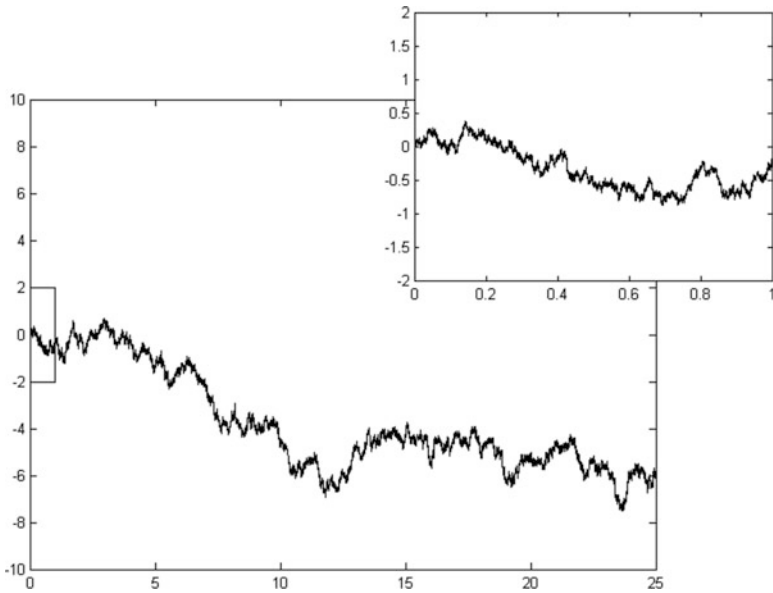


Fig. 2.5: Sample path of a Wiener process

$$dX(t) = \frac{F(X)}{\gamma} dt + \sqrt{2D} dW(t) \quad (2.2.1)$$

where  $X(t)$  is the stochastic position of the particle at time  $t$ ,  $\gamma$  is a drag coefficient, and  $W(t)$  is a so-called Wiener process whose differential  $dW(t)$  is a Gaussian random variable with

$$\langle dW(t) \rangle = 0, \quad \langle dW(t) dW(t') \rangle = \delta(t - t') dt dt', \quad (2.2.2)$$

where  $\delta(t)$  is the Dirac delta function (Box 2B). Mathematically speaking,  $W(t)$  is a continuous but everywhere non-differentiable function of time  $t$  (see Fig. 2.5). For the moment, we simply view  $W(t)$  as a formal representation of the effects of a fluctuating environment. A more mathematical treatment of  $W(t)$  and SDEs is presented in Sect. 2.6, together with methods for numerically simulating an SDE (see also the review by Higham [259]).

Suppose, for the moment, that  $F$  is a constant. Formally integrating Eq. (2.2.1) with  $X(0) = 0$  shows that

$$X(t) = Vt + \sqrt{2D} \int_0^t dW(t')$$

with  $V = F/\gamma$  the terminal velocity. Averaging with respect to the noise then implies that

$$\langle X(t) \rangle = Vt, \quad \langle (X(t) - Vt)^2 \rangle = 2Dt.$$

That is,

$$\begin{aligned} \langle (X(t) - Vt)^2 \rangle &= 2D \langle \int_0^t dW(t') \int_0^t dW(t'') \rangle = 2D \int_0^t \int_0^t \langle dW(t') dW(t'') \rangle \\ &= 2D \int_0^t \int_0^t \delta(t' - t'') dt' dt'' = 2D \int_0^t dt' = 2Dt. \end{aligned}$$

Hence, the MSD about the deterministic trajectory varies as  $2Dt$ , which suggests identifying  $D$  as a diffusion coefficient. Moreover,  $X(t)$  is itself a Gaussian process whose probability density  $p(x, t)$  is given by the Gaussian distribution (2.1.28) assuming the initial condition  $p(x, 0) = \delta(x)$ . Thus, the probability density of a Brownian particle moving under the action of a constant force obeys an advection–diffusion equation of the form (2.1.25). We would like to extend this framework to the case of an  $x$ -dependent force, for which  $p(x, t)$  is known to satisfy a more general Fokker–Planck (FP) equation.

We will consider a derivation of the FP equation applicable for a position-dependent force  $F(x)$  along similar lines to Gardiner [204]. Since  $X(t)$  is a stochastic variable, each simulation of the Langevin equation generates one sample out of the set of all possible trajectories. This motivates an alternative way of thinking about such a stochastic process, namely in terms of the conditional probability density  $p(x, t|x_0, t_0)$  that the particle is at  $x$  at time  $t$ , given that it started at  $x_0$  at time  $t_0$ . Exploiting the fact that the stochastic process is Markovian, that is,  $X(t + \Delta t)$  only depends on the state at the previous time step  $X(t)$ , it follows that  $p(x, t|x_0, t_0)$  satisfies the Chapman–Kolmogorov equation (Sect. 2.6)

$$p(x, t|x_0, t_0) = \int_{-\infty}^{\infty} p(x, t|x', t') p(x', t'|x_0, t_0) dx' \quad (2.2.3)$$



for any  $t' \in [t_0, t]$ . Such an equation is a defining property of a continuous Markov process. Consider an infinitesimal version of this equation by taking  $t \rightarrow t + \Delta t$ ,  $t' \rightarrow t$  and setting  $w(x, t; u, \Delta t) = p(x + u, t + \Delta t | x, t)$ :

$$p(x, t + \Delta t) = \int_{-\infty}^{\infty} w(x - u, t; u, \Delta t) p(x - u, t) du,$$

where the initial argument  $(x_0, t_0)$  has been suppressed. Now suppose that over a sufficiently small time window  $\Delta t$ , large jumps  $u$  in position are highly unlikely, so that  $u$  can be treated as a small variable (It is possible to relax this requirement - one then obtains integral terms in the evolution equation for  $p(x, t)$  that represent finite jumps between states, see also **Ex. 2.3**). Taylor expanding with respect to  $u$  gives

$$p(x, t + \Delta t) = \alpha_0(x, t) p(x, t) - \partial_x [\alpha_1(x, t) p(x, t)] + \frac{1}{2} \partial_{xx}^2 [\alpha_2(x, t) p(x, t)] + \dots \quad (2.2.4)$$

where

$$\alpha_n(x, t) = \int_{-\infty}^{\infty} w(x, t; u, \Delta t) u^n du.$$

The Langevin equation (2.2.1) can be used to calculate the coefficients  $\alpha_n$ . First, rewrite Eq. (2.2.1) in the infinitesimal form

$$X(t + \Delta t) = x + F(x) \Delta t / \gamma + \sqrt{2D} \Delta W(t),$$

given that  $X(t) = x$ . This implies that the transition probability  $w$  can be written as

$$\begin{aligned} w(x, t; u, \Delta t) &= \langle \delta(x + u - X(t + \Delta t)) \rangle \\ &= \langle \delta(u - F(x) \Delta t / \gamma - \sqrt{2D} \Delta W(t)) \rangle, \\ &= \int_{-\infty}^{\infty} \delta(u - F(x) \Delta t / \gamma - \sqrt{2D} \Delta W(t)) p(\Delta W(t)) \end{aligned}$$

where  $p$  is the probability density of  $\Delta W(t)$ . Since

$$\Delta W(t) = \int_t^{t+\Delta t} dW(s)$$

it follows that  $\Delta W(t)$  is a Gaussian random variable with zero mean and variance  $\Delta t$ ; the corresponding probability density is

$$p(\Delta W) = \sqrt{\frac{1}{2\pi\Delta t}} e^{-\Delta W^2 / 2\Delta t}.$$

Hence, averaging with respect to  $\Delta W(t)$ ,

$$w(x, t; u, \Delta t) = \sqrt{\frac{1}{4\pi D \Delta t}} e^{-(u - F(x) \Delta t / \gamma)^2 / 4D \Delta t}.$$

It follows that

$$\alpha_0 = 1, \quad \alpha_1 = F(x)\Delta t/\gamma, \quad \alpha_2 = 2D\Delta t + \alpha_1^2,$$

and  $\alpha_m = \mathcal{O}(\Delta t^2)$  for  $m > 2$ . Substituting these results into Eq. (2.2.4) and taking the limit  $\Delta t \rightarrow 0$  finally leads to the Fokker–Planck (FP) equation

$$\frac{\partial p(x,t)}{\partial t} = -\frac{1}{\gamma} \frac{\partial [F(x)p(x,t)]}{\partial x} + D \frac{\partial^2 p(x,t)}{\partial x^2}. \quad (2.2.5)$$

Note that in the limit  $D \rightarrow 0$ , the FP equation reduces to the so-called Liouville equation. The latter has a general solution of the form

$$p(x,t) = \int_{\mathbb{R}} \delta(x - \phi(t, x_0)) \rho(x_0) dx_0,$$

where  $\phi(t, x_0)$  is the solution to the deterministic equation  $\dot{x} = F(x)/\gamma$  with initial condition  $x(0) = x_0$  and  $\rho(x_0)$  is a probability density over initial conditions. Thus  $p(x,t)$  represents a distribution of deterministic trajectories with  $p(x,0) = \rho(x)$ .

The 1D FP equation (2.2.5) can be rewritten as a probability conservation law according to

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial J(x,t)}{\partial x}, \quad (2.2.6)$$

where

$$J(x,t) = \frac{1}{\gamma} F(x)p(x,t) - D \frac{\partial p(x,t)}{\partial x} \quad (2.2.7)$$

is the probability flux. An equilibrium steady-state solution corresponds to the conditions  $\partial p/\partial t = 0$  and  $J \equiv 0$ . This leads to the first-order ODE for the equilibrium density  $P(x)$ :  $DP'(x) - \gamma^{-1}F(x)P(x) = 0$ , which has the solution

$$P(x) = \mathcal{N} e^{-\Phi(x)/\gamma D}.$$

Here  $\Phi(x) = -\int^x F(y)dy$  is a potential energy function and  $\mathcal{N}$  is a normalization factor (assuming that it exists). Comparison of the equilibrium distribution with the Boltzmann–Gibbs distribution (1.4.5) (see Sect. 1.4) yields the Einstein relation

$$D\gamma = k_B T, \quad (2.2.8)$$

where  $T$  is the temperature (in degrees Kelvin) and  $k_B \approx 1.4 \times 10^{-23} \text{ JK}^{-1}$  is the Boltzmann constant. This formula relates the variance of environmental fluctuations to the strength of dissipative forces and the temperature. In the case of a sphere of radius  $R$  moving in a fluid of viscosity  $\eta$ , Stoke's formula can be used, that is,  $\gamma = 6\pi\eta R$ . For water at room temperature,  $\eta \sim 10^{-3} \text{ kg m}^{-1} \text{ s}^{-1}$  so that a particle of radius  $R = 10^{-9} \text{ m}$  has a diffusion coefficient  $D \sim 100 \mu\text{m}^2 \text{ s}^{-1}$ .

It is straightforward to generalize the Langevin equation (2.2.1) to higher dimensions. Assuming for simplicity isotropic diffusion and friction, Eq. (2.2.1) becomes

$$dX_i = \frac{F_i(\mathbf{X})}{\gamma} dt + \sqrt{2D} dW_i(t), \quad i = 1, \dots, d \quad (2.2.9)$$

with

$$\langle dW_i(t) \rangle = 0, \quad \langle dW_i(t) dW_j(t') \rangle = \delta_{i,j} \delta(t - t') dt dt'. \quad (2.2.10)$$

The corresponding multivariate FP equation is

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = -\frac{1}{\gamma} \nabla \cdot [\mathbf{F}(\mathbf{x}) p(\mathbf{x}, t)] + D \nabla^2 p(\mathbf{x}, t) \quad (2.2.11)$$

and the probability flux is given by the vector field

$$\mathbf{J}(\mathbf{x}, t) = \frac{\mathbf{F}(\mathbf{x})}{\gamma} p(\mathbf{x}, t) - D \nabla p(\mathbf{x}, t). \quad (2.2.12)$$

Here  $\nabla$  denotes the gradient operator, which in Cartesian coordinates  $\mathbf{x} = (x, y, z)$  (for  $d = 3$ ) takes the form

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z},$$

with  $\mathbf{i}$  the unit vector in the  $x$ -direction. Similarly,  $\nabla^2$  is the Laplacian operator

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Note that the Langevin equation (2.2.1) or (2.2.9) represents diffusive-like motion from the probabilistic perspective of a single microscopic particle moving in a fluid medium. However, it is possible to reinterpret Eq. (2.2.5) or (2.2.11) as a deterministic advection-diffusion equation for the concentration  $u(x, t)$  of many particles (see also Sect. 2.1.1). That is, ignoring any interactions or correlations between the particles, set  $u(x, t) = Np(x, t)$  where  $N$  is the total number of particles (assumed large). Multiplying both sides of Eq. (2.2.5) by  $N$  then leads to the corresponding *Smoluchowski* equation for  $u(x, t)$  with  $N\mathbf{J}(x, t)$  interpreted as the particle flux arising from a combination of advection and Fickian diffusion. One example is the well-known Nernst–Planck equation for electrodiffusion (see Ex. 2.4). However, the relationship between macroscopic and microscopic formulations is more complicated when chemical reactions are included. Macroscopically, reactions are described in terms of the deterministic law of mass action (see Sect. 3.1), whereas microscopically they are modeled stochastically using a chemical master equation. Differences between the two levels of modeling become significant when the number of interacting molecules becomes small [651]. From the macroscopic picture of Fickian diffusion, the conservation equation  $\partial_t u = -\nabla \cdot \mathbf{J}$  can lead to two different forms of the diffusion equation, depending on whether  $\mathbf{J}(\mathbf{x}, t) = -\nabla[D(\mathbf{x})u(\mathbf{x}, t)]$  or  $\mathbf{J}(\mathbf{x}, t) = -D(\mathbf{x})\nabla u(\mathbf{x}, t)$ . (These are equivalent when  $D$  is a constant.) In order to distinguish between the two cases, it is necessary to incorporate details regarding the microscopic dynamics using, for example, kinetic theory [77]. The situation is even more complicated in anisotropic heterogeneous media, where it is no longer possible to characterize the rate of diffusion in terms of a single coefficient. One now needs to consider a diffusion tensor; see the example of active transport on

microtubular networks in Sect. 7.4. Irrespective of the particular interpretation of the FP equation (2.2.11), mathematically speaking, it is a deterministic PDE that can be analyzed using the various methods outlined in Sect. 2.5 for the diffusion equation.

### 2.2.2 Boundary Conditions for the FP Equation

In our study of a random walker in Sect. 2.1, we assumed that the domain was unbounded. However, diffusion of a particle within a cell is bounded and often restricted to a subcellular compartment that has a complex geometry. Therefore, it is necessary to specify the domain  $\Omega \subset \mathbb{R}^d$  over which the FPE is defined and to introduce boundary conditions on  $\partial\Omega$ , where  $\partial\Omega$  denotes the boundary of  $\Omega$ . First consider the one-dimensional case ( $d = 1$ ) with the FP equation (2.2.5) defined on the finite interval  $x \in [0, L]$ . This could represent the domain of a narrow ion channel (see Sect. 7.3) or a microtubular filament along which a molecular motor transports cargo (see Chap. 4). The two most common types of boundary condition at the ends  $x = 0, L$  are the *Dirichlet* and *Neumann* boundary conditions. For example, at  $x = 0$

$$p(0, t) = f(t) \text{ (Dirichlet) or } J(0, t) = g(t) \text{ (Neumann),} \quad (2.2.13)$$

where  $J(x, t)$  is the probability flux (2.2.7) and  $f, g$  are prescribed functions of time  $t$ , which could be time-independent. A homogeneous Dirichlet boundary condition ( $f \equiv 0$ ) is often called an *absorbing* boundary condition, whereas a homogeneous Neumann boundary condition ( $g \equiv 0$ ) is often called a *no-flux* or *reflecting* boundary condition. The analogous boundary conditions in higher dimensions ( $d = 2, 3$ ) [see Eq. (2.2.11)], are

$$p(\mathbf{x}, t) = f(\mathbf{x}, t) \text{ (Dirichlet) or } \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = g(\mathbf{x}, t) \text{ (Neumann) for all } \mathbf{x} \in \partial\Omega, \quad (2.2.14)$$

where  $\mathbf{n}(\mathbf{x})$  is the unit outward normal to the boundary at  $\mathbf{x} \in \partial\Omega$ . It is also possible to have mixed boundary conditions, in which  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$  with Dirichlet on  $\partial\Omega_D$  and Neumann on  $\partial\Omega_N$  (see Sect. 7.2). Alternatively, a boundary may be partially absorbing, in which case we have the *Robin* boundary condition

$$p(\mathbf{x}, t) + \alpha \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \alpha > 0.$$

Consider the particular case of a homogeneous Neumann boundary condition. Integrating the FP equation (2.2.11) over the domain  $\Omega$ , and reversing the order of integration and time differentiation, yields

$$\frac{d}{dt} \int_{\Omega} p(\mathbf{x}, t) d\mathbf{x} = - \int_{\Omega} \nabla \cdot \mathbf{J}(\mathbf{x}, t) d\mathbf{x} = - \int_{\partial\Omega} \mathbf{J}(\mathbf{x}, t) \cdot \mathbf{n}(\mathbf{x}) d\mathbf{x} = 0, \quad (2.2.15)$$

where we have used the divergence theorem [395] and imposed the boundary condition. Hence, in the case of a FPE with reflecting boundaries, the total probability

$P = \int_{\Omega} p(\mathbf{x}, t) d\mathbf{x}$  is conserved, that is,  $dP/dt = 0$ , and the system typically converges to a nontrivial stationary state. On the other hand, the total probability is not conserved in the case of an absorbing boundary, which arises in FPT problems (Sect. 2.3). From the PDE perspective, there are well-established methods for solving boundary value problems for the FPE, some of which will be illustrated in Sect. 2.5 for the simpler diffusion equation.

### 2.2.3 The Ornstein–Uhlenbeck Process

Consider the SDE

$$dX = -kXdt + \sqrt{2D}dW(t), \quad (2.2.16)$$

where  $W(t)$  is a Wiener process, and assume a fixed initial condition  $X(0) = X_0$ . (A Gaussian-distributed initial condition is considered in Ex. 2.5.) One way to solve this equation is to perform the change of variables  $Y(t) = X(t)e^{kt}$ . Then

$$\begin{aligned} dY(t) &= Y(X(t+dt), t+dt) - Y(X, t) = Y(X + dX, t+dt) - Y(X, t) \\ &= (X + dX)e^{k(t+dt)} - Xe^{kt} = kXe^{kt}dt + e^{kt}dX \\ &= kXe^{kt}dt + e^{kt}[-kXdt + \sqrt{2D}dW] = \sqrt{2D}e^{kt}dW(t). \end{aligned}$$

The SDE for  $Y(t)$  can now be integrated to give

$$Y(t) = Y_0 + \sqrt{2D} \int_0^t e^{kt'} dW(t'),$$

that is,

$$X(t) = X_0 e^{-kt} + \sqrt{2D} \int_0^t e^{-k(t-t')} dW(t'). \quad (2.2.17)$$

Given the explicit solution for  $X$ , we can now evaluate the mean and variance using properties of the Wiener process. First,

$$\langle X(t) \rangle = X_0 e^{-kt} + \sqrt{2D} \int_0^t e^{-k(t-t')} \langle dW(t') \rangle = X_0 e^{-kt},$$

since  $\langle dW \rangle = 0$ . Similarly,

$$\begin{aligned} \langle [X(t) - \langle X(t) \rangle]^2 \rangle &= 2D \int_0^t e^{-k(t-s)} \int_0^t e^{-k(t-s')} \langle dW(s) dW(s') \rangle \\ &= 2D \int_0^t e^{-2k(t-s)} ds = \frac{D}{k} (1 - e^{-2kt}). \end{aligned} \quad (2.2.18)$$

Note that in the limit  $k \rightarrow 0$ , we recover the MSD of 1D Brownian motion. (The use of a change of variables to solve a Langevin equation is also considered in Ex. 2.6.)

Equation (2.2.5) implies that the FP equation for the OU process is

$$\frac{\partial p(x, t)}{\partial t} = \frac{\partial [kxp(x, t)]}{\partial x} + D \frac{\partial^2 p(x, t)}{\partial x^2}. \quad (2.2.19)$$

Taking a fixed (deterministic) initial condition  $X(0) = x_0$ , the initial condition of the FP equation is

$$p(x, 0) = \delta(x - x_0).$$

Introduce the characteristic function (Fourier transform)

$$\Gamma(z, t) = \int_{-\infty}^{\infty} e^{izx} p(x, t) dx.$$

Fourier transforming the FP equation shows that  $\Gamma$  satisfies the PDE (see Ex. 2.5):

$$\frac{\partial \Gamma}{\partial t} + kz \frac{\partial \Gamma}{\partial z} = -Dz^2 \Gamma. \quad (2.2.20)$$

This can be solved using separation of variables (see Sect. 2.5) or the method of characteristics (see Sect. 3.6). The result is

$$\Gamma(z, t) = \exp \left[ -\frac{Dz^2}{2k} (1 - e^{-2kt}) + izx_0 e^{-kt} \right],$$

so that on applying the inverse Fourier transform (see Box 2A), we obtain the probability density

$$p(x, t) = \frac{1}{\sqrt{2\pi D[1 - e^{-2kt}]/k}} e^{-(x-x_0 e^{-kt})^2 / (2D[1 - e^{-2kt}]/k)}. \quad (2.2.21)$$

Note that

$$\lim_{t \rightarrow \infty} p(x, t) = p_s(x) \equiv \frac{1}{\sqrt{2\pi D/k}} e^{-kx^2/2D},$$

which is the stationary probability density.

Finally, note that the multivariate version of the OU process is given by

$$dX_i = - \sum_{j=1}^N A_{ij} X_j dt + \sum_{j=1}^N B_{ij} dW_j(t), \quad \mathbf{X}(0) = \bar{\mathbf{x}}. \quad (2.2.22)$$

The solution can be expressed formally in the matrix form (see Ex. 2.7)

$$\mathbf{X}(t) = e^{-\mathbf{A}t} \bar{\mathbf{x}} + \int_0^t e^{-\mathbf{A}(t-t')} \mathbf{B} d\mathbf{W}(t').$$

It can then be shown that the covariance matrix  $\Sigma(t)$  with components

$$\Sigma_{ij}(t) = \langle [X_i(t) - \langle X_i(t) \rangle] [X_j(t) - \langle X_j(t) \rangle] \rangle$$

satisfies the matrix equation

$$\frac{d\Sigma(t)}{dt} = -\mathbf{A}\Sigma(t) - \Sigma(t)\mathbf{A}^T + \mathbf{B}\mathbf{B}^T.$$

It follows that if  $\mathbf{A}$  has distinct eigenvalues with positive real part, then  $\Sigma(t) \rightarrow \Sigma_0$  where  $\Sigma_0$  is the stationary covariance matrix satisfying the Riccati equation

$$\mathbf{A}\Sigma_0 + \Sigma_0\mathbf{A}^T = \mathbf{B}\mathbf{B}^T. \quad (2.2.23)$$

The multivariate OU process will play an important role in the analysis of gene networks (see Chap. 6).

### 2.2.4 Multiplicative Noise

So far we have assumed that the diffusion coefficient in the Langevin equation (2.2.1) is position-independent, that is, the noise term is *additive*. The situation is considerably more involved when the term multiplying  $dW(t)$  depends on  $X(t)$ , that is, when the noise term is *multiplicative*. The scalar Langevin equation then takes the form

$$dX(t) = A(X)dt + B(X)dW(t). \quad (2.2.24)$$

The difficulty arises since, in order to construct a solution of the SDE, we have to deal with stochastic integrals of the form  $\int_0^t A(X(t))dW(t)$ . Here we give a heuristic discussion of the issue—a more detailed discussion can be found in Sect. 2.6 (see also [204, 651]). Suppose for the moment that  $X(t)$  and  $W(t)$  are deterministic functions of time, and we can apply the theory of Riemann integration. That is, we partition the time interval  $[0, T]$  into  $N$  equal intervals of size  $\Delta t$  with  $N\Delta t = t$  and identify the value of the integral with the unique limit (assuming it exists)

$$\lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} A([1 - \alpha]X_j + \alpha X_{j+1})\Delta W_j$$

for  $0 \leq \alpha < 1$ , where  $\Delta W_j = W((j+1)\Delta t) - W(j\Delta t)$  and  $X_j = X(j\Delta t)$ . In the deterministic case, the integral is independent of  $\alpha$ . Unfortunately, this is no longer true when we have a stochastic integral. One way to see this is to note that the  $\Delta W_j$  are independent random variables. Hence,  $A$  is only statistically independent of  $\Delta W_j$  when  $\alpha = 0$ , which is the *Ito* definition of stochastic integration. On the other hand, when  $\alpha = 1/2$  we have the *Stratonovich* version. It turns out that the form of the corresponding FP equation also depends on  $\alpha$  (see Sect. 2.6). In the Ito case,

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial[A(x)p(x,t)]}{\partial x} + \frac{1}{2} \frac{\partial^2 B(x)p(x,t)}{\partial x^2}, \quad (2.2.25)$$

whereas in the Stratonovich case

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial[A(x)p(x,t)]}{\partial x} + \frac{1}{2} \frac{\partial}{\partial x} B(x) \frac{\partial}{\partial x} [B(x)p(x,t)]. \quad (2.2.26)$$

*Example 2.1.* Consider the Langevin equation with linear multiplicative noise interpreted in the sense of Ito:

$$dX = X(t)a(t)dt + X(t)b(t)dW(t). \quad (2.2.27)$$

One way to solve this equation is to eliminate the multiplicative factor by performing the change of variables  $Y(t) = \ln X(t)$ . However, care must be taken when calculating infinitesimals, since the normal rules of calculus no longer apply for Ito stochastic variables; they do for Stratonovich variables. In particular, as shown in Sect. 2.6, one has to take into account the fact that “ $dW^2 = dt$ ”.

$$\begin{aligned} dY(t) &= \ln(X(t+dt)) - \ln X(t) = \ln(X(t) + dX(t)) - \ln X(t) = \ln(1 + dX(t)/X(t)) \\ &= \frac{dX(t)}{X(t)} - \frac{dX(t)^2}{2X(t)^2} = a(t)dt + b(t)dW(t) - \frac{1}{2}[a(t)dt + b(t)dW(t)]^2 \\ &= a(t)dt + b(t)dW(t) - \frac{b(t)^2}{2}dt + o(dt). \end{aligned}$$

Integrating this equation gives

$$Y(t) = Y_0 + \int_0^t \left[ a(s) - \frac{1}{2}b(s)^2 \right] ds + \int_0^t b(s)dW(s),$$

and exponentiating

$$X(t) = X_0 \exp \left( \int_0^t \left[ a(s) - \frac{1}{2}b(s)^2 \right] ds + \int_0^t b(s)dW(s) \right).$$

Some examples of 1D stochastic processes with multiplicative noise are considered in Exs. 2.9 and 2.10.

### 2.2.5 Correlations and the Power Spectrum

A very useful quantity is the *power spectrum* of a stationary stochastic process  $X(t)$ , which is defined as the Fourier transform of the autocorrelation function  $C_X(\tau)$ ,

$$S_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega\tau} C_X(\tau) d\tau, \quad C_X(\tau) = \langle X(t)X(t+\tau) \rangle. \quad (2.2.28)$$

Consider the covariance of two frequency components of  $X(t)$ :

$$\begin{aligned} \langle \tilde{X}(\omega) \tilde{X}(\omega') \rangle &= \left\langle \int_{-\infty}^{\infty} e^{i\omega t} X(t) dt \int_{-\infty}^{\infty} e^{i\omega' t'} X(t') dt' \right\rangle \\ &= \int_{-\infty}^{\infty} e^{i\omega t} \int_{-\infty}^{\infty} e^{i\omega' t'} \langle X(t)X(t') \rangle dt' dt \end{aligned}$$



$$\begin{aligned}
&= \int_{-\infty}^{\infty} e^{i\omega t} \int_{-\infty}^{\infty} e^{i\omega' t'} \left[ \int_{-\infty}^{\infty} e^{-i\Omega(t-t')} S_X(\Omega) \frac{d\Omega}{2\pi} \right] dt' dt \\
&= \int_{-\infty}^{\infty} S_X(\Omega) \left[ \int_{-\infty}^{\infty} e^{i(\omega-\Omega)t} dt \right] \left[ \int_{-\infty}^{\infty} e^{i(\omega'+\Omega)t'} dt' \right] \frac{d\Omega}{2\pi},
\end{aligned}$$

assuming that it is possible to rearrange the order of integration. Using the Fourier representation of the Dirac delta function (Box 2B),  $\int_{-\infty}^{\infty} e^{i\omega t} dt = 2\pi\delta(\omega)$ , we have

$$\langle \tilde{X}(\omega) \tilde{X}(\omega') \rangle = \int_{-\infty}^{\infty} S_X(\Omega) \cdot 2\pi\delta(\omega - \Omega) \cdot 2\pi\delta(\omega' + \Omega) \frac{d\Omega}{2\pi}$$

which establishes a version of the Wiener–Khinchin theorem:

$$\langle \tilde{X}(\omega) \tilde{X}(\omega') \rangle = 2\pi S_X(\omega) \delta(\omega + \omega'). \quad (2.2.29)$$

The Fourier transform of a real-valued variable satisfies  $\tilde{X}(-\omega) = \tilde{X}^*(\omega)$  so

$$\langle \tilde{X}(\omega) \tilde{X}^*(\omega') \rangle = 2\pi S_X(\omega) \delta(\omega - \omega'). \quad (2.2.30)$$

In the case of linear SDEs, it is possible to calculate the spectrum explicitly using the notion of a *white noise* process. Although the derivative of the Wiener process  $W(t)$  does not exist, there is a sense in which the autocorrelation of the derivative does exist, which provides a useful calculational tool. For example, consider the Ornstein–Uhlenbeck process

$$dX(t) = -\kappa X(t) + dW(t).$$

Suppose that we formally rewrite this equation in terms of derivatives according to

$$\frac{dX}{dt} + \kappa X = \xi(t), \quad (2.2.31)$$

where

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t) \xi(t') \rangle = \delta(t - t').$$

The term  $\xi(t)$  is known as Gaussian white noise. In order to have a stationary OU process, we take the initial time to be at  $t = -\infty$ . The solution can be expressed formally in terms of the integral solution

$$X(t) = \int_{-\infty}^{\infty} G(\tau) \xi(t - \tau) d\tau, \quad (2.2.32)$$

where  $G(\tau)$  is known as the causal Green's function or linear response function with the important property that  $G(\tau) = 0$  for  $\tau < 0$ . In the case of the OU process

$$G(\tau) = e^{-\tau\kappa} H(\tau),$$

where  $H(t)$  is the Heaviside function. The main point to emphasize is that although  $\xi(t)$  is not a mathematically well-defined object, one still obtains correct

answers when taking expectations. For example, it is clear that in the stationary state  $\langle X(t) \rangle = 0$  and (for  $s > 0$ )

$$\begin{aligned} \langle X(t)X(t+s) \rangle &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tau)G(\tau') \langle \xi(t-\tau)\xi(t+s-\tau') \rangle d\tau d\tau' \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tau)G(\tau') \delta(s+\tau-\tau') d\tau d\tau' = \int_{-\infty}^{\infty} G(\tau)G(\tau+s) d\tau \\ &= \int_0^{\infty} e^{-\kappa(2\tau+s)} d\tau = \frac{1}{2\kappa} e^{-\kappa s}. \end{aligned}$$

This is the expected result for the autocorrelation function of the OU process.

One of the useful features of formally expressing a solution to a linear SDE in the form (2.2.32) is that one can view the dynamical system as acting as a filter of the white noise process. Applying the Wiener–Khinchin theorem to the white noise autocorrelation function, we see that the spectrum is given by the Fourier transform of a Dirac delta function, which is unity. However, once the noise has been passed through a filter with linear response function  $G(t)$ , the spectrum is no longer flat. This follows from applying the convolution Theorem 2.2 of Box 2A to Eq. (2.2.32):

$$\tilde{X}(\omega) = \tilde{G}(\omega) \tilde{\xi}(\omega),$$

so

$$2\pi S_X(\omega) \delta(\omega - \omega') = \tilde{G}(\omega) \tilde{G}^*(\omega') \langle \tilde{\xi}(\omega) \tilde{\xi}^*(\omega') \rangle.$$

Evaluating the various Fourier transforms, we have

$$\tilde{G}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} G(t) dt = \int_0^{\infty} e^{i\omega t} e^{-\kappa t} dt = \frac{1}{\kappa - i\omega}$$

and

$$\langle \tilde{\xi}(\omega) \tilde{\xi}^*(\omega') \rangle = \int_{-\infty}^{\infty} e^{i\omega t} \int_{-\infty}^{\infty} e^{-i\omega' t'} \langle \xi(t) \xi(t') \rangle dt' dt = 2\pi \delta(\omega - \omega').$$

Hence,

$$S_X(\omega) = \frac{1}{\kappa^2 + \omega^2}. \quad (2.2.33)$$

The spectrum can be used to recover the variance by noting that

$$\langle X(t)^2 \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \tilde{X}(\omega) \tilde{X}(\omega') \rangle e^{i\omega t} e^{i\omega' t} \frac{d\omega}{2\pi} \frac{d\omega'}{2\pi} = \int_{-\infty}^{\infty} S_X(\omega) \frac{d\omega}{2\pi}.$$

Substituting for  $S_X(\omega)$  and using the identity

$$\int_{-\infty}^{\infty} \frac{d\omega}{\omega^2 + \kappa^2} = \frac{\pi}{\kappa},$$

we see that

$$\langle X(t)^2 \rangle = \frac{1}{2\kappa}.$$

Finally, note that the formal method based on white noise is only applicable to SDEs with additive noise. In the multiplicative case, one has to use stochastic calculus (see Sect. 2.6). In this book, we will mainly express SDEs in terms of differentials and Wiener processes, restricting the use of white noise to the analysis of spectra in linear SDEs, as in Sects. 6.4 and 9.3.

## 2.3 First Passage Time Density and the Backward FP Equation

One of the most important ways of quantifying the efficiency of diffusive transport is in terms of the FPT to reach a target [204, 523]. In the case of intracellular transport, such a target could represent a substrate for a subsequent biochemical reaction or an exit from some bounded domain such as a chemical synapse. Consider a particle whose position evolves according to the 1D Langevin equation (2.2.1) with motion restricted to the bounded domain  $x \in [0, L]$ . (The FPT problem for a random walk on a lattice is considered in Ex. 2.11.) Suppose that the corresponding FP equation (2.2.5) has a reflecting boundary condition at  $x = 0$  and an absorbing boundary condition at  $x = L$ :

$$J(0, t) = 0, \quad p(L, t) = 0.$$

We would like to determine the stochastic time  $T(y)$  for the particle to exit the right-hand boundary given that it starts at location  $y \in [0, L]$  at time  $t$ . As a first step, we introduce the survival probability  $\mathbb{P}(y, t)$  that the particle has not yet exited the interval at time  $t$ :

$$\mathbb{P}(y, t) = \int_0^L p(x, t|y, 0) dx.$$

It follows that  $\text{Prob}[T(y) \leq t] = 1 - \mathbb{P}(y, t)$  and we can define the FPT density according to

$$f(y, t) = -\frac{\partial \mathbb{P}(y, t)}{\partial t} = -\int_0^L \frac{\partial}{\partial t} p(x, t|y, 0) dx.$$

Using the FP equation written in conservation form (2.2.6), we see that

$$f(y, t) = \int_0^L \frac{\partial J(x, t|y, 0)}{\partial x} dx = J(L, t|y, 0) - J(0, t|y, 0) = J(L, t|y, 0),$$

due to the reflecting boundary condition at  $x = 0$ . Thus the FPT density is equal to the flux through the absorbing boundary at  $x = L$ . In certain simple cases, the flux can be calculated explicitly, as illustrated in Ex. 2.12. However, for more general cases, it is useful to derive explicit differential equations for moments of the FPT density, in particular, the first moment or *mean first passage time* (MFPT).

In order to derive an equation for the MFPT, it is necessary to use the so-called backward FP equation. This can be derived from the Chapman–Kolmogorov equation (2.2.3) by differentiating both sides with respect to the intermediate time  $t'$ :

$$0 = \int_{-\infty}^{\infty} \partial_{t'} p(x, t | x', t') p(x', t' | x_0, t_0) dx' + \int_{-\infty}^{\infty} p(x, t | x', t') \partial_{t'} p(x', t' | x_0, t_0) dx'.$$

Using the fact that  $p(x', t' | x_0, t_0)$  satisfies a forward FP equation,  $\partial_{t'} [p(x', t' | x_0, t_0)]$  can be replaced by terms involving derivatives with respect to  $x'$ . Integrating by parts with respect to  $x'$  then leads to the result

$$0 = \int_{-\infty}^{\infty} [\partial_{t'} p(x, t | x', t') + A(x') \partial_{x'} p(x, t | x', t') + D \partial_{x'x'}^2 p(x, t | x', t')] p(x', t' | x_0, t_0) dx'$$

where  $A(x) = F(x)/\gamma$ . Since  $p$  is positive, it follows that the expression in square brackets vanishes. Using time translation invariance,

$$\partial_{t'} p(x, t | x', t') = \partial_{t'} p(x, 0 | x', t' - t) = -\partial_t p(x, 0 | x', t' - t) = -\partial_t p(x, t | x', t'),$$

then yields the backward FP equation for  $p$ :

$$\partial_t p(x, t | x', t') = A(x') \partial_{x'} p(x, t | x', t') + D \partial_{x'x'}^2 p(x, t | x', t'). \quad (2.3.1)$$

Taking  $x' \rightarrow y$ ,  $t' = 0$  and integrating with respect to  $x$  shows that  $\mathbb{P}(y, t)$ , and hence  $f(y, t)$ , also satisfies a backward FP equation:

$$\frac{\partial \mathbb{P}(y, t)}{\partial t} = A(y) \frac{\partial \mathbb{P}(y, t)}{\partial y} + D \frac{\partial^2 \mathbb{P}(y, t)}{\partial y^2}. \quad (2.3.2)$$

The MFPT  $\tau(y)$  is defined according to

$$\begin{aligned} \tau(y) &= \langle T(y) \rangle \equiv \int_0^{\infty} f(y, t) t dt \\ &= - \int_0^{\infty} t \frac{\partial \mathbb{P}(y, t)}{\partial t} dt = \int_0^{\infty} \mathbb{P}(y, t) dt, \end{aligned}$$

after integration by parts. Hence, integrating both sides of Eq. (2.3.2) shows that the MFPT satisfies the ODE

$$A(y) \frac{d\tau(y)}{dy} + D \frac{d^2 \tau(y)}{dy^2} = -1. \quad (2.3.3)$$

Equation (2.3.3) is supplemented by reflecting and absorbing boundary conditions for the backward FP equation:

$$\tau'(0) = 0, \quad \tau(L) = 0.$$

It is straightforward to solve Eq. (2.3.3) by direct integration [204]. First, introduce the integration factor

$$\psi(y) = \exp\left(\frac{1}{D} \int_0^y A(y') dy'\right) = \exp(-V(y)/k_B T),$$

where  $D^{-1}A(y) = (D\gamma)^{-1}F(y) = -(k_B T)^{-1}V'(y)$  and  $V(y)$  is a potential energy. Equation (2.3.3) becomes

$$\frac{d}{dy} [\psi(y)\tau'(y)] = -\frac{\psi(y)}{D}$$

so that

$$\psi(y)\tau'(y) = -\frac{1}{D} \int_0^y \psi(y') dy',$$

where the boundary condition  $\tau'(0) = 0$  has been used. Integrating once more with respect to  $y$  and using  $\tau(L) = 0$  then gives

$$\tau(y) = \int_y^L \frac{dy'}{\psi(y')} \int_0^{y'} \frac{\psi(y'')}{D} dy''. \quad (2.3.4)$$

This formula will be the starting point for analyzing escape problem in Sect. 3.3.

In the case of pure diffusion ( $A(x) = 0$ ), we have  $\psi(y) = 1$  and  $\tau(y) = (L^2 - y^2)/2D$ . It follows that for any finite  $L - y$ ,  $\tau(y) \rightarrow \infty$  as  $L \rightarrow \infty$ . Thus, although 1D diffusion is recurrent, i.e., the particle surely reaches the origin, the average time it takes is infinite. (This can also be understood in terms of the scaling properties of the FPT density.) Now suppose that  $L$  is finite and the particle starts at the left-hand boundary. The corresponding MFPT is then  $\tau = L^2/D$ . Within the cytosol of cells, macromolecules such as proteins tend to have diffusivities  $D < 1 \mu\text{m}^2 \text{s}^{-1}$ , which is due to effects such as molecular crowding. This implies that the mean time for a diffusing particle to travel a distance  $100 \mu\text{m}$  is at least  $10^4 \text{s}$  (a few hours), whereas to travel a distance  $1 \text{mm}$  is at least  $10^6 \text{s}$  (10 days). Since neurons, for example, which are the largest cells in humans, have axonal and dendritic protrusions that can extend from  $1 \text{mm}$  up to  $1 \text{m}$ , the mean travel time due to passive diffusion becomes prohibitively large, and an active form of transport becomes essential.

It is also possible to extend the above 1D analysis to the case where the particle can exit from either end [204, 523]. It is often of interest to keep track of which end the particle exits, which leads to the concept of a splitting probability. Let  $S_0(x, t)$  denote the probability that the particle exits at  $x' = 0$  after time  $t$ , having started at the point  $x$ . Then

$$S_0(x, t) = - \int_t^\infty J(0, t'|x, 0) dt'$$

with

$$J(0, t|x, 0) = A(0)p(0, t|x, 0) - D \left. \frac{\partial p(y, t|x, 0)}{\partial y} \right|_{y=0}.$$

Differentiating with respect to  $t$  and using the backward FP equation (2.3.1) gives

$$\begin{aligned}\frac{\partial S_0(x, t)}{\partial t} &= J(0, t|x, 0) = - \int_t^\infty \frac{\partial J(0, t'|x, 0)}{\partial t'} dt' \\ &= A(x) \frac{\partial S_0(x, t)}{\partial x} + D \frac{\partial^2 S_0(x, t)}{\partial x^2}.\end{aligned}\quad (2.3.5)$$

The hitting or splitting probability that the particle exits at  $x' = 0$  (rather than  $x' = L$ ) is  $\Pi_0(x) = S_0(x, 0)$ . Moreover, the probability that the particle exits after time  $t$ , conditioned on definitely exiting through  $x' = 0$ , is  $\text{Prob}(T_0(x) > t) = S_0(x, t)/S_0(x, 0)$ , where  $T_0(x)$  is the corresponding conditional FPT. Since the conditional MFPT satisfies

$$\tau_0(x) = - \int_0^\infty t \frac{\partial \text{Prob}(T_0(x) > t)}{\partial t} dt = \int_0^\infty \frac{S_0(x, t)}{S_0(x, 0)} dt.$$

Equation (2.3.5) is integrated with respect to  $t$  to give

$$A(x) \frac{\partial \Pi_0(x) \tau_0(x)}{\partial x} + D \frac{\partial^2 \Pi_0(x) \tau_0(x)}{\partial x^2} = -\Pi_0(x), \quad (2.3.6)$$

with boundary conditions  $\Pi_0(0) \tau_0(0) = \Pi_0(L) \tau_0(L) = 0$ . Finally, taking the limit  $t \rightarrow 0$  in Eq. (2.3.5) and noting that  $J(0, 0|x, 0) = 0$  for  $x \neq 0$ ,

$$A(x) \frac{\partial \Pi_0(x)}{\partial x} + D \frac{\partial^2 \Pi_0(x)}{\partial x^2} = 0, \quad (2.3.7)$$

with boundary conditions  $\Pi_0(0) = 1, \Pi_0(L) = 0$ . A similar analysis can be carried out for exit through the other end  $x' = L$  such that  $\Pi_0(x) + \Pi_L(x) = 1$ .

The construction of the FPT density can also be extended to higher spatial dimensions. Suppose that a particle evolves according to the Langevin equation (2.2.9) in a compact domain  $\Omega$  with boundary  $\partial\Omega$ . Suppose that at time  $t = 0$  the particle is at the point  $\mathbf{y} \in \Omega$  and let  $T(\mathbf{y})$  denote the FPT to reach any point on the boundary  $\partial\Omega$ . The probability that the particle has not yet reached the boundary at time  $t$  is then

$$\mathbb{P}(\mathbf{y}, t) = \int_\Omega p(\mathbf{x}, t|\mathbf{y}, 0) d\mathbf{x},$$

where  $p(\mathbf{x}, t|\mathbf{y}, 0)$  is the solution to the multivariate FP equation (2.2.11) with an absorbing boundary condition on  $\partial\Omega$ . The FPT density is again  $f(\mathbf{y}, t) = -d\mathbb{P}(\mathbf{y}, t)/dt$  which, on using Eq. (2.2.11) and the divergence theorem, can be expressed as

$$f(\mathbf{y}, t) = - \int_{\partial\Omega} [-\mathbf{A}(\mathbf{x}) p(\mathbf{x}, t|\mathbf{y}, 0) + D \nabla p(\mathbf{x}, t|\mathbf{y}, 0)] \cdot d\boldsymbol{\sigma}$$

with  $\mathbf{A} = \mathbf{F}/\gamma$ . Similarly, by constructing the corresponding backward FP equation, it can be shown that the MFPT satisfies the equation

$$\mathbf{A}(\mathbf{y}) \cdot \nabla \tau(\mathbf{y}) + D \nabla^2 \tau(\mathbf{y}) = -1 \quad (2.3.8)$$

with  $\mathbf{n} \cdot \nabla \tau(\mathbf{y}) = 0$  for  $\mathbf{y} \in \partial\Omega$ . Finally, note that an analogous formulation of FPTs can be formulated for discrete Markov processes evolving according to a master equation. (The particular case of a simple birth–death process is presented in Sect. 6.6, within the context of a polymerization model of gene transcription.)

## 2.4 Diffusion-Limited Reaction Rates

### 2.4.1 Smoluchowski Reaction Rate

An important example of a FPT process arises in Smoluchowski rate theory for diffusion-controlled reactions [124, 323, 523, 536, 603]. The simplest version of the theory concerns the bimolecular reaction  $A + B \rightarrow AB$  for which the concentrations evolve according to the following law of mass action (see Sect. 3.1):

$$\frac{d[AB]}{dt} = k[A][B].$$

We assume that an  $A$  molecule and a  $B$  molecule react immediately to form the complex  $AB$  when they encounter each other within a reaction radius, so that the speed of reaction  $k$  is limited by their encounter rate via diffusion. (Note that  $k$  has units of volume  $\text{s}^{-1}$ . Concentrations are typically measured in *molars*  $M$  with 1 molar = 1,000 moles/ $\text{m}^3$  and 1 mole  $\approx 6 \times 10^{23}$  molecules (Avogadro's number).) One can then formulate the problem as an idealized first passage process, in which one  $A$  molecule, say, is fixed and treated as the center of a spherical target domain of reaction radius  $a$ , while the  $B$  molecules diffuse and are absorbed if they hit the boundary of the target domain (see Fig. 2.6a). It is assumed that the density of the particles is sufficiently small, so that reactions with other  $A$  molecules have a negligible effect on the concentration of  $B$  molecules in a neighborhood of the target molecule. The steady-state flux to the target (if it exists) is then identified as the mean reaction rate  $k$  across many targets. Let  $\Omega$  denote the target domain and  $\partial\Omega$  its absorbing boundary. We then need to solve the diffusion equation for the concentration  $c(\mathbf{x}, t)$  of background molecules exterior to the domain  $\Omega$ :

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} = D \nabla^2 c(\mathbf{x}, t), \quad c(\mathbf{x} \in \partial\Omega, t) = 0, \quad c(\mathbf{x}, 0) = c_0,$$

subject to the far-field boundary condition  $c(\mathbf{x}, t) = c_0$  for  $\mathbf{x} \rightarrow \infty$ . The flux through the target boundary is

$$J = D \int_{\partial\Omega} \nabla c \cdot d\mathbf{S}.$$

Note the sign, which is due to the fact that the flux is from the exterior to the interior of the target.

Let  $d$  denote the spatial dimension of the target. For  $d > 2$ , a diffusing particle is transient, which means that there is a nonzero probability of never reaching the target (see Sect. 2.1). Hence, the loss of reactants by target absorption is balanced by their resupply from infinity. It follows that there exists a steady state in which the reaction rate is finite. On the other hand, for  $d \leq 2$ , reactants are sure to hit the target (recurrent diffusion) and a depletion zone continuously develops around the target so that the flux and reaction rate decay monotonically to zero with respect to time. Although a reaction rate does not strictly exist, it is still useful to consider the time-dependent flux as a time-dependent reaction rate. The two-dimensional case is particularly important when considering interactions of molecules embedded in the plasma membrane of a cell or the lipid bilayer surrounding an intracellular compartment.

First consider the case of a spherical target of radius  $a$  ( $d = 3$ ). Exploiting the radial symmetry of the problem, it is possible to set  $u(r, t) = rc(r, t)$  such that the 3D diffusion equation for  $c$  reduces to a 1D diffusion equation for  $u$  [523]:

$$\frac{\partial u(r, t)}{\partial t} = D \frac{\partial^2 u(r, t)}{\partial r^2}$$

with  $u(r, 0) = rc_0$ ,  $u(a, t) = 0$  and  $u(r, t) = rc_0$  as  $r \rightarrow \infty$ . Laplace transforming this equation gives  $s\tilde{u}(r, s) - rc_0 = D\tilde{u}''(r, s)$ , which has the solution

$$\tilde{u}(r, s) = \frac{c_0}{s} \left[ r - ae^{-(r-a)\sqrt{s/D}} \right].$$

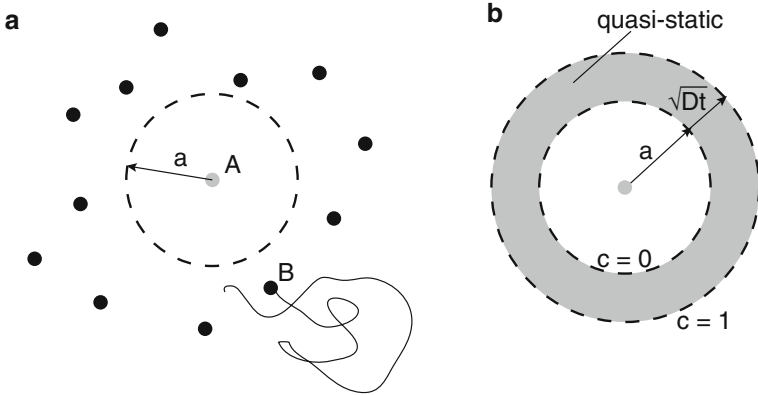


Fig. 2.6: Diffusion-limited reaction rate. (a) Diffusing molecules B in a neighborhood of a fixed target molecule A with reaction radius  $a$ . (b) Quasi-static approximation for calculating time-dependent reaction rate



Since the inverse Laplace transform of  $s^{-1}[1 - e^{-r\sqrt{s/D}}]$  is the error function  $\text{erf}(r/\sqrt{4Dt})$  (see Table 2.1), where

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-r^2} dr,$$

one finds that

$$c(r, t) = c_0 \left(1 - \frac{a}{r}\right) + \frac{ac_0}{r} \text{erf} \left[ \frac{r-a}{\sqrt{4Dt}} \right].$$

It follows that the time-dependent flux is

$$\begin{aligned} J(t) &= 4\pi a^2 D \left. \frac{\partial c}{\partial r} \right|_{r=a} \\ &= 4\pi a D c_0 \left(1 + \frac{a}{\sqrt{\pi Dt}}\right) \xrightarrow{t \rightarrow \infty} 4\pi a D c_0. \end{aligned} \quad (2.4.1)$$

Hence, we obtain the Smoluchowski reaction rate  $k = 4\pi a D$ . As highlighted by Redner [523], it is straightforward to generalize the steady-state result to other three-dimensional targets by making a connection with electrostatics. That is, setting  $\phi(\mathbf{x}) = 1 - c(\mathbf{x})/c_0$  in steady state, it follows that  $\phi$  satisfies Laplace's equation with  $\phi = 1$  on the target boundary and  $\phi = 0$  at infinity, so that  $\phi$  is equivalent to the electrostatic potential generated by a perfectly conducting object  $\Omega$  held at unit potential. Moreover, the steady-state reaction rate  $k = 4\pi D Q$  where  $Q$  is the total charge on the surface of the conductor, which for a unit potential is equal to the capacitance,  $Q = C$ . Thus, determining the reaction rate for a general 3D target is equivalent to finding the capacitance of a perfect conductor with the same shape (see also [107]).

Although it is possible to calculate the exact time-dependent flux for  $d \leq 2$ , a much simpler method is to use a quasi-static approximation [523]. Consider, for example, a target disk of radius  $r = a$ . The region exterior to the disk is divided into a *near zone* that extends a distance  $\sqrt{Dt}$  from the surface and a complementary *far zone* (see Fig. 2.6b). In the near zone, it is assumed that diffusing particles have sufficient time to explore the domain before being absorbed by the target so that the concentration in the near zone can be treated as almost steady or quasi-static. Conversely, it is assumed that the probability of a particle being absorbed by the target is negligible in the far zone, since a particle is unlikely to diffuse more than a distance  $\sqrt{Dt}$  over a time interval of length  $t$ . Thus,  $c(r) \approx c_0$  for  $r > \sqrt{Dt} + a$ . The near zone concentration is taken to be a radially symmetric solution of Laplace's equation, which for  $d = 2$  is  $c(r) = A + B \log r$ . Matching the solution to the boundary conditions  $c(a) = 0$  and  $c(a + \sqrt{Dt}) = c_0$  then gives (for  $\sqrt{Dt} \gg a$ )

$$c(r, t) \approx \frac{c_0 \log(r/a)}{\log(\sqrt{Dt}/a)}.$$

The corresponding time-dependent flux is

$$J(t) \approx \frac{2\pi D c_0}{\log(\sqrt{Dt}/a)}.$$

### 2.4.2 Chemoreception

In the above example, we considered a binary reaction in which the target molecule was a perfect absorber, that is, when a diffusing particle hits the target the reaction occurs immediately. One application of diffusion-limited reactions is to chemoreception. In the case of a bacterium such as *E. coli*, the cell surface is covered in receptors that detect signaling molecules in the surrounding environment (see Fig. 2.7). Treating the cell as a perfect absorber assumes that there is a sufficient number of receptors distributed on the cell surface and that binding of a signaling molecule is instantaneous when it hits the surface. There are two major simplifications of such a model—(i) receptors tend to be nonuniformly distributed on the cell surface and (ii) the rate of receptor/ligand binding  $k_{\text{on}}$  is finite. We shall address the second issue below. The role of receptor clustering in signal amplification will be addressed in Sect. 5.3, where we discuss the biochemical networks involved in bacterial chemotaxis.

Consider a spherical cell with  $M$  receptors distributed uniformly across its surface. Assuming that the concentration  $c(r)$  of signaling molecules around the cell has reached steady state, the number of molecules absorbed per unit time is

$$\frac{dn}{dt} = Mk_{\text{on}}c(a),$$

where  $a$  is the radius of the cell. From mass conservation this must be balanced by the diffusive flux through any virtual sphere of radius  $r$ ,  $r > a$ , centered about the cell:

$$-4\pi r^2 J(r) = 4\pi r^2 D \frac{dc}{dr} = Mk_{\text{on}}c(a),$$

which on integration yields

$$c(r) - c(a) = \int_a^r \frac{Mk_{\text{on}}c(a)}{4\pi Dr^2} dr = \frac{Mk_{\text{on}}c(a)}{4\pi Da} \left( \frac{1}{a} - \frac{1}{r} \right).$$

Finally, using the far-field condition  $c(\infty) = c_0$ , the concentration at the surface is

$$c(a) = \frac{c_0}{1 + Mk_{\text{on}}/(4\pi Da)}.$$

Hence the net absorption rate is

$$k = \frac{4\pi Da Mk_{\text{on}}}{4\pi Da + Mk_{\text{on}}}. \quad (2.4.2)$$

In the limit  $Mk_{\text{on}} \rightarrow \infty$  we recover the result for a perfect absorber with  $c(a) \rightarrow 0$ . On the other hand, if  $Mk_{\text{on}} \ll Da$ , then the depletion rate is so slow that  $c(a) \approx c_0$ , the background concentration.

Chemoreceptors allow motile *E. coli* to detect changes in concentration of a chemoattractant (food source). *E. coli* propels itself by rotating its flagella. In or-

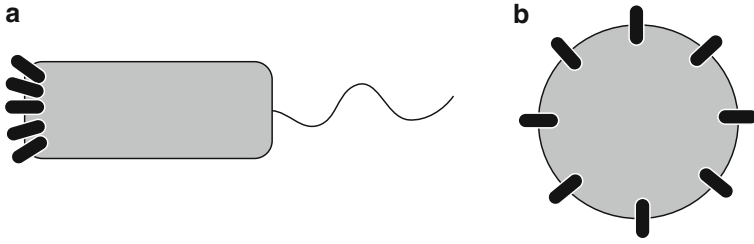


Fig. 2.7: Schematic diagram of (a) a polarized cell such as *E. coli* with a cluster of chemoreceptors and (b) a spherical cell with a uniform distribution of receptors

der to move forward, the flagella rotate together counter clockwise (CCW) enabling the bacterium to “swim” at low Reynolds number. However, when the flagella rotation abruptly changes to clockwise, the bacterium “tumbles” in place and seems incapable of going anywhere. Then the bacterium begins swimming again in some new, random direction. Swimming is more frequent as the bacterium approaches a chemoattractant (food). Tumbling, hence direction change, is more frequent as the bacterium moves away from the chemoattractant. It is the complex combination of swimming and tumbling that keeps them in areas of higher food concentrations. One important issue is why *E. coli* has to move in order to detect changes in concentration rather than simply comparing differences across its body length. The answer is that there are limitations to the sensitivity of chemoreception due to thermal noise, which means that typical concentration changes along a cell body of size  $1\text{ }\mu\text{m}$  are below the signal-to-noise ratio (SNR). This observation was first made in a classical paper of Berg and Purcell [40], whose analysis will be presented in Sect. 5.1. One heuristic way to estimate the sensitivity is to assume that a bacterium integrates signals from chemoreceptors for a mean time  $\tau_{\text{avg}}$ . Assuming a perfect absorber for simplicity, the total number of signaling molecules absorbed is then  $N \sim aDc\tau_{\text{avg}}$ . Based on the law of large numbers, we expect fluctuations in the number of molecules to vary as  $\sqrt{N}$ . Hence,

$$\frac{\delta c}{c} \sim \frac{\delta N}{N} \sim \frac{1}{\sqrt{Dac\tau_{\text{avg}}}}. \quad (2.4.3)$$

Taking  $D \sim 10^{-5}\text{ cm}^2/\text{s}$ ,  $a \sim 1\text{ }\mu\text{m}$  and a typical concentration  $c = 6 \times 10^{11}$  molecules per  $\text{cm}^3$ , we have  $Dac \sim 600\text{ s}^{-1}$ . Assuming that the bacterium integrates for a time  $\tau_{\text{avg}} \sim 1.5\text{ s}$ , then  $\delta c/c \sim 1/30$ . Changes in  $c$  across  $1\text{ }\mu\text{m}$  are just too small to detect. However, since the speed of motion is  $v \sim 10 - 20\text{ }\mu\text{m}/\text{s}$ , it is possible to sample concentration changes of a length scale up to 30 times longer. Note that there is a limit to how large a time  $\tau_{\text{avg}}$  the bacterium can integrate a chemical signal during a run, since rotational diffusion will interfere with the run’s direction over longer time scales. The problem of rotational diffusion is discussed in Ex. 2.13.

## 2.5 Boundary Value Problems

As we have already highlighted, diffusion within the cell is bounded and often restricted to some subcellular compartment with complex geometry. This means that one has to solve an initial boundary value problem for the FPE on a bounded domain  $\Omega \subset \mathbb{R}^d$  with  $d = 1, 2, 3$ . In this section we describe some methods for solving initial boundary value problems in the more specific case of the diffusion equation (see also [395, 554]). The same methods can be applied to the FP equation, although the analysis tends to be considerably more involved when the drift term and diffusivity are space-dependent [204].

### 2.5.1 Eigenfunction Expansions

Let  $u = u(\mathbf{x}, t)$  satisfy the initial boundary value problem

$$\frac{\partial u}{\partial t} = D \nabla^2 u, \quad \mathbf{x} \in \Omega, t > 0, \quad (2.5.1)$$

$$u(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega, t > 0, \quad u(\mathbf{x}, 0) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (2.5.2)$$

where  $\partial\Omega$  denotes the boundary of  $\Omega$ . For the sake of illustration, we consider the Dirichlet boundary condition  $u(\mathbf{x}, t) = 0, \mathbf{x} \in \partial\Omega$ . However, the same methods can be applied to the Neumann or no-flux boundary condition

$$J(\mathbf{x}, t) \equiv -D \mathbf{n}_{\mathbf{x}} \cdot \nabla u(\mathbf{x}, t) = 0, \quad \mathbf{x} \in \partial\Omega, \quad (2.5.3)$$

where  $\mathbf{n}_{\mathbf{x}}$  is the unit normal to the boundary at  $\mathbf{x} \in \partial\Omega$  (with the convention that it points outward from the domain  $\Omega$ ). A standard method for solving this initial boundary value problem is *separation of variables*. The first step is to substitute the solution  $u(\mathbf{x}, t) = U(\mathbf{x})T(t)$  into the diffusion equation to give

$$U(\mathbf{x})T'(t) = DT(t)\nabla^2 U(\mathbf{x}),$$

which we rewrite as

$$\frac{T'(t)}{DT(t)} = \frac{\nabla^2 U(\mathbf{x})}{U(\mathbf{x})} = -\lambda.$$

The essential idea of the method is that  $\lambda$  is a constant, since it cannot be both a function of only  $t$  and only  $\mathbf{x}$ . It follows that we can separate the PDE into a spatial part and a temporal part according to

$$T'(t) = -\lambda DT(t) \quad (2.5.4a)$$

$$-\nabla^2 U(\mathbf{x}) = \lambda U(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad U(\mathbf{x}) = 0, \mathbf{x} \in \partial\Omega, \quad (2.5.4b)$$

where the Dirichlet boundary condition has been imposed on the spatial part.

Equation (2.5.4b) is an example of a boundary value problem for the negative Laplacian  $-\nabla^2$ . For each value of  $\lambda$  for which Eq. (2.5.4b) has a nontrivial solution  $U(\mathbf{x})$ ,  $\lambda$  is called an eigenvalue and  $U(\mathbf{x})$  is the corresponding eigenfunction (defined up to an arbitrary, nonzero, scalar multiplication). More formally,  $\lambda$  is an element of the discrete spectrum of the linear operator  $-\nabla^2$  acting on the given solution domain, which is often taken to be the vector space of functions  $L^2(\Omega)$ . It turns out that the given Dirichlet problem has the following properties [554]:

1. The eigenvalues are real.
2. There are infinitely many eigenvalues that can be ordered as  $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$  with  $\lambda_n \rightarrow \infty$  as  $n \rightarrow \infty$ .
3. Eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the standard inner product on  $\Omega$ , that is,

$$\langle \phi_n | \phi_m \rangle \equiv \int_{\Omega} \phi_n(\mathbf{x}) \phi_m(\mathbf{x}) d\mathbf{x} = 0$$

when  $\lambda_n \neq \lambda_m$ . The number of linearly independent eigenfunctions associated with a degenerate eigenvalue is finite, so that a Schmidt orthogonalization procedure can be used to make them orthogonal to each other, which we assume below.

4. The set of eigenfunction  $\phi_n(\mathbf{x})$  is complete in the sense that any square-integrable function  $F \in L^2(\Omega)$  can be uniquely represented by a generalized Fourier series

$$F(\mathbf{x}) = \sum_{n=1}^{\infty} c_n \phi_n(\mathbf{x}), \quad c_n = \frac{\langle F | \phi_n \rangle}{\|\phi_n\|^2},$$

where  $c_n$  are the generalized Fourier coefficients and the norm is  $\|\phi_n\| = \sqrt{\langle \phi_n | \phi_n \rangle}$ . This means that the truncated Fourier series converges in the  $L^2(\Omega)$  sense,

$$\int_{\Omega} \left( f(\mathbf{x}) - \sum_{n=1}^N c_n \phi_n(\mathbf{x}) \right)^2 d\mathbf{x} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Note that the same properties hold when the Dirichlet boundary condition is replaced by the Neumann boundary condition, except that there now exists a zero eigenvalue  $\lambda_0 = 0$  whose eigenfunction  $\phi_0(\mathbf{x}) = \text{constant}$ . (This reflects the fact that the diffusion equation has a nontrivial steady state in the case of a no-flux boundary condition.)

Returning to Eq. (2.5.4), we immediately see that we can identify the constant  $\lambda$  with one of the eigenvalues  $\lambda_n$  of  $-\nabla^2$ . Solving the equation for  $T$  then shows that we have an infinite set of solutions of the form  $u_n(\mathbf{x}, t) = \phi_n(\mathbf{x}) e^{-D\lambda_n t}$ . Since the diffusion equation is linear, we can apply the principle of superposition to write down the general solution

$$u(\mathbf{x}, t) = \sum_{n=1}^{\infty} c_n \phi_n(\mathbf{x}) e^{-\lambda_n D t}. \quad (2.5.5)$$

Finally, imposing the initial condition requires that

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} c_n \phi_n(\mathbf{x}),$$

so that we can identify the  $c_n$  as the generalized Fourier coefficients of  $f$ . That is,

$$c_n = \frac{1}{\|\phi_n\|^2} \int_{\Omega} f(\mathbf{x}) \phi_n(\mathbf{x}) d\mathbf{x}. \quad (2.5.6)$$

Substituting for  $c_n$  into the general solution and taking the eigenfunctions to have unit normalization ( $\|\phi_n\|^2 = 1$ ) yields

$$u(\mathbf{x}, t) = \sum_{n=1}^{\infty} \left( \int_0^L f(\mathbf{y}) \phi_n(\mathbf{y}) d\mathbf{y} \right) e^{-n^2 \pi^2 D t / L^2} \phi_n(\mathbf{x}).$$

Formally switching the order of summation and integration (which is valid provided that the functions are sufficiently well-behaved), the solution can be reexpressed in the compact form

$$u(\mathbf{x}, t) = \int_0^L K(\mathbf{x}, \mathbf{y}, t) f(\mathbf{y}) d\mathbf{y}, \quad (2.5.7)$$

where

$$K(\mathbf{x}, \mathbf{y}, t) = \sum_{n=1}^{\infty} e^{-n^2 \pi^2 D t / L^2} \phi_n(\mathbf{x}) \phi_n(\mathbf{y}). \quad (2.5.8)$$

Finally, taking the limit  $t \rightarrow 0$ , we deduce the completeness relation

$$\sum_{n=1}^{\infty} \phi_n(\mathbf{x}) \phi_n(\mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}). \quad (2.5.9)$$

*Example 2.2.* Consider the following initial boundary value problem for the 1D diffusion equation:

$$\begin{aligned} \frac{\partial u}{\partial t} &= D \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < L, t > 0. \\ u(0, t) &= 0 = u(L, t), \quad t > 0, \\ u(x, 0) &= f(x), \quad 0 < x < L. \end{aligned}$$

After performing separation of variables, we obtain the eigenvalue problem

$$-U''(x) = \lambda U(x), \quad 0 < x < L, \quad U(0) = U(L) = 0.$$

The eigenvalues and eigenfunctions are thus

$$\lambda_n = \frac{n^2 \pi^2}{L^2}, \quad \phi_n(x) = \sin \frac{n\pi x}{L}, \quad n = 1, 2, \dots$$

It follows that the general solution is given by

$$u(x, t) = \sum_{n=1}^{\infty} c_n e^{-n^2 \pi^2 D t / L^2} \sin \frac{n \pi x}{L}.$$

Comparison with the initial data shows that the  $c_n$  are the Fourier coefficients in the series expansion of  $f(x)$ ,

$$f(x) = \sum_{n=1}^{\infty} c_n \sin \frac{n \pi x}{L},$$

and thus

$$c_n = \int_0^L f(\xi) \sin \frac{n \pi \xi}{L} d\xi / \int_0^L \sin^2 \frac{n \pi x}{L} dx.$$

Evaluating the denominator and substituting for  $c_n$  into the general solution yields

$$u(x, t) = \frac{2}{L} \sum_{n=1}^{\infty} \left( \int_0^L f(\xi) \sin \frac{n \pi \xi}{L} d\xi \right) e^{-n^2 \pi^2 D t / L^2} \sin \frac{n \pi x}{L}.$$

Formally switching the order of summation and integration (which is valid provided that the functions are sufficiently well-behaved), the solution can be reexpressed in the compact form

$$u(x, t) = \int_0^L K(x, \xi, t) f(\xi) d\xi,$$

where

$$K(x, \xi, t) = \frac{2}{L} \sum_{n=1}^{\infty} e^{-n^2 \pi^2 D t / L^2} \sin \frac{n \pi \xi}{L} \sin \frac{n \pi x}{L}.$$

Some further examples of boundary value problems are considered in Exs. 2.13 and 2.14.

### 2.5.2 Green's Functions and Steady-State Analysis

In order for the diffusion equation to have a nontrivial steady state (time-independent solution), it is necessary to include inhomogeneous source terms in the PDE and/or inhomogeneous boundary conditions. Therefore, consider the steady-state equation in  $d = 1, 2, 3$  dimensions

$$\nabla^2 u(\mathbf{x}) = -f(\mathbf{x}), \quad \mathbf{x} \in \Omega \subset \mathbb{R}^d, \quad u(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega. \quad (2.5.10)$$

One way to analyze inhomogeneous equations is to use Green's functions. The Green's function  $G(\mathbf{x}, \mathbf{y})$  for the Dirichlet boundary value problem is defined by the equation

$$\nabla_{\mathbf{y}}^2 G(\mathbf{x}, \mathbf{y}) = -\delta(\mathbf{x} - \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \Omega, \quad G(\mathbf{x}, \mathbf{y}) = 0, \quad \mathbf{y} \in \partial\Omega, \quad (2.5.11)$$

where  $\delta(\mathbf{x})$  is the Dirac delta function in  $\mathbb{R}^d$ . Hence, in Cartesian coordinates for  $d = 3$ ,  $\delta(\mathbf{x}) = \delta(x)\delta(y)\delta(z)$ . Once one has determined the Green's function, the solution of the inhomogeneous boundary value problem can be obtained from the following Green's identity:

$$\begin{aligned} \int_{\Omega} [u(\mathbf{y})\nabla_y^2 G(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y})\nabla_x^2 u(\mathbf{y})] d\mathbf{y} \\ = \int_{\Omega} \nabla_y \cdot [u(\mathbf{y})\nabla_y G(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y})\nabla_y u(\mathbf{y})] d\mathbf{y}. \end{aligned}$$

Applying the steady-state equations to both terms on the left-hand side and using the divergence theorem on the right-hand side shows that

$$-u(\mathbf{x}) + \int_{\Omega} G(\mathbf{x}, \mathbf{y})f(\mathbf{y})d\mathbf{y} = \int_{\partial\Omega} [u(\mathbf{y})\nabla_y G(\mathbf{x}, \mathbf{y}) - G(\mathbf{x}, \mathbf{y})\nabla_y u(\mathbf{y})] \cdot \mathbf{n} d\mathbf{y},$$

where  $\mathbf{n}$  is the outward normal along the boundary  $\partial\Omega$ . Imposing the boundary conditions on  $u$  and  $G$  and rearranging yields the solution

$$u(\mathbf{x}) = \int_{\Omega} G(\mathbf{x}, \mathbf{y})f(\mathbf{y})d\mathbf{y} - \int_{\partial\Omega} \partial_n G(\mathbf{x}, \mathbf{y})g(\mathbf{y})d\mathbf{y}, \quad (2.5.12)$$

where  $\partial_n G$  denotes the normal derivative of  $G$ .

From the properties of the spectrum of the negative Laplacian listed in Sect. 2.5.1, it follows that the Green's function has a formal expansion in terms of the complete set of orthonormal eigenfunctions:

$$G(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x})\phi_n(\mathbf{y})}{\lambda_n}. \quad (2.5.13)$$

This is straightforward to establish, since

$$\begin{aligned} \nabla_y^2 G(\mathbf{x}, \mathbf{y}) &= \nabla_y^2 \left( \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x})\phi_n(\mathbf{y})}{\lambda_n} \right) \\ &= \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x})\nabla_y^2 \phi_n(\mathbf{y})}{\lambda_n} = - \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x})\lambda_n \phi_n(\mathbf{y})}{\lambda_n} \\ &= - \sum_{n=1}^{\infty} \phi_n(\mathbf{x})\phi_n(\mathbf{y}) = -\delta(\mathbf{x} - \mathbf{y}). \end{aligned}$$

We have reversed the order of summation and integration and used the completeness relation (2.5.9). Note that the definition of the Green's function has to be slightly modified in the case of Neumann boundary conditions, since there exists a zero eigenvalue. The so-called generalized or modified Neumann Green's function has the eigenfunction expansion

$$G(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x})\phi_n(\mathbf{y})}{\lambda_n}. \quad (2.5.14)$$



Since the completeness relation (2.5.9) has to be extended to include the constant normalized eigenfunction  $\phi_0(\mathbf{x}) = 1/\sqrt{|\Omega|}$ , where  $|\Omega|$  denotes the volume of the bounded domain  $\Omega$ , we see that the Neumann Green's function satisfies

$$\nabla_y^2 G(\mathbf{x}, \mathbf{y}) = \frac{1}{|\Omega|} - \delta(\mathbf{x} - \mathbf{y}). \quad (2.5.15)$$

One of the significant features of the Dirichlet or Neumann Green's function  $G(\mathbf{x}, \mathbf{y})$  in two and three dimensions is that it is singular in the limit  $\mathbf{x} \rightarrow \mathbf{y}$ . Moreover, these singularities take the specific form

$$G(\mathbf{x}, \mathbf{y}) \sim \ln(|\mathbf{x} - \mathbf{y}|) \text{ (in 2D)}, \quad G(\mathbf{x}, \mathbf{y}) \sim \frac{1}{|\mathbf{x} - \mathbf{y}|} \text{ (in 3D)}. \quad (2.5.16)$$

The nature of these singularities is established in Box 2C.

**Box 2C. The 2D and 3D Green's function for the steady-state diffusion equation.**

Let us begin by considering Laplace's equation in  $\mathbb{R}^2$ :

$$\nabla^2 u(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^2.$$

Since there are no boundaries, this equation is symmetric with respect to rigid body translations and rotations in the plane. This implies that if  $u(\mathbf{x})$  is a solution to Laplace's equation, then so are  $v(\mathbf{x}) = u(\mathbf{x} - \mathbf{a})$  and  $w(\mathbf{x}) = u(\mathbf{R}_\theta \mathbf{x})$ . Here  $\mathbf{a}$  is a constant vector and  $\mathbf{R}_\theta$  is the  $2 \times 2$  rotation matrix about the origin

$$\mathbf{R}_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

This suggests that we look for a radially symmetric solution  $u = u(r)$ . Introducing polar coordinates, Laplace's equation becomes

$$\frac{d^2 u}{dr^2} + \frac{1}{r} \frac{du}{dr} = 0, \quad 0 < r < \infty.$$

The radially symmetric solution is thus of the form

$$u(r) = C_0 \ln(r) + C_1$$

for constants  $C_0, C_1$ . Similarly, radially symmetric solutions in  $\mathbb{R}^3$  satisfy Laplace's equation

$$\frac{d^2 u}{dr^2} + \frac{2}{r} \frac{du}{dr} = 0, \quad 0 < r < \infty,$$

which has the solution

$$u(r) = \frac{C_0}{r} + C_1.$$

For convenience, choosing  $C_1 = 0$ ,  $C_0 = 1/4\pi$  (in 3D), and  $C_0 = -1/2\pi$  (in 2D), we obtain the fundamental solution of the Laplace equation

$$K(\mathbf{x}) = -\frac{1}{2\pi} \ln |\mathbf{x}| \text{ (in 2D)}, \quad K(\mathbf{x}) = \frac{1}{4\pi|\mathbf{x}|} \text{ (in 3D)}.$$

The fundamental solution satisfies Laplace's equation everywhere except the origin, where it is singular. It turns out that  $K$  satisfies the equation

$$\nabla^2 K(\mathbf{x}) = -\delta(\mathbf{x}).$$

We will show this for the 3D case.

Let  $f \in L^2(\mathbb{R}^3)$  be a function that vanishes at  $\infty$ . Define the function

$$u(\mathbf{x}) = \int_{\mathbb{R}^3} K(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) d\mathbf{y} = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{f(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{y}.$$

We will prove that  $\nabla^2 u = -f$  and hence  $\nabla^2 K = -\delta$ . First, it is convenient to rewrite the expression for  $u$  as

$$u(\mathbf{x}) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{f(\mathbf{x} - \mathbf{y})}{|\mathbf{y}|} d\mathbf{y}.$$

Since  $\nabla_x^2 f(\mathbf{x} - \mathbf{y}) = \nabla_y^2 f(\mathbf{x} - \mathbf{y})$ ,

$$\nabla^2 u(\mathbf{x}) = \int_{\mathbb{R}^3} \frac{1}{|\mathbf{y}|} \nabla_y^2 f(\mathbf{x} - \mathbf{y}) d\mathbf{y}.$$

We would like to integrate by parts, but since  $K(\mathbf{y})$  is singular at  $\mathbf{y} = 0$ , we first have to isolate the origin by surrounding it with a small sphere  $B_r(0)$  of radius  $r$ . That is, we write

$$\nabla^2 u(\mathbf{x}) = \left[ \int_{B_r(0)} + \int_{\mathbb{R}^3 \setminus B_r(0)} \right] \frac{1}{4\pi|\mathbf{y}|} \nabla_y^2 f(\mathbf{x} - \mathbf{y}) d\mathbf{y} \equiv I_r + J_r.$$

Here  $\mathbb{R}^3 \setminus B_r(0)$  denotes  $\mathbb{R}^3$  excluding the sphere around the origin. Using spherical polar coordinates,

$$\begin{aligned} |I_r| &\leq \frac{\max |\nabla_y^2 f|}{4\pi} \int_{B_r(0)} \frac{1}{|\mathbf{y}|} d\mathbf{y} = \max |\nabla_y^2 f| \int_0^r \rho d\rho \\ &= \frac{\max |\nabla_y^2 f|}{2} r^2 \rightarrow 0 \text{ as } r \rightarrow 0. \end{aligned}$$

Recalling that  $f$  vanishes at infinity, we can integrate  $J_r$  by parts twice. First,

$$\begin{aligned} J_r &= \frac{1}{4\pi} \int_{\mathbb{R}^3 \setminus B_r(0)} \left[ \nabla_y \cdot \left( \frac{1}{|\mathbf{y}|} \nabla_y f(\mathbf{x} - \mathbf{y}) \right) - \nabla_y \frac{1}{|\mathbf{y}|} \cdot \nabla_y f(\mathbf{x} - \mathbf{y}) \right] d\mathbf{y} \\ &= \frac{1}{4\pi} \int_{\mathbb{R}^3 \setminus B_r(0)} \nabla_y \cdot \left( \frac{1}{|\mathbf{y}|} \nabla_y f(\mathbf{x} - \mathbf{y}) - f(\mathbf{x} - \mathbf{y}) \nabla_y \frac{1}{|\mathbf{y}|} \right) d\mathbf{y} \\ &\quad + \frac{1}{4\pi} \int_{\mathbb{R}^3 \setminus B_r(0)} f(\mathbf{x} - \mathbf{y}) \nabla_y^2 \frac{1}{|\mathbf{y}|} d\mathbf{y}. \end{aligned}$$

Using the fact that  $\nabla_y^2(1/|\mathbf{y}|) = 0$  in  $\mathbb{R}^3 \setminus B_r(0)$ , and applying the divergence theorem, we have

$$J_r = \frac{1}{4\pi} \int_{\partial B_r(0)} \left( \frac{1}{|\mathbf{y}|} \nabla_y f(\mathbf{x} - \mathbf{y}) - f(\mathbf{x} - \mathbf{y}) \nabla_y \frac{1}{|\mathbf{y}|} \right) \cdot \mathbf{n}_y d\mathbf{y}.$$

The first integral vanishes in the limit  $r \rightarrow 0$ , since

$$\frac{1}{4\pi r} \left| \int_{\partial B_r(0)} \nabla_y f(\mathbf{x} - \mathbf{y}) \cdot \mathbf{n}_y d\mathbf{y} \right| \leq r \max |\nabla_y f| \rightarrow 0.$$

On the other hand, since  $\nabla_y(1/|\mathbf{y}|) = -\mathbf{y}/|\mathbf{y}|^3$  and  $\mathbf{n}_y = -\mathbf{y}/r$ , the second integral yields

$$\frac{1}{4\pi} \int_{\partial B_r(0)} f(\mathbf{x} - \mathbf{y}) \nabla_y \frac{1}{|\mathbf{y}|} \cdot \mathbf{n}_y d\mathbf{y} = \frac{1}{4\pi r^2} \int_{\partial B_r(0)} f(\mathbf{x} - \mathbf{y}) d\mathbf{y} \rightarrow f(\mathbf{x}) \text{ as } r \rightarrow 0.$$

We conclude that  $I_r \rightarrow 0$  and  $J_r \rightarrow -f(\mathbf{x})$  as  $r \rightarrow 0$ , which implies that  $\nabla^2 u = -f$  and  $\nabla^2 K = -\delta$ . A similar analysis can be carried out in 2D using the logarithmic fundamental solution. Finally, given the properties of the fundamental solution  $K(\mathbf{x})$ , we can construct the Green's function for a boundary value problem in terms of  $K(\mathbf{x})$  and a non-singular or regular part that satisfies Laplace's equation everywhere, that is,

$$G(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y}) + R(\mathbf{x}, \mathbf{y}).$$

We will use this result in Sect. 7.2, when considering narrow escape problems and diffusion to small targets.

## 2.6 Appendix: An Informal Introduction to Stochastic Calculus

In this appendix we present an informal introduction to stochastic calculus, following along the lines of Jacobs [298]. A more detailed treatment can be found in Gardiner [204], and a rigorous mathematical account can be found in [483].

The basic approach is to treat a continuous-time stochastic process as the limit of a discrete-time process. That is, an SDE prescribes how a stochastic variable  $X(t)$  changes in each infinitesimal time step  $dt$ . Determining changes over finite times then requires evaluating an associated stochastic integral. In order to make sense of this, we discretize time into small, but finite, intervals of duration  $\Delta t$  and consider a corresponding stochastic difference equation for  $X_n = X(n\Delta t)$ . A more abstract formulation of probability theory and discrete-time stochastic processes is presented in Chap. 11.

### 2.6.1 What Is a Continuous Stochastic Process?

Suppose that an experiment is carried out over a time interval of length  $T$  and has a given set of possible outcomes  $\Omega$ . In the case of tracking a single molecule diffusing in the cell membrane,  $\Omega$  could be the set of all possible trajectories. On the other hand, in the case of an ion channel,  $\Omega$  could specify whether the channel is open or closed. Each time the experiment is run, one obtains a particular realization of a continuous-time stochastic process (or random function)  $X(\omega, t)$  with  $\omega \in \Omega$ . (If the time interval is sampled at discrete times, then one has a discrete-time stochastic process or random sequence.) For fixed  $\omega$ ,  $X(\omega, t) = X_\omega(t)$  is a function of time corresponding to a particular trajectory in state space, which is specified by the parameter  $\omega$ . On the other hand, fixing time  $t$  yields a family of random variables  $X(\omega, t) = X_t(\omega)$  that are parameterized by  $t$ . In the case of diffusion  $X_t(\omega)$  is a continuous random variable, whereas for an ion channel  $X_t(\omega)$  is a discrete random variable (see Sect. 1.3). For concreteness, we will focus on the continuous case. In physical and biological applications, the explicit dependence on the events  $\omega$  and the nature of the underlying probability space are ignored, and one simply writes  $X = X(t)$ . The cumulative distribution function of the stochastic process is defined according to

$$P(x, t) = \text{prob}[X(t) \leq x]. \quad (2.6.1)$$

Using the frequency interpretation of probability, this represents the fraction of trials for which the trajectory through state space does not exceed the value  $x$  at time  $t$ . One can then define the corresponding probability density (assuming it exists) according to

$$p(x, t) = \frac{\partial P(x, t)}{\partial x}. \quad (2.6.2)$$

Moreover, as in classical probability theory, we can introduce joint cumulative distributions and densities

$$P(x_1, \dots, x_n; t_1, \dots, t_n) = \text{Prob}[X(t_1) \leq x_1, \dots, X(t_n) \leq x_n].$$

and

$$p(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{\partial^n P(x_1, \dots, x_n; t_1, \dots, t_n)}{\partial x_1 \dots \partial x_n}.$$

A further extension is the conditional probability density, which takes the form (for  $n = 2$ )

$$p(x_2, t_2 | x_1, t_1) = \frac{p(x_1, x_2; t_1, t_2)}{p(x_1, t_1)}. \quad (2.6.3)$$

Thus,  $p(x_2, t_2 | x_1, t_1)$  is the probability density for  $X(t_2)$  conditioned on  $X(t_1) = x_1$ .

Given a probability density  $p$ , one can define various moments of the stochastic process. Some important examples are the mean

$$x(t) = \langle X(t) \rangle = \int_{-\infty}^{\infty} x p(x, t) dx, \quad (2.6.4)$$

and the two-point *autocorrelation function*

$$\langle X(t_1) X(t_2) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 p(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (2.6.5)$$

A related quantity is the *covariance* given by

$$C(t_1, t_2) = \langle (X(t_1) - x(t_1))(X(t_2) - x(t_2)) \rangle = \langle \langle X(t_1) X(t_2) \rangle \rangle. \quad (2.6.6)$$

The equal-time covariance  $C(t, t)$  is the corresponding variance. Double brackets are often used to denote cumulants of the stochastic process. The latter are defined using a generating function:

$$\begin{aligned} & \left\langle \exp \left( -i \int_0^t X(t') dt' \right) \right\rangle \\ &= \exp \left[ \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_0^t \dots \int_0^t \langle \langle X(t_1) X(t_2) \dots X(t_n) \rangle \rangle dt_1 dt_2 \dots dt_n \right]. \end{aligned} \quad (2.6.7)$$

An important concept in stochastic processes is stationarity: a stochastic process  $X(t)$  is *stationary* if every joint probability distribution for finite  $n$  is invariant under a global time shift:

$$P(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau) = P(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$$

for arbitrary  $\tau$ . It follows that  $P(x, t)$  is time-independent and the covariance  $C(t, t') = C(t - t')$ .

A very important type of stochastic process is a *Markov process*, which is defined by the property that the conditional probability density satisfies

$$p(x_n, t_n | x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1}) = p(x_n, t_n | x_{n-1}, t_{n-1}). \quad (2.6.8)$$

In other words, given  $(x_{n-1}, t_{n-1})$ , the process has no “memory” of values at earlier times. It follows that a Markov process is fully determined by the two functions  $p(x_1, t_1)$  and  $p(x_2, t_2 | x_1, t_1)$ . For example,

$$\begin{aligned} p(x_1, x_2, x_3; t_1, t_2, t_3) &= p(x_3, t_3 | x_1, x_2; t_1, t_2) p(x_1, x_2; t_1, t_2) \\ &= p(x_3, t_3 | x_1, x_2; t_1, t_2) p(x_2, t_2 | x_1, t_1) p(x_1, t_1) \\ &= p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1) p(x_1, t_1). \end{aligned} \quad (2.6.9)$$

However, the functions  $p(x_1, t_1)$  and  $p(x_2, t_2 | x_1, t_1)$  cannot be chosen arbitrarily, since they must obey two important identities. The first is obtained by integrating

$$p(x_1, x_2; t_1, t_2) = p(x_2, t_2 | x_1, t_1) p(x_1, t_1)$$

with respect to  $x_1$ :

$$p(x_2, t_2) = \int_{-\infty}^{\infty} p(x_2, t_2 | x_1, t_1) p(x_1, t_1) dx_1. \quad (2.6.10)$$

The other is obtained by integrating Eq. (2.6.9) with respect to  $x_2$ , assuming that  $t_1 < t_2 < t_3$ :

$$p(x_1, x_3; t_1, t_3) = p(x_1, t_1) \int_{-\infty}^{\infty} p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1) dx_2.$$

Since  $p(x_1, x_3; t_1, t_3) = p(x_3, t_3 | x_1, t_1) p(x_1, t_1)$ , this reduces to the Chapman–Kolmogorov equation

$$p(x_3, t_3 | x_1, t_1) = \int_{-\infty}^{\infty} p(x_3, t_3 | x_2, t_2) p(x_2, t_2 | x_1, t_1) dx_2. \quad (2.6.11)$$

### 2.6.2 Ito Stochastic Integrals

Suppose that the time interval  $[0, T]$  is divided into  $N$  increments of size  $\Delta t = T/N$  and set  $t_n = n\Delta t$ . Consider the stochastic difference equation

$$\Delta X(t_n) \equiv X(t_{n+1}) - X(t_n) = \Delta W_n,$$

where  $\Delta W_n$ ,  $n = 0, \dots, N-1$ , are independent and identically distributed Gaussian variables with zero mean and variance  $\sigma^2 = \Delta t$ :

$$P(\Delta W) = \frac{1}{\sqrt{2\pi\Delta t}} e^{-(\Delta W)^2/2\Delta t}. \quad (2.6.12)$$

(Note that a sequence of random variables is independent and identically distributed (i.i.d.) if each random variable has the same probability distribution as the others

and all are mutually independent.) Iterating the difference equation starting from  $X(0) = 0$  yields

$$X_n \equiv X(n\Delta t) = \sum_{j=0}^{n-1} \Delta W_j.$$

Using the fact that the sum of Gaussian random variables is also a Gaussian, it follows that the probability density for  $X_n$  is a Gaussian. Thus, we only need to determine its mean and variance. Since the  $\Delta W_j$  are all independent, we have

$$\langle X_n \rangle = \sum_{j=0}^{n-1} \langle \Delta W_j \rangle = 0, \quad \text{Var}(X_n) = \sum_{j=0}^{n-1} \text{Var}(\Delta W_j) = n\Delta t,$$

and

$$P(X_n) = \frac{1}{\sqrt{2\pi n\Delta t}} e^{-X_n^2/(2n\Delta t)}.$$

We can now construct a corresponding continuous-time process by taking the limit  $N \rightarrow \infty$  such that  $\Delta t \rightarrow 0$  with  $N\Delta t = T$  fixed. In particular,

$$X(T) = \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} \Delta W_j \equiv \int_0^T dW(t) \equiv W(T),$$

where  $W(T)$  is identified as a Wiener process—a Gaussian process with independent and stationary increments. (For the moment, we will not worry about the precise meaning of convergence and limits of stochastic variables—this will be addressed below.)  $X(T)$  is still a Gaussian, whose mean and variance are obtained by taking the limit  $N \rightarrow \infty$  of the results for  $X_n$ . We deduce that  $W(t)$  has the Gaussian probability density

$$P(w(t)) = \frac{1}{\sqrt{2\pi t}} e^{-w(t)^2/2t}.$$

Now consider the modified stochastic difference equation

$$X_{n+1} - X_n = f(t_n)\Delta W_n,$$

where  $f(t)$  is a deterministic function of time. Once again  $X_n$  is a Gaussian random variable, with

$$\langle X_n \rangle = \sum_{j=0}^{n-1} \langle f(t_j)\Delta W_j \rangle = 0, \quad \text{Var}(X_n) = \sum_{j=0}^{n-1} \text{Var}(f(t_j)\Delta W_j) = \sum_{j=0}^{n-1} f(t_j)^2 \Delta t.$$

Taking the continuum limit along identical lines to the previous case yields the continuous-time Gaussian variable

$$X(T) = \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} f(t_j)\Delta W_j \equiv \int_0^T f(t)dW(t), \quad (2.6.13)$$

with zero mean and variance

$$\text{Var}(X(T)) = \int_0^T f(s)^2 ds. \quad (2.6.14)$$

Substituting for  $X(T)$  into this equation gives

$$\left\langle \int_0^T f(t) dW(t) \int_0^T f(s) dW(s) \right\rangle = \int_0^T f(s)^2 ds,$$

which can be captured by the rule

$$\langle dW(t)dW(s) \rangle = \delta(t-s) dt ds. \quad (2.6.15)$$

However, care must be taken with this rule when  $\delta(t-s)$  appears inside an integral having  $t$  or  $s$  as one of its limits. For example, consider the double stochastic integral

$$\int_0^T \left[ \int_0^t f(s) dW(s) \right] g(t) dW(t) \equiv \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} \left[ \sum_{m=0}^{n-1} f(t_m) \Delta W_m \right] g(t_n) \Delta W_n.$$

We see that there are no terms in the double sum on the right-hand side that have a product of Wiener increments in the same time interval. Thus, taking the expectation of both sides,

$$\left\langle \int_0^T \left[ \int_0^t f(s) dW(s) \right] g(t) dW(t) \right\rangle = 0.$$

Hence, we require

$$\int_0^t f(s) \delta(t-s) ds = 0, \quad \int_0^t f(s) \delta(s) ds = f(0). \quad (2.6.16)$$

Following the previous examples, let us turn to a discretized version of the general SDE for  $X(t)$ ,

$$dX = a(X, t) dt + b(X, t) dW(t), \quad (2.6.17)$$

which takes the form

$$X_{n+1} - X_n = a(X_n, t_n) \Delta t + b(X_n, t_n) \Delta W_n. \quad (2.6.18)$$

Iterating this equation starting from a fixed  $X(0) = x_0$  yields

$$X_N = x_0 + \sum_{n=0}^{N-1} a(X_n, t_n) \Delta t + \sum_{n=0}^{N-1} b(X_n, t_n) \Delta W_n.$$

The continuum limit then gives the stochastic integral equation

$$X(T) = x_0 + \int_0^T a(X(t), t) dt + \int_0^T b(X(t), t) dW(t), \quad (2.6.19)$$



with the final term defined as the Ito stochastic integral

$$\int_0^T b(X(t), t) dW(t) = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} b(X_n, t_n) \Delta W_n. \quad (2.6.20)$$

The integral equation is not very useful for generating an explicit solution for  $X(t)$ . However, from the definition of the Ito stochastic integral, it immediately follows that

$$\left\langle \int_0^T b(X(t), t) dW(t) \right\rangle = 0, \quad (2.6.21)$$

since  $X_n$  is a function of previous Wiener increments  $\Delta W_{n-1}, \dots, \Delta W_0$  so it is uncorrelated with  $\Delta W_n$ . The stochastic difference equation (2.6.18) is the starting point for developing numerical schemes for solving an SDE. However, if one is interested in carrying out explicit calculations, it is usually more useful to go to the associated FPE for the probability density. In order to derive the FP equation from the corresponding SDE, we first need to consider the object  $(dW)^2$ .

In terms of Wiener increments,

$$\int_0^T (dW(t))^2 = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} (\Delta W_n)^2.$$

Taking the expectation of both sides and using the fact that each  $\Delta W_n$  is an i.i.d. gives

$$\left\langle \int_0^T (dW(t))^2 \right\rangle = \int_0^T \langle (dW(t))^2 \rangle = \int_0^T dt = T. \quad (2.6.22)$$

What about the variance? Using the Gaussian probability density (2.6.12), it is simple to show that

$$\text{Var}[(\Delta W)^2] = 2(\Delta t)^2 = 2T^2/N^2.$$

Hence,

$$\begin{aligned} \text{Var} \left[ \int_0^T (dW(t))^2 \right] &= \lim_{N \rightarrow \infty} \text{Var} \left[ \sum_{n=0}^{N-1} (\Delta W_n)^2 \right] = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} \text{Var} [(\Delta W_n)^2] \\ &= \lim_{N \rightarrow \infty} \frac{2T^2}{N} = 0. \end{aligned}$$

We thus obtain the surprising result that the integral of  $(dW)^2$  is deterministic and thus equal to its mean:

$$\int_0^T (dW(t))^2 = T = \int_0^T dt. \quad (2.6.23)$$

In other words, we can set  $(dW)^2 = dt$  inside integrals, a result known as Ito's rule (see below). Using the higher moments of Gaussians, it can also be shown that  $dW^m = 0$  for  $m > 2$ . We now consider a result from stochastic calculus, which will be useful when discussing numerical simulations, in particular Milstein's method (see Sect. 2.6.6).

*Example 2.3.*  $\int_{t_0}^t W(t')dW(t')$ . Consider the discrete sum

$$\begin{aligned} S_n &= \sum_{n=0}^{N-1} W_n \Delta W_n \\ &= \frac{1}{2} \sum_{n=0}^{N-1} [(W_n + \Delta W_n)^2 - W_n^2 - (\Delta W_n)^2] \\ &= \frac{1}{2} [W(t) - W(t_0)]^2 - \frac{1}{2} \sum_{n=0}^{N-1} (\Delta W_n)^2, \end{aligned}$$

where  $W_0 = W(t_0)$  and  $W_{N-1} = W(t)$ . We now calculate the mean and variance of the last term. First,

$$\left\langle \sum_{n=0}^{N-1} (\Delta W_n)^2 \right\rangle = \sum_{n=0}^{N-1} \langle (\Delta W_n)^2 \rangle = \sum_{n=0}^{N-1} (t_{n+1} - t_n) = t - t_0.$$

Second,

$$\begin{aligned} &\left\langle \left[ \sum_{n=0}^{N-1} (\Delta W_n)^2 - (t - t_0)^2 \right]^2 \right\rangle \\ &= \left\langle \left[ \sum_{n=0}^{N-1} (\Delta W_n)^4 + 2 \sum_{n < m} (\Delta W_n)^2 (\Delta W_m)^2 - 2(t - t_0) \sum_{n=0}^{N-1} (\Delta W_n)^2 + (t - t_0)^2 \right] \right\rangle. \end{aligned}$$

Since  $\Delta W_n$  and  $\Delta W_m$  are independent Gaussian random variables for  $n \neq m$ , we have

$$\langle (\Delta W_n)^2 (\Delta W_m)^2 \rangle = (t_{n+1} - t_n)(t_{m+1} - t_m),$$

and the fourth moment of a Gaussian is given by

$$\langle (\Delta W_n)^4 \rangle = 3 \langle (\Delta W_n)^2 \rangle^2 = 3(t_{n+1} - t_n)^2.$$

Hence,

$$\begin{aligned} \left\langle \sum_{n=0}^{N-1} (\Delta W_n)^2 \right\rangle &= 2 \sum_{n=0}^{N-1} (t_{n+1} - t_n)^2 + \sum_{n,m} [(t_{n+1} - t_n) - (t - t_0)][(t_{m+1} - t_m) - (t - t_0)] \\ &= 2 \sum_{n=0}^{N-1} (t_{n+1} - t_n)^2 \rightarrow 0 \text{ as } N \rightarrow \infty. \end{aligned}$$

We deduce that

$$\int_{t_0}^t W(t')dW(t') = \frac{1}{2} [W(t)^2 - W(t_0)^2 - (t - t_0)]. \quad (2.6.24)$$

### 2.6.3 Ito's Formula and the Fokker–Planck Equation

The result  $dW(t)^2 = dt$  has important implications for how one carries out a change of variables in stochastic calculus. This is most directly established by considering the SDE for an arbitrary function  $f(X(t))$  with  $X(t)$  evolving according to Eq. (2.6.17):

$$\begin{aligned} df(X(t)) &= f(X(t) + dX(t)) - f(X(t)) \\ &= f'(X(t))dX(t) + \frac{1}{2}f''(X(t))dX(t)^2 + \dots \\ &= f'(X(t))[a(X,t)dt + b(X,t)dW(t)] + \frac{1}{2}f''(X(t))b(X,t)^2dW(t)^2, \end{aligned}$$

where all terms of higher order than  $dt$  have been dropped. Now using  $dW(t)^2 = dt$ , we obtain the following SDE for  $f$ , which is known as Ito's formula:

$$df(X(t)) = \left[ a(X(t),t)f'(X(t)) + \frac{1}{2}b(X,t)^2f''(X(t)) \right] dt + b(X,t)f'(X(t))dW(t). \quad (2.6.25)$$

Hence, changing variables in Ito calculus is not given by ordinary calculus unless  $f$  is a constant or a linear function.

We can now use Ito's formula to derive the FP equation for an Ito SDE. First,

$$\begin{aligned} \frac{\langle df(X(t)) \rangle}{dt} &= \left\langle a(X(t),t)f'(X(t)) + \frac{1}{2}b(X(t),t)^2f''(X(t)) \right\rangle \\ &= \int \left[ a(x,t)f'(x) + \frac{1}{2}b(x,t)^2f''(x) \right] p(x,t)dx, \\ &= \int f(x) \left[ -\frac{\partial}{\partial x}(a(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(b(x,t)^2p(x,t)) \right] dx, \end{aligned} \quad (2.6.26)$$

after integration by parts, where  $p(x,t)$  is the probability density of the stochastic process  $X(t)$  under the initial condition  $X(t_0) = x_0$ . However, we also have

$$\begin{aligned} \frac{\langle df(X(t)) \rangle}{dt} &= \left\langle \frac{df(X(t))}{dt} \right\rangle \\ &= \frac{d}{dt} \langle f(X(t)) \rangle \\ &= \int f(x) \frac{\partial}{\partial t} p(x,t) dx. \end{aligned} \quad (2.6.27)$$

Comparing Eqs. (2.6.26) and (2.6.27) and using the fact that  $f(x)$  is arbitrary, we obtain the Ito version of the FP equation

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x}(a(x,t)p(x,t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(b(x,t)^2p(x,t)). \quad (2.6.28)$$

### 2.6.4 Multiplicative Noise and Stratonovich Integrals

It turns out that there is more than one way to define a stochastic difference equation driven by an incremental Wiener process and thus more than one way to obtain an SDE in the continuum limit. This issue only arises in the case of multiplicative noise, that is, when the term multiplying  $dW(t)$  depends on the state variable  $X(t)$ . Recall that in the Ito integral Eq. (2.6.20), it is the value of  $b(x, t)$  at the start of the  $n$ th time step that multiplies  $\Delta W_n$ , so that there are no contributions of the form  $(\Delta W_n)^2$ . An alternative definition of a stochastic integral is the Stratonovich integral

$$\oint_0^T b(X(t), t) dW(t) = \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} b\left(\frac{X_{n+1} + X_n}{2}, t_n\right) \Delta W_n, \quad (2.6.29)$$

where we have used  $\oint$  to distinguish it from the Ito integral. Now  $b$  depends on the value  $X_{n+1}$  at the end of the  $n$ th time step, which means there will be an extra term involving  $(\Delta W_n)^2$ . In order to compare the Ito and Stratonovich integrals, suppose that  $X_n$  evolves according to the stochastic difference equation (2.6.18). Thus, in the continuum limit,  $X(t)$  is the solution to an Ito SDE. Suppose that we Taylor expand the  $n$ th term in the sum defining the Stratonovich integral about the point  $X_n$  and set  $b_n = b(X_n, t_n)$ :

$$b\left(\frac{X_{n+1} + X_n}{2}, t_n\right) = b_n + \frac{\Delta X_n}{2} \frac{\partial b_n}{\partial x} + \frac{1}{2} \left(\frac{\Delta X_n}{2}\right)^2 \frac{\partial^2 b_n}{\partial x^2} + \dots$$

Substituting for  $\Delta X_n$  using Eq. (2.6.18) and dropping terms that are higher order than  $\Delta t$  shows that

$$b\left(\frac{X_{n+1} + X_n}{2}, t_n\right) = b_n + \left(\frac{a_n}{2} \frac{\partial b_n}{\partial x} + \frac{b_n^2}{8} \frac{\partial^2 b_n}{\partial x^2}\right) \Delta t + \left(\frac{b_n}{2} \frac{\partial b_n}{\partial x}\right) \Delta W_n.$$

Applying this result to the sum appearing in the definition of the Stratonovich integral, Eq. (2.6.29), and again dropping higher-order terms in  $\Delta t$  yields the result

$$\sum_{n=0}^{N-1} b\left(\frac{X_{n+1} + X_n}{2}, t_n\right) \Delta W_n = \sum_{n=0}^{N-1} b_n \Delta W_n + \sum_{n=0}^{N-1} \frac{b_n}{2} \frac{\partial b_n}{\partial x} (\Delta W_n)^2.$$

Finally, taking the continuum limit with  $dW(t)^2 = dt$ , we have

$$\oint_0^T b(X(t), t) dW(t) = \int_0^T b(X(t), t) dW(t) + \frac{1}{2} \int_0^T \frac{\partial b(X(t), t)}{\partial x} b(X(t), t) dt. \quad (2.6.30)$$

Now suppose that  $Y(t)$  is a stochastic process evolving according to the Stratonovich SDE

$$dY = a(Y, t) dt + b(Y, t) dW(t). \quad (2.6.31)$$

This means that the integral equation satisfied by  $Y(t)$  is based on the Stratonovich integral, that is,

$$Y(t) = y_0 + \int_0^t a(Y(s), s) ds + \oint_0^t b(Y(t), t) dW(t). \quad (2.6.32)$$

Using Eq. (2.6.30), we can rewrite the solution in terms of an Ito integral according to

$$Y(t) = y_0 + \int_0^t \left[ a(Y(s), s) + \frac{1}{2} \frac{\partial b(Y(s), s)}{\partial y} b(Y(s), s) \right] ds + \int_0^t b(Y(s), s) dW(s). \quad (2.6.33)$$

The latter is the solution to an equivalent Ito SDE of the form

$$dY = \left[ a(Y(t), t) + \frac{b(Y(t), t)}{2} \frac{\partial b(Y(t), t)}{\partial y} \right] dt + b(Y(t), t) dW(t). \quad (2.6.34)$$

Finally, given that we know the FP equation corresponding to an Ito SDE, we can immediately write down the FP equation corresponding to the Stratonovich SDE (2.6.31):

$$\frac{\partial}{\partial t} p(y, t) = -\frac{\partial}{\partial y} (a(y, t) p(y, t)) + \frac{1}{2} \frac{\partial}{\partial y} \left( b(y, t) \frac{\partial}{\partial y} [b(y, t) p(y, t)] \right). \quad (2.6.35)$$

### 2.6.5 Ito Integration and Convergence

So far we have not been specific about the form of convergence used to take the continuum limit of a discrete sum of random variables in order to define a stochastic integral. Following Gardiner [204], we now revisit some results on Ito calculus using the notion of convergence in the mean-square. That is, we define a random variable  $X$  to be the limit of a sequence of random variables  $\{X_1, X_2, \dots, X_n\}$  if

$$\lim_{n \rightarrow \infty} \langle |X - X_n|^2 \rangle = 0, \quad (2.6.36)$$

that is, for any  $\varepsilon > 0$ , there exists an integer  $N = N(\varepsilon)$  such that for all  $n > N$ ,  $\langle |X - X_n|^2 \rangle < \varepsilon$ . Given this definition of convergence, a stochastic process  $X(t)$  is said to be mean-square integrable on the interval  $(0, t)$  if there exists a random process  $Z(t)$  such that the following limit exists:

$$\lim_{n \rightarrow \infty} \langle (Z_n - Z(t))^2 \rangle = 0, \quad (2.6.37)$$

where

$$Z_n = \Delta t \sum_{j=0}^n X(j\Delta t), \quad n\Delta t = t.$$

We then formally write  $Z(t) = \int_0^t X(s) ds$ .

Suppose that  $G(t)$  is a *non-anticipating function*, that is,  $G(t)$  is statistically independent of  $W(s) - W(t)$  for all  $s > t$ , where  $W(t)$  is a Wiener process. We will show that

$$\int_0^t G(t') [dW(t')]^2 = \int_0^t G(t') dt',$$

in the mean-square sense, that is, Eq. (2.6.37) holds with

$$Z_n = \sum_{j=0}^n G_j [\Delta W_j]^2, \quad Z(t) = \int_0^t G(t') dt',$$

where  $G_j = G(j\Delta t)$  and  $\Delta W_j = W((j+1)\Delta t) - W(j\Delta t)$ . Consider

$$\begin{aligned} I &= \lim_{n \rightarrow \infty} \left\langle \left[ \sum_j G_j (\Delta W_j^2 - \Delta t) \right]^2 \right\rangle \\ &= \lim_{n \rightarrow \infty} \left\langle \sum_j G_j^2 (\Delta W_j^2 - \Delta t)^2 + 2 \sum_{i>j} G_i G_j (\Delta W_j^2 - \Delta t) (\Delta W_i^2 - \Delta t) \right\rangle. \end{aligned}$$

Note that  $G_j^2$  is statistically independent of  $(\Delta W_j^2 - \Delta t)^2$  and  $G_i G_j (\Delta W_j^2 - \Delta t)$  is statistically independent of  $(\Delta W_i^2 - \Delta t)$  for  $j < i$ . Using the Gaussian nature of  $\Delta W_i$ , we have

$$\langle \Delta W_i^2 \rangle = \Delta t, \quad \langle (\Delta W_i^2 - \Delta t)^2 \rangle = 2\Delta t^2.$$

Thus we find that

$$I = 2 \lim_{n \rightarrow \infty} \sum_j G_j^2 \Delta t^2 = 0,$$

assuming that  $G(t)$  is bounded. Thus, for Ito integrals  $dW(t)^2$  acts like  $dt$ .

### 2.6.6 Simulation of Stochastic Differential Equations

Consider the scalar SDE

$$dX = a(X)dt + b(X)dW(t), \quad (2.6.38)$$

where  $W(t)$  is a Wiener process. As with ordinary differential equations, the simplest numerical scheme is to use a direct Euler method. That is, given the solution  $X(t)$  at time  $t$ , the solution at time  $t + \Delta t$  is given by  $X(t + \Delta) = X(t) + \Delta X$ , where  $\Delta X$  is determined explicitly by the equation

$$\Delta X = a(X(t))\Delta t + b(X(t))\Delta W, \quad (2.6.39)$$

for a Gaussian random variable  $\Delta W$  with zero mean and variance equal to  $\Delta t$ . Iterating this process using a random number generator to choose a new value of  $\Delta W$  at each time step  $\Delta t$  results in an approximation of a sample path of the stochastic process  $X(t)$ . Repeating the simulation over many independent trials up to a time  $T$  then generates a histogram of values of  $X(T)$ , which can be used to determine an approximation of the probability density for  $X(T)$  and to estimate the mean and variance. The direct Euler method is easily extended to multivariate SDEs and those with nonautonomous coefficients  $a, b$ .

The accuracy of Euler's method increases with decreasing step size  $\Delta t$ , and the approximate sample path converges in mean-square to the true sample path in the limit  $\Delta t \rightarrow 0$ . For a rigorous discussion of estimating the accuracy of a stochastic numerical algorithm see the book by Kloeden and Platen [338]. Here we give a heuristic definition of the numerical error (see also [204, 298]). Suppose that the time interval  $(0, T)$  is divided into  $N$  infinitesimal subintervals of size  $\tau = T/N$ , so that the stochastic process  $X(t)$  is sampled at the times  $\tau_n$  with  $n = 0, \dots, N$ . Let  $x_n = X(\tau_n)$  be the exact solution on a given sample path and  $y_n$  the corresponding numerical approximation of the solution on the *same* sample path. At the  $n$ th time step let  $e_n = x_n - y_n$  and define the error at time  $T$  to be the root mean-square (RMS) value

$$E(T) = \sqrt{\langle e_N^2 \rangle}. \quad (2.6.40)$$

In the case of the direct Euler method,

$$E(T) \sim \tau^{1/2},$$

and the Euler method is said to be accurate to  $\tau^{1/2}$  or that the order of convergence is  $\tau^{1/2}$ . (In general the order of convergence will depend on how we define the numerical error, that is, the particular measure of convergence. We will restrict ourselves to mean-square convergence.) One practical method for checking the accuracy of a numerical simulation of a given sample path is to repeat the simulation after halving the time step  $\Delta t$ . Suppose that  $T = N\Delta t$  and the sample path is generated by the  $N$  random increments  $\Delta W_n$ ,  $n = 0, \dots, N-1$ . If we then halve the time step, then in order to generate an approximation to the same sample path, it is necessary to produce a set of  $2N$  Gaussian random numbers  $\Delta \hat{W}_m$ ,  $m = 1, \dots, 2N$ , such that

$$\Delta \hat{W}_{2n} + \Delta \hat{W}_{2n+1} = \Delta W_n, \quad n = 0, \dots, N-1.$$

Given the values  $\Delta W_n$ , this can be realized by generating  $N$  random variables  $r_n$  with zero mean and variance  $\Delta t/2$ , and setting

$$\Delta \hat{W}_{2n} = r_n, \quad \Delta \hat{W}_{2n+1} = \Delta W_n - r_n.$$

One can thus successively halve the time step until errors are within acceptable bounds for the given application. The method can also be used to estimate the rate of convergence.

### 2.6.7 Milstein's Method

The direct Euler method is a low-order numerical method for SDEs due to the fact that in approximating an SDE one needs to take into account the fact that  $dW(t)$  is of order  $\sqrt{dt}$ . Suppose that we rewrite the scalar SDE (2.6.38) as the integral equation

$$X(t) = X(t_0) + \int_{t_0}^t a(X(s))ds + \int_{t_0}^t b(X(s))dW(s). \quad (2.6.41)$$

We recover Euler's method by taking  $t = t_0 + \Delta t$ , with  $X(t_0)$  known, and approximating the functions  $a, b$  in the interval  $s \in (t_0, t_0 + \Delta t)$  according to

$$a(X(s)) \approx a(X(t_0)), \quad b(X(s)) \approx b(X(t_0)).$$

In order to obtain a more accurate approximation, we apply Ito's formula (2.6.25) to the functions  $a$  and  $b$ . For example,

$$\begin{aligned} b(X(s)) &= b(X(t_0)) + \int_{t_0}^s \left[ a(X(s'))b'(X(s')) + \frac{1}{2}b(X(s'))^2b''(X(s')) \right] ds' \\ &\quad + \int_{t_0}^s b(X(s'))b'(X(s'))dW(s') \end{aligned} \quad (2.6.42)$$

and similarly for  $a(X(s))$ . Iterating these equations by successively applying Ito's formula to  $a(X(s'))$  and so forth generates an approximation of  $a(X(s))$  and  $b(X(s))$  in terms of  $a(X(t_0)), b(X(t_0))$ , higher-order derivatives of  $a(x), b(x)$  evaluated at  $x_0$ , and a remainder. Substituting such an expansion of  $a(X(s))$  and  $b(X(s))$  into Eq. (2.6.41) generates a higher-order numerical scheme. The Milstein method is the next higher-order approximation to the stochastic integral Eq. (2.6.41) after Euler's method. It is obtained by substituting

$$a(X(s)) \approx a(X(t_0)), \quad b(X(s)) \approx b(X(t_0)) + b(X(t_0))b'(X(t_0)) \int_{t_0}^s dW(s')$$

into Eq. (2.6.41) for  $t = t_0 + \Delta t$  and  $s \in (t_0, t_0 + \Delta t)$ . This yields the following equation for  $\Delta X$ :

$$\begin{aligned} \Delta X &= a(X(t_0)) \int_{t_0}^{t_0+\Delta t} ds + b(X(t_0)) \int_{t_0}^{t_0+\Delta t} dW(s) \\ &\quad + \frac{1}{2}b(X(t_0))b'(X(t_0)) \int_{t_0}^{t_0+\Delta t} \int_{t_0}^s dW(s)dW(s'). \end{aligned}$$

The double integral can be evaluated using Eq. (2.6.24). That is,



$$\begin{aligned}
\int_{t_0}^{t_0+\Delta t} \int_{t_0}^s dW(s)dW(s') &= \int_{t_0}^{t_0+\Delta t} [W(s) - W(t_0)]dW(s) \\
&= \int_{t_0}^{t_0+\Delta t} W(s)dW(s) - W(t_0)W(t_0 + \Delta t) \\
&= \frac{1}{2}[W(t_0 + \Delta t)^2 - W(t_0)^2 - \Delta t] - W(t_0)W(t_0 + \Delta t) \\
&= \frac{1}{2}[(\Delta W)^2 - \Delta t].
\end{aligned}$$

Hence, we arrive at the Milstein algorithm

$$\begin{aligned}
\Delta X &= \left[ a(X(t_0)) - \frac{1}{2}b(X(t_0))b'(X(t_0)) \right] \Delta t + b(X(t_0))\Delta W \\
&\quad + \frac{1}{2}b(X(t_0))b'(X(t_0))\Delta W^2.
\end{aligned} \tag{2.6.43}$$

It turns out that this algorithm has order  $\Delta t$  accuracy, which improves upon the  $\sqrt{\Delta t}$  accuracy of Euler's method.

The complexity of Milstein's method increases when there are multiple noise sources. Consider the multivariate SDE

$$dX_i = a_i(\mathbf{X}(t), t)dt + \sum_{j=1}^M B_{ij}(\mathbf{X}(t), t)dW_j(t),$$

where  $W_i(t)$  are independent Wiener processes. The Milstein approximation of this equation takes the form

$$\Delta X_i = a_i \Delta t + \sum_{j=1}^M B_{ij} \Delta W_j + \sum_{j,k=1}^M \left[ \sum_{m=1}^M B_{mj} \frac{\partial B_{ik}}{\partial X_m} \right] \int_{t_0}^{t_0+\Delta t} \int_{t_0}^s dW_j(t')dW_k(s).$$

Unfortunately, only when  $j = k$  does the double integral reduce to a simple expression involving the discrete stochastic increments  $\Delta W_j$  along the lines of the scalar case. However, it can be shown that the symmetrized integral is also reducible according to

$$\int_{t_0}^{t_0+\Delta t} \int_{t_0}^s [dW_i(s)dW_j(s') + dW_j(s)dW_i(s')] = \Delta W_i \Delta W_j - \delta_{i,j}(t - t_0).$$

It follows that when the matrix  $\mathbf{G}$  satisfies the set of relations (commutative noise)

$$\sum_{m=1}^M B_{mj} \frac{\partial B_{ik}}{\partial x_m} = \sum_{m=1}^M B_{mk} \frac{\partial B_{ij}}{\partial x_m} \tag{2.6.44}$$

for all  $i, j, k$ , the double integral can be symmetrized and Milstein's algorithm becomes

$$\begin{aligned} \Delta X_i = & a_i \Delta t + \sum_{j=1}^M B_{ij} \Delta W_j + \frac{1}{2} \sum_{j,k=1}^M \left[ \sum_{m=1}^M B_{mj} \frac{\partial B_{ik}}{\partial X_m} \right] \Delta W_j \Delta W_k \\ & - \frac{1}{2} \sum_{j,k=1}^M \left[ \sum_{m=1}^M B_{mj} \frac{\partial B_{ij}}{\partial X_m} \right] \Delta t. \end{aligned}$$

### 2.6.8 Runge–Kutter and Implicit Methods

One limitation of the Milstein method is that it requires an evaluation of the first derivative of the function  $b(X)$  or its higher-dimensional matrix version. In a similar fashion to deterministic equations, one can use a Runger–Kutta method to eliminate the need to evaluate any derivatives. A first-order method that builds upon the Milstein algorithm has been developed by Platen [338]. It is based on using the approximation

$$b(X)b'(X) \approx \frac{1}{\sqrt{\Delta t}} [b(\hat{X}) - b(X)],$$

where

$$\hat{X} = X + a\Delta t + b\sqrt{\Delta t}.$$

Substituting into Eq. (2.6.43) yields the Milstein–Platen method

$$\Delta X = a\Delta t + b\Delta W + \frac{1}{2\sqrt{\Delta t}} [b(\hat{X}) - b(X)][(\Delta W)^2 - \Delta t]. \quad (2.6.45)$$

Similarly, for a multivariate process, one substitutes into the Milstein method the approximation

$$\sum_{m=1}^M B_{mj}(\mathbf{X}, t) \frac{\partial B_{ik}(\mathbf{X}, t)}{\partial X_m} \approx \frac{1}{\sqrt{\Delta t}} [B_{ij}(\hat{\mathbf{X}}^{(k)}) - B_{ij}(\mathbf{X})]$$

with

$$\hat{X}_i^{(k)} = X_i + a_i \Delta t + B_{ik} \sqrt{\Delta t}.$$

Another issue that numerical methods for solving SDEs share with their deterministic counterparts is instability. This refers to a rapid, exponential increase in numerical error, which can occur spontaneously even though the algorithm appears to be converging to a numerically accurate solution prior to the instability. This feature is a particular problem for “stiff” differential equations, that is, those that have two or more disparate time scales. Often an instability can be fixed by using an *implicit* rather than an explicit method. For example, consider a simple Euler scheme for a single variable,

$$X(t + \Delta t) = X(t) + \Delta X(t) = X(t) + a(X(t), t)\Delta t + b(X(t), t)\Delta W(t).$$

The implicit version is obtained by replacing  $X(t)$  with  $X(t + \Delta t)$  in the functions  $a, b$ :

$$X(t + \Delta t) = X(t) + \Delta X(t) = X(t) + a(X(t + \Delta t), t)\Delta t + b(X(t + \Delta t), t)\Delta W(t).$$

This is clearly an implicit equation for  $X(t + \Delta t)$ , which can be solved numerically using the Newton–Raphson method.

Finally, note that we have focused on the speed and accuracy of numerical methods for generating sample paths of a SDE. Convergence to a sample path is known as *strong convergence*. If one is only interested in properties of the corresponding probability density such as the mean and variance, then these properties are determined by averaging over many sample paths. For a given numerical method, the rate of convergence to the mean or variance tends to differ from the rate of strong convergence and is thus referred to as *weak convergence*.

## 2.7 Exercises

---

**Problem 2.1 (1D random walk).** Consider the probability distribution for a 1D unbiased random walk

$$P_N(r) = \frac{1}{2^N} \frac{N!}{\left(\frac{N+r}{2}\right)! \left(\frac{N-r}{2}\right)!}.$$

Using Stirling's formula

$$\log N! \approx N \log N - N + \frac{1}{2} \ln(2\pi N),$$

derive the Gaussian approximation

$$P_N(r) \sim \frac{1}{\sqrt{2\pi N}} e^{-r^2/2N}.$$

This result includes a factor of  $1/2$  in order to take into account the fact that  $r$  is even (odd) when  $N$  is even (odd).

**Problem 2.2 (Random walk on a lattice).** Consider a random walker on a 1D lattice with sites  $\ell$  and displacement distribution  $p(\ell)$ . The probability  $P_n(\ell)$  that the walker is at site  $\ell + \ell_0$  after  $n$  steps, starting at  $\ell_0$ , satisfies the recurrence relation

$$P_n(\ell) = \sum_{\ell'} p(\ell - \ell') P_{n-1}(\ell').$$

(For a homogeneous random walk,  $\ell_0$  is arbitrary so we can set  $\ell_0 = 0$ .) Define the generating function  $\Gamma(\ell, z)$  according to

$$\Gamma(\ell, z) = \sum_{n \geq 0} z^n P_n(\ell).$$

(a) Show that the generating function satisfies the equation

$$\Gamma(\ell, z) = \delta_{\ell,0} + z \sum_{\ell'} p(\ell - \ell') \Gamma(\ell', z).$$

(b) Introduce the discrete Fourier transform

$$\hat{\Gamma}(k, z) = \sum_{\ell} e^{ik\ell} \Gamma(\ell, z)$$

and define the structure function of the walk to be

$$\lambda(k) = \sum_{\ell} e^{ik\ell} p(\ell).$$

From part (a), show that

$$\hat{\Gamma}(k, z) = 1 + z\lambda(k)\hat{\Gamma}(k, z),$$

so that

$$\hat{\Gamma}(k, z) = \frac{1}{1 - z\lambda(k)}.$$

(c) For a standard RW with  $p(\ell) = (\delta_{\ell,1} + \delta_{\ell,-1})/2$ , we have  $\lambda(k) = \cos(k)$ . Using the inverse transform

$$\Gamma(\ell, z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\ell} \hat{\Gamma}(k, z) dk,$$

and the result of part (b), evaluate the integral to show that

$$\Gamma(0, z) = (1 - z^2)^{-1/2}.$$

Hint: make the change of variables  $t = \tan(k/2)$ .

(d) For a general structure function

$$\Gamma(\ell, z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-ik\ell}}{1 - z\lambda(k)} dk,$$

divergence of the integral is only possible if  $\lambda(k_0) = 1$  for some  $k = k_0$ . If this holds then the integral will be dominated by the region around  $k_0$ . Show  $e^{ik_0\ell} = 1$  for all  $\ell$  such that  $p(\ell) > 0$  and hence  $\lambda(k) = \lambda(k - k_0)$ . It follows that the local

behavior of  $1 - \lambda(k)$  near  $k_0$  is the same as the local behavior of  $1 - \lambda(k)$  around the origin. Show that for small  $k$  and an unbiased RW,  $\sum_{\ell} \ell p(\ell) = 0$ , we have

$$\lambda(k) \approx 1 - \frac{\sigma^2 k^2}{2}, \quad \sigma^2 = \sum_{\ell} \ell^2 p(\ell).$$

Hence, deduce that an unbiased 1D RW is recurrent when the MSD per step is finite.

**Problem 2.3 (Chapman-Kolmogorov equation).** Consider a stationary stochastic process  $X(t)$  with initial condition  $X(0) = 0$ . Setting  $p(x, t | x', t') = p(x - x', t - t')$ , the Chapman-Kolmogorov equation takes the form

$$p(x, t) = \int_{-\infty}^{\infty} p(x - y, t - \tau) p(y, \tau) dy.$$

- (a) Using Fourier transforms, show that the CK equation is satisfied if the characteristic function  $G(k, t)$  has the general form  $\ln G(k, t) = t g(k)$  for some function  $g(k)$ .
- (b) Suppose that the probability density of the continuous process evolves according to the equation

$$\frac{dp}{dt} = \int_{-\infty}^{\infty} w(x - y) [p(y, t) - p(x, t)] dy,$$

where  $w$  is a transition probability per unit time. That is, we have a jump process for a continuous random variable. Using Fourier transforms, obtain the solution

$$p(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left[ -ikx + t \int_{-\infty}^{\infty} w(x') \left[ e^{ikx'} - 1 \right] dx' \right] dk.$$

Hence determine  $g(k)$  for this process.

**Problem 2.4 (Electrodiffusion).** The flow of ions through channels in the cell membrane is driven by the combination of concentration gradients and electric fields. If interactions between the ions are ignored, then each ion can be treated as an independent Brownian particle moving under the influence of the electric force  $-q\nabla\phi$ , where  $\phi$  is the electrical potential and  $q$  is the charge on the ion. Multiplying the corresponding FP equation by the number  $N$  of ions and using an Einstein relation, we obtain the *Nernst-Planck equation*

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} = -\nabla \cdot \mathbf{J}, \quad \mathbf{J}(\mathbf{x}, t) = -D \left( \nabla c + \frac{qc}{k_B T} \nabla \phi \right),$$

where  $c$  denotes ion concentration. Treating an ion channel as a quasi-one-dimensional domain, this reduces to the 1D equation

$$\frac{\partial c(x, t)}{\partial t} = -\frac{\partial J}{\partial x}, \quad J(x, t) = -D \left( \frac{\partial c}{\partial x} + \frac{qc}{k_B T} \frac{\partial \phi}{\partial x} \right).$$

- (a) Suppose that the cell membrane extends from  $x = 0$  (inside) to  $x = L$  (outside) and denote the extracellular and intracellular ion concentrations by  $c_e$  and  $c_i$ , respectively. Solve the 1D steady-state Nernst–Planck equation to show that there is zero flux through the membrane if the potential difference  $V = \phi_i - \phi_e$  across the membrane is given by the Nernst potential

$$V = \frac{k_B T}{q} \ln \left( \frac{c_e}{c_i} \right).$$

- (b) Now suppose that there is a constant nonzero flux  $J$  of ions through the channel, and assume for simplicity that the electric field is uniform, that is,  $\partial \phi / \partial x = -V/L$ . Solving the steady-state Nernst–Planck equation with boundary conditions  $c(0) = c_i, c(L) = c_e$ , derive the Goldman–Hodgkin–Katz equation for the current density:

$$I \equiv qJ = \frac{D q^2 V}{L k_B T} \frac{c_i - c_e \exp(-qV/k_B T)}{1 - \exp(-qV/k_B T)}.$$

Check that the Nernst potential is recovered when  $J = 0$ .

- (c) Consider two ion species with opposite charges  $q_1 = -q_2 = q$ . Applying part (b) to the current for each ion species, derive an expression for the membrane voltage  $V$  at which the total ionic current is zero.

**Problem 2.5 (Ornstein–Uhlenbeck process).** Consider the Ornstein–Uhlenbeck process

$$dX = -kXdt + \sqrt{D}dW(t),$$

where  $W(t)$  is a Wiener process.

- (a) Using the solution of the SDE,

$$X(t) = X(0)e^{-kt} + \sqrt{D} \int_0^t e^{-k(t-t')} dW(t'),$$

show that if the initial condition  $X(0)$  is Gaussian distributed with zero mean and variance  $\sigma^2$  then

$$\langle X(t) \rangle = \langle X(0) \rangle e^{-kt}, \quad \text{Var}[X(t)] = \sigma^2 e^{-2kt} + \frac{D}{2k} [1 - e^{-2kt}].$$

- (b) The FP equation for the OU process is

$$\frac{\partial p(x,t)}{\partial t} = \frac{\partial [kxp(x,t)]}{\partial x} + \frac{D}{2} \frac{\partial^2 p(x,t)}{\partial x^2}.$$

Taking the fixed (deterministic) initial condition  $X(0) = x_0$ , the initial condition of the FP equation is

$$p(x,0) = \delta(x - x_0).$$

Introducing the characteristic function (Fourier transform)

$$\Gamma(z, t) = \int_{-\infty}^{\infty} e^{izx} p(x, t) dx,$$

show that

$$\frac{\partial \Gamma}{\partial t} + kz \frac{\partial \Gamma}{\partial z} = -\frac{D}{2} z^2 \Gamma.$$

Use separation of variables to obtain a solution of the form

$$\Gamma(z, t) = \Gamma_0(z e^{-kt}) e^{-Dz^2/4k}$$

with  $\Gamma_0$  determined by the initial condition for  $p$ . Hence, obtain the result

$$\Gamma(z, t) = \exp \left[ -\frac{Dz^2}{4k} (1 - e^{-2kt}) + izx_0 e^{-kt} \right].$$

- (c) The probability density  $p(x, t)$  can be obtained from  $\Gamma(z, t)$  using the inverse Fourier transform

$$p(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-izx} \Gamma(z, t) dz.$$

Substituting for  $\Gamma$  using part (e), show that  $p(x, t)$  is a Gaussian with mean and variance

$$\langle X(t) \rangle = x_0 e^{-kt}, \quad \text{Var}[X(t)] = \frac{D}{2k} [1 - e^{-2kt}].$$

- (d) Show that the solution to the steady-state FP equation is

$$p_s(x) = (2\pi D/k)^{-1/2} e^{-kx^2/2D}$$

and that this is consistent with the time-dependent solution in the limit  $t \rightarrow \infty$ .

**Problem 2.6 (Additive noise).** Solve the SDE

$$dX = -\alpha t^2 X dt + dW(t).$$

- (a) Performing the change of variables  $Y(t) = X(t)e^{\alpha t^3/3}$ , show that

$$dY(t) \equiv Y(X(t+dt), t+dt) - Y(X(t), t) = e^{\alpha t^3/3} dW(t).$$

How was the change of variables chosen?

- (b) Use part (a) to obtain the solution

$$X(t) = X_0 e^{-\alpha t^3/3} + \int_0^t e^{-\alpha[t^3-s^3]/3} dW(s).$$

- (c) Determine the mean  $\langle X(t) \rangle$  and variance  $\text{var}[X(t)]$ . In particular show that the variance is a dimensionless function of  $\alpha t^3$ .

**Problem 2.7 (Multivariate Ornstein–Uhlenbeck process).** Consider the multivariate SDE

$$dX_i = - \sum_{j=1}^N A_{ij} X_j dt + \sum_{j=1}^N B_{ij} dW_j(t),$$

where  $W_j(t)$  form a set of independent Wiener process:

$$\langle dW_i(t) \rangle = 0, \quad \langle dW_i(t) dW_j(t') \rangle = \delta_{i,j} \delta(t - t').$$

Assume a deterministic initial condition  $X_j(0) = \bar{x}_j$ .

(a) Show that the solution in vector form is given by

$$\mathbf{X}(t) = e^{-\mathbf{A}t} \bar{\mathbf{x}} + \int_0^t e^{-\mathbf{A}(t-t')} \mathbf{B} d\mathbf{W}(t').$$

(b) Introduce the correlation function  $\mathbf{C}(t, s) = \langle \mathbf{X}(t), \mathbf{X}^T(s) \rangle$  with components

$$C_{ij}(t, s) = \langle X_i(t), X_j(s) \rangle = \langle [X_i(t) - \langle X_i(t) \rangle] [X_j(s) - \langle X_j(s) \rangle] \rangle.$$

Using part (a), show that

$$\mathbf{C}(t, s) = \int_0^{\min(t, s)} e^{-\mathbf{A}(t-t')} \mathbf{B} \mathbf{B}^T e^{-\mathbf{A}^T(s-t')} dt'.$$

(c) Introduce the covariance matrix  $\Sigma(t) = \mathbf{C}(t, t)$  with components

$$\Sigma_{ij}(t) = \langle [X_i(t) - \langle X_i(t) \rangle] [X_j(t) - \langle X_j(t) \rangle] \rangle.$$

Derive the matrix equation

$$\frac{d\Sigma(t)}{dt} = -\mathbf{A}\Sigma(t) - \Sigma(t)\mathbf{A}^T + \mathbf{B}\mathbf{B}^T.$$

Hence, show that if  $\mathbf{A}$  has distinct eigenvalues with positive real part, then  $\Sigma(t) \rightarrow \Sigma_0$  where  $\Sigma_0$  is the stationary covariance matrix satisfying

$$\mathbf{A}\Sigma_0 + \Sigma_0\mathbf{A}^T = \mathbf{B}\mathbf{B}^T.$$

**Problem 2.8 (1D Fokker–Planck equation with space-dependent variance).**

Consider the 1D FPE with a space-dependent variance due to multiplicative noise:

$$\frac{\partial P}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} [D(x)P],$$

with  $x \in [-1, 1]$  and reflecting boundary conditions.

(a) Determine the steady-state probability density for general  $D(x)$ .

(b) Calculate the steady-state probability density when  $D(x) = k(a + |x|)$  for  $k > 0, a > 1$ . What happens when  $a \rightarrow \infty$ ?



**Problem 2.9 (FPE with multiplicative noise).** Write down the (Ito) FP equation for the SDE

$$dX = adt + bxdW,$$

for positive constants  $a, b$  and  $X \in [0, 1]$  with reflecting boundary conditions.

- Solve the steady-state FP equation up to a normalization factor.
- Calculate the steady-state density for  $y = 1/x$  and determine the normalization factor—use the change of random variables formula from Sect. 1.3.
- Determine  $\langle 1/x \rangle$  as a function of  $a$  and  $b$ .

**Problem 2.10 (Power spectrum).** Consider the Langevin equation for a noise-driven, damped harmonic oscillator:

$$m \frac{d^2 X}{dt^2} + \gamma \frac{dX}{dt} + kX(t) = 2D\xi(t),$$

where  $\xi(t)$  is a Gaussian white noise process with zero mean and covariance

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t').$$

- Plot the spectrum of  $X(t)$  as a function of the angular frequency  $\omega$  for  $\omega_0 \equiv \sqrt{k/m} = 1$ ,  $2D/m = 1$  and various values of  $\beta = \gamma/m$ . What happens in the limit  $\beta \rightarrow 0$ ? What is the significance of  $\omega_0$ ?
- The solution to the Langevin equation can be formally written as

$$X(t) = \int_{-\infty}^{\infty} G(\tau)\xi(t - \tau)d\tau,$$

where  $G(\tau)$  is the causal Green's function. Determine the real and imaginary parts of the Fourier transform  $\tilde{G}(\omega)$  and plot them as a function of  $\omega$  for the same parameters as plot (a).

**Problem 2.11 (FPT for random walks on a lattice).** Consider a random walker on a 1D lattice with sites  $\ell$  and displacement distribution  $p(\ell)$ . The probability  $P_n(\ell)$  that the walker is at site  $\ell$  after  $n$  steps starting at  $\ell_0 = 0$  satisfies the recurrence relation (see Ex. 2.2)

$$P_n(\ell) = \sum_{\ell'} p(\ell - \ell') P_{n-1}(\ell').$$

Let  $F_n(\ell)$  denote the probability of arriving at site  $\ell$  for the first time on the  $n$ th step, given that the walker started at  $\ell_0 = 0$ .

- $P_n(\ell)$  and  $F_n(\ell)$  are related according to the recurrence relation

$$P_n(\ell) = \delta_{\ell,0} \delta_{n,0} + \sum_{m=1}^n F_m(\ell) P_{n-m}(0), \quad n \geq 0.$$

Explain what this relation means physically.

(b) Show that the corresponding generating functions are related according to

$$\Gamma_F(\ell, z) = \frac{\Gamma_P(\ell, z) - \delta_{\ell,0}}{\Gamma_P(0, z)},$$

where

$$\Gamma_P(\ell, z) = \sum_{n \geq 0} z^n P_n(\ell), \quad \Gamma_F(\ell, z) = \sum_{n \geq 0} z^n F_n(\ell).$$

Hence, use Ex. 2.2c to show that for a standard, unbiased RW

$$\Gamma_F(0, z) = 1 - \sqrt{1 - z^2}.$$

(c) Let  $R(\ell)$  denote the probability that site  $\ell$  is ever reached by a walker starting at  $\ell_0 = 0$ :

$$R(\ell) = \sum_{n=1}^{\infty} F_n(\ell) \leq 1.$$

Use part (b) to show that  $R(0) = 1$  for an unbiased RW (recurrent rather than transient) while the MFPT  $\tau(0)$  to return to the origin is infinite, where

$$\tau(\ell) = \sum_{n=1}^{\infty} n F_n(\ell).$$

**Problem 2.12 (FPT for a Brownian particle in a semi-infinite domain).** Consider a Brownian particle restricted to a semi-infinite domain  $x \in [0, \infty)$  with an absorbing boundary condition at  $x = 0$ . The FP equation is given by

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2}, \quad 0 < x < \infty,$$

with  $p(0, t) = 0$ .

(a) Check that the solution of the FP equation for the initial condition  $x(0) = x_0$  is

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-(x-x_0)^2/4Dt} - \frac{1}{\sqrt{4\pi Dt}} e^{-(x+x_0)^2/4Dt}.$$

(Such a solution can be derived using the method of images, in which one imagines initially placing a fictitious Brownian particle at the image point  $x = -x_0$ .)

(b) Show that for large times where  $\sqrt{Dt} \gg x_0$ , the probability density can be approximated by

$$p(x, t) \approx \frac{1}{\sqrt{4\pi Dt}} \frac{xx_0}{Dt} e^{-(x^2+x_0^2)/4Dt}.$$

(c) Calculate the FPT density  $f(x_0, t)$  to reach the origin starting from  $x_0$  by calculating the flux through the origin using part (a):

$$f(x_0, t) = D \left. \frac{\partial p(x, t | x_0, 0)}{\partial x} \right|_{x=0}.$$

Hence show that when  $\sqrt{Dt} \gg x_0$ , we have the asymptotic behavior

$$f(x_0, t) \sim \frac{x_0}{t^{3/2}}.$$

Deduce that the MFPT to reach the origin is infinite.

**Problem 2.13 (Rotational diffusion).** Consider a Brownian particle undergoing diffusion on the circle  $\theta \in [-\pi, \pi]$ . This could represent the orientation of a bacterium during a single run (see Sect. 2.4). The corresponding FP equation for  $p(\theta, t)$

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial \theta^2}, \quad -\pi < \theta < \pi, \quad p(-\pi, t) = p(\pi, t), \quad p'(-\pi, t) = p'(\pi, t),$$

where  $D$  is the rotational diffusion coefficient.

- (a) Using separation of variables the initial condition  $p(\theta, 0) = \delta(0)$ , show that the solution of the FP equation is

$$p(\theta, t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in\theta} e^{-Dn^2 t}.$$

- (b) If  $t$  is sufficiently small then  $p(\theta, t)$  is strongly localized around the origin  $\theta = 0$ . This means that the periodic boundary conditions can be ignored and we can effectively take the range of  $\theta$  to be  $-\infty < \theta < \infty$ . That is, performing the rescalings  $x = \theta/\varepsilon$  and  $\tau = \varepsilon^2 t$ , show that  $p(\theta, t)$  can be approximated by a Gaussian  $p(x, t)$  and deduce the small-time approximation

$$\langle \theta^2 \rangle = 2Dt, \quad t \ll \pi^2/D.$$

- (c) What happens in the limit  $t \rightarrow \infty$ ?

**Problem 2.14 (Diffusion in a sphere).** Consider the diffusion equation in a spherical cell of radius  $R$ :

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} = D \nabla^2 u(\mathbf{x}, t), \quad 0 < |\mathbf{x}| < R,$$

with boundary condition  $u(|\mathbf{x}| = R, t) = u_1$  and initial condition  $u(\mathbf{x}, 0) = u_0$  with  $u_0, u_1$  constants.

- (a) Assume a radially symmetric solution  $v(r, t) = u(r, t) - u_1$  so that

$$\frac{\partial v(r, t)}{\partial t} = D \frac{\partial^2 v}{\partial r^2} + \frac{2}{r} D \frac{\partial v}{\partial r}, \quad 0 < r < R,$$

with  $v(R, t) = 0$  and  $v(r, 0) = u_0 - u_1$ . Use separation of variables  $v(r, t) = V(r)T(t)$  to derive the general solution

$$v(r, t) = \sum_{n=1}^{\infty} c_n e^{-t D n^2 \pi^2 / R^2} \frac{1}{r} \sin(n \pi r / R).$$

Hint: in order to solve the boundary value problem for  $V(r)$ , perform the change of variables  $\hat{V}(r) = rV(r)$ .

- (b) Setting  $t = 0$  in the general solution and using  $v(r, 0) = u_0 - u_1$ , determine the coefficients  $c_n$ . Hint: you will need to use the identity

$$\int_0^R \sin(n \pi r / R) \sin(m \pi r / R) dr = \frac{R}{2} \delta_{n,m}.$$

- (c) Determine an approximation for the concentration  $u(0, t)$  at the center of the sphere by taking the limit  $r \rightarrow 0$ , with  $r^{-1} \sin(\theta r) \rightarrow \theta$ . Keeping only the leading order exponential term ( $n = 1$ ), show that the time  $\tau$  for the center to reach a concentration  $u^*$ ,  $u_1 < u^* < u_0$ , is approximately

$$\tau = \frac{R^2}{D \pi^2} \ln \frac{2(u_0 - u_1)}{u^* - u_1}.$$

**Problem 2.15 (Computer simulations: Langevin equation).** Use the algorithms of Sect. 2.6.6 to solve the following problems in MatLab.

- (a) Consider the Ornstein–Uhlenbeck process

$$dX(t) = -\lambda X(t)dt + dW(t), \quad X(0) = x_0,$$

where  $W(t)$  is a Wiener process. Use direct Euler to simulate 1,000 trajectories on the time interval  $[0, 1]$  for  $\lambda = 1/2$ ,  $\Delta t = 0.01$  and  $x_0 = 1$ . Compare the mean and covariance of the trajectories with the theoretical values of Ex. 2.6]

- (b) Use Milstein's method to simulate the following SDE on the time interval  $[0, 1]$

$$dX(t) = -\lambda X(t)dt + \mu X(t)dW(t), \quad X(0) = x_0$$

for  $\lambda = 0.1$ ,  $\mu = 0.1$ , and  $x_0 = 1$ . Compare the cases  $\Delta t = 0.1$ ,  $\Delta t = 0.001$ , and  $\Delta t = 10^{-5}$ . Check that the histogram of values at  $t = 1$  is similar to the histogram obtained by simulating the exact solution

$$X(t) = x_0 \exp \left[ (-\lambda - \mu^2/2)t + \mu W(t) \right].$$



<http://www.springer.com/978-3-319-08487-9>

Stochastic Processes in Cell Biology

Bressloff, P.C.

2014, XVII, 679 p. 206 illus., 90 illus. in color.,

Hardcover

ISBN: 978-3-319-08487-9