Chapter 1

Langevin and Fokker-Planck equations and their applications

"Probability theory is nothing but common sense reduced to calculation." (Pierre Simon Laplace)

In our previous discussion of random walks (say in 1d), we worked with the discrete equation:

$$x_{i+1} = x_i + \xi. \tag{1.1}$$

Let's now consider the continuous version, where:

$$\frac{dx}{dt} = \xi(t), \tag{1.2}$$

with $\xi(t)$ is a noise with vanishing mean and correlation function $\langle \xi(t')\xi(t) \rangle = f(t-t')$ (typically, $\langle \, ; \, \rangle$ brackets will denote averaging over all noise realizations – i.e., ensemble-averaging). This is a simple example of a *Langevin equation*.

In many cases we want to describe the system on timescales much larger than the correlation time of the noise – which means we can replace $f(t)$ by a δ-function, $f(t) = C\delta(t)$. Keep in mind that for certain applications, the noise correlation time could be relevant (see Problem 3.1 for an example).

Consider now the time derivative of $\langle x^2 \rangle$. For a particular realization we may express:

$$x(t) = x_0 + \int_0^t \xi(t')dt'. \tag{1.3}$$

Therefore we find:

$$\frac{d\langle x^2 \rangle}{dt} = 2\langle x(t)\xi(t) \rangle = \langle 2 \left[x_0 + \int_0^t \xi(t')dt' \right] \xi(t) \rangle. \tag{1.4}$$

Since at every t the mean of ξ vanishes, the first term vanishes and we are left with:

$$\frac{d\langle x^2 \rangle}{dt} = 2 \int_0^t \langle \xi(t')\xi(t) \rangle dt' = 2 \int_0^t f(t - t')dt'. \tag{1.5}$$

Since we are integrating over *half* of the δ function, we find that:

$$\frac{d\langle x^2 \rangle}{dt} = C, \tag{1.6}$$

hence $Ct = 2Dt$, and we obtain diffusion with a constant $D = C/2$.

How can we describe the random process in the δ function limit? It turns out that this can be done in several ways, which may lead to different results occasionally – two common methods used are the Itô and Stratonovich versions of stochastic calculus. The issue is that in these descriptions there are two timescales which vanish – the discretization time and the correlation time – and the order in which we send them to zero matters. We will not go deeply into this subtle issue, but for our purposes we will work mostly within the Itô formalism.

> **Itô vs. Stratonovich formalisms for stochastic calculus** The issue of interpretation of the stochastic equations arises when dealing with the Langevin equation. Shortly we will derive the so-called Fokker-Planck equation, which describes the *deterministic* evolution of the probability distribution in time – since it is "just" a PDE, there is no ambiguity of interpretation once the Fokker-Planck equation is written. For most applications we will see, both Itô and Stratonovich lead to the *same* Fokker-Planck equation. In particular, this is true when the noise is additive. However, for cases involving *multiplicative* noise (i.e., where the noise term multiplies a term associated with the variable x rather than being added to it) the resulting PDE takes a different form. In such cases, the correct stochastic interpretation depends on the particular application at hand and the relevant order of limits.

Within the Itô formalism, the correlation time is shorter than the discretization time, so we can think of our continuous random walker as driven by a set of independent, random kicks. It is often useful to think how one would *simulate* such a process. To do this, we can make time discrete, with small timesteps Δt (defining our time resolution), and for each timestep we would like to choose the noise as a Gaussian, independent random variable (we will see shortly why it should be Gaussian), with magnitude W. This recovers Eq. (1.1). After time t, we would have taken $\frac{t}{\Delta t}$ steps, hence:

$$\langle x^2 \rangle = \frac{t}{\Delta t}W^2. \tag{1.7}$$

To get the correct scaling, we must therefore choose: $W = C\sqrt{\Delta t}$. In fact, this simple scheme to simulate stochastic processes is known as the *Euler-Maruyama* method. The discrete Langevin equation corresponding to this scheme in this situation of a simple random walk is:

$$x_{t+\Delta t} = x_t + C\sqrt{\Delta t}. \tag{1.8}$$

(for an interesting pedagogical paper elaborating on the utility of this discrete approach, see Ref. [1]). What would have happened if we would have chosen a non-Gaussian distribution? In that case, if we decrease our time resolution, i.e., make the step size larger, the effective noise in each step would be the sum of a large number of variables, and would therefore converge to a Gaussian (assuming that the distribution has finite variance). On the other hand, Gaussian distributions are infinitely divisible – we can always write $\xi = \xi_1 + \xi_2$, where ξ_1 and ξ_2 are *also* Gaussian variables, with half the variance. So we can think of our Gaussian variable as resulting from a large sum of independent Gaussian variables.

> **Divisibility: can a random variable be expressed as a sum of two independent random variables?** Problem 3.9 deals with the *lack* of divisibility of the uniform distribution – a uniformly chosen random variable cannot be represented as the sum of two other (independent) random variables! In Chapter 6 we will find the most general family of distributions which have the much stronger property which the Gaussian distribution exhibits – whereby the sum of two (or more) variables drawn from the distribution results in a variable corresponding to a probability distribution identical in shape to the original (i.e., the two are equal upon appropriate shifting and scaling).

Eq. (1.2) is the simplest instance of a Langevin equation, first introduced by Langevin to model diffusion of particles in a liquid or gas. We have seen before that the probability distribution is described by the diffusion equation, which is the simplest instance of a Fokker-Planck equation. Next, we will study more general Langevin equations, of the form:

$$\frac{dx}{dt} = -\mu V'(x) + \xi. \tag{1.9}$$

The function $f(x) = -V'(x)$ can be associated with a physical force being applied to the diffusing particles, but in other scenarios it will be associated with a feedback in the system (we will study a biological example of this sort). Moreover, many times the Langevin equation will be defined in a higher dimension, as we shall study when we deal with certain "escape-over-a-barrier" problems.

Example I: particle in a potential driven by random noise

Consider a particle in a potential $V(x)$ driven by random kicks due to its interaction with a thermal bath. For simplicity, let's consider the 1d case. The Langevin equation reads:

$$m\frac{dv}{dt} = -\gamma v + \zeta(t) - V'(x), \tag{1.10}$$

where ζ is the noise and γ the dissipation. In fact, the two originate from the same physical mechanism (interactions with the environment) and so are physically related – as we shall shortly see.

> **The Ornstein-Uhlenbeck process** In the case $V = 0$, Eq. 1.10 is known as the Ornstein-Uhlenbeck process. It's properties are explored within problem 3.3.

One commonly relevant limit is the *overdamped* case, in which the particle's inertia is irrelevant. In this case we can ignore the LHS, and write:

$$\frac{dx}{dt} = \mu f(x) + \xi, \tag{1.11}$$

where $f(x) = -V'(x)$, $\xi = \zeta/\gamma$ and $\mu = 1/\gamma$. This is precisely Eq. (1.9).

In the case where inertia is relevant, it is convenient sometimes to think of the dynamics in a two-dimensional space, by defining a vector $\vec{x}$ with $x_1 \equiv x$, $x_2 \equiv mv$. Therefore:

$$\frac{d\vec{x}}{dt} = -\mu \nabla \tilde{V}(\vec{x}) + \vec{\psi}, \tag{1.12}$$

where in this case μ is a *matrix*:

$$\mu = \begin{pmatrix} 0 & -1 \\ 1 & \gamma \end{pmatrix}, \tag{1.13}$$

$\tilde{V} = \frac{1}{2}mv^2 + V(x)$, and the noise term is given by $\vec{\psi} = (0, \gamma\xi)$. This is a generalization of Eq. (1.9) which we shall use later in the course.

We shall now understand how to go from this Langevin equation to the so-called Fokker-Planck equation, which generalizes our derivation of the diffusion equation. We shall then use the Langevin equation to study a contemporary problem in biology, and will see how analogous equations to the Langevin and Fokker-Planck equations arise in the context of economics. In Chapter 4 we will study a more involved physical problem using Fokker-Planck equations known as "escape-over-a-barrier".

Example II: the isothermal atmosphere, thermal equilibrium arising from superimposing diffusion and drift

This may at first seem off-topic, but will end up helping us develop intuition for deriving the Fokker-Planck equation. Consider a column of gas molecules with molecular mass m, assumed to be at thermal equilibrium at temperature T. According to the Boltzmann distribution, the density of air will depend on the altitude as:

$$n(h) \propto e^{-\frac{mgh}{kT}}, \tag{1.14}$$

with k the Boltzmann constant. Note: this is *not* a good approximation for the Earth's atmosphere!

Note that the above equation is for the spatial dependence of *particle density*. We could have written the same equation for the *probability distribution* of particle position – the two quantities only differ by a normalization constant.

We will now understand how this comes about from the "molecular" picture of Eq. (1.9). The gas molecules are subject to the force of gravity, which, if the density were to be constant, would lead to a constant current of particles downward – in the overdamped limit the particles would move at constant velocity $v = \mu F = \mu m g$. However, this initial drift would lead to a higher concentration of molecules at lower altitude. The gradient of density implies that the *random walk* of the molecules would lead to an effective current *upward*. This current is proportional to the gradient of density (and opposite in sign), a result known as Fick's law. To illustrate this, let us temporarily assume that no external forces are acting on the molecules (i.e., we switched gravity "off" for now). Consider the relation:

$$\vec{j} = -D\nabla n, \tag{1.15}$$

From the continuity equation we have:

$$\nabla \cdot \vec{j} + \frac{\partial n}{\partial t} = 0, \tag{1.16}$$

where $\nabla\cdot$ is the divergence of the current $\vec{j}$ (which is – in the general case – a vector field). Hence:

$$\frac{\partial n}{\partial t} = D\nabla^2 n, \tag{1.17}$$

and we see that the diffusion equation we previously derived indeed arises from Fick's law.

Going back to the "isothermal atmosphere", at steady-state all time derivatives vanish, hence the divergence of the current must vanish. In this effectively one-dimensional problem this implies the current must be constant. Since there can be no flux at the surface, this implies the current must vanish everywhere. In other words, the current due to diffusion (upward) would cancel that of gravity:

$$-Dn' = mgn\mu. \tag{1.18}$$

The solution is indeed that of Eq. (1.14), provided that: $D = kT\mu$. This relation is known as the Einstein relation , and is extremely useful. It implies that both diffusion and μ (which is the reciprocal of the drag coefficient γ) are related – they both arise from the same physical mechanism, collisions with other molecules. This is an example of a fluctuation–dissipation theorem: it relates the diffusion constant (associated with the fluctuations) to the mobility (associated with the dissipation due to the viscous drag).

Going from a Langevin to a Fokker-Planck equation Consider a particle performing a 1d random walk in the overdamped limit – as described by the Langevin equation of Eq. (1.9) (with μ potentially depending on position). The continuity equation (1.16) is generally true, and the current will have, as in the previous

example, two contributions: one from the force ($\vec{j_f}$), and the second from the diffusion term ($\vec{j_d}$). The former can be readily understood in the overdamped limit in the absence of stochasticity: the force would lead the particle to a constant velocity $\vec{v} = \mu \vec{f}$, hence the current will be:

$$\vec{j_f} = -\mu p \nabla V(\vec{r}). \tag{1.19}$$

In the absence of a force, we saw that the diffusion equation arises (following Einstein's derivation). Furthermore, the example in the previous section showed that it may also be interpreted as the continuity equation combined with Fick's law, therefore the diffusive current is:

$$\vec{j_d} = -D\nabla p. \tag{1.20}$$

Plugging in Eqs. (1.19) and (1.20) into the continuity equation $\frac{\partial p}{\partial t} + \nabla \cdot \vec{j} = 0$, we obtain the Fokker-Planck equation:

$$\frac{\partial p}{\partial t} = \nabla \cdot [\mu \nabla V(\vec{r})p] + \nabla[D\nabla p], \tag{1.21}$$

where in general D and μ can be space-dependent, and if we are dealing with a statistical mechanics problem they are related by the Einstein relation. Note that as in Eq. (1.16), $\nabla\cdot$ denotes the divergence of the vector field. This is the general form of the Fokker-Planck equation – and is basically derived by considering the "probability currents", just as we have done when deriving the diffusion equation (the derivation can be done more systematically, within the Itô or Stratonovich stochastic calculus formalism – in this case both lead to the same equation).

Revisiting Einstein's derivation in the presence of a potential

Previously, we derived the diffusion equation following Einstein's insights, while assuming that the particle is diffusing in a uniform potential. We shall now repeat the derivation for a particle in a one-dimensional potential $V(x)$, assuming the overdamped limit (where velocity is always proportional to force). Our starting point will be the discrete Langevin equation (see Eq. (1.8)):

$$x_{t+\Delta t} = \underbrace{x_t + \mu f(x_t)\Delta t}_{\text{Deterministic}} + \underbrace{\xi\sqrt{\Delta t}}_{\text{Stochastic}}, \tag{1.22}$$

with timestep Δt and ξ a Gaussian with vanishing mean and standard deviation σ.

Let us assume that the probability distribution at time t is $P(x,t)$. If the particle was at position $\tilde{x}$ at time t, clearly the probability that it will be at position x at time $t + \Delta t$ is given by the probability that the noise term equals $x - \tilde{x} - \mu f(\tilde{x})\Delta t$; hence it equals:

$$G(x|\tilde{x}) = \frac{1}{\sqrt{2\pi\sigma^2\Delta t}}e^{-\frac{[x-\tilde{x}-\mu f(\tilde{x})\Delta t]^2}{2\sigma^2\Delta t}}. \tag{1.23}$$

We can also approximate $f(\tilde{x}) \approx f(x) + f'(x)(\tilde{x} - x)$ (for small Δt, this first-order Taylor expansion will be a good approximation). This allows us to write an equation for the time evolution of $p(x, t)$, by considering all the possibilities of $\tilde{x}$:

$$p(x, t + \Delta t) = \int G(x|\tilde{x})p(\tilde{x}, t)d\tilde{x}. \tag{1.24}$$

Following a similar logic to our previous derivation of the diffusion equation, we may now expand $p(\tilde{x}, t)$ to second order (around the point x):

$$p(\tilde{x}, t) \approx p(x, t) + p'(x, t)(\tilde{x} - x) + \frac{p''(x, t)}{2}(\tilde{x} - x)^2. \tag{1.25}$$

Plugging this form into Eq. (1.24), we may readily perform the Gaussian integrals and find (to order $O(\Delta t)$):

$$p(x, t + \Delta t) = \frac{p(x, t)}{1 + \mu\Delta t f'} - \mu\Delta t f p'(x, t) + \frac{\sigma^2}{2}p''(x, t). \tag{1.26}$$

Expanding to first order in Δt we have $\frac{1}{1+\mu\Delta t f'} \approx 1 - \mu\Delta t f'$, and taking the continuous limit leads to the one-dimensional version of Eq. (1.21).

Note that in Eq. (1.22) the force is evaluated at x_t. Identical results are obtained when the force is evaluated at $x_{t+\Delta t}$.

Revisiting Einstein's derivation in higher dimensions Try to repeat the derivation for a particle in a potential $V(\vec{r})$, assuming the overdamped limit. You should recover Eq. (1.21)   .

1.1   Application of a discrete Langevin equation to a biological problem

This section deals with the application of a discrete version of a Langevin equation to a problem in contemporary cell biology. It is a problem where the biological data could only be interpreted within the framework of a mathematical model, and in which the quantitative details actually matter and lead to conceptual advances.

A little biological background: bacteria, such as the *E. coli* strains that we all have in our guts, grow such that they double their volume in several tens of minutes (in good nutrient conditions the doubling time can be as short as 20 minutes), after which they divide symmetrically. It turns out that during this period the cell volume grows *exponentially*. This brings about an interesting problem: if we naively assume that the

cell cycle relies on a "timer" (i.e., the duration between two division events is set by a clock), then denoting the cell size at birth by v_b, its size at division will be given by:

$$v_d = v_b 2^{t/t_d},\tag{1.27}$$

hence the newborn size of the daughter cell is :

$$\log_2 v_b^{n+1} = \log_2 v_b^n + t/t_d - 1.\tag{1.28}$$

If $t = t_d$ precisely, we have perfect doubling and perfect division every generation, and all would be fine. However, nothing is perfect, and there must be fluctuations in t around t_d. Defining $\xi = t/t_d - 1$, Eq. (1.28) shows that the *logarithm* of size would perform a random walk. Hence size will be *unbounded*, with some cells becoming very small while others very large – which is not what happens in reality.

This goes to show that there must be a feedback mechanism, that (potentially among other functions) controls size. The goal of this section would be to derive a phenomenological description of this feedback, in the spirit of the Langevin equation. Such a model would go beyond the "timer" described above, converging to a stable size distribution in the same way that the restoring force in the Ornstein-Uhlenbeck process of Eq. 1.10 leads to a stationary velocity distribution (and without it, as we have seen, the velocity will perform a random walk – analogous to the random walks of the logarithm of cell size discussed above). Such a model can make predictions regarding the statistics of the model variables (namely, cell size and time), including correlations between cells in the same lineage and distributions of these variables. These relations can then be empirically tested, to validate or falsify the model.

Formally, we can describe this using a discrete version of the Langevin equation. The notation for the discussion that follows is summarized in the following table.

Table 1: Notations for cell size regulation problem	
Variable	**Definition**
v_b	Size at birth (stochastic variable)
v_d	Size at division (stochastic variable)
v_0	Average size at birth
$f(x)$	Size regulation (deterministic) strategy
t_d	Volume doubling time (assumed constant)
ξ	Time-additive noise
t_a	Deterministic generation time given a size at birth
t_n	Noise added to deterministic generation time
σ_n	Standard deviation of t_n
x	$\log_2[v_b/v_0]$, see Eq. (1.34)
σ_v	Standard deviation of x for the stationary solution
σ_t	Standard deviation of generation time t for the stationary solution
C_{xy}	Pearson correlation coefficient between variables x and y

We will assume that the cell controls size by attempting to divide at a size:

$$v_d = f(v_b). \tag{1.29}$$

The choice of function f determines the regulation strategy used by the cell. For example, the (failed) timer strategy we described above corresponds to $f(v_b) = 2v_b$. An *a-priori* reasonable strategy is $f(v_b) = const$, in which the cell attempts to divide when reaching a critical size (this thresholding strategy is known as a "sizer"). Clearly this will have strong size control (i.e., size distributions will be as narrow as possible), but we shall later claim it is also not consistent with experimental data on *E. coli*. In order to implement a strategy described by a function f, the cell must attempt to grow for a time:

$$t_a = t_d \log_2(f(v_b)/v_b). \tag{1.30}$$

To this "deterministic" component we will add a stochastic term t_n such that the actual generation time t will be given by:

$$t = t_a + t_n. \tag{1.31}$$

This is precisely the philosophy which Langevin used in writing his equation, where the force due to collisions with the molecules was decoupled into a deterministic component (leading to the viscous drag) and a noise term. We will assume the noise t_n to be Gaussian with vanishing mean and standard deviation σ_n (though

this assumption can be relaxed without affecting the main conclusions). This concludes the definition of the model (for now). To make further progress, we will assume that the cell size distribution will ultimately be narrow (an assumption which can be verified experimentally, and also self-consistently within our model). Hence we don't actually care about the behavior of $f(x)$ for any x, but really only for a narrow range around the typical newborn cell size v_0. This implies that we can Taylor expand f around v_0 to find that:

$$f(v_b) \approx 2v_0 + f'|_{v_0}(v_b - v_0), \tag{1.32}$$

where we utilized the fact that for the typical newborn size v_0 we should have $f(v_0) = 2v_0$, such that cells double in volume.

It will be useful to denote $f'|_{v_0} = 2(1 - \alpha)$. In this notation the "timer" strategy corresponds to a slope of 2 and hence $\alpha = 0$, while the thresholding "sizer" strategy corresponds to vanishing slope and hence $\alpha = 1$. For any f, we can find the relevant value of α by evaluating its derivative – for example, if $f(v_b) = v_b + v_0$, then $\alpha = \frac{1}{2}$.

All models with the same value of α will follow approximately the same behavior, since their Taylor expansion agrees to first order. To solve the model, it will be useful to choose a particular function $f(v_b) = 2v_b^{1-\alpha}v_0^\alpha$. You can check that its derivative at v_0 gives $2(1-\alpha)$, hence the α appearing in this definition of f is indeed consistent with our previous definition of α. This function of f was chosen such that the "attempted" generation time will take a particularly simple expression:

$$t_a = t_d - t_d\alpha \log_2[v_b/v_0] \tag{1.33}$$

Defining $x \equiv \log_2[v_b/v_0]$, our stochastic equation becomes:

$$x_{n+1} = x_n(1 - \alpha) + \xi, \tag{1.34}$$

with ξ a Gaussian noise with standard deviation σ_n/t_d. This is a discrete version of the Langevin equation we had before. If $\alpha = 0$, we are back to the random walk scenario. $\alpha > 0$ corresponds to a feedback, which *might* be able to correct the random drift due to the noise term ξ.

It is very easy to simulate this equation – which Problem 3.5 deals with. But it is also easy to guess a solution – since ξ is a Gaussian variable with vanishing mean, if x_n is Gaussian with vanishing mean x_{n+1} will also be such, so we would have a stationary solution if the variances will also be consistent. Denoting the variance of x_n by σ_v^2, we obtain the self-consistent equation:

$$\sigma_v^2 = (1 - \alpha)^2\sigma_v^2 + (\sigma_n/t_d)^2. \tag{1.35}$$

Note that we relied on the fact that the noise term is uncorrelated with the size at the beginning of that generation (x_n), but it *will* be correlated with the size at the end of that generation (x_{n+1}), and the two variables x_n, x_{n+1} may also be correlated.

Solving Eq. 1.35, we find that:

$$\sigma_v^2 = (\sigma_n/t_d)^2/[\alpha(2-\alpha)]. \tag{1.36}$$

Hence we see that as $\alpha \to 0$, the size distribution becomes infinitely broad, as we expect from our consideration of the case $\alpha = 0$. Less obvious *a-priori* is that as $\alpha \to 2$ the same happens. Therefore the regime of stable size distributions is $0 < \alpha < 2$, and it is easy to see that the distribution is narrowest for $\alpha = 1$, which is also intuitive, since that corresponds to the size-thresholding case.

Furthermore, since the logarithm of size was found to be a Gaussian variable, the size distribution will be *log-normal*:

$$p(v_b) \propto e^{-\log_2(v_b/v_0)^2/\sigma_v^2}/v_b. \tag{1.37}$$

Similarly, we can find the time distribution. Since the time is given by:

$$t = t_a + t_d\xi, \tag{1.38}$$

and t_a was found to be Gaussian, the time distribution will also be Gaussian, with a variance of:

$$\sigma_t^2 = \alpha^2 t_d^2 \sigma_v^2 + \sigma_n^2 = \sigma_n^2(1 + \alpha^2/[\alpha(2-\alpha)]) = 2\sigma_n^2/(2-\alpha). \tag{1.39}$$

As we might expect, the time distribution is *narrowest* for $\alpha = 0$ (the "timer" is good at controlling time!), and monotonically increases with α. It is important to note that one stochastic noise term, σ_n, determines the width of both *size and time distributions*. It is useful to quantify the relative width of distributions by their *coefficient of variation* (CV), which is the ratio of their standard deviation to their mean. Using Eqs. (1.37) and (1.39) it is easy to see that the CV of the size distribution is $\log(2)/\sqrt{2\alpha}$ times that of the time distribution.

Armed with this machinery, we can now quantify correlations between various variables. A useful entity is the Pearson correlation coefficient, which is defined as:

$$C_{xy} \equiv \frac{\mathbf{E}[(x-\bar{x})(y-\bar{y})]}{\sigma_x\sigma_y}, \tag{1.40}$$

where $\bar{x}$, $\bar{y}$ are the averages of the variables, and σ_x, σ_y their standard deviations.

Clearly, the Pearson correlation coefficient of a variable with itself is 1, and with its negative is -1. Additionally, the correlation coefficient vanishes for two independent variables. Therefore, its value between -1

and 1 is a measure of how correlated or anti-correlated two variables are. Furthermore, it is clear that this definition is insensitive to adding a constant to the variables, or scaling by a positive factor.

Consider now the size of a cell at birth, and its size at division, $C_{v_b v_d}$. According to our assumption of perfectly symmetric division, the correlation coefficient between these two variables would be the same as that between size at birth and the size of the daughter cell at birth. Plugging into the definition we have:

$$C_{v_b v_d} \equiv \frac{\mathbf{E}[(v_b^n - v_0)(v_b^{n+1} - v_0)]}{\sigma^2}, \tag{1.41}$$

with σ^2 the variance of the birth size. To simplify things, it is useful to note that by Taylor expanding $x = \log_2(v_b/v_0)$ around v_0, the correlation coefficient would be nearly the same for the variables x_n and x_{n+1}, hence:

$$C_{v_b v_d} \approx \frac{\mathbf{E}[x_n x_{n+1}]}{\sigma_v^2}. \tag{1.42}$$

Using Eq. (1.34), and the fact that the noise is uncorrelated with x_n, we find that:

$$C_{v_b v_d} \approx \frac{\mathbf{E}[x_n^2(1 - \alpha)]}{\sigma_v^2} = 1 - \alpha. \tag{1.43}$$

For $\alpha = 1$, the vanishing of the correlation coefficient is very plausible: since in this case thresholding washes away the memory of the initial size. In the case $\alpha = 0$, which has no size control, the cell attempts to double its size, which is why we get a perfect correlation coefficient.

In 1980, this correlation coefficient was experimentally measured [3], and found to be approximately $1/2$, suggesting that $\alpha = 1/2$. As mentioned earlier, this is what we would get if $f(v_b) = v_b + v_0$, i.e., the cell attempts to add a relative volume to its volume at birth. But note that this can only be interpreted within a quantitative model such as the one we described above, which is the reason for which this "interpretation" of the measurement was not provided until 2014, several decades after the experimental work [4]. We can now calculate additional correlation coefficients to further test the theory, for example, that between the size at birth and the time to division. As before, it is useful to replace v_b by x, which hardly affects the coefficient. Thus:

$$C_{v_b t} \approx \frac{\mathbf{E}[x(t - t_d)]}{\sigma_v \sigma_t} = \frac{\mathbf{E}[x(t_a - t_d)]}{\sigma_v \sigma_t}. \tag{1.44}$$

Using $t_a - t_d = -t_d \alpha x$, we find that:

$$C_{v_b t} \approx -\alpha \frac{t_d \mathbf{E}[x^2]}{\sigma_v \sigma_t} = -\alpha \frac{\sigma_v t_d}{\sigma_t}. \tag{1.45}$$
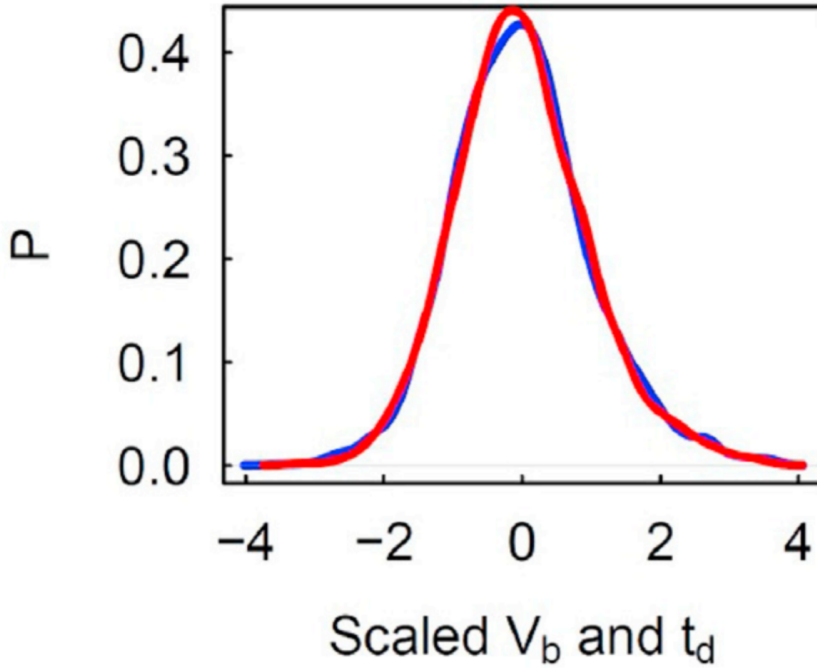
FIGURE 1.1: Data collapse of the scaled size distribution (blue) and generation time distribution (red) for the bacterium *E. coli*, without any fitting parameters. From Ref. [6].

Using Eqs. (1.36) and (1.39) we find that:

$$C_{v_b t} \approx -\alpha t_d \sqrt{\frac{(\sigma_n/t_d)^2/[\alpha(2-\alpha)]}{2\sigma_n^2/(2-\alpha)}}. \tag{1.46}$$

Hence:

$$C_{v_b t} \approx -\sqrt{\alpha/2}, \tag{1.47}$$

and for $\alpha = 1/2$ we expect it to be about $-1/2$. Precisely this value was reported in 2014 [5], corroborating this model. Finally, the fact that one source of stochasticity determines both the size and time distributions allows us to scale both distributions without using any fitting parameters: from Eqs. (1.36) and (1.39), we find that for $\alpha = 1/2$, the distribution of the scaled size variable $\log_2(v_b/v_0)$ and the scaled time variable $(t-\tau)/\tau$ both have vanishing mean and identical variance. This is shown in Fig. 1.1, from Ref. [6]. As a corollary of this, CV of the size distribution is $\log(2)$ smaller than that of the time distribution, which was also corroborated experimentally in Ref. [6].

In fact, the story is more involved, and there are other models which are not "birth-centric" which would give the same correlations, but that is outside the scope of this course. Interested readers may refer to [7] for a recent discussion and more biological insights into this problem, and [8] for further details of mathematical models of cell size control in microbes.

1.2 The Black-Scholes equation: pricing options

Next, we introduce some basic ideas of option pricing in finance, and present a derivation of the Black-Scholes equation as the continuous limit of a discrete-time model. This chapter is inspired by the notes of Terence Tao.[1]

We focus on two primary options, the **call** and **put options**, which are in a sense inverse to each other. We also only study **European options**, which may only be exercised at the expiration time, denoted t_1 (an American option, on the other hand, allows the owner to exercise at any time prior to the expiration time). Throughout, we consider some underlying asset S, with price S_t at time t.

1. **Call option**: Right to purchase S at price P (strike price) from the seller of the option at time t_1.

2. **Put option**: Right to sell S at price P to the seller of the option at time t_1.

The central option pricing problem is as follows: at time $t_0 < t_1$, what is the appropriate price to assign to an option with expiration date t_1 with strike price P?

To solve this problem, we make a number of simplifying assumptions:

1. **No arbitrage**: This is the central assumption we will rely on the computing prices; essentially, it means that we can't make "easy money" – say by buying in euros and selling in dollars etc.

2. **Infinite liquidity**: Cash is not a constraint; participants can buy and sell options at any time and price.

3. **No market power**: The buying/selling of an asset has no effect on its price.

4. **No transaction costs**.

5. **Infinite credit**: Participants can borrow or lend as much as desired at a given interest rate r.

6. **Infinite divisibility**: The asset and option can be sold in as small an increment as desired.

7. **Short-selling**. We can sell stocks that we don't have, and later on purchase them to cover this. In a sense, it will allow us to work with *negative* stock values.

The price of an option within this model will depend on our assumptions regarding the (stochastic) time dynamics of the stocks – as will be discussed below. The following table summarizes the notation we will be using throughout the chapter.

---

[1]T. Tao, "The Black-Scholes Equation." https://terrytao.wordpress.com/2008/07/01/the-black-scholes-equation.

Table 2: Notations for Black-Scholes problem	
Variable	**Definition**
r	Interest rate
S_t	Stock price at time t
μ	Drift of stocks, see Eq. (1.50)
σ	Volatility of stock dynamics, see Eq. (1.50)
$V(t, S)$	Option price at a given time and given the current stock price (for Call or Put option)
t_1	Expiration time, at which Call/Put option are exercised
P	Strike price (for Call or Put option)

1.2.1   Time is money: the time value of money

Our assumption of an interest rate r, together with the assumption of no arbitrage, allows us to find (for example) how much a bond that would pay us 1 dollar at a future time t_1 would be worth today, at time $t_0 = t_1 - t$: Let's call this value X. We could always borrow an amount of money X *today*, and buy such a bond, getting back 1 dollar in the future and pay off the debt then (an amount Xe^{rt}, due to interest) – since we can't make money with this strategy (no arbitrage), it means that $1 - Xe^{rt} \leq 0$, i.e., $X \geq e^{-rt}$. On the other hand, If $X > e^{-rt}$, then I can create such a bond today and sell it, getting X units of cash, and pay a dollar to the buyers at time t_1. This way I made $Xe^{rt} - 1$ amount of cash, which is positive by assumption – hence $X = e^{-rt}$.

1.2.2   Bounding the price of a put option

Let $V_t(S_t)$ denote the price of an option of asset S at time t. For example, consider a put option on S at strike price P with expiration time t_1. To avoid an arbitrage opportunity, the price of the option must reflect both the price of the asset and the time-discounted strike price. If the strike price is high and the option is cheap, we could make money from buying the option and selling the stock – in contrast to the no arbitrage assumption. Let us quantify this:

We would be buying a unit of stock and the option, at the present. The amount of money we would be making is P, but this is at a *future* time t_1, hence when comparing it to the money we pay now we have to multiply by $e^{-r(t_1-t_0)}$ (taking into account the time value of money).

Hence our *gain* is:

$$X = e^{-r(t_1-t_0)}P - S_{t_0} - V_{t_0}(S_{t_0}). \tag{1.48}$$

This *cannot* be positive – otherwise we have arbitrage! Thus we conclude:

$$V_{t_0}(S_{t_0}) \geq e^{-r(t_1-t_0)} P - S_{t_0}. \tag{1.49}$$

We found a lower-bound for the option price, but can we find a precise equation for the option price? Remarkably, the answer is yes – after we make some assumptions regarding the dynamics of the stocks.

1.2.3   Stocks make a multiplicative random walk

For simplicity, let's assume that time is discrete, and call the unit of time dt. To derive the Black-Scholes equation, we will assume that:

$$S_{t+dt} = S_t e^{\mu dt + \xi}, \tag{1.50}$$

where μ is a constant describing the drift of the random walk, and ξ a noise term (e.g.: Gaussian), with standard deviation ξ_0 and vanishing mean. For a given dt, the logarithm of the stock will do a random walk. The size of each step has a variance ξ_0^2, and the number of steps until time t is t/dt. Therefore the standard deviation of the logarithm of the stock will be $\xi_0 \sqrt{t/dt}$. As we have seen before, in order for the limit of $dt \to 0$ to make sense, we need to scale the noise appropriately: $\xi_0 \propto \sqrt{dt}$. For this scaling, the standard deviation of $\log(S)$ will be proportional to $\sqrt{t}$. We will denote the proportionality constant by σ, which is called the **volatility** of the stock.

Note that the noise plus drift term *multiplies* the current value of the stock – hence the noise is multiplicative rather than additive. One benefit of this formulation is that the stock price will always remain positive (since we are always multiplying it by a constant close to 1). This is also known as **geometric Brownian motion** or **geometric random walk**.

We can readily generate a random instance of such dynamics by running the following MATLAB code:

```
dt=0.01; mu=0.1; sigma=1; t=10;
S=1;
for indx=1:(t/dt)
    S(indx+1)=S(indx)*exp(mu*dt+randn*sigma*sqrt(dt));
end;
```

Fig. 1.2 shows the result of this code.

In fact, for the derivation below it will be easier to assume that at each discrete time step the stock price follows one of two *discrete* options, i.e., the noise will be bimodal $\pm\sigma\sqrt{dt}$. We can always "coarse-grain" this model by pooling together $N \gg 1$ steps, leading to a Gaussian noise with the same volatility, so we will not miss much by following this (simpler) route.
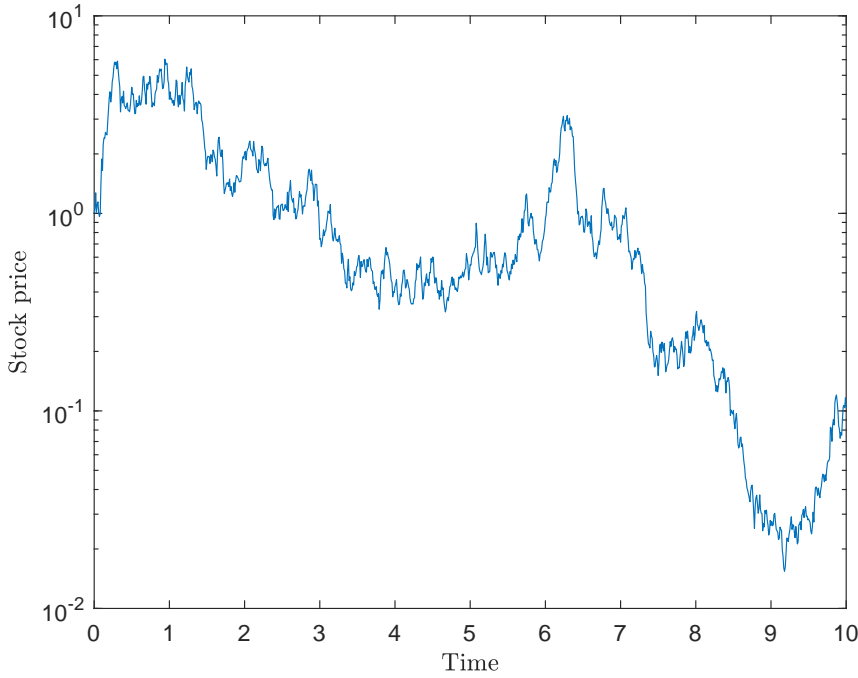
FIGURE 1.2: Stock price is assumed to follow a multiplicative random walk.

1.2.4 Pricing a call option

Suppose we are now at time t_1, and the call option (the right to purchase S at price P) is still available (i.e. suppose we are at 11:59 PM before the date of the option). If $V_{t_1}(S_{t_1}) > S_{t_1} - P > 0$, then anyone can buy S at price S_{t_1}, sell such an option at $V_{t_1}(S_{t_1})$, and sell S back at P (this hinges on the assumption $P < S_{t_1}$, otherwise we won't be able to sell S back). Thus, for the case $P < S_{t_1}$ this arbitrageur ends up with:

$$V_{t_1}(S_{t_1}) - S_{t_1} + P > 0 \tag{1.51}$$

so we must have $V_{t_1}(S_{t_1}) \leq S_{t_1} - P$ by the no-arbitrage condition. But looking at it from the opposite point of view (i.e. short-sell S, buy an option, buy S back at P), we see that $V_{t_1}(S_{t_1}) \geq S_{t_1} - P$. Thus, in fact we must have: $V_{t_1}(S_{t_1}) = S_{t_1} - P$ in the case that $S_{t_1} > P$. But if on the other hand $S_{t_1} \leq P$, then no one would buy the option at any positive price (and option prices must be nonnegative), and so the price of the option must be 0 (trivially). Thus, for a call option:

$$V_{t_1}^{call}(S_{t_1}) = \max(S_{t_1} - P, 0), \tag{1.52}$$

and analogous logic shows that for a put option:

$$V_{t_1}^{put}(S_{t_1}) = \max(P - S_{t_1}, 0). \tag{1.53}$$

1.2.5   Deriving the Black-Scholes equation

Consider a call option to buy a unit of S at price P at time t_1. We will denote the price of the option by $\boldsymbol{V(t, S)}$. The time t can be any time ranging from the present to t_1, and S will be the price of the stock at that time. Our goal will be to show that under our assumptions we can find a unique formula for V. First, consider $\boldsymbol{V(t_1, S)}$. That's precisely the exercise we did in the previous section. Hence:

$$\boldsymbol{V(t_1, S)} = max(S - P, 0). \tag{1.54}$$

We considered above the case when we could buy and exercise an option on the same time, but the more realistic and interesting case is when we purchase an option at some earlier time t_0 and have the right to exercise it at a later time t_1. We will do this recursively by considering one discrete time step back, i.e. by dt; that is, we consider buying the option at $t - dt$, *assuming* that we know how to price the option at time t. Eq. (1.54) will provide the necessary boundary condition to solve the recursive equations we will derive.

Assume that at time $t - dt$, the stock has price S_{t-dt}. According to our random walk formulation, at time t the stock can have one of two prices:

$$\frac{S_t}{S_{t-dt}} = e^{\mu dt \pm \sigma dt^{1/2}}. \tag{1.55}$$

Thus, let us define:

$$S_+ = S_{t-dt} e^{\mu dt + \sigma dt^{1/2}]} \approx S_{t-dt}[1 + \tilde{\mu} dt + \sigma dt^{1/2}], \tag{1.56}$$

and:

$$S_- = S_{t-dt} e^{\mu dt - \sigma dt^{1/2}]} \approx S_{t-dt}[1 + \tilde{\mu} dt - \sigma dt^{1/2}], \tag{1.57}$$

where we expanded the exponential retaining terms of order up to dt, and defined $\tilde{\mu} \equiv \mu + \sigma^2/2$. According to our assumptions, we know that at time t the option will be worth either $\boldsymbol{V(t, S_+)}$ or $\boldsymbol{V(t, S_-)}$ – but we can't be sure which one it will be (and usually, these two will different). Is it possible to take action at time $t - dt$ in a risk-free fashion? i.e., that our financial fate will be the same regardless of whether the stock goes up or down? For the call option under consideration, if we buy the option at $t - dt$ it is in our best interest to have the stock go up in value (since we have the option to buy it at the strike price, regardless of the stock price). How can we "hedge-away" the risk associated with the scenario in which the stock goes *down*? The solution is simple: we need to buy a negative amount of stock (this is possible, since we can short-sell). Let's find out how many negative units of stock we should buy, which we denote by x, to counter-act the price change of one unit of options (we are creating a "portfolio"). If the stock goes down, we lose $\boldsymbol{V(t - dt, S)}(1 + rdt) - \boldsymbol{V(t, S_-)}$, and we gain $x(S(1 + rdt) - S_-)$, where r is the interest rate. Hence our total gain in case that the stock goes down is:

$$G_- = x(S(1 + rdt) - S_-) - (\boldsymbol{V(t - dt, S)}(1 + rdt) - \boldsymbol{V(t, S_-)}). \tag{1.58}$$

Similarly, if the stock goes up our gain will be:

$$G_+ = x(S(1+rdt) - S_+) - (\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S})(1+rdt) - \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_+})). \tag{1.59}$$

Equating the two implies that we have *hedged-away* the risk: no matter what the stock does, our gain will be $G_+ = G_- \equiv G$! We can easily see that this amounts to choosing x obeying the equation:

$$x(S_+ - S_-) = \boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S})(1+rdt) - \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-}) - \boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S})(1+rdt) + \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_+}). \tag{1.60}$$

Clearly, G cannot be positive – since then we violate the no arbitrage condition since we can deterministically make profit. On the other hand, if $G < 0$ we can buy $-x$ amount of stock, and *short-sell* an option at time $t - dt$, thereby making a positive profit of $-G$. The conclusion is that the no arbitrage condition implies that $G = 0$. Since we found x, this allows us to connect $\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S})$ to $\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_+})$ and $\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-})$!

Hence we have:

$$x = \frac{\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_+}) - \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-})}{S_+ - S_-}, \tag{1.61}$$

and:

$$\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S})(1+rdt) = \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-}) + x(S(1+rdt) - S_-). \tag{1.62}$$

The last two equations are already sufficient to recursively find V at an earlier time – and hence we have essentially solved the problem. Here is a short piece of code which computes this, by filling in a full "column" of the matrix $V(t, s)$ based on the previous one (this is, essentially, dynamical programming). Note that for simplicity we chose $r = 0$.

```
dt=0.01; t=0:dt:1; sigma=1; mu=0;
p=1; %stock strike price
s_vec=0:0.01:50; %possible values of stock price
v=zeros(length(t),length(s_vec));
v(1,:)=max(0,s_vec-p); %boundary conditions for solving the backwards recursive relation
for indx=2:length(t)
    for s_indx=1:length(s_vec)
        s=s_vec(s_indx);
        s_p=s*exp(mu*dt+sigma*sqrt(dt)); %if the market goes up
        s_m=s*exp(mu*dt-sigma*sqrt(dt)); %and down
```
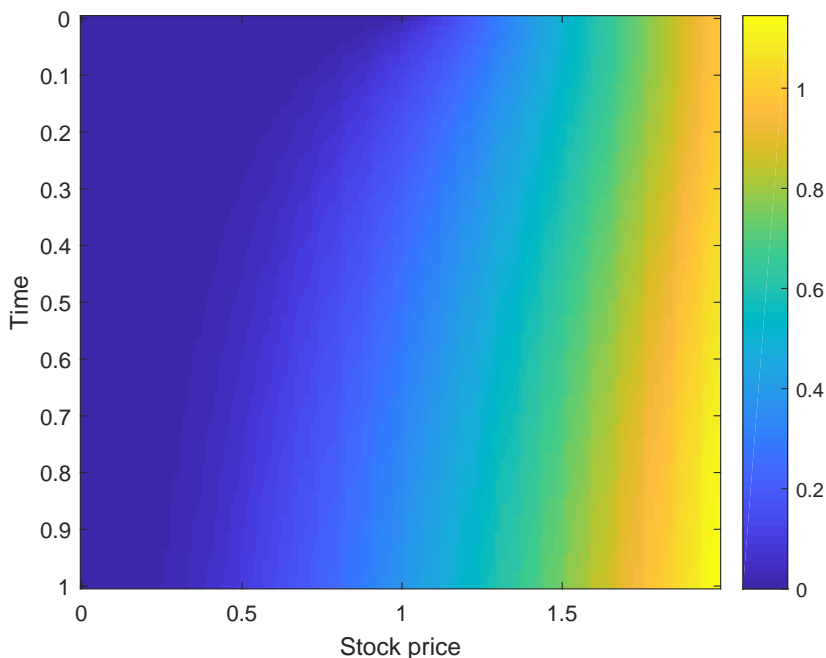
FIGURE 1.3: Option pricing as found using Eqs. (1.54), (1.61) and (1.62)

.

```
        [tmp,indx_p]=min(abs(s_p-s_vec)); %used to round s_p to value in s_vec
        [tmp,indx_m]=min(abs(s_m-s_vec));
        x=(v(indx-1,indx_p)-v(indx-1,indx_m))/(s_p-s_m);
        v(indx,s_indx)=v(indx-1,indx_p)-x*(s_p-s); %option pricing for earlier time
    end;
end;
range=1:200;
imagesc(s_vec(range),t,v(:,range));
ylabel('Time'); xlabel('Stock price');
figure;
plot(s_vec(range),v(1,range),s_vec(range),v(end,range),'linewidth',2);
xlabel('Stock price'); ylabel('Option price'); legend('t=0', 't=1');
```

Figs. 1.3 and 1.4 illustrate the behavior of the option pricing according to these equations, for some arbitrary choice of parameters. Note that we are only plotting a subset of the output matrix – one needs to be careful to use only values for which the results are independent of the cutoff used for the vector of stock price and the time step used. This is why the plots in the above code restrict the range of the stock variable.

We see that as we go back in time, the sharp bilinear behavior that is encapsulated in Eq. (1.54) gets "smeared out". We will shortly understand why this is the case, based on the results of Chapter 2. Plugging Eq. (1.61) into Eq. (1.62) we find that:
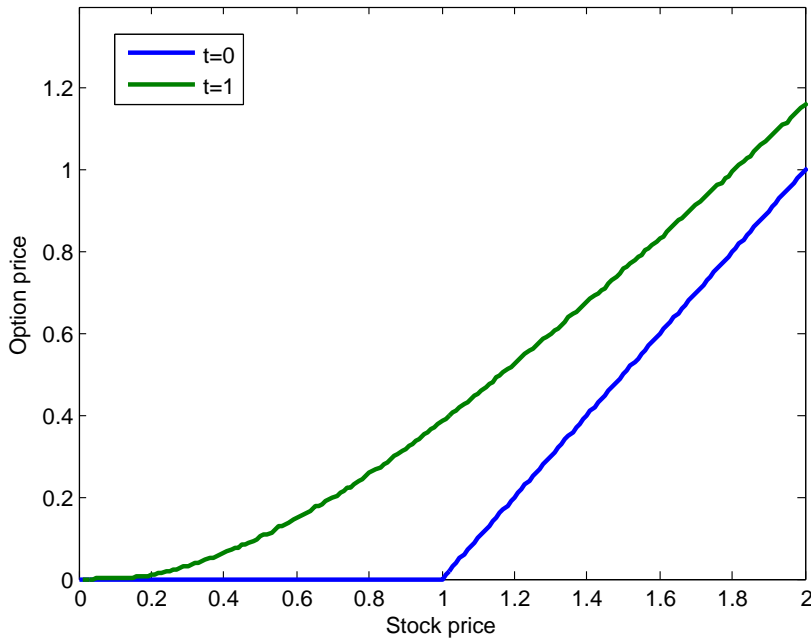
FIGURE 1.4: "Cuts" through the heatmap of Fig. 1.4 showing the pricing of this option at the strike time and 1 unit of time beforehand.

$$\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S}) \approx \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-}) - rdt\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-}) + \frac{\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_+}) - \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S_-})}{S_+ - S_-}[S - S_-/(1 + rdt)]. \qquad (1.63)$$

Hence, defining $dS = S\sqrt{dt}\sigma$ and retaining terms to order $O(dt)$:

$$\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S}) \approx \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) + [S\tilde{\mu}dt - dS]\frac{\partial \boldsymbol{V}}{\partial S}\Big|_{S-dS/2} - rdt\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) + \frac{2dS\frac{\partial \boldsymbol{V}}{\partial S}\big|_S}{2dS}[Srdt + dS - S\tilde{\mu}dt]. \quad (1.64)$$

Note that the derivative is evaluated a the mid-point between S and $S - dS$, making the approximation correct to higher order. This is essential here, since the prefactor of this term is of order $\sqrt{dt}$. Also, note that the two terms corresponding to $\pm\tilde{\mu}dtS\frac{\partial \boldsymbol{V}}{\partial S}$ cancel each other – so that the drift term $\tilde{\mu}$ dropped out (can you see why the fact that they are evaluated at slightly different stock prices does not matter for this cancelation?).

This leads to:

$$\boldsymbol{V}(\boldsymbol{t} - \boldsymbol{dt}, \boldsymbol{S}) \approx \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) - dS\frac{\partial \boldsymbol{V}}{\partial S}\Big|_{S-dS/2} - rdt\boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) + \frac{\partial \boldsymbol{V}}{\partial S}\Big|_S dS + \frac{\partial \boldsymbol{V}}{\partial S}\Big|_S Srdt. \qquad (1.65)$$

We can replace:

$$\frac{\partial \boldsymbol{V}}{\partial S}\Big|_S dS - dS \frac{\partial \boldsymbol{V}}{\partial S}\Big|_{S-dS/2} \approx dS \frac{\partial^2 \boldsymbol{V}}{\partial S^2}\Big|_S \frac{dS}{2} = \sigma^2 dt \frac{S^2}{2} \frac{\partial^2 \boldsymbol{V}}{\partial S^2}\Big|_S, \tag{1.66}$$

Leading to:

$$-\frac{\partial \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S})}{\partial t} = -r \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) + \frac{\partial \boldsymbol{V}}{\partial S}\Big|_S Sr + \frac{\sigma^2 S^2}{2} \frac{\partial^2 \boldsymbol{V}}{\partial S^2}\Big|_S. \tag{1.67}$$

This is the celebrated Black-Scholes equation [9]. It is rather counter-intuitive that the drift term μ does not appear here – only the volatility is important for pricing the option within this framework!

1.2.6 Solving the Black-Scholes equation using the Green's function of the diffusion equation

As shown above, it is straightforward to solve the equation numerically. Having cast it as a PDE, however, allows us to obtain explicit solutions. We would like to make a change of variables to cancel the interest rate r and get rid of the multiplicative nature of the random walk.

We will attempt to make a transformation of the form:

$$s = \log(S); \tilde{t} = -t; \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S}) \rightarrow v(\tilde{t}, s). \tag{1.68}$$

Computing the partial derivatives we find that:

$$\frac{\partial v}{\partial s} = \frac{\partial \boldsymbol{V}}{\partial S} S, \tag{1.69}$$

and:

$$\frac{\partial^2 v}{\partial s^2} = \frac{\partial^2 \boldsymbol{V}}{\partial S^2} S^2 + S \frac{\partial \boldsymbol{V}}{\partial S}. \tag{1.70}$$

The time derivative is:

$$\frac{\partial v}{\partial \tilde{t}} = -\frac{\partial V}{\partial t}. \tag{1.71}$$

Putting it together we find:

$$-\frac{\partial \boldsymbol{V}(\boldsymbol{t}, \boldsymbol{S})}{\partial t} = \frac{\partial v}{\partial \tilde{t}} = -rv + \frac{\partial v}{\partial s} r + \frac{\sigma^2}{2}[\frac{\partial^2 v}{\partial s^2} - \frac{\partial v}{\partial s}]. \tag{1.72}$$

This already is a simplification since all the coefficients are now constant.

The second step will be in replacing $v \to \tilde{v} = e^{r\tilde{t}} v$. This gives:

$$\frac{\partial \tilde{v}}{\partial \tilde{t}} = e^{r\tilde{t}} \frac{\partial v}{\partial \tilde{t}} + r\tilde{v}, \tag{1.73}$$

hence:

$$e^{r\tilde{t}} \frac{\partial v}{\partial \tilde{t}} = \frac{\partial \tilde{v}}{\partial \tilde{t}} - r\tilde{v} = -r\tilde{v} + \frac{\partial \tilde{v}}{\partial s} r + \frac{\sigma^2}{2} [\frac{\partial^2 \tilde{v}}{\partial s^2} - \frac{\partial \tilde{v}}{\partial s}]. \tag{1.74}$$

So we succeeded in eliminating the rv term:

$$\frac{\partial \tilde{v}}{\partial \tilde{t}} = [r - \sigma^2/2] \frac{\partial \tilde{v}}{\partial s} + \frac{\sigma^2}{2} \frac{\partial^2 \tilde{v}}{\partial s^2}. \tag{1.75}$$

Finally, to eliminate the first derivative term, we need to make a less intuitive substitution that will mix the time and stock variables, defining: $x \equiv s + A\tilde{t}$, with $A \equiv r - \sigma^2/2$, $\tilde{v}(\tilde{t}, s) \to \hat{v}(\tilde{t}, x)$, we find:

$$\left.\frac{\partial \tilde{v}(\tilde{t}, s)}{\partial \tilde{t}}\right|_s = \left.\frac{\partial \hat{v}(\tilde{t}, x)}{\partial \tilde{t}}\right|_x \left.\frac{\partial \tilde{t}}{\partial \tilde{t}}\right|_s + \left.\frac{\partial \hat{v}(\tilde{t}, x)}{\partial x}\right|_{\tilde{t}} \left.\frac{\partial x}{\partial \tilde{t}}\right|_s = \frac{\partial \hat{v}(\tilde{t}, x)}{\partial \tilde{t}} + A\frac{\partial \hat{v}}{\partial x}, \tag{1.76}$$

and similarly:

$$\left.\frac{\partial \tilde{v}(\tilde{t}, s)}{\partial s}\right|_{\tilde{t}} = \left.\frac{\partial \hat{v}(\tilde{t}, x)}{\partial x}\right|_{\tilde{t}} \left.\frac{\partial x}{\partial s}\right|_{\tilde{t}} + \left.\frac{\partial \hat{v}(\tilde{t}, x)}{\partial \tilde{t}}\right|_{\tilde{t}} \left.\frac{\partial \tilde{t}}{\partial s}\right|_{\tilde{t}} = \frac{\partial \hat{v}(\tilde{t}, x)}{\partial x}. \tag{1.77}$$

A note on notation: in the above two equations $\Big|_a$ implies keeping the variable a fixed while taking the partial derivative with respect to the other variable – not to be confused with the notation used in our derivation of the Black-Scholes equation where this implied evaluating the partial derivative at a particular value.

Hence:

$$\frac{\partial \hat{v}(\tilde{t}, x)}{\partial \tilde{t}} + A\frac{\partial \hat{v}}{\partial x} = [r - \sigma^2/2] \frac{\partial \hat{v}}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 \hat{v}}{\partial x^2}. \tag{1.78}$$

So we end up with the standard diffusion equation:

$$\frac{\partial \hat{v}(\tilde{t}, x)}{\partial \tilde{t}} = \frac{\sigma^2}{2} \frac{\partial^2 \hat{v}}{\partial x^2}. \tag{1.79}$$

This allows us to use our knowledge and intuition from studying random walks also for this problem! Previously, we solved this equation for a δ-function boundary condition (i.e., the Green's function or propagator). We can easily recast our initial conditions for $V(t_1, S)$ to one in terms of x, $\tilde{V}$, which will denote $V_0(x)$.

The solution for $\hat{v}$ can now be readily expressed in terms of the Green-function:

$$G(t, t', x, x') = \frac{1}{\sqrt{4\pi D|t - t'|}} e^{-\frac{|x-x'|^2}{4D|t-t'|}}. \tag{1.80}$$

Thus:

$$\hat{v}(\tilde{t}, x) = \int G(\tilde{t}, 0, x, x') V_0(x') dx'. \tag{1.81}$$

The integrals can be evaluated in terms of the error function, and the solution is known as the Black—Merton–Scholes formula [9, 10].

1.3   Another example: the "well function" in hydrology

A mathematically similar example where a diffusion equation is solved using the Green's function is found in the field of hydrology [11]: Consider pumping from a deep well, connected to an underwater reservoir, which is (approximately) two-dimensional (i.e., it is thin compared to its lateral extent). Moreover, it turns out that the porous nature of the soil implies that to a good approximation the velocity is proportional to the pressure gradient (this is known as "Darcy's law"). This helps us establish an analogy with the diffusion equation: we saw that we can derive the diffusion equation from the continuity equation supplemented with Fick's law, $\vec{j} = -D\nabla n$. In this example, the current is proportional to the pressure gradient, hence $\vec{j} \propto -\nabla p$. The density of the water is constant, and the height of the water column is what compensates for the non-uniform currents. The height of the column is also proportional to the pressure (up to an additive constant, which is inconsequential since only pressure *gradients* matter here). We therefore find:

$$\frac{\partial h}{\partial t} = D\nabla^2 h, \tag{1.82}$$

where D is a constant (not to be confused with actual diffusion! we nevertheless use this notation to emphasize the formal analogy). If we don't pump water into or out of the well, its level will equilibrate. What happens if we continuously pump water out of the reservoir?

The solution can be written in terms of the Green's function of two-dimensional diffusion, as:

$$H(r, T) \propto -\int_0^T \frac{e^{-\frac{r^2}{4Dt'}}}{4\pi Dt'} dt', \tag{1.83}$$

where the minus sign accounts for the fact that we are pumping water *out* of the well. To proceed, let us define $u \equiv \frac{r^2}{4Dt'} \rightarrow$

$$H(r,T) \propto -\frac{1}{4\pi D} \int_{\frac{r^2}{4DT}}^{\infty} \frac{e^{-u}}{u} du. \tag{1.84}$$

This can be expressed using the *exponential integral function*:

$$E_1(x) = \int_x^{\infty} \frac{e^{-u}}{u} du. \tag{1.85}$$

For small, positive arguments, this function diverges logarithmically. This can be seen by integrating by parts:

$$E_1(x) = \log(u)e^{-u}|_x^{\infty} + \int_x^{\infty} e^{-u} \log(u) du. \tag{1.86}$$

The integral in Eq. (1.86) converges also when we set its lower limit as 0, and the result is known as the Euler-Mascheroni constant, γ_E. Thus we find:

$$E_1(x) \approx -\gamma_E - \log(x). \tag{1.87}$$

This function in fact pops up in numerous applications in physics and maths (and we will see this integral again later on in Chapter 6).

So close to the well (how close?) the height obeys:

$$H \propto \gamma_E + \log\left(\frac{r^2}{4DT}\right). \tag{1.88}$$

Fig. 1.5(a), taken from Ref. [11], shows the excellent agreement of this prediction with the measurements. In fact, this is mathematically analogous to a recent application in the context of quantum optics, see Ref. [12] and Fig. 1.5(b).

1.4   Summary

In this chapter, we first generalized the random walker of Chapter 2 to the case where a force can act on the particle, thereby biasing the random motion in a space-dependent manner – described by the Langevin equation. We introduced an important discretization procedure that allows us to simulate such equations, and is also helpful conceptually. We next derived a PDE describing the time evolution of the probability distribution, the Fokker-Planck equation. In the case of particles diffusing in a potential due to thermal fluctuations, we derived a relation between the diffusion constant and the mobility of the particles, known
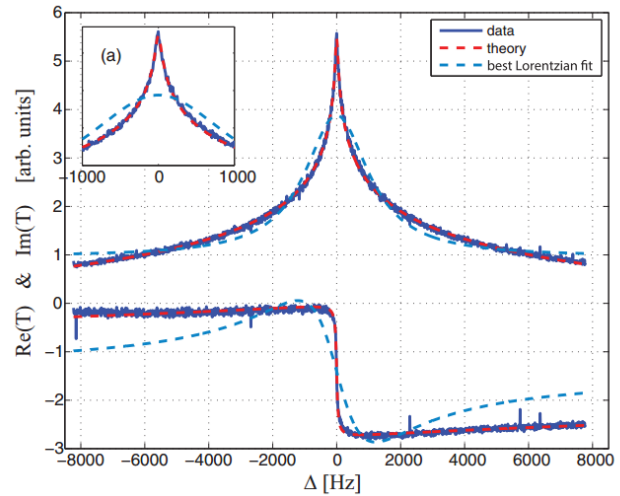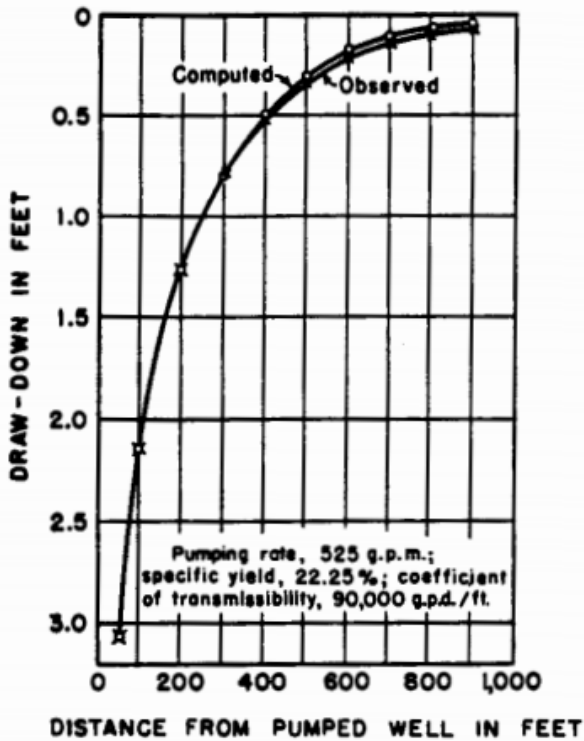
FIGURE 1.5: (a) The "well function" of Eq. (1.84), and the measurements of water height near a well during constant pumping. From Ref. [11]. (b) The same mathematics applies in the context of quantum optics. Figure adapted from Ref. [12].

as the Einstein relation. We next illustrated the philosophy behind the Langevin equation by discussing a discrete version of it, recently utilized in the context of cell size regulation. Similarly, we derived a PDE for the temporal evolution of the prices of options – the Black-Scholes equation – which resembled the Fokker-Planck equation of a freely diffusing particle. Finally, we showed how a similar diffusion equation arises in a very different context – pumping water out of a well – and how the Green's function we derived in the context of the random walk problem can solve this problem and lead to interesting physical insights.

For further reading Refs. [13, 14] discuss Langevin and Fokker-Planck in a far more comprehensive way than that outlined in this chapter, including numerous additional applications. Ref. [15] has an extended discussion of applications to finance going well beyond the discussion in this chapter.

Bibliography

[1] Gillespie, D.T., 1996. The mathematics of Brownian motion and Johnson noise. American Journal of Physics, 64(3), pp.225-240.

[2] Amir, A., 2017. Is cell size a spandrel?. Elife, 6.

[3] Koppes, L. J., et al. "Correlation between size and age at different events in the cell division cycle of Escherichia coli." Journal of bacteriology 143.3 (1980): 1241-1252.

[4] Amir, Ariel. "Cell size regulation in bacteria." Physical Review Letters 112.20 (2014): 208102.

[5] Robert, Lydia, et al. "Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism." BMC biology 12.1 (2014): 17.

[6] Soifer, Ilya, Lydia Robert, and Ariel Amir. "Single-cell analysis of growth in budding yeast and bacteria reveals a common size regulation strategy." Current Biology 26.3 (2016): 356-361.

[7] Ho, Po-Yi, and Ariel Amir. "Simultaneous regulation of cell size and chromosome replication in bacteria." Frontiers in microbiology 6 (2015).

[8] Ho, P.Y., Lin, J. and Amir, A., 2018. Modeling cell size regulation: From single-cell level statistics to molecular mechanisms and population level effects, Annual Reviews of Biophysics 47, 251 (2018).

[9] Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–654.

[10] Merton, R.C., 1973. Theory of rational option pricing. The Bell Journal of economics and management science, pp.141-183.

[11] Theis, C. V. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground water storage. Washington, DC: US Department of the Interior, Geological Survey, Water Resources Division, Ground Water Branch, 1935.

[12] Pugatch, R. et al. "Anomalous symmetry breaking in classical two-dimensional diffusion of coherent atoms." Physical Review A 89.3 (2014): 033807.

[13] Gardiner, C., 2009. Stochastic methods (Vol. 4). Berlin: Springer.

[14] Schuss, Z., 2009. Theory and applications of stochastic processes: an analytical approach (Vol. 170). Springer Science and Business Media.

[15] Paul, W. and Baschnagel, J., 2013. Stochastic processes (Vol. 1). Springer.

1.5   Exercises

3.1 Dephasing in Nuclear Magnetic Resonance (NMR)

A nuclear spin in a constant magnetic field B is described by Hamiltonian $H = - \begin{pmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{pmatrix}$ where $\omega_0 \propto B$ is half the Larmor frequency, and we set $\hbar = 1$. The Schrodinger equation $i\dot{\Psi} = H\Psi$ for the time evolution of wavefunction $\Psi = c_1 \langle \uparrow \rangle + c_2 \langle \downarrow \rangle = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ yields the differential equations $\dot{c}_1 = i\omega_0 c_1$, $\dot{c}_2 = -i\omega_0 c_2$.

To account for fluctuations δB of the magnetic field due to coupling with other spins, the differential equations are modified by a noise term $\delta\omega$: $\dot{c}_1 = i(\omega_0 + \delta\omega) c_1$, $\dot{c}_2 = -i(\omega_0 + \delta\omega) c_2$ and the solutions by a phase factor: $c_1 = c_1(t_0) e^{i(\omega_0 t + \varphi(t))}$, $c_2 = c_2(t_0) e^{-i(\omega_0 t + \varphi(t))}$. The modified differential equations and solutions equation yield a Langevin equation for the time evolution of the phase: $\dot{\varphi} = \delta\omega$ (that is, $\dot{\varphi} = \delta B$).

(a) Given the Langevin equation $\frac{d\varphi}{dt} = \xi$ where ξ is a Gaussian random variable representing fluctuations of the magnetic field with autocorrelation function $C(t) = W^2 e^{-t/\tau}$, calculate $\left\langle e^{i\varphi(t)} e^{-i\varphi(t')} \right\rangle$ which determines the spin relaxation.

(b) What happens for short and long correlation times of the noise?

3.2 Boltzmann distribution

Show that the Boltzmann distribution is a stationary solution of the Fokker-Planck equation for a general $V(x)$, $D(x)$, and $\mu(x)$ under the assumption of the Einstein relation between D and μ.

3.3 Fokker-Planck equation for an Ornstein-Uhlenbeck process

Consider the Langevin equation:

$$\frac{dv}{dt} = \frac{-v}{\tau} + \xi$$

with ξ as white noise $\langle \xi(t) \xi(t') \rangle = A\delta(t - t')$.

(a) What is the corresponding Fokker-Planck equation?

(b) If $v(t=0) = v_0$, what is the velocity distribution at time t? *Hint: you may find it useful to represent v at time t given a noise $\xi(t')$, and rely on the statistics of the noise to infer the distribution.*

(c) What happens as $t \to \infty$?

(d) What is the thermodynamic interpretation of (c)? How does A depend on temperature?

3.4 Diffusion in a logarithmic potential

Consider diffusion of a particle in a logarithmic potential $V(x) = a\log(|x|) + U(x)$, where the correction $U(x)$ is negligible for large $|x|$ and it ensures that $V(x)$ does not diverge at the origin. Assume that the temperature is such that $kT = 1$.

(a) Write the Langevin equation and the corresponding Fokker-Planck equation diffusion in logarithmic potential, assuming that we are in the over-damped limit and that the diffusion constant is D.

(b) For $a > 1$, what is the steady-state solution (away from the origin)? What is the issue for $a < 1$?

(c) For $a < 1$, guess a scaling solution for the dynamics of the probability distribution: $p(x,t) = 1/t^b f(x/t^c)$ Find the value of c for which f obeys a second-order *ordinary* differential equation. Find b from demanding that the probability distribution is normalized.

(d) Write the ODE that f obeys. What are the boundary conditions?

We thank Dr. Ori Hirschberg for help in creating this problem.

3.5 Cell size control

(a) Simulate the following discrete map in MATLAB or the software of your choice:

$$x_{n+1} = (1-\alpha)x_n + \xi,$$

where ξ is a random variable drawn from a Gaussian distribution, with mean 0 and standard deviation of 0.1, and for values of α of 0, 0.1, 0.5 and 1. (b) For each instance, run the map for 100,000 steps, and plot the distribution of x. (c) Calculate numerically the Pearson correlation coefficient between the variables x_n and x_{n+1}.

Hint: the MATLAB commands randn and corrcoeff could be useful.

(d) Scientists in a laboratory at Harvard have discovered two new bacterial species, species A and species B, and they are trying to understand the how each species regulates its size. In particular, they assume that bacteria born at size v_b grow until they reaches a size $f(v_b) = 2v_b^{1-\alpha}v_0^\alpha$, where the parameter α depends

on the species. Note that for the logarithmic of size, x, this dependence implies the discrete map of part (a). They also assume that the bacteria have a doubling time of 20 minutes, with some timing error σ_n that depends on the species. For each species, the scientists have (i) measured the distribution of sizes in a population of 10,000 bacteria, and (ii) made a scatter plot of the size of mother at birth versus the size of the subsequent daughter at birth. Figure 1.6 shows the plots for species A and B. One species has $\alpha = 0.99$; the other has $\alpha = 0.01$. One species has $\sigma_n = 5$ minutes; the other species has $\sigma_n = \frac{1}{2}$ minutes. What are the values of α and σ_n for species A and for species B?
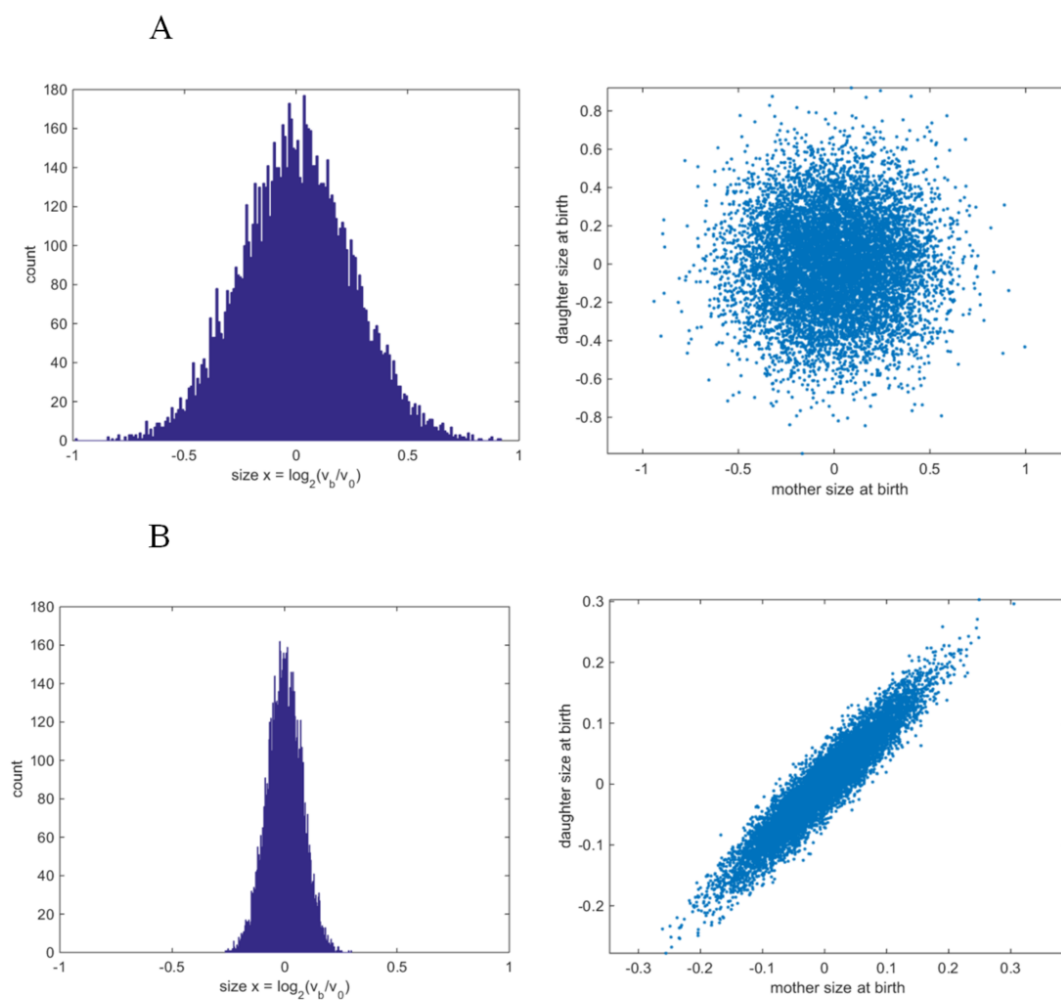

FIGURE 1.6: Histograms and correlations for two fictitious species of bacteria.

3.6 Cell size control*

In the notes, we derived the Pearson correlation coefficient (i) between the volume of the mother at birth and the volume of the daughter at birth, and (ii) between the volume at birth and the interdivision time (the time from birth to division).

(a) Simulate the incremental model ($\alpha = 1/2$) for 10,000 bacteria for one division cycle, and computationally estimate these two Pearson correlation coefficients (i) and (ii).

(b) Analytically derive the Pearson correlation coefficients (iii) between the interdivision times of mothers and daughters, and (iv) between the interdivision times of two cousins. What are the correlations for arbitrary α, and for $\alpha = 1/2$? Compare your results to recent experiments on human HeLa cells [Sandler et al. *Nature* (2015)] that obtained a Pearson correlation coefficient ≈ 0.04 for mother-daughter interdivision times, and ≈ 0.63 for cousin-cousin interdivision times (and ≈ 0.76 for sister-sister interdivision times). Are the model and experiment consistent?

3.7 Black-Scholes when stocks do a normal random walk

(a) For the case of a *multiplicative* random walk (done in class), if the volatility is $\sigma = \frac{1}{\sqrt{day}}$, and the current value of a stock is 10\$, how much would a "call option" be worth to buy that stock at a price of 5\$ in 1, 5 and 10 days from now? Assume that interest rate $r = 0$.

(b) Make the (unrealistic) assumption that the stock market does a random walk rather than a *multiplicative* random walk. Repeat the derivation of the Black-Scholes equation we did in class. What is the (discrete) backwards recursive relation?

(c) By Taylor expanding the equation to the appropriate order, derive the PDE corresponding to the discrete equation you solved in part (b).

3.8 Price of options within Black-Scholes

Prove that the option price increases as the time difference between now (t_0) and the exercising time (t_1, $t_1 > t_0$) increases, for a given strike price at t_1 and stock price S, and under the assumption that the interest rate is positive.

3.9 Divisibility**

Prove that a random variable uniformly distributed in $[0, 1]$ cannot be written as the sum of two independent and identically distributed (i.i.d.) random variables.