Chapter 6

Working with the single-locus diffusion model

The formal results in the previous chapter glossed over an important question: what are the benefits of showing that a particular microscopic model is equivalent to the single-locus diffusion,

$$\frac{\partial f}{\partial t} = \underbrace{sf(1-f)}_{\text{selection}} + \underbrace{\mu(1-f) - \nu f}_{\text{mutation}} + \underbrace{\sqrt{\frac{f(1-f)}{N}} \cdot \eta(t)}_{\text{genetic drift}}, \qquad (6.1)$$

for some effective parameters N, s, μ, and ν?

One benefit is that it is sometimes possible to speed up simulations of a particular model by replacing it with a different member of the same universality class. This is a common strategy in computational studies. For example, the Poisson sampling step in the serial dilution model is slightly faster to simulate on a computer than the equivalent Binomial sampling step in the Wright-Fisher model — our convergence results show that the faster version will still yield similar predictions when the population size is large.

Similar logic applies for analytical results as well. If one model is analytically tractable, it can provide a "mathematical speedup" for all of the other models

that share the same universality class. In this chapter, we will show that the SDE model in Eq. (6.1) constitutes one such example, enabling analytical results that are difficult or impossible to derive in the original Wright-Fisher model (or our serial dilution model from Chapter 4).

To do so, we will need to develop methods for "integrating" stochastic differential equations like Eq. (6.1). This was trivial in the case of the Gaussian random walk in Section 5.2.1, since the step sizes all had the same mean and variance. The task is more challenging for models like Eq. (6.1) where the sizes depend on the current value of $f(t)$. For sufficiently dense sampling of $f(t)$, we could generalize the "path-integral" formula in Eq. (5.16) to write down a similar approximation for the probability of the full trajectory,

$$p(f(\delta t), \ldots, f(t)|f(0)) \approx \prod_{k=0}^{\frac{t}{\delta t}-1} \frac{e^{-\frac{[\delta f(k \cdot \delta t) - \mu(f(k \cdot \delta t)) \cdot \delta t]^2}{2 \cdot \sigma^2(f(k \cdot \delta t)) \cdot \delta t}}}{\sqrt{2\pi\sigma^2(f(k \cdot \delta t))}}, \tag{6.2}$$

This approximation will be valid if the frequency shifts in successive timepoints are small enough that $\mu(f(t + \delta t)) \approx \mu(f(t))$ and $\sigma^2(f(t + \delta t)) \approx \sigma^2(f(t))$. Unfortunately, this path integral formulation is only rarely useful in practice, since it is rare that we will ever have such fine-grained measurements of $f(t)$. Instead, we are often interested in more coarse-grained outcomes: will the mutation eventually fix or go extinct? How much do we expect its frequency to change over a non-infinitesimal time interval? Enumerating and summing over all of the trajectories that contribute to these outcomes using Eq. (6.2) is often as difficult as solving the SDE itself.
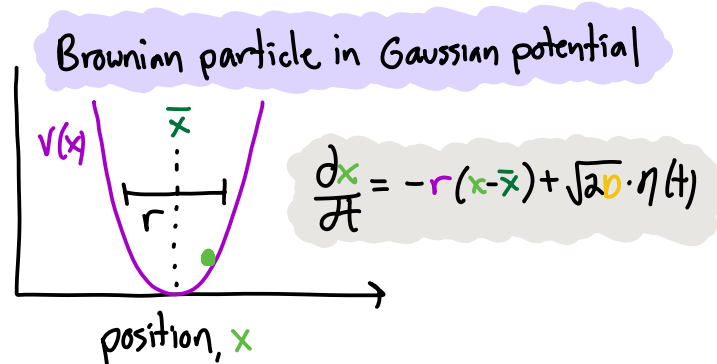
To make progress, we will need to introduce a different methodology for working directly with SDEs like Eq. (6.1). These methods are sometimes known as the *stochastic calculus*. The name may sound intimidating, but we will see that most of these rules can be derived from straightforward applications of the discrete update rule in Eq. (5.35). To illustrate these ideas, it will be helpful to start by considering a simpler SDE that frequently arises in physical settings. We will then show how we can extend these ideas to the single-locus model in Eq. (6.1).

6.1 Detour: Brownian particle in a quadratic potential

We will start by considering the simpler problem of a **_Brownian particle in a quadratic potential_**. This model arises in many different areas of physics, and will turn out to be useful in some evolutionary contexts as well. The goal is to model the position of an ordinary random walker that is subject to a restoring force $F(x) = -\frac{\partial V(x)}{\partial x}$. For a quadratic potential, the SDE can be written in the general form,

$$\frac{\partial x}{\partial t} = -\underbrace{r(x - \overline{x})}_{\frac{\partial V}{\partial x}} + \sqrt{2D} \cdot \eta(t) \tag{6.3}$$

where $\overline{x}$ is the **_equilibrium point_**, r is the **_restoring force_**, and D is the **_diffusion constant_** (which is proportional to the temperature kT in many physical problems).[1] This is the simplest SDE model with non-constant coefficients.

In the absence of noise ($D = 0$), the solution to Eq. (6.3) is given by the deterministic function,

$$x_{\text{det}}(t) = \overline{x} + [x(0) - \overline{x}]e^{-rt} . \tag{6.4}$$

---

[1] If you're curious where this equation comes form, you can think of it as an instantaneous balance between a conservative force $F(x) = -\frac{\partial V(x)}{\partial x}$ and a viscous drag force $F_{\text{drag}} = -\gamma \frac{\partial x}{\partial t}$ when the inertia of the particle is small. This explains why the SDE depends on the first derivative of $x(t)$ rather than the second derivative that you might expect from Newtonian mechanics. For simplicity, we have absorbed the drag coefficient γ into the definition of $V(x)$ in Eq. (6.3).

This shows that the position of the particle approaches the equilibrium point $\overline{x}$ at an exponential rate r. How does this picture change in the presence of noise?

One way to approach this question is to focus on the moments of $x(t)$, e.g. the mean $\langle x(t) \rangle$. We can calculate this quantity by recalling that our SDE notation is shorthand for the discrete update rule in Eq. (5.23). For the SDE in Eq. (6.3), the corresponding update rule is

$$x(t + \delta t) = x(t) - r[x(t) - \overline{x}]\delta t + \sqrt{2D \cdot \delta t} \cdot Z_t \,, \tag{6.5}$$

where Z_t is a Gaussian random variable $\langle Z_t \rangle = 0$ and $\langle Z_t^2 \rangle = 1$. Taking the average of both sides, we find that

$$
\begin{aligned}
\langle x(t + \delta t) \rangle &= \langle x(t) - r[x(t) - \overline{x}]\delta t + \sqrt{D \cdot \delta t} \cdot Z_t \rangle \\
&= \langle x(t) \rangle - r\left[\langle x(t) \rangle - \overline{x}\right]\delta t + 0
\end{aligned}
\tag{6.6}
$$

where we have exploited the linearity of the expectation and the fact that $\langle Z_t \rangle = 0$. This is now a *deterministic* update rule for the mean position $\langle x(t) \rangle$. Taking differences and dividing by δt yields

$$\frac{\langle x(t + \delta t) \rangle - \langle x(t) \rangle}{\delta t} = -r\left[\langle x(t) \rangle - \overline{x}\right] \tag{6.7}$$

which reduces to

$$\frac{\partial \langle x(t) \rangle}{\partial t} = -r\left[\langle x(t) \rangle - \overline{x}\right] \tag{6.8}$$

in the limit that δt is small. We've now reduced our stochastic differential equation to an *ordinary differential equation* for $\langle x(t) \rangle$. The solution is given by

$$\langle x(t) \rangle - \overline{x} = [x(0) - \overline{x}] e^{-rt} \tag{6.9}$$

which is equivalent to the deterministic solution in Eq. (6.4). Thus, in this particular example, the average value of $x(t)$ is unchanged by the addition of noise.

What about the spread around this average value? When $x(0) = \overline{x} = 0$, the variance in the position is given by the second moment, $\langle x(t)^2 \rangle$. We can calculate this quantity using the same basic procedure that we used for the mean:

Step 1: We can start by using the update rule in Eq. (6.5) to write down an expression for $x(t + \delta t)^2$,

$$x(t + \delta t)^2 = \left[x(t) - rx(t)\delta t + \sqrt{2D\delta t}Z_t \right]^2 \qquad (6.10)$$

Taking the average of both sides yields:

$$\langle x(t + \delta t)^2 \rangle = \left\langle \left[x(t) - rx(t)\delta t + \sqrt{2D\delta t}Z_t \right]^2 \right\rangle \qquad (6.11)$$

Step 2: We can then expand the terms in the right hand side as a power series in δt. This yields

$$\langle x(t + \delta t)^2 \rangle = \left\langle x(t)^2 - 2rx(t)^2\delta t + 2x(t)\sqrt{D\delta t}Z_t + 2DZ_t^2\delta t + \ldots \right\rangle$$
$$= \langle x(t)^2 \rangle - 2r\langle x(t)^2 \rangle \delta t + 2D\delta t + \ldots \qquad (6.12)$$

where we have used the fact that $\langle Z_t \rangle = 0$ and $\langle Z_t^2 \rangle = 1$.

Step 3: We can then rewrite this expression as a difference equation:

$$\frac{\langle x(t + \delta)^2 \rangle - \langle x(t)^2 \rangle}{\delta t} = -2r\langle x(t) \rangle + 2D + \ldots \qquad (6.13)$$

where the higher-order terms are all larger powers of δt. When δt is small, this yields an ordinary differential equation for the second moment

$$\frac{\partial \langle x(t)^2 \rangle}{\partial t} = -2r\langle x(t)^2 \rangle + 2D \qquad (6.14)$$

The first term is exactly what we'd expect by applying the chain rule to the deterministic solution,

$$\frac{\partial[x_{\mathrm{det}}(t)^2]}{\partial t} = 2x_{\mathrm{det}}(t) \cdot \partial_t x_{\mathrm{det}}(t) = -2r x_{\mathrm{det}}(t)^2 \qquad (6.15)$$

The second term in Eq. (6.14) is a new contribution that arises purely from the addition of noise. This term arose from taking the average of two correlated noise terms, each of which was $\propto \sqrt{\delta t}$. In this way, we see that SDEs require us to generalize the ordinary chain rule from calculus (this is often known as ***Ito's formula***).

At long times, the variance approaches the steady-state value

$$\frac{\partial\langle x(t)^2\rangle}{\partial t} = 0 \implies \langle x(t)^2\rangle = \frac{D}{r} \qquad (6.16)$$

which constitutes a balance between the diffusive noise D and the deterministic restoring force r. Larger values of D (or smaller values of r) lead to larger deviations from $\overline{x}$, while smaller values of D (or larger values of r) lead to a tighter distribution.

One can repeat the steps above to calculate higher moments of $x(t)$. In this case, it is also possible to calculate the entire distribution of $x(t)$ at long times. We can do this by turning to the Fokker-Planck representation of the SDE (Section 5.2.3). For a general SDE of the form

$$\frac{\partial x}{\partial t} = -\frac{\partial V(x)}{\partial x} + \sqrt{2D} \cdot \eta(t) , \qquad (6.17)$$

the corresponding Fokker-Planck equation is

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[-\frac{\partial V}{\partial x} \cdot p(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[Dp(x,t)\right] . \qquad (6.18)$$

At long times, we expect that the time derivatives will vanish ($\partial_t p = 0$), so that

the steady-state distribution will satisfy

$$0 = -\frac{\partial}{\partial x}\left[-\frac{\partial V}{\partial x} \cdot p(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[Dp(x,t)\right] \tag{6.19}$$

The solution is given by the so-called ***Boltzmann distribution***
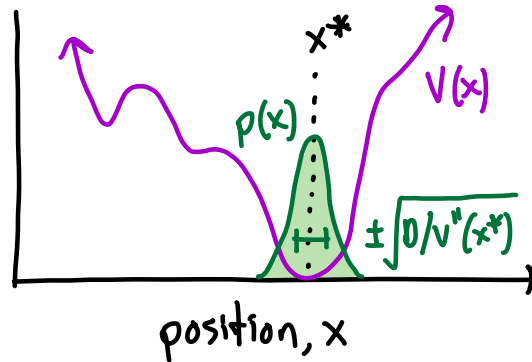
$$p(x) \propto e^{-\frac{V(x)}{D}}, \tag{6.20}$$

which you might recognize from a course in equilibrium statistical mechanics. In the Boltzmann distribution, the most likely position of the particle is the one with the lowest value of $V(x)$. This will often occur at a local minimum of $V(x)$, where

$$\left.\frac{\partial V}{\partial x}\right|_{x=x^*} = 0, \quad \left.\frac{\partial^2 V(x)}{\partial x^2}\right|_{x=x^*} > 0 \tag{6.21}$$

If we expand around this point, the Boltzmann distribution approaches the Gaussian form,

$$p(x) \propto e^{-\frac{V(x^*)+V'(x^*)(x-x^*)+\frac{V''(x^*)}{2}(x-x^*)^2+\dots}{D}} \propto e^{-\frac{V''(x^*)}{2D}(x-x^*)^2} \tag{6.22}$$

with mean $\langle x \rangle = x^*$ and variance $\sigma^2 = D/V''(x^*)$. This approximation is exact in the case of our quadratic potential, where $x^* = \overline{x}$ and $V''(x) = r$. Thus, we can think about the quadratic potential as describing the generic behavior of a Brownian particle near a local energy minimum. This will be a "Case 1" distribution according to our terminology in Section 2.2.1: the particle will stay close to its deterministic expectation (x^*) with some fuzziness due to noise ($\pm\sqrt{D/V''(x^*)}$).

6.2   Back to the single-locus model

We can now try to use the same techniques to analyze the single-locus model in Eq. (6.1). This model has two key differences from the Brownian particle example above:

1. The effective diffusion coefficient now depends on the current value of $f(t)$. This means that the strength of noise will vary for different parts of the mutation trajectory.

2. The selection term in the deterministic portion of the SDE is now a *non-linear* function of $f(t)$.

These may seem like small differences, but we will see that they can have a profound effect on the dynamics of the resulting mutation trajectories.

6.3   Dynamics of the mean and variance

For example, if we want to calculate the average mutation frequency $\langle f(t) \rangle$, starting from a given frequency $f(0)$, we can repeat the procedure above starting

from the analogous discrete update rule:

$$f(t + \delta t) = f(t) + sf(t)[1 - f(t)]\delta t + \mu(1 - f)\delta t$$
$$- \nu f \delta t + \sqrt{\frac{f(t)[1 - f(t)]\delta t}{N}} \cdot Z_t \,. \qquad (6.23)$$

Taking the average of both sides, we find that

$$\langle f(t + \delta t) \rangle = \langle f(t) \rangle + s \left[\langle f(t) \rangle - \langle f(t)^2 \rangle \right] \delta t$$
$$+ \mu(1 - \langle f(t) \rangle)\delta t - \nu \langle f(t) \rangle \delta t \qquad (6.24)$$

or in differential form,

$$\frac{\partial \langle f \rangle}{\partial t} = s[\langle f \rangle - \langle f^2 \rangle] + \mu(1 - \langle f \rangle) - \nu \langle f \rangle \,. \qquad (6.25)$$

Note that the selection term contains a contribution from the *second* moment $\langle f(t)^2 \rangle$, rather than the first moment squared. Thus, in contrast to the example in the previous section, we will generally need to know the behavior of the second moment $\langle f(t)^2 \rangle$ to calculate the dynamics of the mean frequency $\langle f(t) \rangle$.

We can repeat the same procedure to derive a corresponding equation for the second moment $\langle f(t)^2 \rangle$. In this case we find that

$$\frac{\partial \langle f^2 \rangle}{\partial t} = \underbrace{2s \cdot \langle f \cdot f(1 - f) \rangle}_{\substack{\text{from deterministic} \\ \text{part (chain rule)}}} + \underbrace{\frac{f(1 - f)}{N}}_{\substack{\text{collision of correlated} \\ \text{stochastic terms}}} + \dots \qquad (6.26)$$

where we have omitted the contributions from the mutation terms. This equation depends on the first and second moments of $f(t)$, but it also has a contribution from the third moment $\langle f(t)^3 \rangle$. An analogous calculation shows that the third moment depends on the fourth moment, and so on. In this case, we say that the ***the moment hierarchy does not close*** (also known as ***"moment hell"***). This effect will always occur when we are dealing with nonlinear SDEs.

Since the non-linearities are ultimately caused by the selection term in Eq. (6.1), one approach is to restrict our attention to scenarios where natural selection is absent ($s = 0$). In this **neutral limit**, the moment equations do close, and one can derive exact solutions for the dynamics of the mean and variance of the mutation trajectory. (You will carry out this calculation yourself in Problem 3 of Homework 2). Much of the classical population genetics literature has focused this limit, and a sophisticated set of mathematical tools (collectively known as the **neutral theory**) have been developed for analyzing this case. We will revisit this topic later in the course when we talk about multi-site genomes. For the rest of this chapter, we will continue to consider cases where natural selection is present.

6.4   Stationary distribution

We can also try to derive an analogue of the stationary distribution in Eq. (6.20). The Fokker-Planck equation for the single-locus model in Eq. (6.1) is given by

$$\frac{\partial p(f,t)}{\partial t} = -\frac{\partial}{\partial f}\left[(sf(1-f) + \mu(1-f) - \nu f)\, p(f,t)\right] \\ + \frac{\partial^2}{\partial f^2}\left[\frac{f(1-f)}{2N}\cdot p(f,t)\right] \tag{6.27}$$

In this case, one can show[2] that the stationary distribution ($\partial_t p \approx 0$) can be written in the general form,

$$p(f) \propto f^{-1}(1-f)^{-1}e^{-2N\Lambda(f)}\,, \tag{6.28}$$

where the function $\Lambda(f)$ is defined by

$$\Lambda(f) = -\left[sf + \mu \log f + \nu \log(1-f)\right] \tag{6.29}$$

---

[2]For example, one can verify that this is a solution by substituting it into Eq. (6.27). A more constructive derivation is given in the Appendix at the end of this chapter.

Equation (6.28) can be viewed as the frequency-space analogue of the Boltzmann distribution in Eq. (6.20). The function $\Lambda(f)$ plays the role of the potential function $V(x)$. This analogy goes beyond the Boltzmann-like form of Eq. (6.28): the $\Lambda(f)$ function has the special property that

$$\frac{\partial \Lambda(f)}{\partial f} = -\frac{1}{f(1-f)} \left(\frac{\partial f}{\partial t} \right)_{\text{det}}, \qquad (6.30)$$

where $(\partial f / \partial t)_{\text{det}}$ is the deterministic portion of the SDE in Eq. (6.27). Thus, in the absence of noise, the deterministic dynamics of $f(t)$ will act to decrease the value of $\Lambda(f)$:

$$\left(\frac{d\Lambda(f(t))}{dt} \right)_{\text{det}} = \frac{\partial \Lambda}{\partial f} \left(\frac{\partial f}{\partial t} \right)_{\text{det}} = -\frac{1}{f(1-f)} \left(\frac{\partial f}{\partial t} \right)_{\text{det}}^2 \leq 0 \qquad (6.31)$$

similar to the classical minimization of the potential energy. Conversely, the inverse population size $1/2N$ constitutes the analogue of the diffusion coefficient (or temperature) in Eq. (6.20), which parameterizes the overall strength of noise. Large values of N will tend to amplify the importance of the energy function $\Lambda(f)$, while small values of N will make it less important. For the specific form of $\Lambda(f)$ in Eq. (6.29), the stationary distribution reduces to:
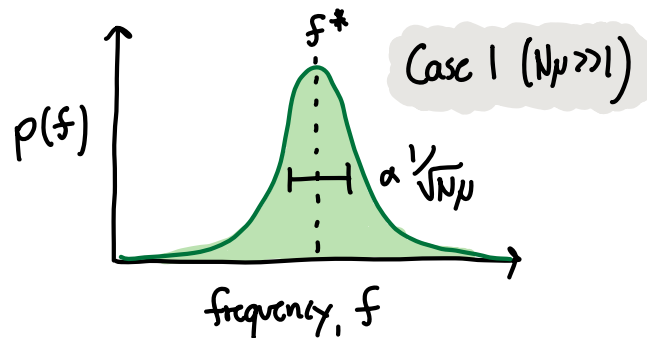
> **Stationary distribution for a single genetic locus**
> (*mutation-selection-drift balance*)
>
> $$p(f) \propto f^{2N\mu-1}(1-f)^{2N\nu-1}e^{2Nsf} \qquad (6.32)$$

This is a celebrated result known as *mutation-selection-drift-balance* (originally derived by Sewall Wright in the 1930's). It describes the long-term distribution of the mutation's frequency under the joint action of natural selection, mutation, and genetic drift. The shape of this distribution will strongly depend on the values of the compound parameters $N\mu$ and $N\nu$. These are sometimes

known as the *population scaled mutation rates* or *mutation supply rates*; they have a natural interpretation as the number of new mutations produced by the population each generation (similar to our Fermi calculation in Chapter 3).

Case 1 ("fuzzy noise"). When $N\mu, N\nu \gg 1$, the stationary distribution in Eq. (6.32) becomes strongly peaked around a characteristic frequency f^*, which minimizes the potential energy function $\Lambda(f)$.

This minimum occurs when

$$\frac{\partial \Lambda}{\partial f} = f^{-1}(1-f)^{-1}\left(\frac{\partial f}{\partial t}\right)_{\text{det}} = 0 \tag{6.33}$$

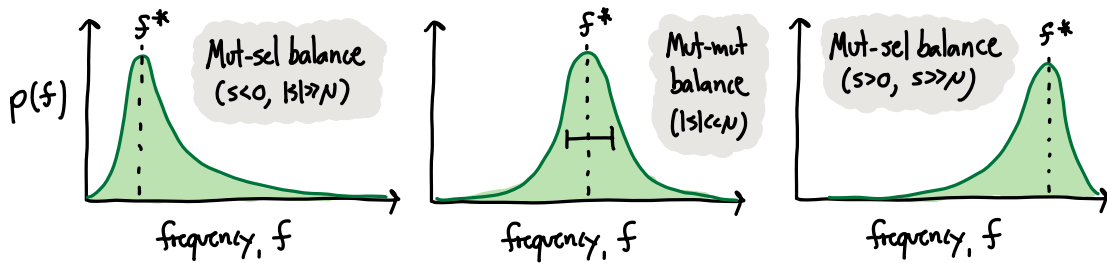which implies that it is an equilibrium solution of the deterministic dynamics,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{det}} = \underbrace{sf^*(1-f^*)}_{\text{selection}} + \underbrace{\mu(1-f^*) - \nu f^*}_{\text{mutation}} = 0 \tag{6.34}$$

This point is often known as *deterministic mutation-selection balance* (or just *deterministic mutation balance* if $s = 0$), since it is independent of genetic drift. We can use the quadratic formula to solve for f^* as a function of s, μ, and ν. However, it can be more instructive to examine its leading-order behavior in different asymptotic limits. Using the dominant balance methods

in Chapter 2, one can show that

$$f^* \approx \begin{cases} \frac{\mu}{|s|} & \text{if } s < 0 \text{ and } |s| \gg \mu, \nu, \\ \frac{\mu}{\mu+\nu} & \text{if } |s| \ll \mu, \nu, \\ 1 - \frac{\nu}{s} & \text{if } s > 0 \text{ and } s \gg \mu, \nu. \end{cases} \tag{6.35}$$

Each of these cases has a simple heuristic interpretation:

The first case represents a balance between forward mutations and negative selection ($s < 0$), which pins the mutant strain at a small but nonzero frequency. This equilibrium frequency is larger for higher mutation rates, and smaller for larger fitness costs. The second case represents a balance between forward and back mutations, leading to a mixture of mutant and wildtype at intermrediate frequencies. The last case represents a balance between positive selection and back mutation, which pins the *wildtype* population at a small but nonzero frequency. Stronger positive selection pushes the mutant closer to fixation, but while higher rates of back mutation push it down.

If we expand around f^*, the stationary distribution again approaches a Gaussian form,
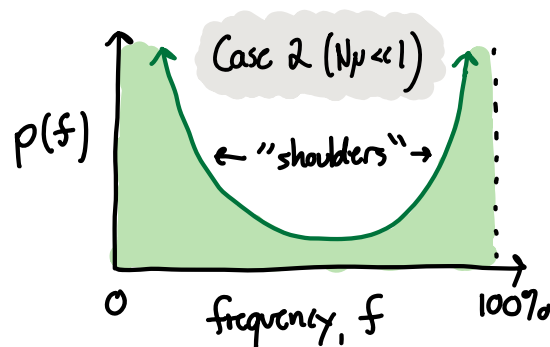
$$p(f) \propto f^{*-1}(1 - f^*)^{-1} e^{-2N\Lambda(f^*) - N\Lambda''(f^*)(f-f^*)^2} \tag{6.36}$$

which has a variance equal to

$$\sigma_f^2 = \frac{1}{2N\Lambda''(f^*)} = \frac{[f^*(1 - f^*)]^2}{2N\left[\mu(1 - f^*)^2 + \nu f^{*2}\right]} \tag{6.37}$$

99

Our Gaussian expansion assumed that this spread is small compared to the peak frequency f^*. Plugging in our expression for f^* in Eq. (6.35), we see that this will be true when $N\mu \gg 1$ and $N\nu \gg 1$, which is exactly the regime we are considering. In this case, we can think of genetic drift as inducing a small amount of spread around the deterministic expectation for $f(t)$ ("fuzzy noise" according to our classification system in Chapter 2).

Case 2 ("jagged noise"). In the opposite case, where $N\mu, N\nu \ll 1$, the $f^{-1}(1-f)^{-1}$ prefactor in Eq. (6.32) causes the stationary distribution to adopt a *U-shaped form* that is qualitatively different from the case above.

This U-shaped distribution has a pair of peaks at $f = 0$ and $f = 1$, with *"shoulders"* that fall of as $\sim 1/f$ and $\sim 1/(1 - f)$ respectively. The relative heights of the shoulders are (roughly) controlled by the exponential factor e^{2Nsf}, which increases by a factor of e^{2Ns} between $f \approx 0$ and $f \approx 1$. It is clear that these U-shaped distributions cannot be interpreted as a small perturbation away from the deterministic expectation above — even when the population size is very large. As long as the corresponding mutation rates are small enough that $N\mu < 1$, then genetic drift will always play a non-negligible role. Interestingly, we will see in later chapters that empirical mutation frequency distributions (e.g. in humans) are often of the U-shaped variety, suggesting that this will be an important case to consider when we are interested in making connections to data.

6.5 Extinction and fixation probabilities

While the stationary distributions were more analytically tractable than the moment equations in Section 6.3, they raise several natural questions that are difficult to answer from these formal solutions alone. What's going on "behind the scenes" to generate the stationary distributions above? i.e. what do the individual mutation trajectories look like? What do the the "shoulders" of the U-shaped distribution correspond to? And perhaps most importantly, how long does it take for the population to reach this long-term steady-state? Will it ever be relevant in practice (e.g. when making comparisons with data)?

We can start to gain some insight into these questions by considering a final stationary distribution scenario. If we consider a single-locus model in the absence of new mutations ($\mu = \nu = 0$),

$$\frac{\partial f}{\partial t} = \underbrace{sf(1-f)}_{\text{selection}} + \underbrace{\sqrt{\frac{f(1-f)}{N}}\eta(t)}_{\text{genetic drift}} \tag{6.38}$$
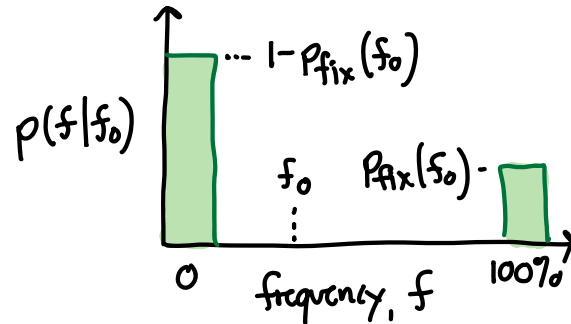
then at long times we know that the mutation must either fix or go extinct. The stationary distribution will therefore consist of two spikes at $f = 0$ and $f = 1$:

$$p(f|f_0) = p_{\text{fix}}(f_0) \cdot \delta(f-1) + (1 - p_{\text{fix}}(f_0)) \cdot \delta(f) \tag{6.39}$$

whose relative heights are specified by a single number

$$p_{\text{fix}}(f_0) = \Pr[f(\infty) = 1 | f(0) = f_0] \tag{6.40}$$

that depends on the initial frequency f_0.

This number can be interpreted as the ***fixation probability*** of a mutation that started at an initial frequency f_0. The dependence on the initial condition is critical — in the absence of further mutations, a mutant that started at 100% frequency will remain there 100% of the time, while a mutant with $f_0 = 0$ is guaranteed to stay extinct. Due to this dependence on the initial condition, the stationary distribution in Eq. (6.39) is capturing a fundamentally out-of-equilibrium process, even though it is phrased in terms of a long-term equilibrium measurement.

The Fokker-Planck equation is not very useful for discrete distributions like Eq. (6.39), since the derivatives with respect to f are not really well-defined. However, the generating function,

$$H(z, t) \equiv \langle e^{-zf(t)} \rangle = \int e^{-zf} p(f, t | f_0) \, df \,, \qquad (6.41)$$

is still very useful in this case. Since the generating function is just a moment of $f(t)$, we can derive a (deterministic) equation for $H(z, t)$ using the same approach we used for $\langle f(t) \rangle$ and $\langle f(t)^2 \rangle$ in Section 6.3.

Deriving an equation of motion for $H(z, t)$:

Using the infinitesimal version of Eq. (6.38),

$$f(t + \delta t) = f(t) + sf(t)[1 - f(t)]\delta t + \sqrt{\frac{f(t)[1 - f(t)\delta t]}{N}} Z_t \qquad (6.42)$$

102

we can express the generating function in the next timestep as

$$H(z, t + \delta t) \equiv \left\langle e^{-zf(t+\delta t)} \right\rangle$$

$$= \left\langle e^{-z\left[f(t)+sf(t)[1-f(t)]\delta t + \sqrt{\frac{f(t)[1-f(t)]\delta t}{N}} Z_t\right]} \right\rangle \qquad (6.43)$$

Expanding the right hand side in powers of δt, we obtain

$$H(z, t + \delta t) \approx \left\langle e^{-zf(t)} \left[1 - zsf(t)[1 - f(t)]\delta t \right. \right.$$

$$\left. \left. - \sqrt{\frac{f(t)[1 - f(t)]\delta t}{N}} Z_t + \frac{z^2 f(1 - f) Z_t^2}{2N} \delta t \right] \right\rangle \qquad (6.44)$$

$$H(z, t + \delta t) \approx \left\langle e^{-zf(t)} \left[1 - zsf(t)[1 - f(t)]\delta t \right. \right.$$

$$\left. \left. - \sqrt{\frac{f(t)[1 - f(t)]\delta t}{N}} Z_t + \frac{z^2 f(1 - f) Z_t^2}{2N} \delta t \right] \right\rangle$$

$$\approx \langle e^{-zf(t)} \rangle - \left(zs - \frac{z^2}{2N} \right) \langle f(t)[1 - f(t)]e^{-zf(t)} \rangle \delta t \qquad (6.45)$$

where we have exploited the linearity of the expectation and the fact that Z_t is an independent random variable with $\langle Z_t \rangle = 0$ and $\langle Z_t^2 \rangle = 1$. Note that the second term can be rewritten in terms of the generating function by taking derivatives with respect to the (deterministic) variable z:

$$\langle f(1 - f)e^{-zf} \rangle = - \left\langle \left(\frac{\partial}{\partial z} + \frac{\partial^2}{\partial z^2} \right) e^{-zf} \right\rangle = - \left(\frac{\partial H}{\partial z} + \frac{\partial^2 H}{\partial z^2} \right) \qquad (6.46)$$

Substituting this result into Eq. (6.45), we obtain a partial differential equation for the moment generating function,

$$\frac{\partial H(z, t)}{\partial t} = \left[sz - \frac{z^2}{2N} \right] \left[\frac{\partial H(z, t)}{\partial z} + \frac{\partial^2 H(z, t)}{\partial z^2} \right] . \qquad (6.47)$$
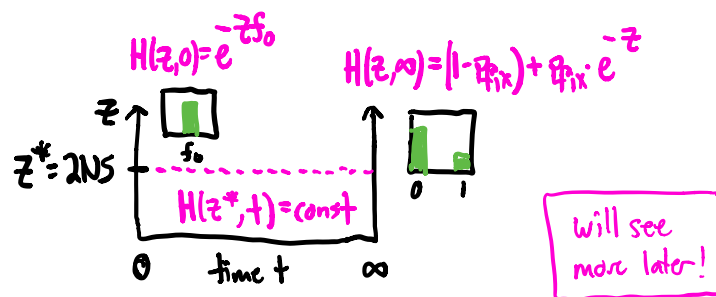
We could have also derived this equation by taking the Laplace transform of the Fokker-Planck equation in Eq. (5.49), emphasizing that these two formulations are equivalent representations of $p(f, t)$. We can see any discreteness in the probability density will still produce a continuous dependence on z, so Eq. (6.47) is still well-defined even for spiky distributions like Eq. (6.39).

Back to the fixation probability

The PDE in Eq. (6.47) are still difficult to solve in the general case. However, there is a particular value of the dummy variable z where the dynamics greatly simplify. We note that if we take $z^* = 2Ns$, then the prefactor on the right hand side of Eq. (6.47) vanishes, so that

$$\frac{\partial H(z^*, t)}{\partial t} = 0 \implies H(z^*, t) = \text{const} \tag{6.48}$$

This allows us to connect the value of $H(z^*, t)$ at the initial and final timepoints:

If the mutation frequency starts at $f = f_0$, then

$$H(z^*, 0) = \langle e^{-z^* f_0} \rangle = e^{-z^* f_0} \tag{6.49}$$

while the stationary distribution (Eq. 6.39) yields

$$H(z^*, \infty) = \langle e^{z^* f(\infty)} \rangle = p_{\text{fix}} e^{-z^* \cdot 1} + (1 - p_{\text{fix}}) e^{-z^* \cdot 0} \tag{6.50}$$

Equating these two expressions yields a formula for the fixation probability of the mutant lineage as a function of the population size N, the fitness advantage (or disadvantage) s, and the initial frequency f_0:

Fixation probability (Kimura formula)

$$p_{\text{fix}}(N, s, f_0) = \frac{1 - e^{-2Nsf_0}}{1 - e^{-2Ns}} \qquad (6.51)$$

This is a celebrated result that is sometimes known as the **_"Kimura formula"_**. We can see that the population size N and the selection strength s always enter through their product Ns (often known as the **_scaled selection strength_**). This signals that fixation emerges from a battle between the forces of natural selection and genetic drift. There will be two characteristic regimes that depend on the overall magnitude of Ns:

Weak-selection regime. When $N|s| \ll 1$, the fixation probability reduces to the neutral limit,

$$p_{\text{fix}} \approx f_0 \,, \qquad (6.52)$$

where mutations fix in proportion to their initial frequencies. This tells us that the neutral behavior from Chapter 4 applies not just in the extreme case where $s = 0$, but rather for any value of $|s| \ll 1/N$. In this weak selection regime, genetic drift will always dominate over natural selection.

Strong-selection regime. In the opposite case, where $N|s| \gg 1$, the fate of the mutation will strongly depend on the *sign* of s and the initial frequency f_0. For beneficial mutations ($s > 0$), Eq. (6.51) reduces to

$$p_{\text{fix}} \approx 1 - e^{-2Nsf_0} \qquad (6.53)$$

This equation exhibits two characteristic regimes that depend on the compound parameter Nsf_0:

$$p_{\text{fix}} \approx \begin{cases} 1 & \text{if } Nsf_0 \gg 1, \\ 2Nsf_0 & \text{if } Nsf_0 \ll 1. \end{cases} \tag{6.54}$$

In the first case ($Nsf_0 \gg 1$), the mutation will fix nearly 100% of the time. This matches our intuitive picture of natural selection, which acts to drive beneficial variants to fixation. We can therefore think of this case as one where natural selection dominates over genetic drift.

In the opposite case ($Nsf_0 \ll 1$), the fixation probability becomes much smaller than one, and the mutant will go extinct most of the time. This behavior is most extreme in the case of a new mutation, where $f_0 = 1/N$. Substituting this value into Eq. (6.53) leads the simple formula for the fixation probability of a new mutation:

Fixation probability of a new beneficial mutation

For a strongly beneficial mutation ($Ns \gg 1$) with an initial frequency $f_0 = 1/N$, the fixation probability reduces to

$$p_{\text{fix}} \approx 2s \,, \tag{6.55}$$

which is independent of the population size N. This is a celebrated result known as *Haldane's formula*.

We can gain some intuition for this result by plugging in some concrete numbers: a new mutation with a fitness benefit of $s \approx 0.01$ in a population of size $N = 10^6$ would be predicted to fix only 2% of the time. In the other 98% of cases, the mutation will drift to extinction. Thus, genetic drift has a profound effect on the long-term fate of the mutation, even when Ns is much larger than one. At

the same time, we have seen from our simulations in Chapter 4 that the *same mutation* mixed at a 50-50 ratio will rapidly and consistently take over. How can we explain these divergent outcomes?

Since the strength of genetic drift is proportional to $1/N$, we might naively expect that for large N, the genetic drift term would eventually constitute only a small perturbation to Eq. (6.38). This would suggest that the dominant balance is given by the deterministic terms,
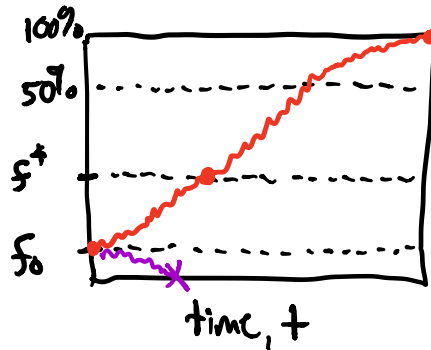
$$\frac{\partial f}{\partial t} \approx s f (1 - f) \,, \tag{6.56}$$

which produces the logistic trajectory,

$$f(t) = \frac{f_0 e^{st}}{1 - f_0 + f_0 e^{st}} \,. \tag{6.57}$$

However, this deterministic expression predicts that a beneficial mutation will always approach fixation, regardless of its initial frequency f_0. This suggests that something goes wrong with our dominant balance argument, at least when $N s f_0 \ll 1$. This behavior is sometimes known as a ***singular limit*** — it refers to a scenario where arbitrarily small values of $\epsilon = 1/N$ produce *qualitatively* different behavior than we observe if we directly set $\epsilon = 0$. We will see that such singular limits will often arise in evolutionary problems (often in a similar context related to the overall population size N). How can we understand this behavior in the present case?

One approach is to note that according to Eq. (6.53), the fate of a mutation is only uncertain once $f_0 \lesssim 1/2Ns \ll 1$. Mutations with $f_0 \gg 1/2Ns$ will nearly always fix, even when f_0 itself is very small. This suggests that it would be interesting to focus on mutations with $f_0 \ll 1/2Ns$, and then break the trajectory into two parts: an initial phase from f_0 to an intermediate frequency $f^* \ll 1$, and a second phase from $f^* \to 1$.

Since our diffusion model doesn't allow for large jumps, we know that all successful mutations must eventually pass through f^* before they are able to fix. This means that we can decompose the fixation probability into two pieces,

$$p_{\text{fix}}(f_0) = \Pr[f_0 \to f^*] \cdot p_{\text{fix}}(f^*), \tag{6.58}$$

where

$$\Pr[f_0 \to f^*] = \frac{p_{\text{fix}}(f_0)}{p_{\text{fix}}(f^*)} = \frac{2Nsf_0}{p_{\text{fix}}(f^*)} \tag{6.59}$$

is the probability that the mutation ever reaches f^* before it goes extinct. By varying the intermediate frequency f^*, we can gain some additional insight into the source of the uncertainty in the mutation's long-term probability of fixation. Given the functional form of Eq. (6.53), we expect that the behavior will strongly depend on the compound parameter Nsf^*:

- If $f^* \gg 1/2Ns$, then $p_{\text{fix}}(f^*) \approx 1$, and the probability of reaching f^* becomes

$$\Pr[f_0 \to f^*] \approx \frac{2Nsf_0}{1} \approx p_{\text{fix}}(f_0) \tag{6.60}$$
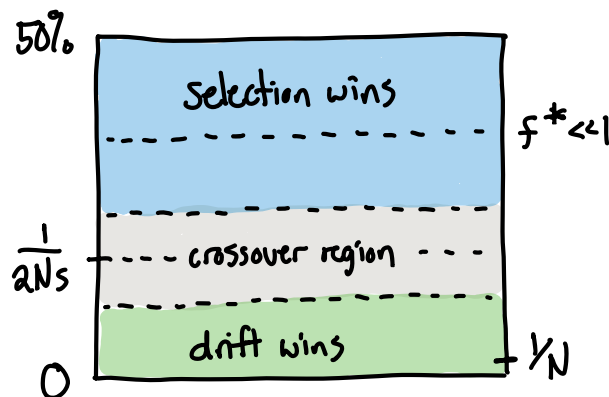
This shows that all the uncertainty in the mutation's fate plays out while it transits through frequencies less than f^*. In other words, selection always dominates over genetic drift once the mutation frequency first exceeds $f^* \gg 1/2Ns$.

- On the other hand, if $f^* \ll 1/2Ns$, then the probability of reaching f^* becomes

$$\Pr[f_0 \to f^*] \approx \frac{2Nsf_0}{2Nsf^*} = \frac{f_0}{f^*} \qquad (6.61)$$

This is independent of the selection coefficient s, and matches the probability that a neutral mutation reaches f^* before it goes extinct. This suggests that the dynamics of $f(t)$ will be similar to a neutral mutation when $f(t) \ll 1/2Ns \ll 1$.

These considerations hint at an interesting partitioning of frequency space:

There is a region where $f \ll 1/2Ns$ where genetic drift always dominates over natural selection. Conversely, there is also a region $1/2Ns \ll f \ll 1$ where natural selection dominates over genetic drift, even while the mutant itself remains rare. There is also a crossover region around $1/2Ns$ where both drift and selection are relevant. However, Eq. (6.53) shows that the width of this crossover region is relatively narrow compared to the other two regimes above — a few multiples of $1/2Ns$ are sufficient to enter the selection dominated region, while the converse is true for the drift dominated region.

This partitioning of frequency space helps explain the divergent outcomes as a function of the mutation's initial frequency. The drift-dominated region is

irrelevant for mutations that are initialized in a 50-50 mixture, so we expect it to play a major role in the competition assays that are used to measure s. However, the existence of the drift dominated region is extremely important for evolution, because all new mutations will arise at an initial frequency that is always much smaller than $1/Ns$. These mutations must always pass through the drift dominated region before natural selection is able to take over. This explains why genetic drift can never be fully neglected, even in asymptotically large populations ($Ns \gg 1$).

Next steps: Since all the interesting parts of this process take place at low frequencies, it suggests that we can obtain a complete picture of the dynamics by focusing on the limit where the frequencies are all much smaller than 1. This will lead to crucial simplifications of the single-locus model in Eq. (6.1) that will enable significant analytical progress. We will consider this case in the next chapter. We will then see how we can patch this solution onto the deterministic dynamics that take place at higher frequencies, using a technique known as *asymptotic matching*. Together, these analyses will allow us to obtain a complete picture of the trajectories of new mutations in the single-locus model in Eq. (6.1). We will then be in a position to compare these predictions to DNA sequencing data.

6.6   Appendix

6.6.1   Solving for the stationary distribution

One can verify that our solutions for the stationary distribution in Eqs. (6.20) and (6.28) are correct by substituting them into their corresponding Fokker-Planck equations. This emphases how solving differential equations really only requires a lucky guess. Nevertheless, it can sometimes be useful to see a more "constructive" derivation to motivate where these solutions come from. We present one such argument below.

Brownian particle example

For a Brownian particle in a general potential, the stationary Fokker-Planck equation can be written as

$$\frac{\partial}{\partial x}\left[\frac{V'(x)}{D}p(x) + \frac{\partial p(x)}{\partial x}\right] = 0 \tag{6.62}$$

Integrating once over x yields

$$\frac{V'(x)}{D}p(x) + \frac{\partial p(x)}{\partial x} = C_1 \tag{6.63}$$

where C_1 is a constant of integration. Multiplying both sides by the integrating factor $e^{\frac{V(x)}{D}}$ yields

$$\frac{\partial}{\partial_x}\left[p(x)e^{\frac{V(x)}{D}}\right] = C_1 e^{\frac{V(x)}{D}} \tag{6.64}$$

Integrating one more time over x yields

$$p(x) = C_2 e^{-\frac{V(x)}{D}} + C_1 e^{-\frac{V(x)}{D}}\int_0^x e^{\frac{V(x)}{D}}. \tag{6.65}$$

where C_2 is another integration constant. To ensure that $p(x)$ remains finite and nonnegative as $|x| \to \infty$, we can see that the first integration C_1 must vanish. This yields the Boltzmann equation in Eq. (6.20) above.

Single-locus model of mutation frequencies

A similar argument can be used for the single-locus model in Eq. (6.1). In this case, the stationary Fokker-Planck equation can be written as

$$0 = \frac{\partial}{\partial f}\left[2N\left(sf + \frac{\mu}{f} - \frac{\nu}{1-f}\right)\cdot f(1-f)p(f)\right.$$
$$\left. + \frac{\partial}{\partial f}\left(f(1-f)p(f)\right)\right] = 0 \tag{6.66}$$

$$\frac{\partial}{\partial f}\left[\left(s+\frac{\mu}{f}-\frac{\nu}{1-f}\right)\cdot f(1-f)p(f)+\frac{1}{2N}\frac{\partial}{\partial f}\left(f(1-f)p(f)\right)\right]=0$$

$$\tag{6.67}$$

If we define

$$\tilde{p}(f)(f)=f(1-f)p(f) \tag{6.68}$$

and

$$\Lambda(f)=-\int\left(s+\frac{\mu}{f}-\frac{\nu}{1-f}\right)df=-\left[sf+\mu\log f+\nu\log(1-f)\right]$$

$$\tag{6.69}$$

then we can rewrite this as

$$\frac{\partial}{\partial f}\left[2N\Lambda'(f)p(f)+\frac{\partial p(f)}{\partial f}\right]=0 \tag{6.70}$$

This is the same form as the example above with the replacements $p(x)\to\tilde{p}(f)(f)$, $D\to 1/2N$ and $V'(x)\to\Lambda'(f)$. A similar argument therefore yields

$$\tilde{p}(f)f=f(1-f)p(f)\propto e^{-2N\Lambda(f)}. \tag{6.71}$$

which is equivalent to the stationary distribution in Eq. (6.28) in the main text.

6.6.2   Formal solutions for the time-dependent case

This approach can be extended to obtain a formal solution for the full time-dependent Fokker-Planck equation. The basic idea is to utilize the separation of variables technique, writing a particular solution in the factorized form

$$p(f,t)=\phi(f)\cdot\psi(t) \tag{6.72}$$

Substituting this ansatz into Eq. (6.27), we see that a general solution can be written as an eigenfunction decomposition,

$$p(f,t) = \sum_{n=0}^{\infty} c_n e^{-\lambda_n t} \phi_n(f) \tag{6.73}$$

where λ_n and $\phi_n(f)$ are the solutions to the eigenvalue problem,

$$\begin{aligned} -\lambda_n \phi_n(f) = &-\frac{\partial}{\partial f} \left[\left(sf(1-f) + \mu(1-f) - \nu f \right) p(f,t) \right] \\ &+ \frac{\partial^2}{\partial f^2} \left[\frac{f(1-f)}{2N} \cdot p(f,t) \right] \end{aligned} \tag{6.74}$$

given the boundary conditions on $p(f,t)$. Kimura showed that the solutions to this eigenvalue problem can be written as solutions of the oblate spheroidal equation.[3] While some progress can be made by starting from this formal solution, we will see that significantly more intuition about the time-dependent behavior can be obtained by leveraging the branching process approximations described in Chapter 7. For this reason, the series solution in Eq. (6.73) will be of limited use for us in this course.

---

[3]The details can be found in Kimura, "Stochastic processes and distribution of gene frequencies under natural selection," *Cold Spring Harb Symp Quant Biol* (1955).