CHAPTER 15

# Continuous stochastic variables

*"...now suppose that the masses are initially at rest, and examine the manner in which they acquire velocity under the impact of the projectiles. ...the general equation, applicable not merely to the initial and final, but to all stages of the acquirement of motion. ...$df/dt = d^2 f/du^2 + 2hd(uf)/du$."*

***Rayleigh, Philos. Mag. (1891)***

Periodic boundary conditions, however elegant, are an artificial construct. Every real system is part of a larger extended system, and at the boundary of every subsystem there are inevitable interactions with the surroundings. It can be shown [1] that even the slightest random interactions with the bath suffice to create a Boltzmann distribution from a subsystem with otherwise Hamiltonian dynamics. Therefore, real systems are intrinsically stochastic even if we ignore quantum mechanics.

This chapter begins with the historically and practically important Langevin equation, the first and simplest stochastic equation of molecular motion [2]. We will examine its properties and show how it can be equivalently studied using the parallel language of Fokker-Planck equations [2,3]. We will further show how the Langevin equation behaves in the "overdamped" limit of large friction [3]. We outline how discrete stochastic models from Chapter 14 can be converted into Fokker-Planck equations when all transitions are between closely neighboring states [4]. Finally, we revisit the discussion of spectral theory, now starting from a Fokker-Planck equation.

## 15.1 Inertial Langevin dynamics

Langevin's stochastic equation of motion began as an effort to model Brownian diffusion [5]. His original equation looked somewhat different from the one that bears his name today. There was no potential of mean force (PMF), but the central new elements were already present: friction and random forces. In modern applications, the Langevin equation is used to model a diverse range of phenomena in physics, biology, chemistry, and even economics. To model the dynamics of a small system coupled to its environment the Langevin equation needs a PMF, friction, and random forces. If the small system is the scalar reaction coordinate $q$, then [2]

$$m\ddot{q} = -\frac{\partial F}{\partial q} - m\gamma\dot{q} + R(t) \tag{15.1.1}$$

where $F(q)$ is the PMF, $\gamma$ is a drag (friction) coefficient, and $R(t)$ is a randomly fluctuating force due to interactions of $q$ with its environment. For now, assume that the random force $R(t)$ is perfectly Markovian, i.e. that the random force has no memory of its own history nor of the history of $q(t)$. The random forces on average are zero,

$$\langle R(t)\rangle = 0. \tag{15.1.2}$$

The fluctuation-dissipation theorem requires an additional relationship between the random kicks, the friction, and the temperature. Intuitively there are several reasons to suspect such a relationship. First, the random kicks and dissipative forces are both associated with coupling to the bath, and the bath properties depend on the temperature. Second, friction in the absence of random kicks would eventually drain the system of all energy until it reached an energy minimum corresponding to $T = 0$ K. To reach a thermal equilibrium, the random kicks must somehow compensate for the tendency of friction to dissipate thermal energy. To balance the friction, the strength of the kicks must increase in proportion to the friction and the temperature. The random forces must also depend on the mass of the particle because equipartition requires $\langle m(dq/dt)^2\rangle = k_B T$. The specific relationship required to balance fluctuations and dissipation is

$$\langle R(0)R(t)\rangle = 2m\gamma k_B T\delta[t] \tag{15.1.3}$$

Note that equation (15.1.1) uses $-\partial F/\partial q$ instead of the more typical notation $-\partial V/\partial q$. Assuming that the bath response is fast and independent of $q$ is tantamount to assuming that the force is averaged over the rapid fluctuations in the bath. Therefore the PMF, especially for condensed phase processes, is more appropriate for the internal forces term. Accordingly, a Langevin model based on simulation data should be constructed from the free energy profile or PMF along the coordinate $q$ and not from raw instantaneous forces.[1]

Equation (15.1.3) describes an infinitely fast noise, but in a simulation, the random force cannot decorrelate any faster than the integration time step. The usual practice is to generate the random force as a Gaussian random variable with mean zero and variance $2m\gamma k_B T/\Delta t$ [6]. Upon integration over one time step, the effect is the same as an instantaneous impulse of size $2m\gamma k_B T$. The Langevin equation is then numerically solved as a pair of coupled first order equations. A crude algorithm is

[1] Differences between the PMF and the Landau free energy can be important, especially when the variable $q$ is not slow compared to other variables. For example, the PMF as a function of distance $r$ between two molecules approaches zero as $r \to \infty$, while the free energy diverges to $-\infty$ as $r \to \infty$. Because diffusion along $r$ is no slower than diffusion on the sphere of constant $r$, a thermodynamically and dynamically correct model should use the PMF and also include the "angle" variables in the spherical coordinate system.

$$
\begin{aligned}
v(t+\Delta t) &= v(t) - \gamma \Delta t\, D \frac{\partial \beta F}{\partial q} - \gamma \Delta t\, v(t) + \sqrt{2m^{-1} k_B T \gamma \Delta t} \cdot \xi \\
q(t+\Delta t) &= q(t) + v(t)\Delta t
\end{aligned}
$$

where $D = k_B T/m\gamma$ and $\xi$ is a Gaussian random number with zero mean and unit variance. In general, Langevin dynamics algorithms, like deterministic molecular dynamics algorithms, use forces at the current position to extrapolate the new positions and new velocities after a finite time step. The dimensionless time step $\gamma \Delta t$ should be small enough that the deterministic part of the force changes only slightly between time steps [7]. Langevin dynamics may be applied for a single degree of freedom (as described above), or for an assembly of particles (as described below). The Langevin dynamics algorithm below is from Allen and Tildesley [8].

**Algorithm: Langevin dynamics**

First define parameters that depend on the timestep $\Delta t$, friction $\gamma$, mass $m$, and temperature $T$:

$D = k_B T/m\gamma$

$c_0 = exp[-\gamma \Delta t]$

$c_1 = (1.0 - c_0)/(\gamma \Delta t)$

$c_2 = (1.0 - c_1)/(\gamma \Delta t)$

$\sigma_r^2 = \Delta t\, D\,[2 - (3 - 4c_0 + c_0^2)/(\gamma \Delta t)]$

$\sigma_v^2 = \gamma D(1 - c_0^2)$

$c_{rv} = D(1 - c_0)^2/(\sigma_r \sigma_v)$

Then, for each timestep,

and for each coordinate,

$r_1 = ran(0, 1)$

$r_2 = ran(0, 1)$

$g_1 = (-2lnr_1)^{1/2} cos(2\pi r_2)$

$g_2 = (-2lnr_1)^{1/2} sin(2\pi r_2)$

$g_r = \sigma_r g_1$

$g_v = \sigma_v (c_{rv} g_1 + (1 - c_{rv}^2)^{1/2} g_2)$

$f_j = -\partial V/\partial q_j$

$\Delta q_j = c_1 \Delta t\, v_j + c_2 \Delta t^2 f_j/m + g_r$

$\Delta v_j = (c_0 - 1.0) v_j + c_1 \Delta t\, f_j/m + g_v$

update positions and velocities

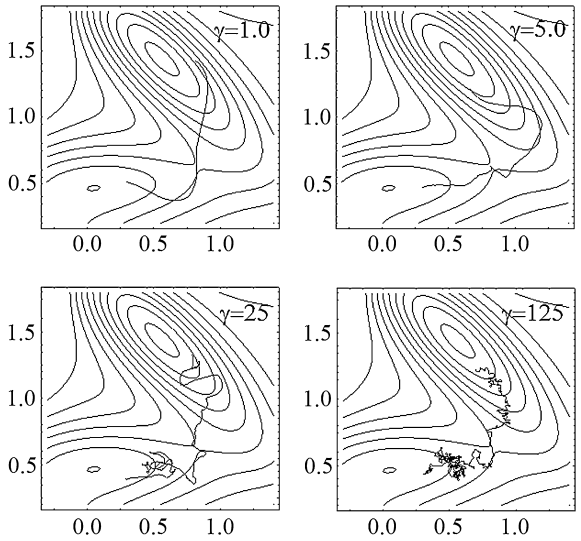
$$q_j = q_j + \Delta q_j$$
$$v_j = v_j + \Delta v_j$$

Langevin dynamics can be used in atomistic simulations to generate trajectories that sample the canonical distribution. Shirts [10] and Leimkuhler [7] examined the extent to which various Langevin dynamics algorithms and other thermostats sample the correct states when used with finite timesteps. Leimkuhler and coworkers have also developed metrics to gauge the degree to which thermostatted dynamics deviate from undisturbed energy conserving microcanonical trajectories [7]. Their analyses have identified more accurate (and still relatively simple) Langevin dynamics algorithms. Note that hydrodynamic coupling between particles is not included in simple Langevin dynamics algorithms, so the Langevin friction cannot accurately replace that from a real solvent. See Ermak and McCammon for an algorithm that includes hydrodynamic interactions [9].

We have seen that the fluctuation-dissipation theorem requires a special balance between friction and random forces. However, it does not prescribe an appropriate value of the friction constant. As the example below shows, the value chosen for the friction has a strong influence on the nature of dynamical trajectories.

### Example: Langevin dynamics on a model surface

The four panels show Langevin dynamics trajectories initiated from the saddle point of the Muller-Brown model potential energy surface at four values of the friction coefficient $\gamma$. Paths were propagated forward and backward in time until they reached a critical distance from the minimum energy locations. All other parameters are unchanged between the four cases.



As $\gamma$ increases the dynamics gradually lose their inertial characteristics and begin to resemble a random walk (diffusion) on the free energy landscape.

So what is the most appropriate Langevin friction? The Langevin thermostat provides a mechanism for heat transfer to/from the simulation box. However, dissipation by the Langevin thermostat does not occur via heat conduction or convection through the boundaries as in a natural system. Instead, dissipation occurs via a drag on every atom throughout the bulk simulation volume. Therefore the Langevin thermostat cannot precisely mimic real constant temperature processes with any choice of the friction constant. To model the real dynamics as closely as possible, some advocate using no friction at all. Deterministic dynamics are indeed necessary for testing certain dynamical properties [11,12], but recall from the opening remarks of this chapter that perfectly deterministic trajectories are also unrealistic. Most commercial and open source molecular dynamics codes recommend a Langevin friction that is independent of the simulation box size. However, the most natural choice for the dissipation rate depends on the thermal conductivity, heat capacity, and dimensions of the simulation box. For example, the friction should ideally be tuned so that the kinetic temperature fluctuations decay on time scale $\Delta t \sim \ell^2 C\rho/k$ where $k$ is the thermal conductivity, $\rho$ is the density, $C$ is the heat capacity, and $\ell$ is the box side length. This prescription ensures that dissipation of spontaneous fluctuations occurs at a rate that matches the natural rate of heat dissipation by conduction. According to this prescription, the ideal friction constant should scale as $\gamma \sim \ell^{-2}$, consistent with intuition that thermostats are unnecessary for extremely large systems.

The Langevin equation (15.1.1) assumes that the bath decorrelation is instantaneous, an assumption that limits the types of processes that it can model. In the gas phase, the frequency of kicks from the bath is the collision frequency. One might model gas phase dynamics by setting the Langevin friction so that velocity correlations decay on the collision frequency time scale. However, the relaxation towards thermal equilibrium would then occur slowly and gradually, whereas the real system would evolve microcanonically between occasional abrupt collisions [13]. Clearly, Langevin dynamics is not realistic as a kinetic theory of gases even though it would generate the proper canonical ensemble [14].

More typically, the Langevin equation is used to model condensed phase dynamics. The fastest bath time scales are set by the intermolecular distance [$\sim$Å] and the thermal root mean square velocity [$(2k_BT/m)^{1/2} \sim 100$ m/s for small molecules at room temperature]. The typical collision time in the condensed phase is therefore $\sim 10^{-12}$ s. Whether the *ps* time scale can be considered fast depends on the process being modeled. 1.0 ps is extremely fast relative to the conformational transition time of a large biomolecule, so a Langevin model would be justified. However, 1.0 ps exceeds molecular vibrational periods and relaxation times. Therefore a simple Langevin model will not accurately describe the dynamics of bond breaking and bond

formation in solution. On the other hand, chemical reaction dynamics can be modeled using a *generalized* Langevin equation in which the bath forces persist over a non-zero memory time. Generalized Langevin equations and the Grote-Hynes theory of reactions in solution are discussed in Chapter 17.

The solution to equation (15.1.1) for the special case of a constant force $\partial F/\partial q$ is particularly illuminating. Let us anticipate the emergence of a drift velocity[2]

$$v_D = -D\frac{\partial \beta F}{\partial q} \tag{15.1.4}$$

where the diffusion constant is that given by Einstein [15]

$$D = \frac{k_B T}{m\gamma} \tag{15.1.5}$$

Note that for constant $\partial F/\partial q$, there is no $q$-dependence in the Langevin equation. Therefore, in this special case, it can be written as a first order equation for the velocity

$$\dot{v} = -\gamma(v - v_D) + R(t)/m$$

where $v = \dot{q}$. The equation can be solved using the usual techniques for first order linear differential equations. The solution is

$$v(t) = v_0 e^{-\gamma t} + v_D(1 - e^{-\gamma t}) + \int_0^t dt' e^{-\gamma(t-t')} R(t')/m \tag{15.1.6}$$

Averaging over realizations of the random noise gives

$$\langle v(t)\rangle_R = v_D + (v_0 - v_D)e^{-\gamma t}$$

where the subscript $R$ indicates an average over realizations of the random force, but not over the initial velocity. The velocity never actually settles to a constant average, but in the long time limit it fluctuates around $v_D$.

What is the mean squared displacement as a function of time? Recall that we must account for the non-zero limiting drift velocity. Starting from $\delta q(t) = \int_0^t v(t')dt'$, using equation (15.1.6), gives

$$\delta q(t) = v_D t + (v_0 - v_D)\gamma^{-1}(1 - e^{-\gamma t}) + \int_0^t dt' \int_0^{t'} dt'' e^{-\gamma(t'-t'')} R(t'')/m$$

[2] When $D$ depends on $q$, the drift velocity will become $v_D = -D_q \partial \beta F/\partial q + \partial D_q/\partial q$.

Now to obtain the mean squared displacement as a function of time, $\delta q(t)$ must be squared and averaged over both the noise $R(t)$ and over the initial conditions $v_0$. The $\langle\cdot\rangle_R$ average can be evaluated using the random force properties: $\langle R(t)\rangle = 0$ and $\langle R(0)R(t)\rangle = 2m\gamma k_B T\delta[t]$. Several terms vanish because they linear in $R(t)$ and/or in $v_0$. However, a quadruple integral involving $R(t'')R(\tau'')$ must be completed. Fortunately, the same integral was done by Uhlenbeck and Ornstein in their analysis of the special $v_D = 0$ case [16]. Their noise averaged result, generalized to the case of a constant but non-zero force, is

$$\begin{aligned}\left\langle (\delta q(t))^2 \right\rangle_R &= \left\{ v_D t + (v_0 - v_D)\gamma^{-1}(1 - e^{-\gamma t}) \right\}^2 \\ &\quad + \frac{k_B T}{m\gamma^2}\left\{ 2\gamma t - 3 + 4e^{-\gamma t} - 2e^{-2\gamma t} \right\}\end{aligned}$$

which has not yet been averaged over initial velocities. The first term arises from initial velocities and deterministic forces, while the second term arises from random forces. Regardless of the initial velocity, the drift velocity will dominate the mean squared displacement at long times. For the special case where the drift velocity is zero, the average over $v_0$ gives [16]

$$\left\langle (\delta q)^2 \right\rangle = \frac{2k_B T}{m\gamma^2}\left(\gamma t - 1 + e^{-\gamma t}\right) \tag{15.1.7}$$

Thus in the absence of a potential energy gradient, the Langevin equation predicts ballistic (inertial) motion for times $t < \gamma^{-1}$ and diffusion with $D = k_B T/m\gamma$ for longer times. (See Figure 15.1.1.)

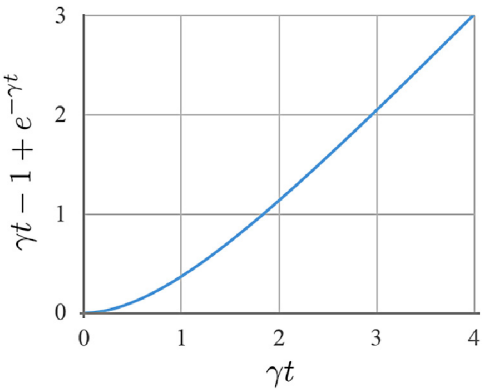



**Figure 15.1.1: The Langevin equation gives ballistic motion for $\gamma t < 1$ and the linear relationship between mean squared displacement and time is recovered for $\gamma t > 1$.**

## 15.2 Overdamped Langevin dynamics

For a flat potential energy surface, the Langevin equation *with any non-zero friction* gives trajectories that resemble Brownian motion when observed over times $t \gg \gamma^{-1}$. The magnitude

of the friction only changes the effective size of the steps in the random walk, i.e. friction controls the time scale and the ballistic distance traveled during velocity-velocity decorrelation.[3] The typical distance that is traveled in a ballistic fashion before the system 'forgets' its initial velocity is

$$(\delta q)_{bal.} = \gamma^{-1}\sqrt{k_B T/m}. \tag{15.2.1}$$

As friction increases, the ballistic distance shrinks.

For a PMF with hills and valleys, what happens when the ballistic distance becomes too short to appreciably change the force? This question is particularly important in reaction rate theory. According to transition state theory, ballistic motion along the reaction coordinate should persist from the moment of barrier crossing to a point far enough down from the barrier top to make the recrossing probability negligible. At the other extreme are systems in which typical reactive trajectories cross the dividing surface many times as they diffuse over the barrier top. These limiting dynamical regimes are both addressed by the Kramers theory (see Chapter 16). The important point for our present discussion is that the relative sizes of $(\delta q)_{bal.}$ and length scales like the width of the barrier top determine whether we should think of the dynamics as inertial or diffusive.

Now let us develop quantitative guidelines for omitting the inertial term in the Langevin dynamics. Displacement by a distance $(\delta q)_{bal.}$ changes the (deterministic) forces by an amount $|(\delta q)_{bal.}\partial^2 F/\partial q^2|$. Meanwhile the same displacement sees the friction forces change by an approximate amount $m\gamma\sqrt{k_B T/m}$. The deterministic acceleration term in the Langevin equation can be omitted when, for all locations $q$, the friction forces change much more than the systematic PMF forces do, i.e. when $m\gamma\sqrt{k_B T/m} \gg |(\delta q)_{bal.}\partial^2 F/\partial q^2|$. After simplification, the criterion becomes

$$\text{if} \quad \frac{m\gamma^2}{|\partial^2 F/\partial q^2|} \gg 1, \quad \text{the dynamics are overdamped.}$$

For example, in a harmonic vibrational well with free energy $F(q) = \frac{1}{2}m\omega^2 q^2$, the inertial Langevin equation can be replaced with an overdamped Langevin equation if $\gamma^2 \gg \omega^2$. The overdamped oscillator will not exhibit any oscillatory behavior, whereas trajectories in the inertial regime will exhibit oscillations that gradually dephase.

As friction increases, the typical magnitude of the acceleration term $m\ddot{q}$ becomes quite large (not small as sometimes argued). However, the sign of the acceleration is rapidly changing because of the strong random forces and friction so that all inertial character is damped out. To model this overdamped motion, we simply omit the acceleration term from the Langevin

[3] From equation (15.1.6), a flat PMF gives $\langle v(t)v(0)\rangle = \left\langle v(0)^2 \right\rangle e^{-\gamma t}$.

equation. We *cannot* eliminate the random forces because all trajectories from the equation $dq/dt = -D\partial\beta F/\partial q$ would deterministically evolve toward minima. In the high friction limit, the overdamped Langevin equation takes the form

$$\frac{dq}{dt} = -D\frac{\partial \beta F}{\partial q} + \widetilde{R}(t) \tag{15.2.2}$$

where $\langle \widetilde{R} \rangle = 0$ and

$$\langle \widetilde{R}(0)\widetilde{R}(t) \rangle = 2D\delta[t] \tag{15.2.3}$$

Note that the random forces have been renormalized to $\widetilde{R}(t) = R(t)/\gamma m$ and that the separate $m$, $\gamma$, and $k_B T$ parameters are now collapsed into the one parameter $D = k_B T/m\gamma$. The mass/inertia now appears only in the diffusivity.

Up to this point, our discussion has focused entirely on cases with coordinate independent friction and diffusion. These textbook situations are rarely encountered in practice. What is a coordinate dependent diffusivity? Consider for example, models of nucleation where the nucleus size $n$ is the stochastic variable of interest. The diffusivity along coordinate $n$, i.e. $D_n$, is the frequency of monomer attachment and detachment events. It often scales with nucleus surface area, so that $D_n \sim n^{2/3}$. Coordinate dependent diffusion is also common in biomolecular conformational transitions [17,18]. Coordinate dependent diffusivities and random forces lead to versions of equations (15.2.2) and (15.2.3) where $D$ and $\widetilde{R}(t)$ are replaced by $D_q$ and $\widetilde{R}_q(t)$. The Euler-Maruyama algorithm (below) for simulating overdamped dynamics is applicable to coordinate dependent diffusion.

**Algorithm: Euler-Maruyama integration**

Overdamped trajectories are generated by the extremely simple algorithm [19]

$$q(t + \Delta t) = q(t) - D_q \frac{\partial \beta F}{\partial q}\Delta t + \sqrt{2D_q \Delta t} \cdot \xi \tag{15.2.4}$$

where at each time step $\xi$ is a Gaussian random number with zero mean and unit variance. The timesteps must be small, otherwise the force will change between steps and then the assumptions of our overdamped equation will be violated. In some cases, the forces and friction explicitly depend on time as well as position. These can also be integrated with an Euler-Maruyama algorithm: $q(t + \Delta t) = q(t) + v_D(q(t), t)\Delta t + \sqrt{2D(q(t), t)\Delta t} \cdot \xi$ where $D(q, t)$ is the effective diffusivity and $v_D(q(\tau), \tau)$ is the systematic drift velocity. An appropriate time step [20] should satisfy $\Delta t \cdot |\partial v_D/\partial q| < \varepsilon$ and $\Delta t^{1/2} \cdot |\partial\sqrt{2D(q, t)}/\partial q| < \varepsilon$ where $\varepsilon$ is an error tolerance. For accurate results the trajectories need not appear continuous when plotted, but visually continuous trajectories also require that $\Delta t \leq$

$2\varepsilon^2 D(q,t)/v_D^2$ [20]. See Leimkuhler and Matthews for a similarly simple algorithm that has superior accuracy vs. time step properties [21].

## 15.3 Fokker-Planck equations

The Langevin equation describes a single trajectory $q(t)$ under the influence of random forces and friction. Each Langevin dynamics trajectory is different, so time averages and correlation functions must be computed from a long Langevin dynamics trajectory or from an ensemble of Langevin dynamics trajectories. Today, stochastic trajectories can be generated and numerically analyzed with little difficulty, but early pioneers of stochastic processes did not have this luxury. Instead of simulating trajectories, they [22–25] developed Fokker-Planck equations to directly model the time evolution of a probability density.

The Fokker-Planck equation is an approximation, but it is highly accurate for Markov processes that have small individual jumps [20,26–28]. In the overdamped Langevin equation jumps are infinitesimally small ($-v_D(q)dt + \sqrt{2D_q dt}\cdot\xi$), and accordingly the Fokker-Planck equation for overdamped Langevin dynamics is exact [31]. Master equations for nucleation assume that nucleus size evolves by attachment and detachment of single molecules (jumps of size one), and accordingly these master equations can be *approximately* reformulated as Fokker-Planck equations [29,30].

The Fokker-Planck equation for a single variable $q$ is

$$\frac{\partial \rho(q,t)}{\partial t} = -\frac{\partial}{\partial q}\left[v_D(q)\rho(q,t) - \frac{\partial}{\partial q}\left(D_q \rho(q,t)\right)\right] \tag{15.3.1}$$

It describes the evolution of the probability density $\rho(q,t)$ from an ensemble of stochastic trajectories initiated with density $\rho(q,0)$. The Fokker-Planck equation does not explicitly invoke random forces, but their effects are evident in the diffusion term which drives sharp initial conditions to decay toward a stationary distribution at long times. In writing a Fokker-Planck equation (or a Langevin equation) for the variable $q$ we have implicitly assumed that

(i) $q(t)$ evolves continuously or by small jumps in the state space for $q$,
(ii) $q(t)$ is dynamically separable from all other variables in the system, and
(iii) the dynamics of $q(t)$ are not only separable, but also Markovian.

Most arbitrarily selected variables will not satisfy (i), (ii), and (iii), so the Fokker-Planck equation will not describe their dynamics correctly. Variables that do fulfill requirements (i), (ii), and (iii) are very special. If an appropriate variable $q$ can be identified (see Chapter 20), then its dynamics within a complex multibody dynamical system are conveniently reduced

to equation (15.3.1). Coarse graining the dynamics down to a single variable provides a concise description of the stationary distribution and the relaxation dynamics from any initial distribution.

As we already saw for the (overdamped) Langevin equation, the coefficients in the Fokker-Planck equation can be interpreted as a drift velocity

$$\lim_{\Delta t \to 0} \frac{\langle \Delta q \rangle_q}{\Delta t} = v_D(q) \tag{15.3.2}$$

and a diffusivity

$$\lim_{\Delta t \to 0} \frac{\langle (\Delta q)^2 \rangle_q}{2\Delta t} = D_q \tag{15.3.3}$$

Because it accounts for coordinate dependent diffusion, definition 15.3.2 is more general than (15.1.4). Also note that equations (15.3.2) and (15.3.3) are quite useful simplifications. The complete transition matrix may be terribly complicated, but the drift velocity and diffusivity are all that remains of the dynamics after conversion from a master equation to a Fokker-Planck equation.

### Fokker-Planck equation from the master equation

The Fokker-Planck equation can be derived via the master equation and via the Langevin equation. Both are useful transformations in practice, so we briefly outline them here. Start from the master equation (in integral form) [25]

$$\frac{\partial \rho}{\partial t} = \int w_{q \leftarrow q+\Delta q} \rho(q+\Delta q, t) d\Delta q - \int w_{q+\Delta q \leftarrow q} \rho(q,t) d\Delta q$$

where $w_{q \leftarrow q+\Delta q}$ and $w_{q+\Delta q \leftarrow q}$ are continuous counterparts of the $\mathbf{W}_{i \leftarrow j}$ rate matrix elements. The two integrals in this master equation represent flow into and flow out of state $q$. The transition rates are written in terms of $q$ and $\Delta q$ (as opposed to $w_{q' \leftarrow q}$) to exploit property (i): if only small jumps in $q$ are allowed, then $w_{q \leftarrow q+\Delta q}$ and $w_{q+\Delta q \leftarrow q}$ should have compact support near $\Delta q = 0$. By comparison, $w_{q \leftarrow q+\Delta q}$, $w_{q+\Delta q \leftarrow q}$, and $\rho(q,t)$ must vary more gradually with changes in $q$. Let us Taylor expand the $q$-dependence in the "flow into $q$" integral of the master equation:

$$\begin{aligned}
&\int w_{q \leftarrow q+\Delta q} \rho(q+\Delta q, t) d\Delta q \\
&\quad = \int w_{q-\Delta q \leftarrow q} \rho(q,t) d\Delta q - \int \Delta q \frac{\partial}{\partial q} [w_{q-\Delta q \leftarrow q} \rho(q,t)] d\Delta q \\
&\qquad + \frac{1}{2} \int (\Delta q)^2 \frac{\partial^2}{\partial q^2} [w_{q-\Delta q \leftarrow q} \rho(q,t)] d\Delta q + \cdots
\end{aligned}$$

The zeroth order integral in the Taylor expansion is exactly the same as the "flow out of $q$" integral. Therefore the master equation simplifies to

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\frac{\partial}{\partial q}\left(\rho(q,t)\int \Delta q\, w_{q-\Delta q\leftarrow q} d\Delta q\right) \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial q^2}\left(\rho(q,t)\int (\Delta q)^2\, w_{q-\Delta q\leftarrow q} d\Delta q\right) \end{aligned} \tag{15.3.4}$$

where $\rho(q,t)$ and $\partial/\partial q$ have been pulled outside of the integrals over $\Delta q$. Terms of order $(\Delta q)^3$ and higher have been neglected.

The integrals that remain in equation (15.3.4) are surely complicated, but let us consider what they mean. One is the average displacement per time from location $q$, i.e. the drift velocity

$$v_D(q) \equiv \int \Delta q\, w_{q+\Delta q\leftarrow q} d\Delta q$$

The other is the average squared displacement per time from location $q$, i.e. twice the diffusivity

$$2D_q \equiv \int (\Delta q)^2\, w_{q+\Delta q\leftarrow q} d\Delta q$$

Inserting these identities[4] for the drift velocity and diffusivity into equation (15.3.4) gives the nonlinear Fokker-Planck equation (15.3.1). If we had retained the higher order terms we would have obtained the (exact) Kramers-Moyal expansion instead of the (approximate) Fokker-Planck equation.

### *Detailed balance and the Smoluchowski equation*

In many applications to chemistry and physics, the dynamics obey a detailed balance relation. Detailed balance relates transition rates to *equilibrium* probability distributions, so a pedagogical discussion on detailed balance in *non-equilibrium* rate processes is warranted. Detailed balance does not imply equilibrium, nor stationarity, nor an absence of sinks, sources, and probability current. It is a relationship between transition rates and *hypothetical* equilibria. In some cases, the hypothetical equilibria are impossible to attain, e.g. the supersaturated conditions that drive nucleation are incompatible with a normalized equilibrium distribution of nuclei. Additionally, many rate calculations involve probability sources and sinks at boundaries that create non-equilibrium currents. Nevertheless, we can invoke detailed balance between hypothetical equilibrated microstates for all states except those whose microscopic transition rates are directly perturbed at the boundaries.

[4] Both identities require additional changes of sign before we can use them.

We have already seen that the Fokker-Planck equation no longer includes the detailed state-to-state interconversion rates. They are instead coarse grained to a simpler description of drift velocities and diffusivities at each point on the $q$-axis. So how can we still incorporate detailed balance? The Fokker-Planck equation is a continuity equation for probability,

$$\partial\rho/\partial t = -\partial j/\partial q,$$

where $j \equiv v_D(q)\rho(q,t) - \partial[D_q\rho(q,t)]/\partial q$. At equilibrium, all probability currents should vanish. Therefore detailed balance is imposed by requiring that $v_D(q)$, $D_q$, and the equilibrium distribution satisfy [28]

$$v_D(q)\rho_{eq}(q) - \frac{d}{dq}\left(D_q\rho_{eq}(q)\right) = 0 \tag{15.3.5}$$

Inserting the detailed balance relationship into the nonlinear Fokker-Planck equation (15.3.1) yields the Smoluchowski equation [3,32],

$$\frac{\partial\rho(q,t)}{\partial t} = \frac{\partial}{\partial q}\left[\rho_{eq}(q)D_q\frac{\partial}{\partial q}\left(\frac{\rho(q,t)}{\rho_{eq}(q)}\right)\right].$$

As noted in the discussion above, the hypothetical equilibrium distribution invoked in $\rho_{eq}$ need not be attainable or normalizable. Thus in some cases it is more than a convenience to write the Smoluchowski equation as

$$\frac{\partial\rho(q,t)}{\partial t} = \frac{\partial}{\partial q}\left[e^{-\beta F(q)}D_q\frac{\partial}{\partial q}\left(e^{\beta F(q)}\rho(q,t)\right)\right] \tag{15.3.6}$$

where we have used $\rho_{eq}(q) \propto exp[-\beta F(q)]$. In contrast to equation (15.3.1), note that equation (15.3.6) has $D$ *between* the derivatives as a consequence of detailed balance. The Smoluchowski equation is a natural starting point for analyses of overdamped barrier crossings (see Chapter 18).

A useful generalization of equation (15.3.6) begins with a *system* of overdamped Langevin equations and results in a multidimensional Smoluchowski equation with a vector force and a diffusion tensor:

$$\frac{\partial\rho}{\partial t} = \frac{\partial}{\partial\mathbf{q}}\cdot\left[e^{-\beta F(\mathbf{q})}\mathbf{D}(\mathbf{q})\frac{\partial}{\partial\mathbf{q}}\left(e^{\beta F(\mathbf{q})}\rho(\mathbf{q},t)\right)\right] \tag{15.3.7}$$

Many interesting phenomena can arise depending on the nature of the diffusion tensor $\mathbf{D}$. For example, a highly anisotropic $\mathbf{D}$ indicates that some coordinates have very high mobilities relative to others. In other cases, diffusion tensors with non-zero off-diagonal coupling elements indicate motions of coordinates that are correlated to each other. Chapter 18 will explore the structure of $\mathbf{D}$ and its effects on barrier crossings.

### Fokker-Planck equations from simulation data

Analyses in the physics and chemistry literature often begin with a Fokker-Planck equation based on a simple phenomenological model for the free energy surface and the diffusivity. These studies are powerful sources of insight and they often yield testable predictions. On the other hand, Fokker-Planck equations that are constructed in this way ultimately contain the physics that is built into them.

An alternative strategy is to start from a complex many-body dynamical system (an MD simulation), identify an appropriately Markovian and separable variable $q$, and then use simulation data to construct its Fokker-Planck equation. Assuming that an appropriate variable $q$ can be identified, the basic idea is to extract the local drift velocity and diffusivity from swarms of short trajectories initiated at each position $q_0$ along the $q$-axis. Several computational frameworks have implemented this strategy [33–35]. Estimators for the drift velocity and diffusivity are loosely based on equations (15.3.2) and (15.3.3), but some modifications are necessary. Simulation trajectories have a finite duration, so we cannot simply square the displacement and take the $\Delta t \to 0$ limit to estimate $D_{q_0}$ (as equation (15.3.3) might seem to suggest). We must account for the drifting mean position in the swarm of trajectories at each $q_0$. Define the drift corrected displacement:

$$\delta q(t) \equiv q(t) - \langle q(t) \rangle_{q_0}$$

where the average is over a swarm of trajectories initiated at location $q(0)$. The drift velocity is that of the moving average [33,36]

$$v_D(q_0) = \frac{d}{dt} \langle q(t) \rangle_{q_0} \tag{15.3.8}$$

and the diffusivity is computed from the same swarm of trajectories as

$$2D_{q_0} = \frac{d}{dt} \left\langle (\delta q(t))^2 \right\rangle_{q_0} \tag{15.3.9}$$

The quantities on the left in equations (15.3.8) and (15.3.9) are the numerically averaged rates at which a swarm of trajectories drifts and spreads when started at $q_0$. To construct the Fokker-Planck equation, initiate swarms of trajectories at a series of $q_0$-values. For each $q_0$, compute $d \langle q(t) \rangle_{q_0} / dt$ and $d \left\langle (\delta q(t))^2 \right\rangle_{q_0} / dt$ from the trajectory swarm data. Then solve (15.3.8) and (15.3.9) for $v_D(q)$ and $D_q$ at each of the $q_0$-values. In some cases, the data can be used to identify unknown parameters in simple models for $F(q)$ and $D_q$. For example, the appropriate variable for studies of nucleation is nucleus size $n$ and physically motivated models of the free energy and the diffusivity (attachment frequency) have the form $F(n) = -n\Delta\mu + \phi\gamma\, n^{2/3}$ and $D_n = D_1 n^{2/3}$. The seeding approach to nucleation uses short trajectory swarms to identify the unknown parameters $\phi\gamma$ and $D_1$, and thereby parameterizes the Zeldovich-Frenkel (Fokker-Planck) equation [34,35,37,38].

There are some important technical issues to consider in choosing an appropriate swarm duration. First, equations for the overdamped regime are being used, so the duration should be significantly longer than the velocity decorrelation time. Second, the duration should be short so that the force $\partial\beta F/\partial q$ does not change during the swarm evolution. In practice, it can be difficult to straddle these two opposing requirements, and one must often adopt a swarm duration that is too long to ignore changes in $\partial\beta F/\partial q$. Therefore, Hummer and Kevrekidis [33] recommend alternative versions of equations (15.3.8) and (15.3.9). Specifically, they estimate $D(\overline{q})$ and $\partial\beta F/\partial q|_{\overline{q}}$ instead of $D(q_0)$ and $\partial\beta F/\partial q|_{q0}$ where $\overline{q}$ is the time-averaged position of $q$ over the duration of the swarm.

## 15.4 From discrete models to Fokker-Planck equations

We have been considering stochastic processes with continuous state spaces, but some processes with intrinsically discrete states are more easily modeled using continuous variables. Here we show, by way of example, how models with discrete states can be transformed to models with continuous variables while retaining nearly all behaviors of the discrete model. The tools which are illustrated in this section are applicable when transitions occur only between closely neighboring states with similar properties. The following shows a familiar biased random walk example.

**Example: A biased and unbounded random walk**

A random walker takes steps to the right on a one-dimensional lattice with rate constant $k_+$ and steps to the left with rate constant $k_-$. The master equation for this process is

$$\frac{dp_n}{dt} = k_+ p_{n-1} + k_- p_{n+1} - (k_+ + k_-) p_n$$

The matrix $\mathbf{W}$ in this case is of infinite dimension. $\mathbf{W}$ would be tridiagonal, and thus easy enough to write down, but still difficult to analyze. Quite often we can instead regard $n$ as a continuous variable and introduce raising and lowering operators for functions of $n$ as

$$e^{m\partial/\partial n} f(n) = f(n) + m\frac{\partial f}{\partial n} + \frac{1}{2}m^2\frac{\partial^2 f}{\partial n^2} + \cdots = f(n+m)$$

Now writing $p_{n-1}$ and $p_{n+1}$ as expansions around $p_n$ gives

$$\begin{aligned}\frac{\partial p_n}{\partial t} &= -k_+(p_n - p_{n-1}) + k_-(p_{n+1} - p_n) \\ &= -k_+(1 - e^{-\partial/\partial n})p_n + k_-(e^{\partial/\partial n} - 1)p_n\end{aligned}$$

If changes in $p_n$ are sufficiently gradual, then the expansion of $e^{\partial/\partial n}$ can be truncated at the second order. Discussions on the limitations and validity of these truncated expansions can be found elsewhere [2,28,39]. In this case, $m$ is one and $\delta p$ between neighboring states is clearly much smaller than one. Thus using

$$1 - e^{\pm\partial/\partial n} \approx \mp\frac{\partial}{\partial n} - \frac{1}{2}\frac{\partial^2}{\partial n^2}$$

gives

$$\frac{\partial p_n}{\partial t} = (k_- - k_+)\frac{\partial p_n}{\partial n} + \frac{1}{2}(k_- + k_+)\frac{\partial^2 p_n}{\partial n^2}$$

The right side of this partial differential equation describes drift and diffusion, respectively, on the discrete lattice. Note that the differential equation which has emerged has the form of a Fokker-Planck equation.

To eliminate the lattice entirely, we can reintroduce the physical lattice spacing $\Delta q$ by the relation $q = n\Delta q$. Then $dq = \Delta q\, dn$ and $\rho(q)dq = p_n$ is the continuous probability density as a function of $q$. By correspondence with the more familiar drift-diffusion equation,

$$\frac{\partial \rho}{\partial t} = -v_D\frac{\partial \rho}{\partial q} + D\frac{\partial^2 \rho}{\partial q^2}, \tag{15.4.1}$$

we see that

$$v_D = -\Delta q \cdot (k_- - k_+) \quad \textit{and} \quad D = \frac{(\Delta q)^2}{2}(k_- + k_+) \tag{15.4.2}$$

which gives a clear interpretation for drift and diffusion in terms of discrete step sizes, directions, and frequencies.

The above example shows how transition rates in a lattice simulation can be chosen to match diffusion coefficients and drift velocities for an off-lattice model. The discrete random walk model also illustrates that Fokker-Planck equations are special types of continuum master equations that result from processes that only involve local transitions. In other words Fokker-Planck equations correspond to master equations with nearly diagonal transition matrices.

Let us briefly take stock of the different methods we have already seen for studying reaction networks. In Chapter 3 we learned to formulate, simplify, and solve the deterministic rate equations for a complex reaction network. In Chapter 14 we learned about chemical master equations and kinetic Monte Carlo simulations that capture fluctuations around the deterministic solutions and spontaneous switches between stable states. We saw that fluctuations and spontaneous transi-

tions between stable states are often important for reactions in small systems. Because each reaction event changes the species populations by only one or two molecules, all transitions in a chemical master equation are between closely neighboring states. Thus it should (usually) be possible to convert a chemical master equation to a Fokker-Planck description. The next two examples return to the chemical master equation of Schlogl to illustrate how Fokker-Planck equations can be constructed and used to understand fluctuations and multiple steady states in chemical reaction networks.

**Example: Chemical Fokker-Planck equation for the Schlogl system**

The Schlogl system is a biochemical network model with kinetics

$$\begin{array}{lcll} \varnothing & \rightarrow & \mathbf{A} & r_1 \\ \mathbf{A} & \rightarrow & \varnothing & r_{-1} = k_{-1}n \\ 3\mathbf{A} & \rightarrow & 2\mathbf{A}+\mathbf{B} & r_2 = k_2 n(n-1)(n-2) \\ 2\mathbf{A}+\mathbf{B} & \rightarrow & 3\mathbf{A} & r_{-2} = k_{-2}n(n-1) \end{array}$$

where $n$ is the number of **A** molecules in the system. Details about the birth, death, and **A**, **B** interconversion kinetics in this model are given in section 14.4. The master equation governing $n$ is

$$\begin{aligned} \frac{dp_n}{dt} = {} & r_1 p_{n-1} - r_1 p_n \\ & + k_{-1}(n+1)p_{n+1} - k_{-1}np_n \\ & + k_2(n+1)n(n-1)p_{n+1} - k_2 n(n-1)(n-2)p_n \\ & + k_{-2}(n-1)(n-2)p_{n-1} - k_{-2}n(n-1)p_n \end{aligned}$$

It is convenient to abbreviate the master equation using the definitions

$$\begin{aligned} h_1(n) &= (r_1 + k_{-2}n(n-1))p_n(t) \\ h_2(n) &= (k_{-1}n + k_2 n(n-1)(n-2))p_n(t) \end{aligned}$$

Then the master equation takes the compact form

$$\frac{dp_n}{dt} = h_2(n+1) - h_2(n) + h_1(n-1) - h_1(n)$$

Next, expand the off-diagonal terms in the master equation to second order in the variable $n$.

$$h_1(n-1) = h_1(n) - \frac{\partial h_1}{\partial n} + \frac{1}{2}\frac{\partial^2 h_1}{\partial n^2}$$

$$h_2(n+1) = h_2(n) + \frac{\partial h_2}{\partial n} + \frac{1}{2}\frac{\partial^2 h_2}{\partial n^2}$$

Insert these expressions into the master equation to obtain

$$\frac{\partial p_n}{\partial t} = -\frac{\partial}{\partial n}\left[h_1(n) - h_2(n)\right] + \frac{\partial^2}{\partial n^2}\left[\frac{h_1(n) + h_2(n)}{2}\right]$$

Technically, the transformation to a continuum stochastic partial differential equation is done. However, it is useful to write the Fokker-Planck equation in terms of the drift and diffusion rates. The drift rate is

$$\begin{aligned} v_D(n) &= [h_1(n) - h_2(n)]/p_n(t) \\ &= -n(n-1)(n-2)k_2 + n(n-1)k_{-2} - nk_{-1} + r_1 \end{aligned}$$

and the coordinate-dependent diffusion rate is

$$\begin{aligned} D(n) &= [h_1(n) + h_2(n)]/(2p_n(t)) \\ &= \frac{1}{2}\{n(n-1)(n-2)k_2 + n(n-1)k_{-2} + nk_{-1} + r_1\} \end{aligned}$$

Now the Fokker-Planck equation takes the usual form

$$\frac{\partial p_n}{\partial t} = -\frac{\partial}{\partial n}\left[v_D(n)p_n\right] + \frac{\partial^2}{\partial n^2}\left[D(n)p_n\right] \tag{15.4.3}$$

The operator approach in this section can be used to create Fokker-Plank equations for many discrete systems where the transitions are local and where the stationary probabilities vary slowly between states. Stochastic integral equations can be obtained from discrete systems with highly non-local transitions, but these require techniques beyond the scope of this book [28]. Armed with some ways of deriving Fokker-Planck equations (where they are applicable), we are ready to examine the properties of their solutions.

## 15.5 Stationary solutions of Fokker-Planck equations

Chapter 14 included a few examples of discrete master equations with steady-state distributions. Fokker-Planck equations with time-independent drift and diffusion coefficients also yield stationary distributions in many familiar contexts. Chapters 16–18 will make extensive use of non-equilibrium steady-state solutions to Fokker-Planck equations. Diffusion on a bounded free energy landscape (equation (15.3.6)) gives the equilibrium distribution at long times. Here we

show more generally that when the drift and diffusion rates are time independent, and when there are no sources or sinks, then the stationary distribution resembles an equilibrium distribution for some effective potential and density of states. We begin with the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial q}\left\{ v_D(q)\rho - \frac{\partial}{\partial q}\left(D(q)\rho\right)\right\} \tag{15.5.1}$$

where $D(q)$ is a coordinate-dependent diffusivity and $v_D(q)$ is a force-induced drift rate. Recall that the Fokker-Planck equation is a probability continuity equation, i.e. $\partial\rho/\partial\tau = -\partial j/\partial q$ where the flux $j$ is the part of equation (15.5.1) within curly brackets. Steady-state distributions $\rho_{SS}(q)$ must satisfy

$$j_{SS} = v_D(q)\rho_{SS}(q) - \frac{d}{dq}\left(D(q)\rho_{SS}(q)\right) \tag{15.5.2}$$

where $j_{SS}$ is a constant. In later chapters, several rate theories will be developed by making $j_{SS}$ a nonzero constant.

If there are no sources or sinks (even at $\pm\infty$) then $j_{SS}$ must be zero. Let $y(q) = D(q)\rho_{SS}(q)$ to obtain the separable equation $y\, v_D/D = y'$. Then integrate to obtain $ln\, y = \int (v_D/D)dq + C$. Finally, solve for $\rho_{SS}$ and choose the integration constant so that the distribution is normalized. The result is

$$\rho_{SS}(q) = Q^{-1}\frac{1}{D(q)} exp \int_{-\infty}^{q} \left[v_D(q')/D(q')\right] dq' \tag{15.5.3}$$

with

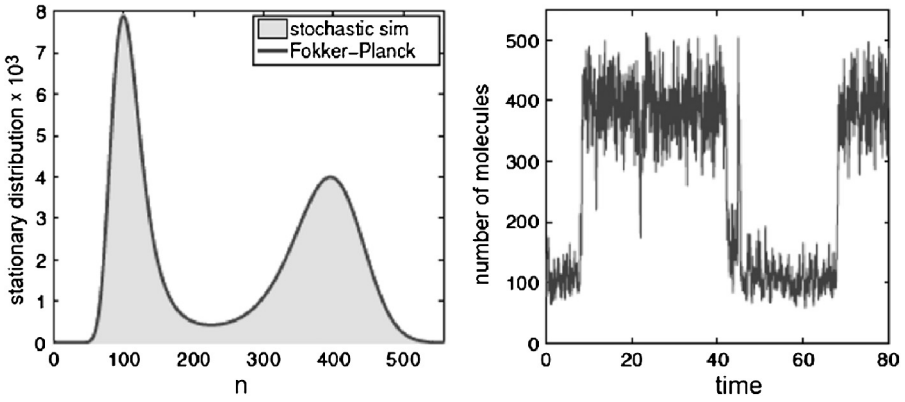
$$Q = \int_{-\infty}^{\infty} dq \frac{1}{D(q)} exp \int_{-\infty}^{q} dq' \left[v_D(q')/D(q')\right] \tag{15.5.4}$$

Note how the distribution $\rho_{SS}(q)$ resembles an equilibrium distribution even if the underlying dynamics is a driven non-equilibrium system. Specifically, equation (15.5.3) can be interpreted as an effective equilibrium distribution for a system with potential $\varphi(q) = -\int (v_D/D)dq$ and density of states $\Omega(q) = 1/D(q)$. If $v_D(q) = -D\partial\beta F/\partial q$ with constant $D$, then equation (15.5.3) yields a proper equilibrium distribution, $\rho_{SS}(q) = Q^{-1}exp[-\beta F(q)]$.

### Example: Chemical Fokker-Planck equation for Schlogl system

Again consider the Schlogl system as an example with the random variable of interest being $n$, the number of **A** molecules. Section 14.4 presented stochastic simulation (kinetic Monte Carlo) results from Erban et al. [40] with values $r_1 = 2250$, $k_{-1} = 37.5$, $k_2 = 0.18$, and $k_{-2} = 2.5\ 10^{-4}$. These parameters gave rise to two steady deterministic states: $n = 100$ or $n = 400$. From a long simulation run, Erban et al. computed the stationary probability of finding $n$ molecules of type **A** in the system.

It should also be possible to predict the probability to find $n$ molecules of type **A** from a Fokker-Planck equation without performing the kinetic Monte Carlo simulation. The Fokker-Planck equation for $p(n,t)$ was given in equation (15.4.3). The figure below, adapted from Erban et al., compares a portion of the stochastic simulation results (right) to the stationary distribution obtained from the Fokker-Planck equation.




The stationary distribution predicted from the Fokker-Planck equation perfectly matches the histogram of kinetic Monte Carlo results (gray) from a long stochastic simulation. Figures from Erban et al. *SIAM J. Appl. Math.* 70, 984–1016 (2009).

The Schlogl system has bimodal stationary distribution with infrequent transitions between two metastable states. We have seen how spectral analyses of the master equation can predict the time scale associated with these transitions. Similar spectral analyses can be performed by starting from the Fokker-Planck equation.

## 15.6 Spectral theory revisited

For ergodic systems with a bounded partition function, the Fokker-Planck equation gives rise to a discrete spectrum of eigenvalues and eigenfunctions.[5] The spectral representation is useful for developing several formal expressions. This section outlines the spectral representation for

[5] The results of this section are not limited to Fokker-Planck equations. They more generally apply to any continuous master equation

$$\frac{\partial p(q,t)}{\partial t} = \int \left\{ \omega(q|q')p(q',t) - \omega(q'|q)p(q,t) \right\} dq'$$

with transition probabilities that obey detailed balance. The Fokker-Planck is a special limiting case that arises when the $\omega(q|q')$ only have support over $|q-q'|$ distances that are small compared to the widths of peaks and

overdamped dynamics in a continuous configuration space. The discussion is terse because the derivations closely parallel the spectral rate theory for discrete master equations (Chapter 14).

Let $L$ be the operator in the Smoluchowski (forward overdamped Fokker-Planck) equation so that equation (15.3.7) becomes

$$\partial\rho/\partial t = L\rho$$

Additionally, let the operator $L$ have "right" eigenfunctions ($\psi_m^R$) and eigenvalues ($-1/\tau_m$): $L\psi_m^R(\mathbf{q}) = -\tau_m^{-1}\psi_m^R(\mathbf{q})$. The first eigenfunction is the equilibrium distribution $\psi_1^R(\mathbf{q}) \sim exp[-\beta F(\mathbf{q})]$ with eigenvalue $1/\tau_1 = 0$. The right eigenfunctions are orthogonal with weight function $1/\psi_1^R(\mathbf{q})$, i.e.

$$\int d\mathbf{q}\psi_m^R(\mathbf{q})\psi_\ell^R(\mathbf{q})/\psi_1^R(\mathbf{q}) = \delta_{m\ell}$$

which assumes an appropriate normalization of each $\psi_m^R(\mathbf{q})$. All higher eigenvalues are negative, so their contributions to the time-dependent solution will gradually decay. As we did for the discrete master equation, we can adopt bra-ket notation for the right and left eigenfunctions. Let the right eigenfunction be the ket

$$|\psi_m\rangle = \psi_m^R(\mathbf{q})$$

and let the left eigenfunction of $L$ be the bra

$$\langle\psi_m| = \psi_m^R(\mathbf{q})/\psi_1^R(\mathbf{q})$$

In bra-ket notation, nearly all formulas from section 14.6 directly extend to the case of a continuous master equation. For example, the orthogonality relation is $\langle\psi_m|\psi_\ell\rangle = \delta_{m\ell}$. Time-dependent solutions can again be expanded as $|\rho(\mathbf{q},t)\rangle = \sum_m e^{-t/\tau_m}|\psi_m(\mathbf{q})\rangle\langle\psi_m(\mathbf{q}_0)\times|\rho(\mathbf{q}_0,0)\rangle$ where the sum now runs over an infinite series of eigenfunctions. For the sharp initial condition $\rho(\mathbf{q},0) = \delta[\mathbf{q}-\mathbf{q}_0]$, the time dependent solution becomes the Greens function

$$\rho(\mathbf{q},\, t\,|\mathbf{q}_0,0) = \sum_{m=1}^{\infty} a_m e^{-t/\tau_m}|\psi_m(\mathbf{q})\rangle \tag{15.6.1}$$

with expansion coefficients

$$a_m = \langle\psi_m(\mathbf{q})|\delta[\mathbf{q}-\mathbf{q}_0]\rangle = \psi_m^R(\mathbf{q}_0)/\psi_1^R(\mathbf{q}_0)$$

minima in the equilibrium distribution [41]. We have focused on Fokker-Planck equations and Langevin equations because they are more important for the other topics in this book.

The completeness relation for the eigenfunctions of a continuous Fokker-Planck equation is

$$\sum_m |\psi_m(\mathbf{q})\rangle \langle \psi_m(\mathbf{q}')| = \delta[\mathbf{q} - \mathbf{q}']$$

Exact solutions of the Fokker-Planck equation for a high dimensional system are impossibly difficult, but new computational methods can approximate the slowest eigenfunctions directly from simulation data. The first (right) eigenfunction is the stationary equilibrium distribution. The rest of the eigenfunctions correspond to non-equilibrium relaxation processes with ever faster relaxation times. In many applications, there are one or more spectral gaps corresponding to a clear separation of time scales:

$$\cdots \tau_{k-1} > \tau_k \gg \tau_{k+1} > \tau_{k+2} \cdots$$

Spectral gaps allow an accurate low dimensional model to be constructed by simply discarding everything beyond the first $k$ eigenfunctions. The result is a simple model that accurately describes the $k$ slowest modes and their dynamics. The slowest (left) eigenfunction is particularly important because it provides a numerically exact reaction coordinate that increases monotonically as one moves through the high dimensional space from reactants to products [42–46]. The algorithm below shows how the eigenfunctions can be constructed using molecular simulation data.

### Algorithm: Diffusion map

Let $S = \{\mathbf{x}_i\}_{i=1,\ldots,n}$ be a large collection of configurations sampled from a long molecular dynamics trajectory. The trajectory should be sufficiently long to see several examples of the slowest transition. Diffusion map results depend critically on a distance threshold parameter $\varepsilon$. It should be just small enough that no configurations within a distance $\varepsilon$ are separated by a slow transition. See Singer for additional recommendations on the choice of $\varepsilon$ [43,47].

1. For each $\mathbf{x}_i \in S$ compute

$$\rho_\varepsilon(\mathbf{x}_i) = \sum_j K(\mathbf{x}_i, \mathbf{x}_j)$$

where [42]

$$K(\mathbf{x}, \mathbf{y}) = exp[-||\mathbf{x} - \mathbf{y}||^2/2\varepsilon^2]$$

$\rho_\varepsilon(\mathbf{x}_i)$ quantifies the local density of the sample $S$ at location $\mathbf{x}_i$.

2. Construct the symmetric matrix [42]

$$\widetilde{\mathbf{K}}_{ij} = \frac{K(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{\rho_\varepsilon(\mathbf{x}_i)\rho_\varepsilon(\mathbf{x}_j)}}$$

   To the extent that nearby configurations have similarly fast interconversion rates, the matrix $\widetilde{\mathbf{K}}_{ij}$ is analogous to the symmetrically renormalized transition matrix from the discrete spectral theory.
3. For each $\mathbf{x}_i \in S$, compute the diagonal matrix with elements $\mathbf{D}_{ii} = \sum_j \widetilde{\mathbf{K}}_{ij}$. $\mathbf{D}_{ii}$ indicates the total (renormalized) mobility in and out of state $i$. Use the $\mathbf{D}_{ii}$ to compute the symmetrized Markov matrix [42]

$$\mathbf{M} = \mathbf{D}^{-1/2}\widetilde{\mathbf{K}}\mathbf{D}^{-1/2}$$

4. Compute the first few eigenvalues and eigenvectors of $\mathbf{M}$, ideally up to a natural spectral gap. The actual "diffusion maps" are eigenvectors of $\mathbf{D}^{-1/2}\mathbf{M}\mathbf{D}^{1/2}$, but the eigenvectors of the symmetric matrix $\mathbf{M}$ are easier to obtain with numerical methods. The $m^{th}$ diffusion map eigenvector is obtained from $\psi_m^L \approx \mathbf{D}^{-1/2}\psi_m^S$ where $\psi_m^S$ is the $m^{th}$ eigenvector of the symmetric matrix $\mathbf{M}$ [42].

The right eigenvectors are obtained by multiplying the left eigenvectors with the equilibrium distribution. The accuracy of the eigenvectors depends on the initial amount of data in $S$ and an appropriate choice of parameter $\varepsilon$.

If the diffusion map procedure has worked correctly, then the following interpretations can be ascribed to the diffusion map eigenvectors.

- The first eigenvector corresponds to equilibrium: $1/\tau_1 = 0$ and $\psi_1^L(\mathbf{x}) \approx const$.
- The reaction coordinate is $\psi_2^L$. Reactant configurations are at $\psi_2^L(\mathbf{x}) < 0$ and products are at $\psi_2^L(\mathbf{x}) > 0$. Transition states are configurations with $\psi_2^L(\mathbf{x}) = 0$ and the rate of reactant-product interconversion is $1/\tau_2$. For barriers that are much larger than $k_B T$, the value of $\psi_2^L$ changes rapidly as one moves across the transition state. Plateau regions where $\psi_2^L(\mathbf{x})$ is approximately constant correspond to reactants, products, or perhaps to metastable intermediate basins.
- $\psi_3^L(\mathbf{x})$ is the reaction coordinate for the next process in the time scale hierarchy. Often the $\psi_3^L(\mathbf{x})$ process is a relaxation within the reactant basin or within the product basin.

The free energy as a function of $\psi_2^L$ or as a function of $\psi_2^L$ and $\psi_3^L$ provides a visual map of the free energy landscape as a function of the slowest dynamical modes. Figure 15.6.1 shows

the diffusion map coordinates from analysis of conformational transitions of a Beta3s peptide [48].

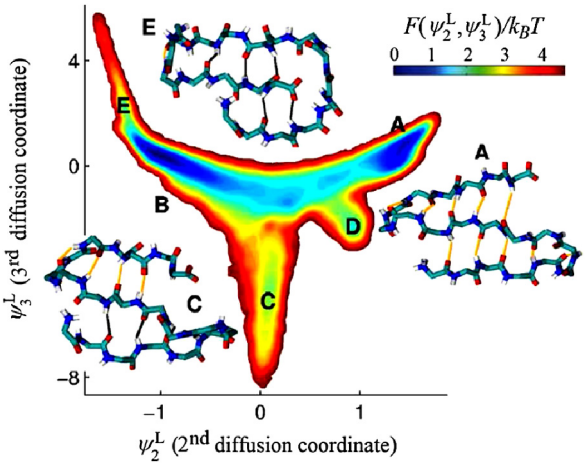



**Figure 15.6.1: The (locally scaled) diffusion map approach was applied to conformational transitions in the Beta3s peptide. Key structures are shown along with the free energy $(-ln\psi_1^R)$ as a function of the first and second diffusion map coordinates. [Modified with permission from Zheng et al. *J. Phys. Chem. B*. (2011).]**

Diffusion maps and MSM eigenvectors are closely related, but there are differences in their computational details and conceptual starting points. Diffusion maps emerged from Fokker-Planck equations for diffusion on a high dimensional continuous manifold. The $\psi_m^L$ coordinates that emerge from MSMs began from a discrete picture. Operationally, both approaches ultimately use discretized states. The MSM approach is probably more robust because it directly counts transition frequencies rather relying on a geometric proximity criterion. If $\varepsilon$ is not chosen correctly, then diffusion map may assume facile interconversion between structures which are dynamically far apart. MSM will correctly describe all interconversions on time scales longer than the time-lag parameter. Of course most MSM implementations also employ a preliminary clustering procedure based on geometric distance, and therefore a poorly chosen distance metric can also corrupt MSM results.

Eigenfunction-based reaction coordinates have many of the properties of an ideal reaction coordinate (see Chapter 20). Furthermore, the eigenfunctions can be systematically generated without *a priori* defined reactant and product basins in configuration space. An eigenfunction-based reaction coordinate accurately summarizes the time scales and transitions observed during long molecular dynamics trajectories. However, eigenfunctions also have some practical disadvantages as reaction coordinates.

- In most implementations, eigenfunctions are constructed from long unbiased trajectories that must spontaneously cross over all significant barriers. Thus, applications to processes with barriers much larger than $k_BT$ are still a major challenge. Methods are being developed that use importance sampling methods in the construction of diffusion maps and MSMs [49–51].

- Eigenfunction-based reaction coordinates only exist for systems that have a well-defined reactant-product equilibrium. They do not exist for processes like nucleation in which the product free energy diverges [52].
- There is (as yet) no reliable way to relate the eigenfunctions to physical variables that have a clear mechanistic interpretation [53]. However, there are ongoing efforts to build the eigenfunctions from bases of physically meaningful collective variables [54,55].

Because of these limitations, spectral analyses based on MD trajectory data at one condition (temperature, ionic strength, etc.) cannot yet predict results across different conditions or across a family of related reactions. Reaction coordinates with a clear physical interpretation are essential ingredients of models that predict such trends [53]. It seems likely that some combination of dimensionality reduction techniques [42,43,45,46], high throughput hypothesis testing techniques [53,56,57], and spectral analyses [58–60] will yield practical reaction coordinates for the most complex transitions in the near future.

## Exercises

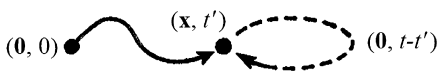
1. Compute the correlation function $\langle q(t)q(0)\rangle_R$ as a function of the initial velocity $v(0)$ and position $q(0)$ for a damped harmonic oscillator with (inertial) Langevin equation $m\ddot{q} = -m\omega^2 q - m\gamma\dot{q} + R(t)$.
2. Compute $\langle q(t)q(0)\rangle_R$ as a function of the initial position $q(0)$ for a harmonic oscillator with overdamped Langevin equation $\dot{q} = -D\beta m\omega^2 q + \widetilde{R}(t)$.
3. A particle diffuses freely from the origin in one dimension. The probability density to find the particle at location $x$ at time $t$, satisfies

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2}$$

with $P(x,0) = \delta[x]$. For any $x \neq 0$ and $t > 0$, $P(x,t)$ also satisfies the integral equation

$$P(x,t) = \int_0^t F(x,t')P(0,t-t')dt'$$

where $F(x,t)$ is the probability that the particle *first* visits $x$ at time $t'$ with $t' < t$ [61].




What is $\hat{F}(x,s)$, i.e. the Laplace transform of $F(x,t)$. Find, (contour) plot, and interpret $F(x,t)$.

4. Show that the Fokker-Planck equation for free inertial Brownian motion is

$$\frac{\partial \rho}{\partial t} = \gamma \frac{\partial}{\partial v}\left\{ v\rho + \frac{k_B T}{m}\frac{\partial \rho}{\partial v}\right\}$$

and that

$$\rho(v,t) = \left(2\pi k_B T\,(1-e^{-2\gamma t})/m\right)^{-1/2} exp[-\frac{mv^2(1-(v/v_0)e^{-\gamma t})^2}{2k_B T\,(1-e^{-2\gamma t})}]$$

is a valid solution.

5. Use the method of characteristics to obtain the solution in problem (4). Hint: first Fourier transform the equation to obtain a first order PDE.
6. This exercise is a guided derivation of the Smoluchowski equation from the overdamped Langevin equation.
   (a) Starting from $\partial\rho/\partial t + \partial j/\partial q = 0$, show that $\partial\rho/\partial t + \hat{A}_q\rho = \hat{B}_q(t)\rho$ where $\hat{A}_q\rho = \frac{\partial}{\partial q}\left(v_D(q)\rho\right)$ and $\hat{B}_q(t)\rho = -\frac{\partial}{\partial q}\widetilde{R}_q(t)\rho$.
   (b) Obtain an implicit solution to $\partial\rho/\partial t + \hat{A}_q\rho = \hat{B}_q(t)\rho$ in the form of an integral equation:

$$\rho(q,t) = e^{-\hat{A}_q t}\rho(q,0) + \int_0^t dt' e^{-\hat{A}_q(t-t')}\hat{B}_q(t')\rho(q,t')$$

   Hint: $y'(t) + ay(t) = b(t)$ has solution $y = e^{-at}y_0 + \int_0^t dt' e^{-a(t-t')}b(t')$.
   (c) Use the implicit solution to show that

$$\frac{\partial\rho(q,t)}{\partial t} + \hat{A}_q\rho(q,t) = \hat{B}_q(t)e^{-\hat{A}_q t}\rho(q,0) + \hat{B}_q(t)\int_0^t dt' e^{-\hat{A}_q(t-t')}\hat{B}_q(t')\rho(q,t')$$

   (d) Now average over the random forces to obtain

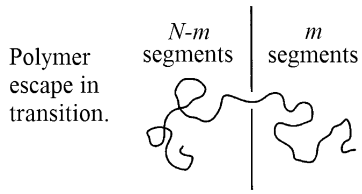
$$\frac{\partial\rho(q,t)}{\partial t} = -\frac{\partial}{\partial q}\left[v_D(q)\rho\right] + \frac{\partial}{\partial q}\left[D_q\frac{\partial\rho}{\partial q}\right]$$

   Comment on some of the subtle aspects of the noise averaging steps. What happens to $\hat{B}_q(t)e^{-\hat{A}_q t}\rho(q,0)$? Why does $\int_0^t dt'\widetilde{R}_q(t)\widetilde{R}_q(t') = \mathbf{1}\times D_q$? Why does noise that influenced $\rho(q,t')$ at times before $t'$ not overlap with the noise at $\widetilde{R}_q(t)$? Finally, why doesn't the derivative in $\int_0^t dt'\widetilde{R}(t)\partial_q[\widetilde{R}_q(t')\rho(q,t')]$ lead to an additional term?
7. Solve the Smoluchowski equation

$$\frac{\partial\rho}{\partial t} = D\frac{\partial}{\partial q}\left\{\frac{\partial\rho}{\partial q} + \rho\frac{\partial\beta F}{\partial q}\right\}$$

for a harmonic oscillator potential of mean force $F(q) = m\omega^2 q^2/2$. First, solve for the initial condition $\rho(q,0) = \delta[0]$, then try the initial condition $\rho(q,0) = \delta[q-q_0]$. Hint: use the equilibrium solution to transform this into an eigenvalue problem.

8. Solve the Smoluchowski equation of problem (7) for a parabolic barrier potential of mean force $\beta F = -m\omega^2 q^2/2$ where $\omega^2 > 0$. First, solve for the initial condition $\rho(q,0) = \delta[0]$, then try the initial condition $\rho(q,0) = \delta[q - q_0]$.
9. Multiply the Smoluchowski equations of problems (7) and (8) by $q^n$ and integrate to obtain a series of ordinary differential equations for the moments, i.e. construct a series of ODEs to obtain $\langle q \rangle_t = \int \rho(q,t) q dq$, $\left\langle (q - \langle q \rangle_t)^2 \right\rangle_t = \int \rho(q,t)(q - \langle q \rangle_t)^2 dq$, etc.
10. A Fokker-Planck equation is always linear in the mathematical sense. However, we refer to them as linear when $v_D(q)$ is linear in $q$ and $D(q)$ is a constant in $q$, and nonlinear otherwise. Extend exercise 9 to show that in the general nonlinear case $d\langle q \rangle_t / dt = \langle v_D(q) \rangle_t$. Then show that for a linear Fokker-Planck equation the first moment evolves according to $d\langle q \rangle_t / dt = v_D(\langle q \rangle_t)$. Note that the same result emerged for the discrete birth-death process in section 14.2.
11. An example in Chapter 14 analyzed a master equation for the discrete random variable $n$ in a birth-death process. $n$ became a Poisson distributed random variable in the stationary limit. Convert the discrete master equation to a Fokker-Planck equation, solve for the stationary distribution, and compare to the exact Poisson result. What parameter controls the accuracy of the results from the transformation to a Fokker-Planck equation?
12. Polymer translocation through a narrow pore involves intermediate low-entropy polymer configurations as shown in the figure below [adapted from Muthukumar, J. Chem. Phys. 111, 10371 (1999)].




The free energy is $\beta F(m) = 2^{-1} ln[m(N-m)]$ where $N$ is the number of segments on the polymer chain and $m$ is the number of segments on the right side of the orifice. Read and reproduce the analysis by Muthukumar [62] to obtain the Fokker-Planck equation for $\rho(m,t)$:

$$\frac{\partial \rho}{\partial t} = k_0 \frac{\partial}{\partial m} \left\{ \frac{\partial \beta F}{\partial m} \rho + \frac{\partial}{\partial m} \rho \right\}$$

where $k_0$ is the $m$-independent frequency at which segments jump rightward across the orifice.

(a) Consider the alternative "linear free energy relationship" model for the transition frequencies from state $m$ to $m+1$,

$$\omega_{m \to m+1} = k_0 exp[-\beta \Delta F_{m \to m+1}/2]$$

where $\Delta F_{m\to m+1} = F(m+1) - F(m)$. Compute the transition frequency from $m+1$ to $m$ to show that detailed balance is "built in" to this construction.

(b) Show that

$$k_0^{-1}\frac{\partial p_m}{\partial t} = (e^{-\partial/\partial m} - 1)\, p_m e^{-\beta \Delta F_{m\to m+1}/2} + (e^{\partial/\partial m} - 1)\, p_m e^{-\beta \Delta F_{m\to m-1}/2}$$

Group the first derivative terms and second derivative terms to identify the drift velocity and diffusivity as functions of $m$.

(c) Show that

$$\lim_{\delta\to 0}(e^{-\beta \Delta F_{m\to m-\delta}/2} - e^{-\beta \Delta F_{m\to m+\delta}/2})/\delta = \frac{d\beta F}{dm}$$

and that for small $\delta$ the diffusivity term can be simplified using $(e^{-\beta \Delta F_{m\to m-\delta}/2} + e^{-\beta \Delta F_{m\to m+\delta}/2})/2 \approx 1$.

(d) Use the results from parts (a)–(c) to obtain the Fokker-Plank equation for a continuous $m$ and large $N$. Is your Fokker-Planck equation the same or different from that of Muthukumar? Explain.

(e) Repeat the analysis in steps (a)–(d) for the general case: $\omega_{m\to m+1} = k_0 exp[-\alpha\beta\Delta F_{m\to m+1}]$ where $0 \le \alpha \le 1$. Do the results depend on $\alpha$?

13. Analyze the Schlogl reaction system as follows.

(a) For the parameters given in the example of section 15.5, write down the macroscopic deterministic rate equations for a constant volume batch reactor. Find the two nullclines, i.e. along one curve in [**A**], [**B**] space $d[\mathbf{A}]/dt = 0$, and along the other $d[\mathbf{B}]/dt = 0$. Characterize the stationary points, i.e. the points where nullcline curves intersect. Use linear stability analysis (or sketch flowlines if preferred) to characterize the stability of each stationary point.

(b) Use the change of variables $y(n,t) = D(n)p(n,t)$ to convert the Fokker-Planck equation (15.4.3) into a Smoluchowski equation. Then make a second change of variables to $\varphi(n,t) = y(n,t)/\sqrt{y_{SS}(n)}$. Show that the operator which results is a self-adjoint Schrodinger-type equation (with imaginary time). You may find some hints in exercise 15.

(c) Explain how these results could be used to estimate the time to relax from any initial condition to the steady state $p_{SS}(n)$. Also explain how you might determine whether it was okay to develop and use a Fokker-Planck equation that ignores the number of [**B**] molecules.

14. Consider a property $A(\mathbf{q})$ and denote its time-evolving average started from initial condition $\mathbf{q}_0$ as $\bar{A}(\mathbf{q}_0, t)$, i.e.

$$\bar{A}(\mathbf{q}_0, t) = \langle A(\mathbf{q}(t))\rangle_{\mathbf{q}(0)=\mathbf{q}_0}$$

where the average is over all realizations of the stochastic noise. Let the probability distribution $\rho(\mathbf{q},t)$ evolve according to a Fokker-Planck equation $\partial\rho/\partial t = L\rho$. Show that $\bar{A}(\mathbf{q}_0,t)$ can be obtained $(i)$ from $\langle A(\mathbf{q})\psi_1(\mathbf{q})|\rho(\mathbf{q},t)\rangle$ where $|\rho(\mathbf{q},t)\rangle$ is given by equation (15.6.1) or $(ii)$ from $\left\langle\psi_1(\mathbf{q})|\hat{A}(t)|\delta[\mathbf{q}-\mathbf{q}_0]\right\rangle$ where $\hat{A}(t) = \hat{A}\sum_m |\psi_m\rangle\langle\psi_m| e^{-t/\tau_m}$. Compare routes (i) and (ii) to the Schrodinger and Heisenberg approaches in quantum mechanics, respectively.

15. Consider the Smoluchowski equation $\partial\rho/\partial t = L\rho(q,t)$ where

$$\hat{L}\rho \equiv D\frac{\partial}{\partial q}\left\{e^{-\beta F(q)}\frac{\partial}{\partial q}\left(e^{\beta F(q)}\rho\right)\right\}$$

(a) Let $\rho(q,t) = e^{-\beta F(q)/2}\phi(q,t)$. Show that $\hat{L}\rho = -De^{-\beta F(q)/2}\hat{L}_S\phi$ where $\hat{L}_S$ is the self-adjoint Schrodinger-like operator [63]

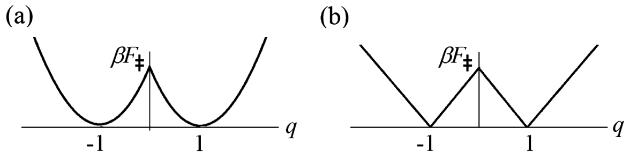
$$\hat{L}_S\phi \equiv -\frac{\partial^2\phi}{\partial q^2} + V(q)\phi$$

where the effective potential $V(q)$ is related to the original $F(q)$ by

$$V(q) = -\frac{1}{2}\frac{\partial^2\beta F}{\partial q^2} + \frac{1}{4}\left(\frac{\partial\beta F}{\partial q}\right)^2. \tag{15.6.2}$$

(b) Show that eigenfunctions of $\hat{L}$ are the functions $\psi_i(q) = e^{-\beta F(q)/2}\phi_i(q)$ where the $\phi_i(q)$ are eigenfunctions of the self-adjoint operator $\hat{L}_S$. Verify that the first eigenfunction of $\hat{L}_S$ is $\phi_1(q) = e^{-\beta F(q)/2}$. Is this consistent with the result in part (a)? [74]

(c) The two free energy surfaces below are both symmetric with cusp-like barriers of the same height and minima at the same location. Also assume the same diffusivity $D$.




Before working out the two solutions, which one do you think will have the faster relaxation rate? Why? Find the equilibrium eigenfunction and the most slowly relaxing eigenfunction for each of these two state systems.

## References

[1] L.P. Kadanoff, Statistical Physics: Statics, Dynamics, and Renormalization, World Scientific, Singapore, 2000.
[2] W.A. Coffey, Y.P. Kalmykov, J.T. Waldron, The Langevin Equation, World Scientific, Singapore, 2004.

[3] A. Nitzan, Dynamics in Condensed Phases: Relaxation, Transfer, and Reactions in Condensed Molecular Systems, Oxford University Press, Oxford, 2006.
[4] R. Erban, S.J. Chapman, Phys. Biol. 6 (2009) 46001.
[5] P. Langevin, C. R. Acad. Sci. Paris 146 (1908) 530–533.
[6] M. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, Oxford University Press, Oxford, UK, 2010.
[7] B. Leimkuhler, Molecular Dynamics: With Deterministic and Stochastic Numerical Methods, Springer, Berlin, Heidelberg, 2015.
[8] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon Press, Oxford, UK, 1987.
[9] D.L. Ermak, J.A. McCammon, J. Chem. Phys. 69 (1978) 1352.
[10] M.R. Shirts, J. Chem. Theory Comput. 9 (2013) 909–926.
[11] M. Grunwald, C. Dellago, P.L. Geissler, J. Chem. Phys. 129 (2008) 194101.
[12] R.G. Mullen, J.-E. Shea, B. Peters, J. Chem. Phys. 140 (2014) 41104.
[13] H.C. Andersen, J. Chem. Phys. 72 (1980) 2384.
[14] P. Hanggi, P. Jung, Adv. Chem. Phys. 89 (1995) 239–326.
[15] A. Einstein, Investigations on the Theory of the Brownian Movement, Courier Corporation, 1956.
[16] G.E. Uhlenbeck, L.S. Ornstein, Phys. Rev. 36 (1930) 823.
[17] G. Hummer, New J. Phys. 7 (2005) 34.
[18] R.B. Best, G. Hummer, Proc. Natl. Acad. Sci. USA 107 (2010) 1088–1093.
[19] D.J. Higham, SIAM Rev. 43 (2001) 525–546.
[20] D.T. Gillespie, Markov Processes: An Introduction for Physical Scientists, Academic Press, Boston, 1992.
[21] B. Leimkuhler, C. Matthews, Appl. Math. Res. Express 2013 (2013) 34–56.
[22] L. Rayleigh, Philos. Mag. 32 (1891) 424–444.
[23] A.D. Fokker, Ann. Phys. 348 (1914) 810–820.
[24] M. von Smoluchowski, Z. Phys. 17 (1916) 557–585.
[25] M. Planck, Sitz.ber. Preuss. Akad. Wiss. 24 (1917).
[26] C.W. Gardiner, Handbook of Stochastic Methods, 2nd ed., Springer-Verlag, Berlin, Heidelberg, 1985.
[27] R. Zwanzig, Nonequilibrium Statistical Mechanics, Oxford University Press, Oxford, 2001.
[28] N.G. van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier, Amsterdam, 2007.
[29] Y.B. Zeldovich, Acta Physicochim. URSS 18 (1943) 1–22.
[30] J. Frenkel, Kinetic Theory of Liquids, Oxford University Press, London, 1946.
[31] A. Kolmogorov, Math. Ann. 104 (1931) 415–458.
[32] R. Zwanzig, Phys. Rev. 124 (1961) 983.
[33] G. Hummer, I.G. Kevrekidis, J. Chem. Phys. 118 (2003) 10762.
[34] B.C. Knott, V. Molinero, M.F. Doherty, B. Peters, J. Am. Chem. Soc. 134 (2012) 1–6.
[35] J.R. Espinosa, C. Vega, C. Valeriani, E. Sanz, J. Chem. Phys. 144 (3) (2016) 034501.
[36] W. Im, B. Roux, J. Mol. Biol. 319 (2002) 1177–1197.
[37] E. Sanz, C. Vega, J.R. Espinosa, R. Caballero-Bernal, J.L.F. Abascal, C. Valeriani, J. Am. Chem. Soc. 135 (2013) 15008–15017.
[38] N.E.R. Zimmermann, B. Vorselaars, D. Quigley, B. Peters, J. Am. Chem. Soc. 137 (2015) 13352–13361.
[39] I. Oppenheim, Adv. Chem. Phys. 15 (1969) 1–11.
[40] R. Erban, S.J. Chapman, I.G. Kevrekidis, T. Vejchodsky, SIAM J. Appl. Math. 70 (2009) 984–1016.
[41] N.G. Van Kampen, Adv. Chem. Phys. 15 (1969) 65–77.
[42] R.R. Coifman, I.G. Kevrekidis, S. Lafon, M. Maggioni, B. Nadler, SIAM J. Multiscale Model. Simul. 7 (2008) 842–864.
[43] A. Singer, R. Erban, I.G. Kevrekidis, R.R. Coifman, Proc. Natl. Acad. Sci. USA 106 (2009) 16090–16095.
[44] P.J. Ledbetter, C. Clementi, J. Chem. Phys. 135 (2011) 44116.
[45] A.L. Ferguson, A.Z. Panagiotopoulos, I.G. Kevrekidis, P.G. Debenedetti, Chem. Phys. Lett. 509 (2011) 1–11.
[46] M.A. Rohrdanz, W. Zheng, C. Clementi, Annu. Rev. Phys. Chem. 64 (2013) 295–316.

[47] A. Singer, Appl. Comput. Harmon. Anal. 21 (2006) 128–134.
[48] W. Zheng, B. Qi, M.A. Rohrdanz, A. Caflisch, A.R. Dinner, C. Clementi, J. Phys. Chem. B 115 (2011) 13065–13074.
[49] X. Huang, G.R. Bowman, S. Bacallado, V.S. Pande, Proc. Natl. Acad. Sci. USA 106 (2009) 19765–19769.
[50] K.A. Beauchamp, G.R. Bowman, T.J. Lane, L. Maibaum, I.S. Haque, V.S. Pande, J. Chem. Theory Comput. 7 (2011) 3412–3419.
[51] W. Zheng, M.A. Rohrdanz, C. Clementi, J. Phys. Chem. B 117 (2013) 12769–12776.
[52] J.S. Langer, Systems Far from Equilibrium, Springer, Berlin, Heidelberg, 1980, pp. 12–47.
[53] B. Peters, J. Phys. Chem. B 119 (2015) 6349–6356.
[54] C.R. Schwantes, V.S. Pande, J. Chem. Theory Comput. 11 (2015) 600–608.
[55] F. Vitalini, F. Noe, B.G. Keller, J. Chem. Theory Comput. 11 (2015) 3992–4004.
[56] B. Peters, B.L. Trout, J. Chem. Phys. 125 (2006) 54108.
[57] B. Peters, Mol. Simul. 36 (2010) 1265–1281.
[58] K.E. Shuler, Phys. Fluids 2 (1959) 442.
[59] N.-V. Buchete, G. Hummer, J. Phys. Chem. B 112 (2008) 6057.
[60] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J.D. Chodera, C. Schutte, F. Noe, J. Chem. Phys. 134 (2011) 174105.
[61] S. Redner, A Guide to First-Passage Processes, Cambridge University Press, Boston, 2001.
[62] M. Muthukumar, J. Chem. Phys. 111 (1999) 10371–10374.
[63] N.G. van Kampen, J. Stat. Phys. 17 (1977) 71–88.