

Chapter 1

Stochastic Processes and Brownian Motion

Equilibrium thermodynamics and statistical mechanics are widely considered to be core subject matter for any practicing chemist [1]. There are plenty of reasons for this:

- A great many chemical phenomena encountered in the laboratory are well described by equilibrium thermodynamics.
- The physics of chemical systems at equilibrium is generally well understood and mathematically tractable.
- Equilibrium thermodynamics motivates our thinking and understanding about chemistry away from equilibrium.

This last point, however, raises a serious question: how well does equilibrium thermodynamics really motivate our understanding of nonequilibrium phenomena? Is it reasonable for an organometallic chemist to analyze a catalytic cycle in terms of rate-law kinetics, or for a biochemist to treat the concentration of a solute in an organelle as a bulk mixture of compounds? Under many circumstances, equilibrium thermodynamics suffices, but a growing number of outstanding problems in chemistry – from electron transfer in light-harvesting complexes to the chemical mechanisms behind immune system response – concern processes that are fundamentally out of equilibrium.

This course endeavors to introduce the key ideas that have been developed over the last century to describe nonequilibrium phenomena. These ideas are almost invariably founded upon a statistical description of matter, as in the equilibrium case. However, since nonequilibrium phenomena contain a more explicit time-dependence than their equilibrium counterparts (consider, for example, the decay of an NMR signal or the progress of a reaction), the probabilistic tools we develop will require some time-dependence as well.

In this chapter, we consider systems whose behavior is inherently nondeterministic, or *stochastic*, and we establish methods for describing the probability of finding the system in a particular state at a specified time.

1.1 Markov Processes

1.1.1 Probability Distributions and Transitions

Suppose that an arbitrary system of interest can be in any one of N distinct states. The system could be a protein exploring different conformational states; or a pair of molecules oscillating between a “reactants” state and a “products” state; or any system that can sample different states over time. Note here that N is finite, that is, the available states are discretized. In general, we could consider systems with a continuous set of available states (and we will do so in section 1.3), but for now we will confine ourselves to the case of a finite number of available states. In keeping with our discretization scheme, we will also (again, for now) consider the time evolution of the system in terms of discrete timesteps rather than a continuous time variable.

Let the system be in some unknown state m at timestep s , and suppose we’re interested in the probability of finding the system in a specific state n , possibly but not necessarily the same as state m , at the next timestep $s + 1$. We will denote this probability by

$$P(n, s + 1)$$

If we had knowledge of m , then this probability could be described as the probability of the system being in state n at timestep $s + 1$ *given that* the system was in state m at timestep s . Probabilities of this form are known as *conditional probabilities*, and we will denote this conditional probability by

$$Q(m, s | n, s + 1)$$

In many situations of physical interest, the probability of a transition from state m to state n is time-independent, depending only on the nature of m and n , and so we drop the timestep arguments to simplify the notation,

$$Q(m, s | n, s + 1) \equiv Q(m, n)$$

This observation may seem contradictory, because we are interested in the time-dependent probability of observing a system in a state n while also claiming that the transition probability described above is time-independent. But there is no contradiction here, because the transition probability Q – a conditional probability – is a different quantity from the time-dependent probability P we are interested in. In fact, we can express $P(n, s + 1)$ in terms of $Q(m, n)$ and other quantities as follows:

Since we don’t know the current state m of the system, we consider *all* possible states m and multiply the probability that the system is in state m at timestep s by the probability of the system being in state n at timestep $s + 1$ *given that* it is in state m at timestep s . Summing over all possible states m gives $P(n, s_1)$ at timestep $s + 1$ in terms of the corresponding probabilities at timestep s .

Mathematically, this formulation reads

$$P(n, s + 1) = \sum_m P(m, s)Q(m, n) \tag{1.1}$$

We’ve made some progress towards a practical method of finding $P(n, s + 1)$, but the current formulation Eq.(1.1) requires knowledge of both the transition probabilities $Q(m, n)$ and the probabilities

$P(m, s)$ for all states m . Unfortunately, $P(m, s)$ is just as much a mystery to us as $P(n, s + 1)$. What we usually know and control in experiments are the initial conditions; that is, if we prepare the system in state k at timestep $s = 0$, then we know that $P(k, 0) = 1$ and $P(n, 0) = 0$ for all $n \neq k$. So how do we express $P(n, s + 1)$ in terms of the initial conditions of the experiment?

We can proceed inductively: if we can write $P(n, s + 1)$ in terms of $P(m, s)$, then we can also write $P(m, s)$ in terms of $P(l, s - 1)$ by the same approach:

$$P(n, s + 1) = \sum_{l,m} P(l, s - 1)Q(l, m)Q(m, n) \quad (1.2)$$

Note that Q has two parameters, each of which can take on N possible values. Consequently we may choose to write Q as an $N \times N$ matrix \mathbf{Q} with matrix elements $(\mathbf{Q})_{mn} = Q(m, n)$. Rearranging the sums in Eq.(1.2) in the following manner,

$$P(n, s + 1) = \sum_l P(l, s - 1) \sum_m Q(l, m)Q(m, n) \quad (1.3)$$

we recognize the sum over m as the definition of a matrix product,

$$\sum_m (\mathbf{Q})_{lm}(\mathbf{Q})_{mn} = (\mathbf{Q}^2)_{ln} \quad (1.4)$$

Hence, Eq.(1.2) can be recast as

$$P(n, s + 1) = \sum_l P(l, s - 1)(\mathbf{Q}^2)_{ln} \quad (1.5)$$

This process can be continued inductively until $P(n, s + 1)$ is written fully in terms of initial conditions. The final result is:

$$P(n, s + 1) = \sum_m P(m, 0)(\mathbf{Q}^{s+1})_{mn} \quad (1.6)$$

$$= P(k, 0)(\mathbf{Q}^{s+1})_{kn} \quad (1.7)$$

where k is the known initial state of the system (all other m do not contribute to the sum since $P(m, 0) = 0$ for $m \neq k$). Any process that can be described in this manner is called a *Markov process*, and the sequence of events comprising the process is called a *Markov chain*.

A more rigorous discussion of the origins and nature of Markov processes may be found in, *e.g.*, de Groot and Mazur [2].

1.1.2 The Transition Probability Matrix

We now consider some important properties of the transition probability matrix \mathbf{Q} . By virtue of its definition, \mathbf{Q} is not necessarily Hermitian: if it were Hermitian, every conceivable transition between states would have to have the same forward and backward probability, which is often not the case.

Example: Consider a chemical system that can exist in either a reactant state A or a product state B, with forward reaction probability p and backward reaction probability $q = 1 - p$,



The transition probability matrix \mathbf{Q} for this system is the 2×2 matrix

$$\mathbf{Q} = \begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

To construct this matrix, we first observe that the given probabilities directly describe the off-diagonal elements Q_{AB} and Q_{BA} ; then we invoke conservation of probability. For example, if the system is in the reactant state A, it can only stay in A or react to form product B; there are no other possible outcomes, so we must have $Q_{AA} + Q_{AB} = 1$. This forces the value $1 - p = q$ upon Q_{AA} , and a similar argument yields Q_{BB} .

Clearly this matrix is not symmetric, hence it is not Hermitian either, thus demonstrating our first general observation about \mathbf{Q} .

The non-Hermiticity of \mathbf{Q} implies also that its eigenvalues λ_i are not necessarily real-valued. Nevertheless, \mathbf{Q} yields two sets of eigenvectors, a left set χ_i and a right set ϕ_i , which satisfy the relations

$$\chi_i \mathbf{Q} = \lambda_i \chi_i \quad (1.8)$$

$$\mathbf{Q} \phi_i = \lambda_i \phi_i \quad (1.9)$$

The left- and right-eigenvectors of \mathbf{Q} are orthonormal,

$$\langle \chi_i | \phi_j \rangle = \delta_{ij} \quad (1.10)$$

and they form a complete set, hence there is a resolution of the identity of the form

$$\sum_i |\phi_i\rangle \langle \chi_i| = \mathbf{1} \quad (1.11)$$

Conservation of probability further restricts the elements of \mathbf{Q} to be nonnegative with $\sum_n \mathbf{Q}_{nm} = 1$. It can be shown that this condition guarantees that all eigenvalues of \mathbf{Q} are bounded by the unit circle in the complex plane,

$$|\lambda_i| \leq 1, \forall i \quad (1.12)$$

Proof of Eq.(1.12): The i^{th} eigenvalue of \mathbf{Q} satisfies

$$\lambda_i \phi_i(n) = \sum_m Q_{nm} \phi_i(m)$$

for each n . Take the absolute value of this relation,

$$|\lambda_i \phi_i(n)| = \left| \sum_m Q_{nm} \phi_i(m) \right|$$

Now we can apply the triangle inequality to the right hand side of the equation:

$$\left| \sum_m Q_{nm} \phi_i(m) \right| \leq \sum_m |Q_{nm} \phi_i(m)|$$

Also, since all elements of \mathbf{Q} are nonnegative,

$$|\lambda_i \phi_i(n)| \leq \sum_m Q_{nm} |\phi_i(m)|$$

Now, the $\phi_i(n)$ are finite, so there must be some constant c such that

$$|\phi_i(n)| \leq c$$

for all n . Then our triangle inequality relation reads

$$c|\lambda_i| \leq c \sum_m Q_{nm}$$

Finally, since $\sum_m Q_{nm} = 1$, we have the desired result,

$$|\lambda_i| \leq 1$$

Another key feature of the transition probability matrix \mathbf{Q} is the following claim, which is intimately connected with the notion of an equilibrium state:

$$\mathbf{Q} \text{ always has the eigenvalue } \lambda = 1 \quad (1.13)$$

Proof of Eq.(1.13): We refer now to the left eigenvectors of \mathbf{Q} : a given left eigenvector χ_i satisfies

$$\chi_i(n)\lambda_i = \sum_m \chi_i(m)Q_{mn}$$

Summing over n , we find

$$\sum_n \chi_i(n)\lambda_i = \sum_n \sum_m \chi_i(m)Q_{mn} = \sum_m \chi_i(m)$$

since $\sum_m Q_{nm} = 1$. Thus, we have the following secular equation:

$$(\lambda_i - 1) \sum_n \chi_i(n) = 0$$

Clearly, $\lambda = 1$ is one of the eigenvalues satisfying this equation.

The decomposition of the secular equation in the preceding proof has a direct physical interpretation: the eigenvalue $\lambda_1 = 1$ has a corresponding eigenvector which satisfies $\sum_n \chi_1(n) = 1$; this stationary-state eigensolution corresponds to the steady state of a system $\chi_1(n) = P_{st}(n)$. It then follows from the normalization condition that $\phi_1(n) = 1$. The remaining eigenvalues $|\lambda_j| < 1$ each satisfy $\sum_n \chi_j(n) = 0$ and hence correspond to zero-sum fluctuations about the equilibrium state.

In light of these properties of \mathbf{Q} , we can define the time-dependent evolution of a system in terms of the eigenstates of \mathbf{Q} ; this representation is termed the *spectral decomposition* of $P(n, s)$ (the set of eigenvalues of a matrix is also known as the spectrum of that matrix). In the basis of left and right eigenvectors of \mathbf{Q} , the probability of being in state n at timestep s , given the initial state as n_0 , is

$$P(n, s) = \langle n_0 | \mathbf{Q}^s | n \rangle = \sum_i \langle n_0 | \phi_i \rangle \lambda_i^s \langle \chi_i | n \rangle \quad (1.14)$$

If we (arbitrarily) assign the stationary state to $i = 1$, we have $\lambda_1 = 1$ and $\chi_1 = P_{st}$, where P_{st} is the steady-state or equilibrium probability distribution. Thus,

$$P(n, s) = P_{st}(n) + \sum_{i \neq 1} \phi_i(n_0) \lambda_i^s \chi_i(n) \quad (1.15)$$

The spectral decomposition proves to be quite useful in the analysis of more complicated probability distributions, especially those that have sufficiently many states as to require computational analysis.

Example: Consider a system which has three states with transition probabilities as illustrated in Figure 1.1. Notice that counterclockwise and clockwise transitions have differing probabilities, which allows this system to exhibit a net current or flux. Also, suppose that $p + q = 1$ so that the system must switch states at every timestep.

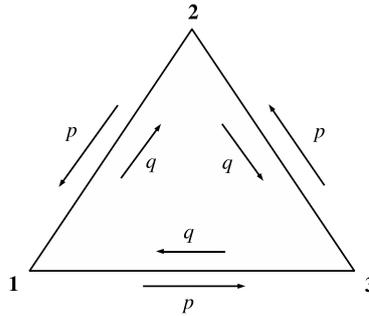


Figure 1.1: A simple three-state system with nonzero flux

The transition probability matrix for this system is

$$\mathbf{Q} = \begin{pmatrix} 0 & p & q \\ q & 0 & p \\ p & q & 0 \end{pmatrix}$$

To determine $P(s)$, we find the eigenvalues and eigenvectors of this matrix and use the spectral decomposition, Eq.(1.14). The secular equation is

$$\text{Det}(\mathbf{Q} - \lambda\mathbf{I}) = 0$$

and its roots are

$$\lambda_1 = 1, \quad \lambda_{\pm} = -\frac{1}{2} \pm \frac{1}{2}\sqrt{3(4pq - 1)}$$

Notice that the nonequilibrium eigenvalues are complex unless $p = q = \frac{1}{2}$, which corresponds to the case of vanishing net flux. If there is a net flux, these complex eigenvalues introduce an oscillatory behavior to $P(s)$.

In the special case $p = q = \frac{1}{2}$, the matrix \mathbf{Q} is symmetric, so the left and right eigenvectors are identical,

$$\begin{aligned} \chi_1 &= \phi_1^T = \frac{1}{\sqrt{3}}(1, 1, 1) \\ \chi_2 &= \phi_2^T = \frac{1}{\sqrt{6}}(1, 1, -2) \\ \chi_3 &= \phi_3^T = \frac{1}{\sqrt{2}}(1, -1, 0) \end{aligned}$$

where T denotes transposition. Suppose the initial state is given as state 1, and we're interested in the probability of being in state 3 at timestep s , $P_{1 \rightarrow 3}(s)$. According to the spectral decomposition formula

Eq.(1.14),

$$\begin{aligned}
 P_{1 \rightarrow 3}(s) &= \sum_i \phi_i(1) \lambda_i^s \chi_i(3) \\
 &= \left(\frac{1}{\sqrt{3}}\right) (1^s) \left(\frac{1}{\sqrt{3}}\right) \\
 &\quad + \left(\frac{1}{\sqrt{6}}\right) (1) \left(-\frac{1}{2}\right)^s \left(\frac{1}{\sqrt{6}}\right) (-2) \\
 &\quad + \left(\frac{1}{\sqrt{2}}\right) (1) \left(-\frac{1}{2}\right)^s \left(\frac{1}{\sqrt{2}}\right) (0) \\
 P_{1 \rightarrow 3}(s) &= \frac{1}{3} - \frac{1}{3} \left(-\frac{1}{2}\right)^s
 \end{aligned}$$

Note that in the evaluation of each term, the first element of each left eigenvector χ and the third element of each right eigenvector ϕ was used, since we're interested in the transition from state 1 to state 3. Figure 1.2 is a plot of $P_{1 \rightarrow 3}(s)$; it shows that the probability oscillates about the equilibrium value of $\frac{1}{3}$, approaching the equilibrium value asymptotically.

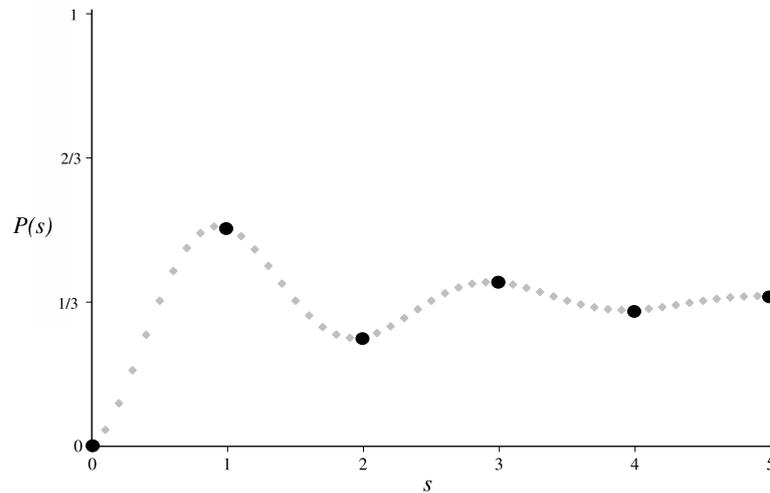


Figure 1.2: Probability of a transition from state 1 to state 3 vs. number of timesteps. Black points correspond to actual timesteps; grey points have been interpolated to emphasize the oscillatory nature of $P(s)$.

1.1.3 Detailed Balance

Our last topic of consideration within the subject of Markov processes is the notion of detailed balance, which is probably already somewhat familiar from elementary kinetics. Formally, a Markov process with transition probability matrix \mathbf{Q} satisfies detailed balance if the following condition holds:

$$P_{\text{st}}(n)Q_{nm} = P_{\text{st}}(m)Q_{mn} \quad (1.16)$$

And this steady state defines the equilibrium distribution:

$$P_{\text{eq}}(n) = P_{\text{st}}(n) = \lim_{t \rightarrow \infty} P(n, t)$$

This relation generalizes the notion of detailed balance from simple kinetics that the rates of forward and backward processes at equilibrium should be equal: here, instead of considering only a reactant state and a product state, we require that all pairs of states be related by Eq.(1.16).

Note also that this detailed balance condition is more general than merely requiring that \mathbf{Q} be symmetric, as the simpler definition from elementary kinetics would imply. However, if a system obeys detailed balance, we can describe it using a symmetric matrix via the following transformation: let

$$V_{nm} = \frac{\sqrt{P_{\text{st}}(n)}}{\sqrt{P_{\text{st}}(m)}} Q_{nm} \quad (1.17)$$

If we make the substitution $P(n, s) = \sqrt{P_{\text{st}}(n)} \cdot \tilde{P}(n, s)$, some manipulation using equations (1.1) and (1.17) yields

$$\frac{d\tilde{P}(n, t)}{dt} = \sum_m \tilde{P}(m, t) V_{mn} \quad (1.18)$$

The derivative $\frac{d\tilde{P}(n, s)}{dt}$ here is really the finite difference $\tilde{P}(n, s+1) - \tilde{P}(n, s)$ since we are considering discrete-time Markov processes, but we have introduced the derivative notation for comparison of this formula to later results for continuous-time systems.

As we did for \mathbf{Q} , we can set up an eigensystem for \mathbf{V} , which yields a spectral decomposition similar to that of \mathbf{Q} with the exception that the left and right eigenvectors ψ of \mathbf{V} are identical since \mathbf{V} is symmetric; in other words, $\langle \psi_i | \psi_j \rangle = \delta_{ij}$. Furthermore, it can be shown that all eigenvalues *not* corresponding to the equilibrium state are either negative or zero; in particular, they are real. The eigenvectors of \mathbf{V} are related to the left and right eigenvectors of \mathbf{Q} by

$$|\phi_i\rangle = \frac{1}{\sqrt{P_{\text{st}}}} |\psi_i\rangle \quad \text{and} \quad \langle \chi_i | = \sqrt{P_{\text{st}}} \langle \psi_i | \quad (1.19)$$

Example: Our final model Markovian system is a linear three-state chain (Figure 1.3) in which the system must pass through the middle state in order to get from either end of the chain to the other. Again we require that $p + q = 1$. From this information, we can construct \mathbf{Q} ,

$$\mathbf{Q} = \begin{pmatrix} q & p & 0 \\ q & 0 & p \\ 0 & q & p \end{pmatrix}$$

Notice how the difference between the three-site linear chain and the three-site ring of the previous example is manifest in the structure of \mathbf{Q} , particularly in the direction of the zero diagonal. This structural difference carries through to general N -site chains and rings.

To determine the equilibrium probability distribution P_{st} for this system, one could multiply \mathbf{Q} by itself many times over and hope to find an analytic formula for $\lim_{s \rightarrow \infty} \mathbf{Q}^s$; however, a less tedious and more intuitive approach is the following:

Noticing that the system cannot stay in state 2 at time s if it is already in state 2 at time $s - 1$, we conclude that $P(1, s + 2)$ depends only on $P(1, s)$ and $P(3, s)$. Also, the conditional probabilities $P(1, s + 2 | 1, s)$ and $P(1, s + 2 | 3, s)$ are both equal to q^2 . Likewise, $P(3, s + 2 | 1, s)$ and $P(3, s + 2 | 3, s)$ are both equal to p^2 . Finally, if the system is in state 2 at time s , it can only get back to state 2 at time $s + 2$ by passing through either state 1 or state 3 at time $s + 1$. The probability of either of these occurrences is pq .

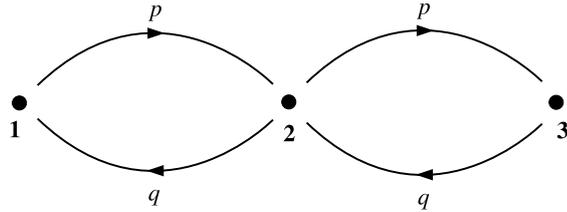


Figure 1.3: A three-state system in which sites are no longer identical

So the ratio $P_{\text{st}}(1) : P_{\text{st}}(2) : P_{\text{st}}(3)$ in the equilibrium limit is $q^2 : qp : p^2$. We merely have to normalize these probabilities by noting that $q^2 + qp + p^2 = (q+p)^2 - qp = 1 - qp$. Thus, the equilibrium distribution is

$$P(s) = \frac{1}{1 - qp}(q^2, qp, p^2)$$

Plugging each pair of states into the detailed balance condition, we verify that this system satisfies detailed balance, and hence all of its eigenvalues are real, even though \mathbf{Q} is not symmetric.

1.2 Master Equations

1.2.1 Motivation and Derivation

The techniques developed in the basic theory of Markov processes are widely applicable, but there are of course many instances in which the discretization of time is either inconvenient or completely unphysical. In such instances, a master equation (more humbly referred to as a rate equation) may provide a continuous-time description of the system that is in keeping with all of our results about stochastic processes.

To formalize the connection between the discrete and continuous formulations, we consider the probability Eq.(1.1) in the limit of small timesteps. Let $t = s\Delta$; Eq.(1.1) then reads

$$P(n, s + 1) = \sum_m P(m, s)Q_{mn}(\Delta) \quad (1.20)$$

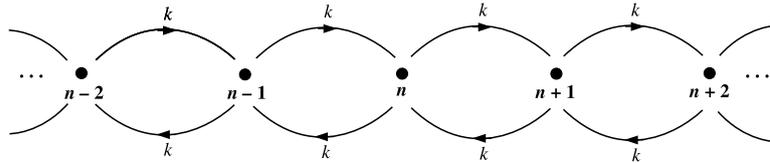
$$P_n(t + \Delta) = \sum_m P_m(t)Q_{mn}(\Delta) \quad (1.21)$$

Here we have switched the labeling of the state whose probability we're interested in from an argument of P to a subscript; this notation is more common in the literature when used in Master equations. The Master equation is the small- Δ limit of Eq.(1.21); in this limit, we can treat the dependence of Q on Δ as a linear dependence, i.e. a Taylor expansion to first order, with the constant term set to zero since the system can't evolve without the passage of time:

$$Q_{mn}(\Delta) = \begin{cases} W_{mn}\Delta, & m \neq n \\ 1 - \sum_k W_{mk}\Delta, & m = n \end{cases}$$

Then in the limit $\Delta \rightarrow 0$, one can verify that Eq.(1.21) becomes

$$\frac{dP_n(t)}{dt} = \sum_m (P_m(t)W_{mn} - P_n(t)W_{nm}) \quad (1.22)$$

Figure 1.4: Infinite lattice with transition rate k between all contiguous states

Eq.(1.22) is a master equation. As the derivation suggests, \mathbf{W} plays the role of a transition probability matrix in this formulation. You may notice that the master equation looks structurally very similar to rate equations in elementary kinetics; in fact, the master equation is a generalization of such rate equations, and the derivation above provides some formal justification for the rules we learn in kinetics for writing them down. The matrix \mathbf{W} is analogous to the set of rate constants indicating the relative rates of reaction between species in the system, and the probabilities P_n are analogous to the relative concentrations of these species.

Example: Consider a random walk on a one-dimensional infinite lattice (see Figure 1.4). As indicated in the figure, the transition probability between a lattice point and either adjacent lattice point is k , and all other transition probabilities are zero (in other words, the system cannot “hop” over a lattice point without first occupying it). We can write down a master equation to describe the flow of probability among the lattice sites in a manner analogous to writing down a rate law. For any given site n on the lattice, probability can flow into n from either site $n - 1$ or site $n + 1$, and both of these occur at rate k ; likewise, probability can flow out of state n to either site $n + 1$ or site $n - 1$, both of which also happen at rate k . Hence, the master equation for all sites n on the lattice is

$$\dot{P}_n = k(P_{n+1} + P_{n-1} - 2P_n)$$

Now we define the average site of occupation as a sum over all sites, weighted by the probability of occupation at each site,

$$\bar{n} = \sum_{n=-\infty}^{\infty} nP_n(t)$$

Then we can compute, for example, how this average site evolves with time,

$$\sum_{n=-\infty}^{\infty} n\dot{P}_n(t) = \dot{\bar{n}} = 0$$

Hence the average site of occupation does not change over time in this model, so if we choose the initial distribution to satisfy $\bar{n} = 0$, then this will always be the average site of occupation.

However, the mean square displacement \bar{n}^2 is not constant; in keeping with our physical interpretation of the model, the mean square displacement increases with time. In particular,

$$\sum_{n=-\infty}^{\infty} n^2\dot{P}_n(t) = \dot{\bar{n}^2} = 2k$$

If the initial probability distribution is a delta function on site 0, $P_n(0) = \delta_0$, then it turns out that Fourier analysis provides a route towards a closed-form expression for the long-time limit of $P_n(t)$:

$$P_n(t) = \frac{1}{2\pi} \int_0^{2\pi} e^{inz} e^{-2k(1-\cos z)t} dz$$

$$\lim_{t \rightarrow \infty} P_n(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{inz} e^{-2D\left(\frac{z^2}{2}\right)t} dz$$

$$\lim_{t \rightarrow \infty} P_n(t) = \sqrt{\frac{1}{4\pi Dt}} e^{-\frac{n^2}{4Dt}}$$

In the above manipulations, we have replaced k with the *diffusion constant* D , the long-time limit of the rate constant (in this case, the two are identical). Thus the probability distribution for occupying the various sites becomes Gaussian at long times.

1.2.2 Mean First Passage Time

One of the most useful quantities we can determine from the master equation for a random walk is the average time it takes for the random walk to reach a particular site n_s for the first time. This quantity, called the *mean first passage time*, can be determined via the following trick: we place an absorbing boundary condition at n_s , $P_{n_s}(t) = 0$. Whenever the walk reaches site n_s , it stays there for all later times. One then calculates the *survival probability* $S(t)$, that is, the probability that the walker has not yet visited n_s at time t ,

$$S(t) = \sum_{n \neq n_s} P_n(t) \quad (1.23)$$

The mean first passage time $\langle t \rangle$ then corresponds to the time-averaged survival probability,

$$\langle t \rangle = \int_0^\infty S(t) dt \quad (1.24)$$

Sometimes it is more convenient to write the mean first passage time in terms of the probability density of reaching site n_s at time t . This quantity is denoted by $f(t)$ and satisfies

$$f(t) = -\frac{dS(t)}{dt} = \sum_{n \neq n_s} P_n(t) W_{nn_s} \quad (1.25)$$

In terms of $f(t)$, the mean first passage time is given by

$$\langle t \rangle = \int_0^\infty t f(t) dt \quad (1.26)$$

The mean first passage time is a quantity of interest in a number of current research applications. Rates of fluorescence quenching, electron transfer, and exciton quenching can all be formulated in terms of the mean first passage time of a stochastic process.

Example: Let's calculate the mean first passage time of the three-site model introduced in Figure 1.1, with all transition rates having the same value k . Suppose the system is prepared in state 1, and we're interested in knowing the mean first passage time for site 3. Applying the absorbing boundary condition at site 3, we derive the following master equations:

$$\begin{cases} \dot{P}_1 = -2kP_1 + kP_2 \\ \dot{P}_2 = kP_1 - 2kP_2 \\ \dot{P}_3 = kP_1 + kP_2 \end{cases}$$

The transition matrix \mathbf{W} corresponding to this system would have a zero column since P_3 does not occur on the right hand side of any of these equations; hence the sink leads to a zero eigenvalue that we can ignore. The relevant submatrix

$$\mathbf{W}_{1,2} = \begin{pmatrix} -2k & k \\ k & -2k \end{pmatrix}$$

has eigenvalues $\lambda_1 = -k$, $\lambda_2 = -3k$. Using the spectral decomposition formula, we find that the survival probability is

$$S(t) = \sum_{i,n} \langle 1|\psi_i\rangle e^{\lambda_i t} \langle \psi_i|n\rangle = e^{-kt}$$

Hence, the previously defined probability density $f(t)$ is given by $f(t) = ke^{-kt}$, and the mean first passage time for site 3 is

$$\langle t \rangle = \frac{1}{k}$$

1.3 Fokker-Planck Equations and Diffusion

We have already generalized the equations governing Markov processes to account for systems that evolve continuously in time, which resulted in the master equations. In this section, we adapt these equations further so that they may be suitable for the description of systems with a continuum of states, rather than a discrete, countable number of states.

1.3.1 Motivation and Derivation

Consider once again the infinite one-dimensional lattice, with lattice spacing Δx and timestep size Δt . In the previous section, we wrote down the master equation (discrete sites, continuous time) for this system, but here we will begin with the Markov chain expression (discrete sites, discrete time) for the system,

$$P(n, s+1) = \frac{1}{2} (P(n+1, s) + P(n-1, s)) \quad (1.27)$$

In terms of Δx and Δt , this equation is

$$P(x, t+\Delta t) = \frac{1}{2} [P(x+\Delta x, t) + P(x-\Delta x, t)] \quad (1.28)$$

Rearranging the previous equation as a finite difference, as in

$$\frac{P(x, t+\Delta t) - P(x, t)}{\Delta t} = \frac{(\Delta x)^2}{2\Delta t} \cdot \frac{P(x+\Delta x, t) + P(x-\Delta x, t) - 2P(x, t)}{(\Delta x)^2} \quad (1.29)$$

and taking the limits $\Delta x \rightarrow 0$, $\Delta t \rightarrow 0$, we arrive at the following differential equation:

$$\frac{\partial}{\partial t} P(x, t) = D \frac{\partial^2}{\partial x^2} P(x, t) \quad (1.30)$$

where $D = \frac{(\Delta x)^2}{2\Delta t}$. This differential equation is called a *diffusion equation* with diffusion constant D , and it is a special case of the Fokker-Planck equation, which we will introduce shortly. The most straightforward route to the solution of the diffusion equation is via spatial Fourier transformation,

$$\tilde{P}(k, t) = \int_{-\infty}^{\infty} P(x, t) e^{ikx} dx \quad (1.31)$$

In Fourier space, the diffusion equation reads

$$\frac{\partial}{\partial t} \tilde{P}(k, t) = -Dk^2 \tilde{P}(k, t) \quad (1.32)$$

and its solution is

$$\tilde{P}(k, t) = \tilde{P}(k, 0)e^{-Dk^2t} \quad (1.33)$$

If we take a delta function $P(x, 0) = \delta(x - x_0)$ centered at x_0 as the initial condition, the solution in x -space is

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-x_0)^2}{4Dt}} \quad (1.34)$$

Thus the probability distribution is a Gaussian in x that spreads with time. Notice that this solution is essentially the same as the long-time solution to the spatially discretized version of the problem presented in the previous example.

We are now in a position to consider a generalization of the diffusion equation known as the Fokker-Planck equation. In addition to the diffusion term $D\frac{\partial^2}{\partial x^2}$, we introduce a term linear in the *first* derivative with respect to x , which accounts for drift of the center of the Gaussian distribution over time.

Consider a diffusion process on a three-dimensional potential energy surface $U(\mathbf{r})$. Conservation of probability requires that

$$\dot{P}(\mathbf{r}, t) = -\nabla \cdot \mathbf{J} \quad (1.35)$$

where \mathbf{J} is the probability current, $\mathbf{J} = -D\nabla P + \mathbf{J}_U$, and \mathbf{J}_U is the current due to the potential $U(\mathbf{r})$. At equilibrium, we know that the probability current $\mathbf{J} = 0$ and that the probability distribution should be Boltzmann-weighted according to energy, $P_{\text{eq}}(\mathbf{r}) \propto e^{-\beta U(\mathbf{r})}$. Therefore, at equilibrium,

$$-D\beta\nabla U(\mathbf{r})P(\mathbf{r})_{\text{eq}} + \mathbf{J}_U = 0 \quad (1.36)$$

Solving Eq.(1.36) for \mathbf{J}_U and plugging the result into Eq.(1.35) yields the Fokker-Planck equation,

$$\dot{P}(\mathbf{r}, t) = D\nabla [\nabla P(\mathbf{r}, t) + \beta\nabla U(\mathbf{r})P(\mathbf{r}, t)] \quad (1.37)$$

1.3.2 Properties of Fokker-Planck Equations

Let's return to one dimension to discuss some salient features of the Fokker-Planck equation.

- First, the Fokker-Planck equation gives the expected results in the long-time limit:

$$\lim_{t \rightarrow \infty} P = P_{\text{eq}} \quad \text{with} \quad \dot{P} = 0 \quad (1.38)$$

- Also, if we define the average position $\bar{x} = \int_{-\infty}^{\infty} xP(x) dx$, then the differential form of the Fokker-Planck equation can be used to verify that

$$\dot{\bar{x}} = D\beta \left(-\frac{\partial}{\partial x} \overline{U(x)} \right) \quad (1.39)$$

Since the quantity in parentheses is just the average force \bar{F} , Eq.(1.39) can be combined with the Einstein relation $D\beta\zeta = 1$ (see section 1.4) to justify that $\zeta\bar{v} = \bar{F}$; the meaning and significance of this equation, including the definition of ζ , will be discussed in section 1.4.

- The Fokker-Planck equation is linear in the first and second derivatives of P with respect to x ; it turns out that any spatial operator that is a linear combination of $\frac{\partial}{\partial x}$, $x\frac{\partial}{\partial x}$, and $\frac{\partial^2}{\partial x^2}$ will define a Gaussian process when used to describe the time evolution of a probability density. Thus, both the diffusion equation and the more general Fokker-Planck equation will generally always describe a Gaussian process.

- One final observation about the Fokker-Planck equation is that it is only analytically solvable in a small number of special cases. This situation is exacerbated by the fact that it is not of Hermitian (self-adjoint) form. However, we can introduce the change of variable $P = e^{-\frac{\beta U}{2}} \Phi$; in terms of Φ , the Fokker-Planck equation is Hermitian,

$$\frac{\partial \Phi}{\partial t} = D [\nabla^2 \Phi - U_{\text{eff}} \Phi] \quad (1.40)$$

where $U_{\text{eff}} = \frac{(\beta \nabla U)^2}{4} - \frac{\beta \nabla^2 U}{2}$. This transformed Fokker-Planck equation now bears the same functional form as the time-dependent Schrödinger equation, so all of the techniques associated with its solution can likewise be applied to Eq.(1.40).

Example: One of the simplest, yet most useful, applications of the Fokker-Planck equation is the description of the diffusive harmonic oscillator, which can be treated analytically. Here we solve the Fokker-Planck equation for the one-dimensional diffusive oscillator with frequency ω . The differential equation is

$$\frac{\partial P}{\partial t} = D \frac{\partial^2}{\partial x^2} P + \gamma \frac{\partial}{\partial x} (xP)$$

where $\gamma = m\omega^2 D\beta$. We can solve this equation in two steps: first, solve for the average position using Eq.(1.39),

$$\dot{\bar{x}} = -\gamma \bar{x}$$

Given the usual delta function initial condition $P(x, 0) = \delta(x - x_0)$, the average position is given by

$$\bar{x}(t) = x_0 e^{-\gamma t}$$

Thus, memory of the initial conditions decays exponentially for the diffusive oscillator.

Then, since the Fokker-Planck equation is linear in P and bilinear in x and $\frac{\partial}{\partial x}$, the full solution must take the form of a Gaussian, so we can write

$$P(x_0, x, t) = \frac{1}{\sqrt{2\pi\alpha(t)}} \exp \left[-\frac{(x - \bar{x}(t))^2}{2\alpha(t)} \right]$$

where $\bar{x}(t)$ is the time-dependent mean position and $\alpha(t)$ is the time-dependent standard deviation of the distribution. But we've already found $\bar{x}(t)$, so we can substitute it into the solution,

$$P(x_0, x, t) = \frac{1}{\sqrt{2\pi\alpha(t)}} \exp \left[-\frac{(x - x_0 e^{-\gamma t})^2}{2\alpha(t)} \right]$$

Finally, from knowledge that the equilibrium distribution must satisfy the stationary condition

$$P_{\text{eq}}(x) = \int_{-\infty}^{\infty} P(x_0, x, t) P_{\text{eq}}(x_0) dx_0$$

we can determine that

$$\alpha(t) = \frac{1 - e^{-2\gamma t}}{m\omega^2 \beta}$$

Thus the motion of the diffusive oscillator is fully described.

The long and short-time limits of $P(x_0, x, t)$ are both of interest to us. At short times,

$$\lim_{t \rightarrow 0} P(x_0, x, t) = \sqrt{\frac{1}{4\pi Dt}} \exp \left[-\frac{(x - x_0)^2}{4Dt} \right]$$

and the evolution of the probability looks like that of a random walk. In the long-time limit, on the other hand, we find the equilibrium probability distribution

$$\lim_{t \rightarrow \infty} P(x_0, x, t) = \sqrt{\frac{m\omega^2\beta}{2\pi}} \exp\left[-\frac{1}{2}m\omega^2\beta x^2\right]$$

which is Gaussian with no mean displacement and with variance determined by a thermal parameter and a parameter describing the shape of the potential. A Gaussian, Markovian process that exhibits exponential memory decay, such as this diffusive oscillator, is called an *Ornstein-Uhlenbeck process*.

1.4 The Langevin Equation

Our focus in this chapter has been on the description of purely stochastic processes. However, a variety of interesting and important phenomena are subject to combinations of deterministic and stochastic processes. We concern ourselves now with a particular class of such phenomena which are described by *Langevin equations*. In its simplest form, a Langevin equation is an equation of motion for a system that experiences a particular type of random force. The archetypal system governed by a Langevin equation is a *Brownian particle*, that is, a particle undergoing *Brownian motion*. (For a brief description of the nature and discovery of Brownian motion, see the Appendix).

The Langevin equation for a Brownian particle in a one-dimensional fluid bath is

$$m\dot{v}(t) + \zeta v(t) = f(t) \quad (1.41)$$

where $v(t) = \dot{x}(t)$ is the velocity of the Brownian particle, ζ is a coefficient describing friction between the particle and the bath, m is the mass of the Brownian particle, and $f(t)$ is a random force. Though it is random, we can make a couple of useful assumptions about $f(t)$:

- The random force is equally likely to push in one direction as it is in the other, so the average over all realizations of the force is zero,

$$\langle f(t) \rangle_f = 0$$

- The random force exhibits no time correlation but has a characteristic strength factor g that does not change over time,

$$\langle f(t_1)f(t_2) \rangle_f = g\delta(t_1 - t_2)$$

Random forces that obey these assumptions are called *white noise*, or more precisely, *Gaussian white noise*. In this case, *all* odd moments of f will vanish, and all even moments can be expressed in terms of two-time correlation functions: for example, the fourth moment is given by

$$\begin{aligned} \langle f(t_1)f(t_2)f(t_3)f(t_4) \rangle_f &= \langle f(t_1)f(t_2) \rangle_f \langle f(t_3)f(t_4) \rangle_f \\ &\quad + \langle f(t_1)f(t_3) \rangle_f \langle f(t_2)f(t_4) \rangle_f \\ &\quad + \langle f(t_1)f(t_4) \rangle_f \langle f(t_2)f(t_3) \rangle_f \end{aligned}$$

In general, complex systems may exhibit time-dependent strength factors $g(t)$, but we will work with the more mathematically tractable white noise assumption for the random force.

The formal solution to the Langevin equation Eq.(1.41) is

$$v(t) = v(0)e^{-\frac{\zeta}{m}t} + \frac{1}{m} \int_0^t e^{-\frac{\zeta}{m}(t-\tau)} f(\tau) d\tau \quad (1.42)$$

In computing the average velocity under the white noise assumption, the second term of Eq.(1.42) vanishes thanks to the condition $\langle f(t) \rangle_f = 0$. So the average velocity is simply

$$\langle v(t) \rangle_f = v(0)e^{-\frac{\zeta}{m}t} \quad (1.43)$$

Of special interest is the velocity-velocity correlation function

$$C(t_1 - t_2) = \langle v(t_1)v(t_2) \rangle_f \quad (1.44)$$

which can also be computed from Eq.(1.42). Invoking the white noise condition for $\langle f(t_1)f(t_2) \rangle_f$, we find that

$$\langle v(t_1)v(t_2) \rangle_f = \left(v(0)^2 - \frac{g}{2m\zeta} \right) e^{-\frac{\zeta}{m}(t_1+t_2)} + \frac{g}{2m\zeta} e^{-\frac{\zeta}{m}(t_2-t_1)} \quad (1.45)$$

So far, we have only performed an average over realizations of the random force, denoted by $\langle \dots \rangle_f$; to proceed, we may also take a thermal average $\langle \dots \rangle_\beta$, that is, the average over realizations of different initial velocities at inverse temperature β . Equipartition tells us that $\langle v_0^2 \rangle_\beta = \frac{1}{m\beta}$; if we use Eq.(1.45) to write down an expression for $\langle \langle v(t_1)v(t_2) \rangle_f \rangle_\beta$ and apply equipartition, we arrive at the conclusion that

$$g = \frac{2\zeta}{\beta} \quad (1.46)$$

which is a manifestation of the *fluctuation-dissipation theorem* (the fluctuations in the random force, described by g , are proportional to the dissipation of energy via friction, described by ζ).

The properties of the velocity variable v enumerated above imply that the distribution of velocities is Gaussian with exponential memory decay, like the diffusive oscillator in section 1.3, and so we can also think of this type of Brownian motion as an Ornstein-Uhlenbeck process. In particular, the probability distribution for the velocity is

$$P(v_0, v, t) = \sqrt{\frac{m\beta}{2\pi(1 - e^{-2\gamma t})}} \exp \left[-\frac{m\beta(v - v_0 e^{-\gamma t})^2}{2(1 - e^{-2\gamma t})} \right] \quad (1.47)$$

We now have a thorough description of the Brownian particle's velocity, but what about the particle's diffusion? We'd like to know how far away the Brownian particle can be expected to be found from its initial position as time passes. To proceed, we calculate the mean square displacement of the particle from its initial position,

$$R^2(t) = \langle (x(t) - x(0))^2 \rangle \quad (1.48)$$

$$= \int_0^t \int_0^t \langle v(\tau_1)v(\tau_2) \rangle d\tau_2 d\tau_1 \quad (1.49)$$

$$= 2 \int_0^t (t - \tau)C(\tau) d\tau \quad (1.50)$$

At long times, the mean square displacement behaves as

$$R^2(t) = 2t \int_0^\infty C(t) dt \quad (1.51)$$

This linear scaling with time is the experimentally observed behavior of Brownian particles, where the proportionality constant is called the diffusion constant D ; hence, we have found an expression for the macroscopic diffusion constant D in terms of the correlation function,

$$D = \int_0^{\infty} C(t) dt \quad (1.52)$$

Eq.(1.52) is known as the *Green-Kubo relation*, and it implies that the mean square displacement at long times is simply

$$\lim_{t \gg 1} R^2(t) = 2Dt \quad (1.53)$$

This result for the mean square displacement also scales linearly with the dimensionality of the system (i.e. in three dimensions, $R^2(t) = 6Dt$).

To determine the behavior of $R^2(t)$ at short times, note that $v(t) \approx v(0)$ for short times, so that $R^2(t) = (\int v(t) dt)^2 \approx \langle v_0^2 \rangle t^2$. Therefore, the short-time limit of the mean square displacement is

$$\lim_{t \ll 1} R^2(t) = \frac{1}{m\beta} t^2 \quad (1.54)$$

For times in between these extremes, the formal solution to the Langevin equation for the velocity would have to be integrated. This can be done; sparing the details, the result after thermal averaging is

$$R^2(t) = \frac{2}{\beta\zeta} \left[t - \frac{1}{\gamma} (1 - e^{-\gamma t}) \right] \quad (1.55)$$

where $\gamma = \frac{\zeta}{m}$.

As a final note, the Langevin equation as presented in this section is often modified to describe more complex systems. The most common modifications to the Langevin equation are:

- The replacement of the friction coefficient ζ with a memory kernel $\gamma(t)$ that allows the system to have some memory of previous interactions.
- The addition of a deterministic mean force $F = -\nabla U$, which permits the system to respond to forces beyond those due to interactions with the bath.

Such modified Langevin equations, also known as Generalized Langevin equations or GLEs, will be explored in further detail in Chapter 4. The Langevin equation and its generalized counterparts provide the basis for a number of successful models of stochastic processes in chemical physics.[3]

1.5 Appendix: Applications to Brownian Motion

Brownian motion is one of the simplest physical examples of a system whose description necessitates a nonequilibrium statistical description. As such, it is the token example that unifies all of the topics in this course, from Markov processes (Ch. 1) and response functions (Ch. 2) to diffusion constants (Ch. 3) and generalized Langevin equations (Ch. 4). In this appendix, the salient features of Brownian motion and the key results about Brownian motion that will be developed during the course are explicated together as a handy reference. Some basic properties of relevant integral transformations are also included in this Appendix.

The discovery of Brownian motion predates the development of statistical mechanics and provided important insight to physicists of the early twentieth century in their first formulations of an atomic description of matter. A fine example of the importance of keeping an eye open for the unexpected in experimental science, Brownian motion was discovered somewhat serendipitously in 1828 by botanist Robert Brown while he was studying pollen under a microscope. Though many others before him had observed the jittery, random motion of fine particles in a fluid, Brown was the first to catalogue his observations[4] and use them to test hypotheses about the nature of the motion.

Interest in the phenomenon was revived in 1905 by Albert Einstein, who successfully related observations about Brownian motion to underlying atomic properties. Einstein's work on Brownian motion[5] is perhaps the least well known of the four paradigm-shifting papers he published in his "Miracle Year" of 1905, which goes to show just how extraordinary his early accomplishments were (the other three papers described the photoelectric effect, special relativity, and mass-energy equivalence)! Einstein determined that the diffusion of a Brownian particle in a fluid is proportional to the system temperature and inversely related to a coefficient of friction ζ characteristic of the fluid,

$$D = \frac{1}{\beta\zeta}$$

Any physical description of Brownian motion will boil down to an equation of motion for the Brownian particle. The simplest way, conceptually, to model the system is to perform Newtonian dynamics on the Brownian particle and N particles comprising the fluid, with random initial conditions (positions and velocities) for the fluid particles. By performing such calculations for all possible initial configurations of the fluid and averaging the results, we can obtain the correct picture of the stochastic dynamics. This procedure, however, is impossibly time-consuming in practice, and so a number of statistical techniques, such as Monte Carlo simulation, have been developed to make such calculations more practical.

Alternatively, we can gain qualitative insight into Brownian dynamics by mean-field methods; that is, instead of treating each particle in the fluid explicitly, we can devise a means to describe their average influence on the Brownian particle, circumventing the tedium of tracking each particle's trajectory independently. This approach gives rise to the Langevin equation of section 1.4, under the assumption that the fluid exerts a random force $f(t)$ on the Brownian particle that obeys the conditions of Gaussian white noise.

For instantaneous (gas-phase) collisions of the fluid and Brownian particle, a Langevin equation with constant frictional coefficient ζ suffices,

$$m\dot{v}(t) + \zeta v(t) = f(t)$$

However, if fluid-particle collisions are correlated, which is the case for any condensed-phase system, this correlation must be taken into account by imbuing the Brownian particle with memory of its previous interactions, embodied by a memory kernel γ ,

$$m\dot{v}(t) + m \int_0^t \gamma(t - \tau)v(\tau) d\tau = f(t)$$

where $\gamma(t) \rightarrow \zeta\delta(t)$ in the limit of uncorrelated collisions.

We now present some of the key features of Brownian motion. Some of these results are derived in section 1.4; others are presented here for reference. Please consult the references at the end of this chapter for further details about the derivation of these properties.

- **Fick's Law:** The spreading of the Brownian particle's spatial probability distribution over time is governed by Fick's Law,

$$\frac{\partial}{\partial t}P(\mathbf{r}, t) = -D\nabla^2P(\mathbf{r}, t)$$

- **Green-Kubo relation:** The diffusion constant D is tied to the particle's velocity-velocity correlation function $C(t)$ by the Green-Kubo relation,

$$D = \int_0^\infty C(t) dt$$

This essentially means that the diffusion constant is the area under the velocity-velocity correlation curve across all times $t > 0$.

- **Solution of the Langevin Equation:** All of the information we require from the Langevin equation is contained in the correlation function. Multiplication of the Langevin equation for $v(t_1)$ by the velocity $v(t_2)$ yields a differential equation for the correlation function,

$$\dot{C} + \int_0^t \gamma(t - \tau)C(\tau) d\tau = 0$$

The Laplace transform of this equation,

$$s\hat{C}(s) - C(0) + \hat{\gamma}(s)\hat{C}(s) = 0$$

has as its solution

$$\hat{C}(s) = \frac{C(0)}{s + \hat{\gamma}(s)}$$

where $C(0)$ is the non-transformed velocity-velocity correlation function at $t = 0$ and s is the Laplace variable.

- **Einstein relation:** The solution to the Langevin equation tells us that

$$\hat{C}(0) = \frac{C(0)}{\hat{\gamma}(0)}$$

Additionally, a comparison of the Green-Kubo relation to the formula for the Laplace transform indicates that $\hat{C}(0) = D$. Finally, we can conclude from the equipartition theorem that $C(0) = \frac{1}{m\beta}$. Combining this information together, we arrive at Einstein's relation,

$$D = \frac{1}{m\beta\hat{\gamma}(0)}$$

In Chapter 4, the behavior of the velocity-velocity correlation function is explored for the cases in which the fluid is a bath of harmonic oscillators, a simple liquid, and an elastic solid. Their general functional forms are summarized here; further details can be found in Chapter 4.

- **Harmonic oscillators:** $C(t)$ is periodic with amplitude $C(0)$ and frequency Ω_0 (the Einstein frequency), where $\Omega_0^2 = \gamma(0)$.
- **Liquids:** $C(t)$ exhibits a few oscillations while decaying, eventually leveling out to zero.
- **Solids:** Like a liquid, $C(t)$ will be damped, but like the harmonic oscillator model, the periodic structure of the solid will prevent $C(t)$ from decaying to zero; some oscillation at the Einstein frequency will continue indefinitely.

Finally, we summarize the response of a Brownian particle to an external force F . The modified Langevin equation for this situation is

$$\dot{v}(t) + \gamma v(t) = \frac{f(t)}{m} + \frac{F(t)}{m}$$

In general, this Langevin equation is difficult to work with, but many forces of interest (such as EM fields) are oscillatory, so we assume an oscillatory form for the external force,

$$F(t) = F_\omega e^{-i\omega t}$$

Then we can use the techniques developed in Chapter 2 to determine that the velocity in Fourier space is given by

$$\tilde{v}(\omega) = \chi(\omega)\tilde{F}(\omega)$$

Finally, from this information it can be determined that the response function $K(t)$ is (see Chapter 2)

$$K(t) = \frac{1}{2\pi} \int_0^\infty \frac{e^{-i\omega t}}{-i\omega + \gamma} d\omega = e^{-\gamma t} \theta(t)$$

These formulas are the basis for the Debye theory of dipole reorganization in a solvent, in the case where F corresponds to the force due to the electric field $E(\omega)$ generated by the oscillating dipoles.

Integral Transformations: We conclude with a summary of the Laplace and Fourier transforms, which are used regularly in this course and in chemical physics generally to solve and analyze differential equations.

1. *Laplace transform:* The Laplace transform of an arbitrary function $f(t)$ is

$$\hat{f}(s) = \int_0^{\infty} e^{-st} f(t) dt$$

Both the Laplace and Fourier transforms convert certain types of differential equations into algebraic equations, hence their utility in solving differential equations. Consequently, it is often useful to have expressions for the first and second derivatives of $\hat{f}(s)$ on hand:

$$\hat{f}^{(1)}(s) = s\hat{f}(s) - f(0)$$

$$\hat{f}^{(2)}(s) = s^2\hat{f}(s) - sf(0) - \hat{f}^{(1)}(0)$$

A convolution of two functions

$$F(t) = \int_0^t f(t)g(t-\tau) d\tau$$

is also simplified by Laplace transformation; in Laplace space, it is just a simple product,

$$\hat{F}(s) = \hat{f}(s)\hat{g}(s)$$

2. *Fourier transform:* The Fourier transform of an arbitrary function $f(t)$ is

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(t) dt$$

Its derivatives are even simpler in structure than those of the Laplace transform:

$$\tilde{f}^{(1)}(\omega) = -i\omega\tilde{f}(\omega)$$

$$\tilde{f}^{(2)}(\omega) = -\omega^2\tilde{f}(\omega)$$

For an even function $f(t)$, the relationship between the Fourier and Laplace transforms can be determined by taking a Laplace transform of f at $s = i\omega$, from which we discover that

$$\tilde{f}(\omega) = 2 \operatorname{Re} \hat{f}(-i\omega)$$

References

- [1] The American Chemical Society. *Undergraduate Professional Education in Chemistry: ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs*. ACS Committee on Professional Training, Spring 2008.
- [2] S. R. De Groot and P. Mazur. *Non-Equilibrium Thermodynamics*. New York: Dover, 1984.
- [3] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North Holland, 2007.
- [4] Robert Brown. A brief account of microscopical observations made in the months of june, july and august, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine*, 4:161–173, 1828.
- [5] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 17:549–560, 1905.

MIT OpenCourseWare
<http://ocw.mit.edu>

5.72 Statistical Mechanics
Spring 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.