

Lecture 11: Non-Equilibrium Statistical Physics

In the previous notes, we studied how systems in thermal equilibrium can change in time even though equilibrium averages of quantities like energy and total spin are time-independent. The main result of those notes was a computational algorithm that allowed us to simulate such systems. In these notes, we extend a result derived in that previous discussion and thereby build a formalism to study these time-dependent problems analytically.

1 Out of equilibrium

Say we have a lattice system consisting of N spins. The spins have magnetic dipole moment μ and are in an external magnetic field H . We initially prepare the system so that all spins are pointed upwards. The system is in contact with a large surrounding environment at some temperature T , but the system is not yet at equilibrium because this initial spin configuration does not match that predicted by equilibrium statistical physics. Instead, the system only becomes analyzable with the methods of equilibrium statistical physics as we allow it to evolve for infinite time.

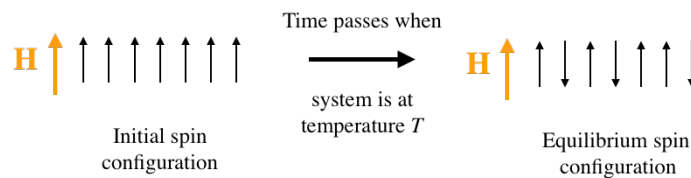


Figure 1

The spins have an energy governed by the energy function $E = (\{s_i\}) = -\mu H \sum_{i=1}^N s_i$. Previously, we devised a computational algorithm to model not only this system's equilibrium distribution of microstates, but also how such a system which is not initially at thermally equilibrium can approach equilibrium over time. We seek to extend the framework developed in the previous notes in order to study—analytically, rather than computationally—how the system depicted in Fig. 1 with energy $E = (\{s_i\}) = -\mu H \sum_{i=1}^N s_i$ reaches thermal equilibrium.

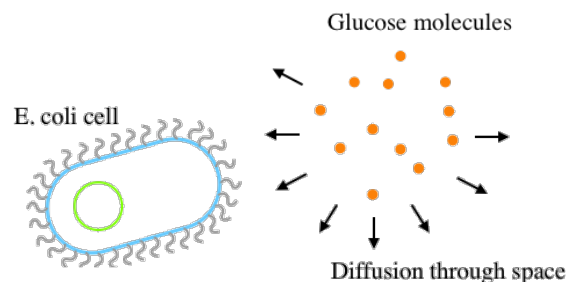


Figure 2: Glucose diffusing through space. The time evolution of the positions of the glucose molecules does not obey Newtonian mechanics, but are rather described by a mathematical model of diffusion.

As we extend our framework we will even be able to encompass systems we have not studied before. *E. coli* is a species of bacteria of roughly a micrometer (i.e., 10^{-6} meters) in length that lives in the human gut.

On the length scale and in the environment in which *E. coli* lives, standard Newtonian dynamics does not apply and the molecules that *E. coli* uses for food do not drift towards the bacterium at a constant velocity. Instead, they move away from their initial location in a process called **diffusion**. It is possible to describe diffusion mathematically using continuous versions of the equations which model how spin systems relax towards equilibrium.

Devising ways to analytically model the time evolution of a spin system and the diffusion of glucose molecules will comprise the two objectives of these notes.

Framing Questions

How can we model the time-dependent relaxation of a thermodynamic system as evolves towards equilibrium? How can we model probabilistic processes that depend on a continuous random variable (e.g., the probabilistic movement of *E. coli* nutrients)?

2 The Master Equation

The fortunate thing is that we have already made great strides towards answering our two framing questions. In the previous notes, prior to laying out a computational algorithm for simulating a thermodynamic system, we derived an equation relating probabilities, time, and transition probabilities:

$$p_j(t + \Delta t) = p_j(t) + \sum_k \left(p_k(t)\pi_{k \rightarrow j} - p_j(t)\pi_{j \rightarrow k} \right). \quad (1)$$

We interpreted Eq.(1) as follows: Given the spectrum of probabilities $p_j(t)$ for a system to be in a microstate j at time t , and the matrix $\pi_{j \rightarrow k}$ defining the probability of moving from a state j to a state k in a time step Δt , we can then determine the spectrum of probabilities to be in the various microstates at a future time $t + \Delta t$. In deriving Eq.(1), we took time to be measured in discrete steps of size Δt because we were primarily concerned with simulating thermodynamic systems, and such a simulation has to use a discrete notion of time. However, in extending this formalism analytically, we are bound by no such constraint. Indeed, we can use the methods of calculus to convert Eq.(1) into a differential equation. To do so, we define the **transition rate** $W_{k \rightarrow j}$ as the probability per unit time to move from microstate k to microstate j . Analytically, $W_{k \rightarrow j}$ is defined as

$$W_{k \rightarrow j} \equiv \lim_{\Delta t \rightarrow 0} \frac{\pi_{k \rightarrow j}}{\Delta t}, \quad (2)$$

where the limit is finite because $\pi_{k \rightarrow j}$ has an implicit first-order dependence on the transition time Δt ¹. Subtracting $p_j(t)$ from both sides of Eq.(1), dividing by Δt , taking $\Delta t \rightarrow 0$, and using the definition of the partial derivative, we obtain

$$\boxed{\frac{\partial}{\partial t} p_j(t) = \sum_k \left(p_k(t)W_{k \rightarrow j} - p_j(t)W_{j \rightarrow k} \right). \quad \text{[Master Equation]}} \quad (3)$$

Eq.(3) is called the **master equation**. It often exists as the first step in studying the time-evolution of time-dependent probabilistic processes.

We should note that although $\sum_k \pi_{j \rightarrow k} = 1$, $W_{j \rightarrow k}$ does not have an identical normalization and in fact $\sum_k W_{j \rightarrow k} \neq 1$. This is because while $\pi_{j \rightarrow k}$ defines a dimensionless probability, $W_{j \rightarrow k}$ defines a probability *per unit time*, and therefore does not have the same normalization requirements.

Conceptually, Eq.(3) represents the idea that the time-rate change of the probability to be in a microstate j is equal to the sum of all the probability currents flowing into j minus the sum of all the probability currents

¹We can expect $\pi_{k \rightarrow j}$ to go to zero as Δt to go to zero because the probability to make a transition should go to zero as the time over which the transition happens goes to zero. The limit is just the claim that $\pi_{j \rightarrow k}$ has at least a first-order dependence on Δt .

flowing out of j (See Fig. 3)

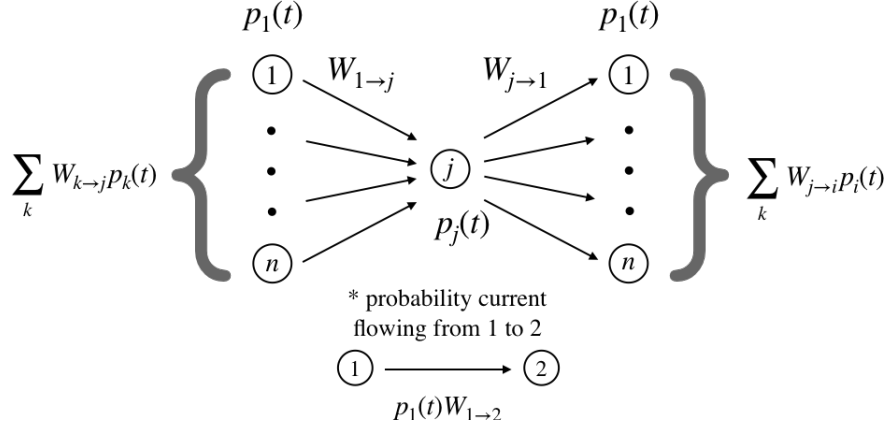


Figure 3: Visual depiction of master equation

We can use Eq.(3) to model how a system with a discrete spectrum of microstates relaxes to equilibrium. We can even extend it to consider a continuous spectrum of microstates. Our systems of interest are a spin lattice and diffusing molecules, but before we consider such systems we can practice applying Eq.(3) to a simpler scenario.

2.1 Poisson process and master equation

We are sitting at the tables outside Bosworth's cafe in MIT's lobby 7 when we decide to devise a time-dependent probabilistic model of people entering the infinite corridor. We want to model the probability that n people enter the infinite corridor from Lobby 7 in a time t . We can use the master equation to construct this model. First we make two assumptions about this system:

1. People enter the infinite corridor one at a time and at the average rate λ .
2. Each person entering the infinite corridor does so independently of every other person.

As is true for many initial models these assumptions do not have perfect fidelity to their real-life counterparts, but they suffice for first-pass model building. The quantity $p_j(t)$ denotes the probability that j people have entered the infinite corridor by time t . We will use the master equation to derive a differential equation for $p_j(t)$ which can then be solved to obtain an explicit expression for $p_j(t)$. To derive the form of in Eq.(3) applicable to this system, we only need to determine the transition rates $W_{k→j}$ for this system.

Because people in this model only enter the infinite corridor one at a time, the transition rate $W_{k→j}$ is only non-zero if k and j differ by 1. Moreover, since we are only counting the people entering the infinite corridor (and not the people exiting it), we can only have a transition from k to j if $k = j - 1$. Finally, we know that people enter the lobby at the average rate λ . Consequently, the rate of transition from $j - 1$ to j is λ , and thus the nonzero elements of $W_{k→j}$ should be λ . From these inferences, we have

$$W_{k→j} = \begin{cases} \lambda, & \text{if } j = k + 1 \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

With Eq.(4), the master equation in Eq.(3) becomes

$$\begin{aligned} \frac{\partial}{\partial t} p_j(t) &= p_{j-1}(t)W_{j-1→j} - p_j(t)W_{j→j+1} \\ &= \lambda(p_{j-1}(t) - p_j(t)). \end{aligned} \quad (5)$$

In order to find the probability that j people have entered the infinite corridor by time t (under the assumptions of this process), we need to find the function $p_j(t)$ that satisfies Eq.(5). The standard method of solving such equations requires techniques we have not developed (and will not subsequently need), so we will quote the solution and check that it satisfies Eq.(5)². The solution is

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}. \quad (6)$$

Calculating the time-derivative of Eq.(6), we find

$$\begin{aligned} \frac{\partial}{\partial t} p_j(t) &= \lambda j \frac{(\lambda t)^{j-1}}{j!} e^{-\lambda t} - \lambda \frac{(\lambda t)^j}{j!} e^{-\lambda t} \\ &= \lambda \left(\frac{(\lambda t)^{j-1}}{(j-1)!} e^{-\lambda t} - \frac{(\lambda t)^j}{j!} e^{-\lambda t} \right) \\ &= \lambda (p_{j-1}(t) - p_j(t)), \end{aligned} \quad (7)$$

as desired. The solution Eq.(6) appears to be a strange one, but it is actually an important discrete probability distribution in probability and statistics. It is called the **Poisson distribution**, and it represents the probability that a certain number of independently-occurring events occur in a given time interval. The most salient statistical property of the Poisson distribution is that the mean is equal to the variance:

$$\langle j \rangle = \sigma_j^2 = \lambda t, \quad (8)$$

as can easily be shown from Eq.(6). The fact that the mean number of people who have entered the infinite corridor is $\langle j \rangle = \lambda t$ makes sense given that λ is the average rate at which people enter the corridor. The fact that the standard deviation of the number of people who have entered the corridor is $\sigma_j = \sqrt{\lambda t}$ indicates that as time goes on the width of our probability distribution increases but does so faster at a rate less than the increase in the mean. Therefore, the relative width around the mean is $\langle j \rangle / \sigma_j = \sqrt{\lambda / t}$ and goes to zero as time goes to infinity.

Having worked through a mostly mathematical example of using Eq.(3) to solve for the time-dependent probability of a system, we can now turn to the physical problem which motivated these notes: The equilibrium relaxation of a spin system.

3 Time-evolution of spin systems

Now, we turn towards trying to develop a time-dependent description of a spin lattice. We will start with the master equation, but we will modify the notation slightly in order to better describe the system of interest. We will replace the indices j and k (which represent different general microstates of the system) in Eq.(3) with \mathcal{S} and \mathcal{S}' which represent different spin microstates. The master equation then becomes

$$\frac{\partial}{\partial t} p(\mathcal{S}, t) = \sum_{\{\mathcal{S}'\}} \left(p(\mathcal{S}', t) W(\mathcal{S}' \rightarrow \mathcal{S}) - p(\mathcal{S}, t) W(\mathcal{S} \rightarrow \mathcal{S}') \right). \quad (9)$$

The quantity $p(\mathcal{S}, t)$ is the probability that the system is in the microstate $\mathcal{S} = \{s_1, \dots, s_N\}$ at a time t , and $W(\mathcal{S}' \rightarrow \mathcal{S})$ is the rate of transitions from microstate $\mathcal{S}' = \{s'_1, \dots, s'_N\}$ to the microstate \mathcal{S} . The summation $\sum_{\{\mathcal{S}'\}}$ represents a sum over all microstates for the system.

In solving Eq.(9), we could seek an explicit form for the time-dependent probability $p(\mathcal{S}, t)$. Computing such a probability is notoriously difficult even with the master equation, however, fortunately, we are often not interested in the explicit time dependence of the probabilities to be in various microstates but rather

²The forward-reasoning solution is found in [1].

the time-dependence of quantities like average spin. Therefore, rather than attempting to find an explicit expression for $p(\mathcal{S}, t)$, we will derive an equation for the time-dependence of the total average spin $\langle \sum_{k=1}^N s_k \rangle$ of the system, and we will solve this equation in order to precisely define how the average spin evolves in time toward equilibrium.

Given the probability $p(\mathcal{S}, t)$ of being in a microstate \mathcal{S} at a time t , we can define time-dependent averages of arbitrary quantities which are functions of the microstate. Generally, for a microstate-dependent quantity $A(\mathcal{S})$, we have the time-dependent average

$$\langle A \rangle = \sum_{\{\mathcal{S}\}} A(\mathcal{S}) p(\mathcal{S}, t), \quad (10)$$

where $\sum_{\{\mathcal{S}\}}$ defines a summation over microstates. Our goal is to use Eq.(9) and Eq.(10) to find a time-dependent description of our lattice system. To find such a description will require us to first choose $W(\mathcal{S} \rightarrow \mathcal{S}')$, but, before we do so, we can use the general properties that we expect $W(\mathcal{S} \rightarrow \mathcal{S}')$ to have in order to derive a time-dependent equation for $\langle s_k \rangle$, the average spin at lattice site k . We first consider the time-dependence of a single spin s_k . By the definition of average given in Eq.(10), we have

$$\begin{aligned} \frac{d}{dt} \langle s_k \rangle &= \frac{d}{dt} \sum_{\{\mathcal{S}\}} s_k p(\mathcal{S}, t) \\ &= \sum_{\{\mathcal{S}\}} s_k \frac{\partial}{\partial t} p(\mathcal{S}, t), \end{aligned} \quad (11)$$

where in the second line we moved the time-derivative within the summation and converted it to a partial time-derivative to acknowledge the fact that we are differentiating a multivariable function. Using Eq.(9), we then have

$$\begin{aligned} \frac{d}{dt} \langle s_k \rangle &= \sum_{\{\mathcal{S}\}} s_k \frac{\partial}{\partial t} p(\mathcal{S}, t) \\ &= \sum_{\{\mathcal{S}\}} s_k \sum_{\{\mathcal{S}'\}} \left(p(\mathcal{S}', t) W(\mathcal{S}' \rightarrow \mathcal{S}) - p(\mathcal{S}, t) W(\mathcal{S} \rightarrow \mathcal{S}') \right) \\ &= \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} s_k W(\mathcal{S}' \rightarrow \mathcal{S}) p(\mathcal{S}', t) - \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} s_k W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t). \end{aligned} \quad (12)$$

The summations $\sum_{\{\mathcal{S}\}}$ and $\sum_{\{\mathcal{S}'\}}$ represent two different summations over the same set of microstates. The microstate label \mathcal{S} also acts as a sort of dummy microstate which we use to denote terms in the summation. Therefore, much in the same way that we can relabel the variables which parameterize integrations and summations without changing the final result, we can also relabel the microstate summations in the above equation without changing the result. We relabel these summations by switching the \mathcal{S} and \mathcal{S}' labels (and the s_k and s'_k labels) in the first term of Eq.(12). We then obtain

$$\begin{aligned} \frac{d}{dt} \langle s_k \rangle &= \sum_{\{\mathcal{S}'\}} \sum_{\{\mathcal{S}\}} s'_k W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t) - \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} s_k W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t) \\ &= \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} s'_k W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t) - \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} s_k W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t) \\ &= \sum_{\{\mathcal{S}\}} \sum_{\{\mathcal{S}'\}} (s'_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t). \end{aligned} \quad (13)$$

In the second line, we switched the order of the summations in the first term, an operation which is valid

because each microstate summation is independent of the other. To further simplify Eq.(13), we recall the Monte Carlo methods of the previous notes. In defining transitions between microstates, we only allowed microstates that differed by a single spin flip to transition into each other. We can represent this result analytically by requiring the transition rate $W(\mathcal{S} \rightarrow \mathcal{S}')$ to be nonzero only for those microstates \mathcal{S}' which differ from \mathcal{S} by a single spin flip. We will denote the set of microstates which differ from \mathcal{S} by a spin flip at lattice site j as \mathcal{S}_j . Thus, \mathcal{S} and \mathcal{S}_j are identical except at lattice site j where they have opposite spins. With this restriction on the microstates that can transition into one another, we can reduce the \mathcal{S}' summation to one only running over the microstates \mathcal{S}_j . There are N microstates of the form \mathcal{S}_j (one for each possible value of j) for every \mathcal{S} microstate, so the summation over all microstates can simply be defined as a summation $j = 1$ to $j = N$. This redefinition yields

$$\sum_{\{\mathcal{S}'\}} (s'_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}') p(\mathcal{S}, t) = \sum_{j=1}^N (s'_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}_j) p(\mathcal{S}, t). \quad (14)$$

The spin s'_k is the k th spin of the microstate \mathcal{S}_j . As we noted, \mathcal{S}_j is identical to \mathcal{S} except at the lattice site j . Therefore, we have

$$s'_k = \begin{cases} s_k, & \text{for } k \neq j, \\ -s_k, & \text{for } k = j, \end{cases} \quad (15)$$

and we can write

$$\begin{aligned} \sum_{j=1}^N (s'_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}_j) p(\mathcal{S}, t) &= \sum_{j \neq k}^N (s_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}_j) p(\mathcal{S}, t) + (-s_k - s_k) W(\mathcal{S} \rightarrow \mathcal{S}_k) p(\mathcal{S}, t) \\ &= -2s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) p(\mathcal{S}, t). \end{aligned} \quad (16)$$

Using this result in Eq.(14) and substituting back into Eq.(13), we have

$$\frac{d}{dt} \langle s_k \rangle = -2 \sum_{\{\mathcal{S}\}} s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) p(\mathcal{S}, t). \quad (17)$$

Finally, using the definition of the average in Eq.(10), gives us

$$\frac{d}{dt} \langle s_k \rangle = -2 \langle s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) \rangle. \quad (18)$$

Eq.(18) is what we were looking for: a differential equation defining the time-evolution of the average spin. In order to use Eq.(18) to solve for the explicit time-dependence of $\langle s_k \rangle$, we need to find an explicit expression for $W(\mathcal{S} \rightarrow \mathcal{S}_k)$ in terms of the spin at lattice site s_k and then use the expression to compute $\frac{d}{dt} \langle s_k \rangle = -2 \langle s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) \rangle$. We turn to this task now. '

3.1 Detailed balance

We want to determine an explicit expression for $W(\mathcal{S} \rightarrow \mathcal{S}_k)$. We begin by deriving a form of the detailed balance result for our continuous-in-time system. As $t \rightarrow \infty$ the time-dependent probability distribution $p(\mathcal{S}, t)$ must reach its time-independent thermal equilibrium value $p^{\text{eq}}(\mathcal{S})$. Therefore, for $t \rightarrow \infty$, Eq.(9) becomes

$$0 = \sum_{\{\mathcal{S}'\}} (p^{\text{eq}}(\mathcal{S}) W(\mathcal{S}' \rightarrow \mathcal{S}) - p^{\text{eq}}(\mathcal{S}) W(\mathcal{S} \rightarrow \mathcal{S}')). \quad (19)$$

Under the assumption that the individual terms in Eq.(19) vanish independently of one another, we then have

$$\frac{W(S' \rightarrow S)}{W(S \rightarrow S')} = \frac{p^{\text{eq}}(S)}{p^{\text{eq}}(S')} = \frac{\exp[-\beta E(S)]}{\exp[-\beta E(S')]}, \quad (20)$$

where in the final equality, we substituted in the equilibrium distribution $p^{\text{eq}}(S) \propto \exp[-\beta E(S)]$ for which $E(S)$ is the energy of microstate S . Eq.(20) is the **detailed balance** condition for this system of spins. In order to make further progress in defining the transition rate, we relabel S' as S_j to represent the fact that S' can only transition to or from S if they differ by a spin flip at some site j :

$$\frac{W(S_j \rightarrow S)}{W(S \rightarrow S_j)} = \frac{\exp[-\beta E(S)]}{\exp[-\beta E(S_j)]}. \quad (21)$$

Next, we specify how our energy depends on the microstates of the system. For a system of spins, each of which has magnetic moment μ and all of which are in a magnetic field H , the energy for a given microstate is

$$E(S) = -\mu H \sum_{i=1}^N s_i. \quad (22)$$

Inserting Eq.(22) into Eq.(21), and noting that S and S_j only differ by a spin flip at j gives us

$$\frac{W(S_j \rightarrow S)}{W(S \rightarrow S_j)} = \frac{e^{\beta\mu H s_j}}{e^{-\beta\mu H s_j}}, \quad (23)$$

where s_j is the j th spin of S , but $-s_j$ is the j th spin of S_j . From here, we can guess a general form for $W(S \rightarrow S_j)$, and then use Eq.(23) to determine any undetermined functions in this guess. Because $S \rightarrow S_j$ simply consists of a spin flip at lattice site j and, from Eq.(22), the spins do not interact with each other, we would expect $W(S \rightarrow S_j)$ to only be a function of s_j , the spin at lattice site j . Thus, we can guess that $W(S \rightarrow S_j)$ has the form

$$W(S \rightarrow S_j) = \frac{1}{\tau} \left[\frac{1}{2} + s_j f(\beta\mu H) \right], \quad (24)$$

where $f(x)$ is a function that goes to zero as $x \rightarrow 0$. The parameter τ is a constant with units of time that is introduced to ensure that W has the proper dimensions. The first term in Eq.(24) is included to account for the case when there is no magnetic field. In such a scenario, we would expect an equal probability of having or not having a spin flip during a microstate transition. Thus, in the $H \rightarrow 0$ limit there is only a 1/2 chance of having a spin flip, and a transition rate of $1/2 \times 1/\tau = 1/2\tau$. The second term in Eq.(24) is the most general expression for a function of s_j and $\beta\mu H$ which goes to zero as $H \rightarrow 0$. The spin s_j is a coefficient in Eq.(24), rather than an argument of the function because, given $s_j^2 = 1$ for all j and the fact that $f(\beta\mu H) \rightarrow 0$ as $H \rightarrow 0$, any general function of s_j and $\beta\mu H$ must reduce³ to a function of $\beta\mu H$ times s_j . Also, s_j only appears linearly in Eq.(24) because, given that $s_j = +1$ or -1 , we have

$$(s_j)^k = \begin{cases} s_j, & \text{for } k \text{ odd} \\ 1, & \text{for } k \text{ even.} \end{cases} \quad (25)$$

We can determine $f(x)$ in Eq.(24) by returning to Eq.(23). We will write the exponentials in the equation in the form

$$e^{\beta\mu H s_j} = \frac{e^{\beta\mu H s_j} + e^{-\beta\mu H s_j}}{2} + \frac{e^{\beta\mu H s_j} - e^{-\beta\mu H s_j}}{2}$$

³This result could be proved in a special case presuming $f(s_j x)$ can be expanded as a Taylor series about $x = 0$. The only way $f(x) \rightarrow 0$ as $x \rightarrow 0$ is if the even powers of x in the expansion of $f(x)$ go to zero. We are left with an expansion which only has odd powers of x which, for odd n , have coefficients of $s_j^n = s_j$.

$$\begin{aligned}
&= \frac{e^{\beta\mu H} + e^{-\beta\mu H}}{2} + s_j \frac{e^{\beta\mu H} - e^{-\beta\mu H}}{2} \\
&= \cosh(\beta\mu H) + s_j \sinh(\beta\mu H).
\end{aligned} \tag{26}$$

In the second to last line, we used Eq.(25), the fact that $e^x + e^{-x}$ only consists of even powers of x , and the fact that $e^x - e^{-x}$ only consists of odd powers of x . Returning to Eq.(25) and using the expanded form of $e^{\beta\mu H s_j}$ in Eq.(26), we have

$$\begin{aligned}
\frac{W(\mathcal{S}_j \rightarrow \mathcal{S})}{W(\mathcal{S} \rightarrow \mathcal{S}_j)} &= \frac{\cosh(\beta\mu H) + s_j \sinh(\beta\mu H)}{\cosh(\beta\mu H) - s_j \sinh(\beta\mu H)} \\
&= \frac{1 + s_j \tanh(\beta\mu H)}{1 - s_j \tanh(\beta\mu H)}.
\end{aligned} \tag{27}$$

Returning to the guess Eq.(24), we see that Eq.(27) suggests that $f(x) = -\tanh(x)/2$. Therefore, the transition rate is

$$W(\mathcal{S} \rightarrow \mathcal{S}_j) = \frac{1}{2\tau} \left[1 - s_j \tanh(\beta\mu H) \right]. \tag{28}$$

We note that $\tanh(x) \rightarrow 0$ as $x \rightarrow 0$, yielding $W(\mathcal{S} \rightarrow \mathcal{S}_j) = 1/2\tau$ for zero magnetic field. With the transition rate fully defined in Eq.(28), we can return to Eq.(18) and find the evolution equation for $\langle s_k \rangle$.

3.2 A differential equation for average spin

Returning to Eq.(18), and using our new found expression for $W(\mathcal{S} \rightarrow \mathcal{S}_j)$, we can now compute $\langle s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) \rangle$. We obtain

$$\begin{aligned}
\langle s_k W(\mathcal{S} \rightarrow \mathcal{S}_k) \rangle &= \left\langle s_k \frac{1}{2\tau} \left[1 - s_k \tanh(\beta\mu H) \right] \right\rangle \\
&= \frac{1}{2\tau} \langle s_k - s_k^2 \tanh(\beta\mu H) \rangle \\
&= \frac{1}{2\tau} \langle s_k \rangle - \frac{1}{2\tau} \tanh(\beta\mu H),
\end{aligned} \tag{29}$$

where we used $s_k^2 = 1$ in the last line. Substituting this result back into Eq.(18) gives us the differential equation

$$\tau \frac{d}{dt} \langle s_k \rangle = -\langle s_k \rangle + \tanh(\beta\mu H). \tag{30}$$

We then sum over the index k to obtain

$$\boxed{\tau \frac{d}{dt} \langle M \rangle = -\langle M \rangle + N \tanh(\beta\mu H)}, \tag{31}$$

where we defined M as the total spin according to

$$M \equiv \sum_{i=1}^N s_i. \tag{32}$$

At this point, we are mostly done. Eq.(31) gives us an equation defining how $\langle M \rangle$ evolves in time under the microstate transition assumption encoded into the definition of $W(\mathcal{S} \rightarrow \mathcal{S}')$. Using Eq.(31) to solve for $\langle M \rangle$, we can obtain an explicit expression for $\langle M \rangle$ as a function of time and thereby be able to precisely describe

how $\langle M \rangle$ relaxes to its equilibrium value. Solving Eq.(31), we find

$$\langle M \rangle_t = \langle M \rangle_0 e^{-t/\tau} + N \tanh(\beta\mu H) \left(1 - e^{-t/\tau}\right), \quad (33)$$

where $\langle M \rangle_0$ is the average total spin at $t = 0$. Eq.(33) represents the complete time-dependent description of how the total average spin changes in time. In previous work, we showed that when we consider this system of spins in an external field at thermal equilibrium, then the mean of a single spin site is $\langle s \rangle^{\text{eq}} = \tanh(\beta\mu H)$, and thus the mean of the sum of all spin sites is $\langle M \rangle^{\text{eq}} = N \langle s \rangle^{\text{eq}} = N \tanh(\beta\mu H)$. We see that the $t \rightarrow \infty$ limit of Eq.(33) reproduces this equilibrium behavior. The consistency between the long-time behavior of Eq.(33) and the results of equilibrium statistical physics stems from the detailed balance condition Eq.(21) that we used to derive the transition rate W . Because these transition rates were chosen to be related to the Boltzmann distribution through detailed balance, then all time-dependent averages derived from these transition rates must have $t \rightarrow \infty$ behavior which matches the result computed from the Boltzmann distribution.

We can now use the explicit form of $\langle M \rangle_t$ to answer a potential question of interest about this system.

Given that we begin in a macrostate with total average spin $\langle M \rangle_0$, how much time does it take to be within a fractional error f of the thermal equilibrium value $N \tanh(\beta\mu H)$?

Noting that the thermal equilibrium value of $\langle M \rangle$ is given by $N \tanh(\beta\mu H)$, and assuming we get within a fraction f of this value at t_f , Eq.(33) gives us

$$\begin{aligned} (1 + f)N \tanh(\beta\mu H) &= \langle M \rangle_0 e^{-t_f/\tau} + N \tanh(\beta\mu H)(1 - e^{-t_f/\tau}) \\ f N \tanh(\beta\mu H) &= [\langle M \rangle_0 - N \tanh(\beta\mu H)] e^{-t_f/\tau} \\ e^{t_f/\tau} &= \frac{N \tanh(\beta\mu H) - \langle M \rangle_0}{f N \tanh(\beta\mu H)}, \end{aligned} \quad (34)$$

which, upon taking the absolute value⁴ and the logarithm of both sides, yields

$$t_f = \tau \ln \left| \frac{N \tanh(\beta\mu H) - \langle M \rangle_0}{f N \tanh(\beta\mu H)} \right|. \quad (35)$$

Eq.(35) answers a question we side stepped when we discussed ways to computationally simulate spin lattice systems. Back then we said we need to implement the Monte Carlo algorithm for a time "long enough" for the system to reach equilibrium. The result above makes this notion of "long enough" more precise by telling us for how much time we need to simulate the system for in order for the average total spin $\langle M \rangle$ to be within a fraction f of the thermal equilibrium value. One thing we should note about this result is that it predicts that the system will never reach thermal equilibrium in finite time. We see this from the fact that as we take $f \rightarrow 0$ (i.e., the fractional error when we reach perfect equilibrium) in Eq.(35), then the time it takes to achieve this perfect equilibrium is $t_f \rightarrow \infty$. Therefore, for finite times, the system will never have fully reached thermal equilibrium.

4 Diffusion Equation and Continuous Systems

Now that we have answered one of our framing questions, we now turn to the other: How can we find a time-dependent description of probabilities when the random variables those probabilities describe is continuous? In order to develop a formalism that can encompass this objective, we will make some assumptions which will sacrifice generality for solubility.

The precise process we would like to model is the continuous version of a **random walk**. In the typical example of a random walk—and let's consider one-dimensions for simplicity—a particle has a 50–50 chance

⁴We take the absolute value to account for the possibility that $\langle M \rangle_0 > N \tanh(\beta\mu H)$.

of moving either to the left or to the right after each time step. For such a random walk, it is easy to show that, after taking N steps, the probability to be a distance L from the starting point is

$$P_N(L) = \frac{1}{2^N} \frac{N!}{\left[\frac{1}{2}(N+L)\right]! \left[\frac{1}{2}(N-L)\right]!}. \quad (36)$$

What we would like to find now is a formula analogous to Eq.(36) for a process in which the time steps are continuous. Such a process well models how glucose molecules move away from their starting points in an aqueous solution. To derive such a formula we start with the master equation for a discrete set of microstates labeled by j :

$$\frac{\partial}{\partial t} p_j(t) = \sum_k \left(p_k(t) W_{k \rightarrow j} - p_j(t) W_{j \rightarrow k} \right). \quad (37)$$

We can translate Eq.(37) into something applicable to a continuous set of microstates x by promoting the discrete summations to integrals, and by making our probability and transition rate functions of x :

$$\frac{\partial}{\partial t} p(x, t) = \int_{-\infty}^{\infty} dx' \left[p(x', t) W(x' \rightarrow x) - p(x, t) W(x \rightarrow x') \right]. \quad (38)$$

In Eq.(38), x is the random variable, $p(x, t)$ is the time dependent probability density for the random variable, and $W(x \rightarrow x')$ is the transition rate between values of the random variable. We note that to say that $p(x, t)$ is a probability density is to say it is not dimensionless but has units which are the inverse of the units of x . The normalization condition for $p(x, t)$ is given by

$$\int_{-\infty}^{\infty} dx p(x, t) = 1. \quad (39)$$

Eq.(38) is the version of the master equation which is applicable to transitions between continuous microstates. Now, we want to specify Eq.(38) to the case of a continuous random walk where x denotes the position of a particle and $p(x, t)$ is the probability density for that position at time t . The quantity $W(x' \rightarrow x)$ denotes the transition rate to move from the position x' to the position x . Reasonably, we can take such a transition rate to depend only on the difference $x - x'$ between these two positions. We are essentially taking the evolution of this system to be **memory-less**, that is, the probability to transition to a new microstate depends only on the current microstate and not the trajectory in reaching the current microstate.

Moreover, since for the discrete random walk, there is equal probability of moving to the left and to the right, $W(x' \rightarrow x)$ as a function of $x - x'$ should yield the same results if we make it a function of $x' - x$. This is to say $W(x' \rightarrow x)$ is an even function of $x' - x$.

Finally, we expect the transition rate for moving towards far away distances to be much less than the transition rate for moving towards nearby distances. Thus $W(x' \rightarrow x)$ should be a monotonically decreasing function of $|x - x'|$, and should, moreover, decrease rapidly as $|x - x'|$ increases in order to ensure transitions to faraway positions are unlikely.

We can encapsulate all of these properties into the function $\chi(u)$ as follows

$$\begin{aligned} W(x_1 \rightarrow x_2) &= \frac{1}{\tau} \chi(x_2 - x_1) \\ \chi(u) &= \chi(-u) \\ \chi(u) &\text{ decreases rapidly with } u \end{aligned} \quad (40)$$

The parameter τ is a time constant which is necessary to ensure W has the correct units. A plot of $\chi(u)$ which satisfies the stated properties is shown in Fig. 4. With our definition of $\chi(u)$, we can write Eq.(38) as

$$\tau \frac{\partial}{\partial t} p(x, t) = \int_{-\infty}^{\infty} dx' \left[p(x', t) \chi(x - x') - p(x, t) \chi(x' - x) \right]. \quad (41)$$

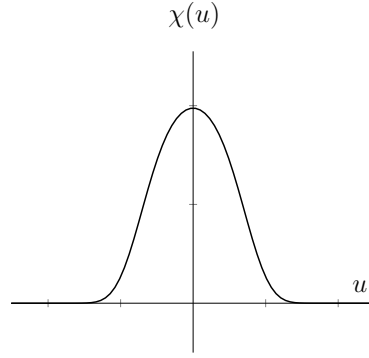


Figure 4: Plot of possible $\chi(u)$. We note that $\chi(u)$ is symmetric about the point $x = 0$, highly peaked at $u = 0$, and decreases rapidly for larger $|u|$.

Now, we make the change of variables $\ell = x' - x$. With the fact that $\chi(\ell) = \chi(-\ell)$, we obtain

$$\tau \frac{\partial}{\partial t} p(x, t) = \int_{-\infty}^{\infty} d\ell \left[p(x + \ell, t) \chi(\ell) - p(x, t) \chi(\ell) \right]. \quad (42)$$

Given that $\chi(\ell)$ is highly peaked near $\ell = 0$ and goes to zero quickly as ℓ moves away from the origin, the product $\chi(\ell)g(\ell)$ for any function $g(\ell)$ should be similarly dominated by its $\ell = 0$ values⁵. Therefore, we can Taylor expand any function of ℓ which multiplies $\chi(\ell)$ about $\ell = 0$ and still recover the pre-dominant behavior of the full function. For the first term in the integrand of Eq.(42), we have

$$\chi(\ell)p(x + \ell, t) = \chi(\ell) \left[p(x, t) + \ell \frac{\partial}{\partial x} p(x, t) + \frac{\ell^2}{2} \frac{\partial^2}{\partial x^2} p(x, t) + \mathcal{O}(\ell^3) \right], \quad (43)$$

and the master equation becomes

$$\begin{aligned} \tau \frac{\partial}{\partial t} p(x, t) &= \int_{-\infty}^{\infty} d\ell \left[\ell \chi(\ell) \frac{\partial}{\partial x} p(x, t) + \frac{\ell^2}{2} \chi(\ell) \frac{\partial^2}{\partial x^2} p(x, t) + \mathcal{O}(\ell^3) \right] \\ &= \frac{\partial}{\partial x} p(x, t) \int_{-\infty}^{\infty} d\ell \ell \chi(\ell) + \frac{1}{2} \frac{\partial^2}{\partial x^2} p(x, t) \int_{-\infty}^{\infty} d\ell \ell^2 \chi(\ell) + \dots \end{aligned} \quad (44)$$

Where the \dots stand in for higher-order terms which are sub-dominant when $\chi(\ell)$ is highly peaked at $\ell = 0$. Because $\chi(\ell)$ is an even function, we have

$$\int_{-\infty}^{\infty} d\ell \ell \chi(\ell) = 0. \quad (45)$$

Also, we define $\overline{\ell^2}$ as

$$\overline{\ell^2} \equiv \int_{-\infty}^{\infty} d\ell \ell^2 \chi(\ell), \quad (46)$$

which is the variance of $\chi(\ell)$ (given that it has a mean of zero). Therefore, to the approximation where we only keep the first nonzero term in Eq.(44), we find

$$\frac{\partial}{\partial t} p(x, t) = \frac{\overline{\ell^2}}{2\tau} \frac{\partial^2}{\partial x^2} p(x, t). \quad (47)$$

⁵Provided, of course, that $g(\ell)$ is not the inverse of $\chi(\ell)$.

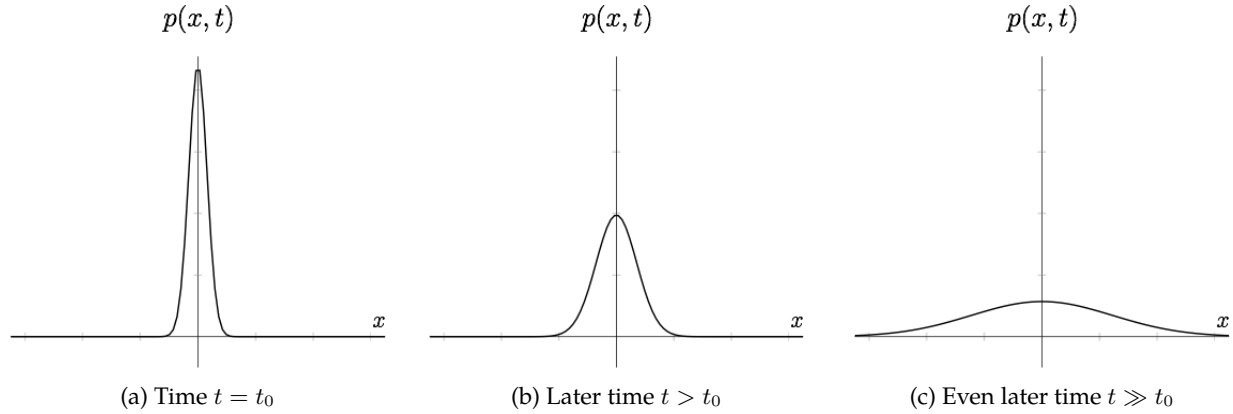


Figure 5: Time evolution of Eq.(50). As the system evolves the Gaussian distribution flattens out, and it becomes more likely to find the particle away from the origin. The archetypal feature of diffusive processes is that the standard deviation in the position x varies as the square root of time t .

Eq.(47) is called the **diffusion equation**. It is a partial differential equation defining the time evolution of the probability density of a continuous-time random walk. In its more typical form we define

$$D \equiv \frac{\overline{\ell^2}}{2\tau}, \quad (48)$$

as the **diffusion coefficient** of this process and we write Eq.(47) as

$$\frac{\partial}{\partial t} p(x, t) = D \frac{\partial^2}{\partial x^2} p(x, t). \quad (49)$$

In order to completely answer the framing question of these notes, we need to solve Eq.(49) and thereby find the explicit form of $p(x, t)$. Such a solution requires methods that extend beyond the prerequisites of this course (but are found in [1]), so we will quote the most important specific solution to Eq.(49). Let us say $p(x, t)$ represents the probability density that a particle (maybe a glucose molecule) is at a position x at time t . Say the particle starts (i.e., at $t = 0$) at the origin $x = 0$, so that the probability density to be anywhere else at this time is zero⁶. Then it is possible to show that for $t > 0$, $p(x, t)$ is given by

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}. \quad (50)$$

Eq.(50) is the probability density that results from taking the $N \gg L \gg 1$ limit of Eq.(36). It is straightforward to check that Eq.(50) satisfies Eq.(49) and the normalization condition Eq.(39) for all times. Moreover, as a Gaussian function of position we can easily state its properties. From Eq.(49), we see that the mean and variance in position are given by

$$\langle x \rangle = 0, \quad \sigma_x^2 = 2Dt. \quad (51)$$

The variance of Eq.(50) increases in time which implies that the probability density gets wider and wider as the system evolves. Therefore, as time goes on, it becomes more likely to find the particle away from the origin. Secondly, from Eq.(50), we see that the height of the peak of the Gaussian decreases in time. This means that as time goes on, it becomes less likely to find the particle at the origin. These two results are necessarily coincident since conservation of probability requires decreases in the probability density at any

⁶We will not get into the fact that the probability density to be at $x = 0$ is not one. Since $p(x, t)$ is a density rather than a raw probability, the density at $x = 0$ is actually infinite.

part of the distribution to be associated with increases in the density at other parts. Plots of Eq.(50) showing this behavior are provided in Fig. 5.

Physically, the importance of Eq.(49) and Eq.(50) exists in their ubiquity. These equations provide good quantitative descriptions for many continuous in time random processes ranging from the diffusion of molecules in water (**Fick's Law of Diffusion**), to the changes in the genetic diversity in a population of species (**Genetic Drift**), and even to the stock market (**Black Scholes Equation**).

References

- [1] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.