Chem 163, Lecture 3

Adam E. Cohen

9/8/2022

1 Review from last class

Last class we discussed the Central Limit Theorem.

The central limit theorem states that if you take N steps from a random distribution with mean μ and variance σ^2, then the distribution after N steps is a Gaussian:

$$P(x) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left[-\frac{(x - N\mu)^2}{2N\sigma^2}\right]. \tag{1}$$

The average displacement is:

$$\langle x \rangle = N\mu, \tag{2}$$

and the variance in x is:

$$\langle \Delta x^2 \rangle = N\sigma^2. \tag{3}$$

We applied this formula to the motions of a diffusing particle.

2 Introduction

There are three pillars of statistical mechanical thinking. One is the central limit theorem. We already discussed that. Two is the equipartition theorem. It's in the appendix of the lecture notes. And three is the fluctuation-dissipation theorem, a special case of which called the Stokes Einstein relation we'll discuss today. After that we'll be thoroughly equipped to tackle a wide variety of interesting and subtle problems.

3 Stokes Einstein Relation

The Equipartition theorem describes the energetics of particles subject to an external force. Now let's talk about the dynamics. We will see that the diffusion coefficient, which we've taken as a given parameter, is actually simply related to other things we know.

Let's assume that the force is a gradient of a potential. That is:

$$\mathbf{F} = -\nabla U. \tag{4}$$

The force imparts a velocity to the particles given by

$$\mathbf{v} = \frac{1}{\gamma}\mathbf{F}, \tag{5}$$

where γ is the drag coefficient. The total flux of particles is:

$$\begin{aligned} \mathbf{j} &= c\mathbf{v} \\ &= -\frac{c}{\gamma}\nabla U. \end{aligned} \tag{6}$$

We've been ignoring the fact that the particle can diffuse too. There's no cross-talk between force-induced drift and diffusion. These two terms simply add together to make the total flux:

$$\mathbf{j} = -D\nabla c - \frac{c}{\gamma}\nabla U. \tag{7}$$

The rate of change of concentration is still given by the negative divergence of the flux:

$$\frac{\partial c}{\partial t} = D\nabla^2 c + \nabla \cdot (\frac{c}{\gamma}\nabla U). \tag{8}$$

This is called the Smoluchowski Equation, which is a special case of the Fokker Planck Equation.

[SHOW EXAMPLES OF PARTICLES MOVING UNDER ELECTRIC FIELDS, AND BAND BROADENING IN ELECTROPHORESIS]

It's important to distinguish two different forms of steady state that are often confused with each other. The term "Steady State" refers to the condition $\partial c/\partial t = 0$, i.e. that the concentrations everywhere are not changing with time. If you have sources and sinks, you can still have a flux, even in a steady state. A more restrictive condition is what we call "thermal equilibrium". Thermal equilibrium must satisfy the condition of *detailed balance*, i.e. that the net flux between any two micro-states is exactly zero. If you have some flux from A to B, you must have equal and opposite flux from B to A. You can't achieve steady state of B by siphoning off particles to some third state C. The Venn diagram is that all thermal equilibria are steady states, but not all steady states are thermal equilibria.

Now we're about to discover something amazing and unexpected. Suppose we're at equilibrium (not just steady state, but *equilibrium*, which means $\mathbf{j} = 0$). Then we know the distribution of concentration must be independent of time and obey:

$$c(\mathbf{r}) = c_0 \exp[-U(\mathbf{r})/k_B T]. \tag{9}$$

We can plug this equation into the one for $\mathbf{j}$ to get:

$$0 = Dc_0 \frac{\nabla U}{k_B T} \exp[-U(\mathbf{r})/k_B T] - \frac{1}{\gamma}c_0 \exp[-U(\mathbf{r})/k_B T]\nabla U. \tag{10}$$

Now we cancel almost everything, and are left with:

$$D = k_B T/\gamma. \tag{11}$$

2

This equation is hugely important. Memorize it. It's called the Stokes-Einstein formula, after its discoverers. The Stokes-Einstein formula is a special case of something called the Fluctuation-Dissipation theorem, which we'll get to in a minute.

But first let's talk about what the Stokes-Einstein formula tells us. On the one hand, it says that the diffusion coefficient is proportional to $k_B T$. That's not so surprising, because we surmised that diffusion was due to thermal motion of molecules. More surprising, perhaps, is that D is inversely proportional to the drag coefficient of the particle in the solution. This is a macroscopic property of the particle, calculable from its geometry and a knowledge of hydrodynamics. For instance, I'll state without proof that the drag coefficient of a solid sphere in water is:

$$\gamma = 6\pi\eta a, \tag{12}$$

where η is the viscosity of the solution and a is the radius of the sphere. So for a sphere we have:

$$D = \frac{k_B T}{6\pi\eta a}. \tag{13}$$

This equation tells us that smaller things diffuse faster, and that increasing the viscosity of a solution can slow down diffusion. In general, the diffusion coefficient of any object scales inversely with its size. So if you have some complicated shaped object and you want to know its diffusion coefficient in water, you can always build a scale model out of wood or something, measure its drag in honey, and then scale your result to the radius of the real particle and the viscosity of water.

By the way, the viscosity of pure water at $20\,^\circ$C is: 1.0020×10^{-3} Pa·s.

There's another important consequence of the Stokes-Einstein formula. Up until the time of Einstein's 1905 paper on diffusion, nobody knew the value of Avogadro's number. Everywhere in classical thermodynamics, Avogadro's number always appears in the combination $R = N_A k_B$. People didn't know the magnitude of Boltzmann's constant either. If you take a particle of known size in a solution of known viscosity, and you track its motion, you can calculate D. Then the only unknown in the Stokes-Einstein equation is k_B. You solve for k_B, and then you get N_A for free. This is a really nice example of how you can learn more about a system by watching its spontaneous fluctuations than you can learn from just its average behavior.

4   Applications of Stokes Einstein

So far we've mostly been talking about translational diffusion, that is things moving from one place to another. Now we'll talk about some other forms of diffusion.

4.1 Fluorescence depolarization

When a small object is placed in solution, not only does it jiggle, but it also tumbles head over heels. Let's make an estimate of how long this tumbling takes. The rotational drag coefficient of a sphere of radius r is:

$$\gamma_r = 8\pi\eta r^3. \tag{14}$$

That is, if you are spinning the sphere at angular velocity ω, then the torque is:

$$\mathbf{T} = \gamma_r\omega. \tag{15}$$

Note that this has different units from the translational drag on a sphere. A simple way to think about this is that Torque $\sim$ Force $\times$ Distance, and $\omega = v/r$, where v is the velocity at the surface. This is where you get two extra factors of r in the denominator. The only thing this simple argument misses is the factor of 8 for rotation, compared to 6 for translation. To get this factor you'd have to do the hydrodynamics. We can use the formula for the rotational drag on a sphere to get the rotational diffusion coefficient, by exactly the same argument that led to the Stokes Einstein formula:

$$D_r = \frac{k_B T}{8\pi\eta r^3}. \tag{16}$$

This rotational diffusion coefficient has units of s^{-1}. For small rotations, θ, the mean square angle through which a diffusing particle rotates is:

$$\langle \Delta\theta^2 \rangle = 4D_r t. \tag{17}$$

Can you say why there is a 4 in front, rather than a 2?

One often works with the inverse of this quantity, $t_r \equiv 1/D_r$, called the rotational correlation time. The rotational correlation time has an interesting property: $t_r \propto r^3$, so for a series of objects of constant density, $t_r \propto M$, where M is the mass of the object. Most proteins have roughly the same density, $\rho \approx 1.3\mathrm{g/cm}^3$ (remember, steak sinks). So the rotational correlation time of a protein is roughly proportional to its molecular weight. What's the constant of proportionality? Well, it works out that if we measure the molecular weight in Daltons, and the rotational correlation time in picoseconds, then the constant is 1. So a molecule with a molecular weight of 1000 Daltons has a rotational correlation time of ~ 1 nanosecond.

The correspondence between molecular weight and rotational correlation time is the basis of the technique of fluorescence depolarization analysis for determining the weight of a protein or a protein complex. The idea is that you attach a fluorophore to the protein, and excite the fluorophore with linearly polarized light. The fluorophore gets excited, and then re-emits the photon a short while later. If the excited state lifetime of the fluorophore is long compared to the rotational diffusion time, then the polarization of the emitted photon is uncorrelated with the polarization of the incident photon. On the other hand, if

the excited state lifetime is short, then the two polarizations are correlated. By measuring the polarization of the emitted photons, you can get a pretty good estimate of the molecular weight of the protein. You can choose fluorophores with convenient lifetimes (typically ranging from a few ns to microseconds), and can adjust the viscosity of the solvent to accommodate proteins with a wide range of molecular weights. If fluorescence depolarization is performed in a microscope setup, one can image the distribution of micro-scale viscosity.

4.2 First passage problems

One often wishes to calculate the rate of some diffusion-limited process, e.g. a bimolecular reaction; or a molecule diffusing out of a spine head into a parent dendrite; or glucose being consumed by a bacterial colony. There is a nice trick for doing this. You arbitrarily select a molecule, and imagine following it until it reacts. The moment it reacts, you introduce another molecule, and so on. This process will set up a nonequilibrium steady-state concentration profile, with an absorbing boundary condition on the reacting surface. You can then calculate the rate of the reaction by looking at the gradient in concentration at the surface.

As a simple example, consider a 1-D problem with a reflecting boundary at $x = 0$ and an absorbing boundary at $x = x_0$. We release particles at $x = 0$ and we want to calculate the mean time for a particle to be absorbed at x_0. Every time the particle reacts at x_0, we release a new particle at $x = 0$. So there is always precisely one particle in the system. I propose that the steady-state concentration profile is given by a straight line of height $2/x_0$ at $x = 0$ and 0 at $x = x_0$. Let's check:

1. $\frac{\partial^2 p}{\partial x^2} = 0$, so this is a steady-state solution ($\frac{\partial p}{\partial t} = 0$).

2. $p(x_0) = 0$, satisfying the boundary condition at $x = x_0$

3. $\int p(x)dx = 1$, so the probability is normalized

The flux at the absorbing boundary (actually, everywhere!) is $j(x_0) = -D\frac{dp}{dx} = 2D/x_0^2$. The units of flux are 'particles per second', so to find the mean time per particle, we just take $1/j$, and get $t_0 = x_0^2/2D$. You'll test this prediction numerically on the homework.

Whereas the linear concentration profile might not be intuitive, similar results in other domains might be more familiar. In heat-flow, if you have a 1-D material that is heated at one end and cooled at the other, you get a linear temperature profile inside. If you have a 1-D region of space where one end is held at one potential and the other at a different potential, then the voltage between follows a line profile (e.g. in a parallel plate capacitor).

4.3 Diffusion-limited reaction rates

Another interesting application of the theory of diffusion is to the question of how quickly two particles in a solution can find each other, if they rely on

diffusion alone. This sets an upper speed limit on many biological processes. Consider two particles with diffusion coefficients D_A and D_B, and radii r_A and r_B. The 1-d distance between them varies with a diffusion coefficient given by $D_A + D_B$ (you might prove this on the next homework). Suppose that each of these species has some concentration in the solution, given by $[A]$ and $[B]$, and that if two of these ever bump into each other, they react. What's the reaction rate?

We go into the frame of reference of one of the A molecules. We imagine that this is an absorbing boundary, and we calculate the steady-state concentration profile of the other reactant around it. Then the flux of B molecules to it is, as with the bacterium eating oxygen in a pond,

$$\frac{dn_B}{dt} = -4\pi(D_A + D_B)(r_A + r_B)[B]. \tag{18}$$

This is gives the rate of disappearance of B due to one molecule of A. The total number of molecules of A is $n_A = [A] \times V$, where V is the volume of the container, so we multiply the right hand side by $[A] \times V$ to get the number of B disappearing per second. Then we divide by V to get the rate of change of concentration of B:

$$\frac{d[B]}{dt} = -4\pi(D_A + D_B)(r_A + r_B)[A][B]. \tag{19}$$

But something interesting happens if we plug in the Stokes Einstein relation for D_A and D_B:

$$\begin{aligned} \frac{d[B]}{dt} &= -\frac{4k_BT}{6\eta}\left(\frac{1}{r_A} + \frac{1}{r_B}\right)(r_A + r_B)[A][B] \\ &= -\frac{2k_BT}{3\eta}\frac{(r_A + r_B)^2}{r_Ar_B}[A][B]. \end{aligned} \tag{20}$$

Note that for aggregation of spheres of the same size (e.g. oligomerization), the diffusion limited rate is independent of the particle size!

$$\frac{d[B]}{dt} = -\frac{8k_BT}{3\eta}[A][B]. \tag{21}$$

Larger spheres diffuse more slowly, but they have a larger capture radius. Of course, if you want to keep the concentration of spheres constant, while increasing their radius, the volume fraction of the spheres will go up too. This shows the surprising fact that for like-sized particles, the diffusion-limited reaction rate constant is independent of particle radius R.

For water, the diffusion-limited reaction speed limit is:

$$k_{max} \approx 7 \times 10^9 \text{s}^{-1}\text{Molar}^{-1}. \tag{22}$$

Let's review the relation between reaction rates and steady state concentration profiles: The rate per molecule of A of B reacting with A is given by the

steady state solution to the diffusion equation, with A as an absorbing boundary. The argument was that the rate per molecule if you follow an individual molecule should be the same as the ensemble average rate. This is subtle, so think about it for a while. The key point is that even when an absorbing species has depleted the local concentration, the remaining molecules are still totally unbiased in their random walks.

A   Appendix: Equipartition Theorem

The equipartition theorem states that for any degree of freedom where the energy depends quadratically on the coordinate, at thermal equilibrium the mean energy is $\frac{1}{2}k_B T$.

Here's the proof. The energy is:

$$U(q) = \frac{1}{2}kq^2, \tag{23}$$

where k is some constant, and q is the coordinate of the particle. We all know that the probability of a particle occupying a state with energy $U(q)$ is:

$$P(q) \propto \exp[-U(q)/k_B T], \tag{24}$$

where the constant of proportionality is whatever it takes to make

$$\int_{-\infty}^{\infty} P(U(q))dq = 1. \tag{25}$$

So for our quadratic degree of freedom we have:

$$P(q) \propto \exp[-\frac{kq^2}{2k_B T}]. \tag{26}$$

We can get the constant of proportionality by noting that this probability distribution is a Gaussian. So:

$$P(q) = \sqrt{\frac{k}{2\pi k_B T}} \exp[-\frac{kq^2}{2k_B T}]. \tag{27}$$

Now let's evaluate the expectation value of the energy. This is given by:

$$\langle U \rangle = \int_{-\infty}^{\infty} \frac{1}{2}kq^2 P(q)dq. \tag{28}$$

Well, we know that the expectation value $\langle q^2 \rangle$ for a Gaussian probability distribution is just the variance:

$$\langle q^2 \rangle = \frac{k_B T}{k}, \tag{29}$$

from which we see that

$$\langle U \rangle = \frac{1}{2}k_B T. \tag{30}$$

There you have it: the Equipartition Theorem.

A.1 Applications of equipartition

Let's talk about some places where this shows up. If you have a Brownian particle on a spring, its mean *potential* energy is $\frac{1}{2}k_B T$ in one dimension. In two dimensions, it has $\frac{1}{2}k_B T$ along each dimension, so the energy is just $k_B T$, and in three dimension it's $\frac{3}{2}k_B T$. Note that the spring constant dropped out of the calculation. It doesn't matter how tightly the particle is bound: it *always* has $\frac{1}{2}k_B T$ of potential energy.

The kinetic energy of a particle is $KE = \frac{1}{2}mv^2$. So in one dimension the mean kinetic energy is $\frac{1}{2}k_B T$, and in 2-D it's $k_B T$, and in 3-D $\frac{3}{2}k_B T$.

The equipartition theorem has multitudinous important consequences. The molar specific heat at constant volume of a monatomic ideal gas is:

$$
\begin{aligned}
c_v &= \frac{\partial E}{\partial T} \\
&= \frac{3}{2}R.
\end{aligned}
\tag{31}
$$

(recall that $R \equiv N_A k_B$). If you have diatomic molecules that can rotate about two axes perpendicular to the axis of symmetry, then $c_v = \frac{5}{2}R$. If, furthermore, the temperature is high enough to excite the internal vibrations along the bond between the atoms, then you get two more degrees of freedom (a kinetic energy and a potential energy) and $c_v = \frac{7}{2}R$. In a free gas the atoms or molecules only have kinetic energy, no potential energy.

In a solid crystal, each atom has three translational degrees of freedom, each of which has a kinetic and a potential energy. Total number of quadratic coordinates: 6. So the molar specific heat should be $3R$. In fact, at high temperatures this is very nearly the case. This is called the law of Dulong and Petit. At lower temperatures (which for many materials includes room temperature), the fact that the vibrations are quantized becomes important, and this analysis breaks down. As a side note, Einstein tried to develop a quantum theory of the specific heat, but basically got it wrong, and later Debye got it right.

The Equipartition Theorem is the reason why $k_B T$ is used as the universal energy scale: everything has approximately $k_B T$ of energy (ok, actually $\frac{1}{2}k_B T$ per quadratic degree of freedom).

Of course energy doesn't always depend quadratically on a coordinate. But for any system at equilibrium, we can say that the energy is at a local minimum (if it weren't, the system would move until it were). This means that $\partial U/\partial q = 0$. So the first non-zero term in the Taylor series is the quadratic term, to which we can apply the equipartition theorem.