Chemistry 163, Lecture 2
Brownian motion and diffusion

Adam E. Cohen

Fall 2022

1 Intro

Today we'll talk about Brownian motion. Brownian motion occurs at thermal equilibrium. Life isn't at equilibrium. In fact, there's a saying the equilibrium equals death.

But nonetheless Brownian motion is very important. It is the dominant mechanism by which small things move around. Brownian motion is the default option for how any molecule gets from A to B in a cell. Many molecules move around by other means, so we will have to understand Brownian motion in order to identify when motion is consistent with this mechanism, and when we need to invoke or invent something else.

The first thing we learn in mechanics class is "Every object in a state of uniform motion tends to remain in that state of motion unless an external force is applied to it." This applies to macroscopic objects in the regime of low friction. It's a really good way of calculating the motion of planets, cannon balls, and falling apples. But this isn't at all a useful description of how things move on the micro-scale. Every physical theory has its range of validity, and Newton's laws are often not the most useful language for describing motion of microscopic particles, just as they are not the correct language for describing motion of relativistic particles.

For a long time people thought that only living things moved spontaneously. In 1828, Robert Brown decided to look at pollen grains under his spiffy new microscope. He saw pollen grains jiggling wildly, and he naturally assumed that they were swimming. So then he tested pollen from plants that had been dead for 100 years. This also jiggled. He thought, "Maybe this is a property of pollen from flowering plans". But pollen from moss also jiggled. He enumerates a huge range of materials he tested, eventually landing on the oldest and deadest thing he could lay his hands on: a piece of the Sphinx. Everything jiggled, so eventually Brown concluded that he had discovered a fundamental property of matter.

An interesting side-note is that Brown found a certain minimum apparent size of his particles in the microscope. Nothing was smaller than this minimum size. He thought he had discovered "molecules", the fundamental building

blocks of matter. Here he was wrong. The minimum size he could resolve was set by the optical resolution of his microscope, and was hundreds of times bigger than most molecules.

I posted Brown's paper on the web. It makes for an amusing read.

It took almost 80 years until Einstein explained the origin of Brownian motion as a consequence of the thermal motion of molecules of the solvent. I also posted Einstein's paper on the web. Next week we'll go over his derivation. For now, we'll take the motion as a given, and calculate some of its consequences. Next week we'll consider the microscopic origin.

SHOW VIDEO OF BROWNIAN MOTION

When I was a graduate student, I built a machine to stop the Brownian motion of a single molecule. I'll be happy to tell anybody who is interested about this during office hours sometime.

There's a saying that power corrupts, and powerpoint corrupts absolutely. Today I'll try using the blackboard. We'll see how it goes.

I'll start by going over some mathematical facts about probability distributions. Then we'll state, but not prove the Central Limit Theorem. Then we'll talk about macroscopic diffusion and some examples.

2 Brief review of probability

Consider a particle that takes steps along the x-axis, where each step, s_i is selected from a probability distribution $w(s)$. This distribution has finite mean and variance, but otherwise can be anything. Because $w(s)$ is a probability, we have:

$$\int_{-\infty}^{\infty} w(s)ds = 1. \tag{1}$$

For any quantity $F(s)$ we write its expectation value as $\langle F \rangle$, which is defined as

$$\langle F \rangle \equiv \int_{-\infty}^{\infty} F(s)w(s)ds. \tag{2}$$

Here is the physical meaning of the expectation value:

- You repeatedly select values of s from the distribution $w(s)$;
- For each value of s you calculate $F(s)$;
- You average the values of $F(s)$ over all these trials.

There are a few functions F whose expectation values we want to calculate so often that we have given these functions special names. The mean of the distribution is $\langle s \rangle$, defined as:

$$\mu = \int_{-\infty}^{\infty} sw(s)ds. \tag{3}$$

The variance is $\langle (s-\mu)^2 \rangle$, defined as:

$$\begin{aligned}
\sigma^2 &= \int_{-\infty}^{\infty} (s-\mu)^2 w(s) ds \\
&= \int_{-\infty}^{\infty} (s^2 - 2\mu s + \mu^2) w(s) ds \\
&= \int_{-\infty}^{\infty} s^2 w(s) ds - \mu^2 \qquad (4) \\
&= \langle s^2 \rangle - \langle s \rangle^2. \qquad (5)
\end{aligned}$$

These two properties of a distribution, the mean and variance, occur all over the place. You can also calculate higher moments like $\langle (s-\mu)^3 \rangle$ and $\langle (s-\mu)^4 \rangle$, which go by funny names like *skewness* and *kurtosis*, but these show up less often. The reason for this will become apparent when we get to the Central Limit Theorem.

3 Statistics of a random walk

Let's go back to our particle that has just taken one random step. Now suppose we take many of these steps in succession. What can we say about where the particle will be after N steps? As warm-up exercises we'll calculate the mean and variance after N steps. Then we'll discuss the whole distribution.

The position of the particle after N steps is:

$$x = \sum_{j=1}^{N} s_j \qquad (6)$$

The mean value of x is:

$$\begin{aligned}
\langle x \rangle &= \langle \sum_{j=1}^{N} s_j \rangle \\
&= \sum_{j=1}^{N} \langle s_j \rangle \\
&= N\mu. \qquad (7)
\end{aligned}$$

And the variance in its position is $\langle (x - N\mu)^2 \rangle$, which we'll write as $\langle \Delta x^2 \rangle$

$$\begin{aligned} \langle \Delta x^2 \rangle &= \langle \left(\sum_{j=1}^{N} s_j - \mu \right)^2 \rangle \\ &= \langle \left(\sum_{j=1}^{N} \Delta s_j \right)^2 \rangle \\ &= \langle \sum_{j,k=1}^{N} \Delta s_j \Delta s_k \rangle. \end{aligned} \tag{8}$$

Let's break the double sum into two parts: where $j = k$ and where $j \neq k$.

$$\langle \Delta x^2 \rangle = \sum_j \langle (\Delta s_j)^2 \rangle + \sum_{j \neq k} \langle \Delta s_j \Delta s_k \rangle. \tag{9}$$

The second term in the sum vanishes because for uncorrelated random variables we have:

$$\begin{aligned} \langle \Delta s_j \Delta s_k \rangle &= \langle \Delta s_j \rangle \langle \Delta s_k \rangle \\ &= 0 \times 0. \end{aligned} \tag{10}$$

So we're left with:

$$\begin{aligned} \langle \Delta x^2 \rangle &= N \langle (\Delta s_j)^2 \rangle \\ &= N\sigma^2. \end{aligned} \tag{11}$$

You see from this that the mean and *variance* of the distribution both grow linearly with N. The variance has units of length^2; to characterize the width of the distribution we typically consider the standard deviation:

$$\sqrt{\langle \Delta x^2 \rangle} = \sqrt{N}\sigma \tag{12}$$

This fact is extremely important whenever you're considering something that can be written as the sum of many random contributions.

Now for the whole distribution. To start, consider a baby random walk of just two steps. How can we calculate the probability distribution after these two steps? I claim that the probability distribution is:

$$P_2(x) = \int_{-\infty}^{\infty} ds_1 \int_{-\infty}^{\infty} ds_2 w(s_1) w(s_2) \delta(x - (s_1 + s_2)). \tag{13}$$

Where did this come from? Without the δ-function, this would be an integral over all possible values of the first step, and all possible values of the second step, so the integral would be 1. The δ-function constrains the values of s_1 and s_2 so that $s_1 + s_2 = x$. We can immediately evaluate the integral over s_2 to get:

$$P_2(x) = \int_{-\infty}^{\infty} ds_1 w(s_1) w(x - s_1). \tag{14}$$

If your distribution is discrete (e.g. coin tosses), then the integral is replaced by a sum.

Let's take a moment to discuss this expression. Expressions like this occur so often that we've given it a name: it's called a convolution. Convolutions have many wonderful and interesting properties. A convolution takes two functions as inputs, and gives you a third function as output. A shorthand way to write a convolution is:

$$f \otimes g.$$

Convolution has many of the properties of multiplication: it is commutative, associative, and obeys the distributive law. Let's consider a little example. Suppose your steps are determined by a coin toss, with equal probability of going to the left or the right.

[DRAW PROBABILITY DISTRIBUTION AFTER 1 AND 2 STEPS].

As you keep on adding more coin tosses, the distribution becomes increasingly smooth. The central limit theorem (CLT) states that the sum of a large number, N, of independent random variables each with finite mean μ_1 and variance σ_1 will follow a Gaussian distribution:

$$P_N(x) = \frac{1}{\sqrt{2\pi N\sigma_1^2}} \exp\left[-(x - N\mu_1)^2/2N\sigma_1^2\right]. \tag{15}$$

So the mean after N steps is $N\mu$ and the variance is $N\sigma_1^2$. This theorem is not obvious, but is hugely important. In case you're curious, I put a proof of this theorem in the Appendix of the lecture notes.

What's the point? Whatever your distribution of individual steps (be they coin tosses, drunkards steps, molecular motions, bacterial saccades), after enough steps your probability distribution is Gaussian, with mean and variance proportional to the mean and variance of the single steps, and both linearly increasing with time. Alternatively, this says that if you convolve a function with itself enough times, it starts to look like a Gaussian (see homework).

4 Diffusion

Now consider a particle diffusing in solution. We divide the path into short steps, each of length Δt. If the particle is diffusing for a total time t, the number of steps is $N = t/\Delta t$. Now each of these steps might be composed of a vast number of smaller steps, but the beauty of the central limit theorem is that we'll get the same statistical properties of the motion for *any* time Δt long compared to the momentum relaxation time of the particle. As we make Δt shorter, the length of each step also gets shorter, as:

$$\langle \Delta s^2 \rangle = 2D\Delta t, \tag{16}$$

where D is a number, called the diffusion coefficient, that describes how wiggly the motion is. For now we'll take D as an experimentally measured quantity. Later I'll show you how to calculate it in a simple and elegant way. That factor

of 2 is just a convention. We put it in now in order to get rid of some factors of 2 in equations we'll come to in a bit. The units of D are length^2/time. The point of dividing things up this way is that now the statistics are independent of our choice of Δt. To see this, consider the variance in the position after N steps:

$$\begin{aligned} \langle (\Delta x)^2 \rangle &= N \langle \Delta s^2 \rangle \\ &= 2DN\Delta t \\ &= 2Dt. \end{aligned} \tag{17}$$

In the absence of external forces, the average position of a particle does not move, so the probability distribution for a Brownian particle released at the origin and diffusing in one dimension is:

$$P(x) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{x^2}{4Dt}\right] \tag{18}$$

The variance of this distribution is:

$$\langle \Delta x^2 \rangle = 2Dt. \tag{19}$$

This is probably the most important equation in the entire course. Memorize it.

If you have diffusion in more than one dimension, the diffusion along different axes is statistically independent. So the total probability distribution in 2-D, for instance is:

$$P(x, y) = P(x)P(y). \tag{20}$$

In two dimensions,

$$\begin{aligned} \langle \Delta \mathbf{r^2}_{2D} \rangle &= \langle \Delta x^2 \rangle + \langle \Delta y^2 \rangle \\ &= 4Dt. \end{aligned} \tag{21}$$

Similarly, for 3-D diffusion $\langle \Delta \mathbf{r^2}_{3D} \rangle = 6Dt$.

5 Examples

Let's calculate a little example. Diffusion limits the size of many organisms. Prokaryotes (bacteria) don't have internal motors to carry things around–they rely entirely on diffusion. A typical bacterium is 1 μm in size. A typical diffusion coefficient for a protein in cytoplasm is $D \approx 0.25 \mu m^2/s$. So the time for a protein to diffuse across a bacterium is of order 2 s. If bacteria were much larger it would take too long for proteins to get from one side to the other. If a bacterium were 10 microns in diameter, it would take a protein 200 s to diffuse across.

But the cells in Eukaryotes (us) are much bigger than a micron–typically they're more like 10 μm. One way that Eukaryotic cells solved this transport problem was by inventing molecular motors, which carry protein cargoes around

the cell. Another approach was by inventing organelles. These are smaller subcompartments within the cell where different biomolecules can be localized, or have enhanced concentrations, facilitating certain reactions. When biomolecules stick to each other or to other structural elements in the cell, this can enhance local concentrations. For instance, some reactions happen in membrane-associated proteins. Adhesion to a 2-D membrane can dramatically enhance reaction rates. Other reactions happen on DNA. Adhesion to DNA can also enhance rates of encounter between DNA-binding proteins. (Though when you have adhesion to a membrane, a piece of DNA, or a cytoskeletal element, one has to balance the increased local concentration against the decreased diffusion of the partially stuck components). There is a lot of current interest in so-called liquid drops in cells: liquid-liquid phase separations where some molecules (typically proteins and RNA) assemble into small liquid droplets bound together by weak intermolecular interactions.

Nonetheless, Eukaryotic cells still rely on diffusion to transport small molecules, glucose and oxygen. These molecules have diffusion coefficients of $D \approx 100 \mu m^2/s$ in cytoplasm. The time for a small molecule to diffuse across a Eukaryotic cell is roughly 0.5 s.

Insects rely on diffusion to carry oxygen through their spiricles to their internal organs. The diffusion coefficient of oxygen in air is $D \approx 0.25 cm^2/s$. So if a typical insect is 1 cm in size, it takes about 2 s for oxygen to get inside. If they were much bigger they'd have trouble getting enough oxygen.

We (meaning vertebrates), solved this problem by having heart and lungs to pump oxygen around the body. The spacing of capillaries in our tissues is determined by balancing the rate of oxygen consumption by tissue with the rate of diffusion of oxygen through the intervening tissue to the blood vessel. In cancer, one cell type starts to proliferate out of control. Once the cell mass exceeds a certain size, the center becomes starved of oxygen. This can lead to necrosis, or cell death in the center of the tissue, and limits the size of the cancer. On the other hand, the body has mechanisms to detect oxygen-starved tissue, and with some tumors will build blood vessels to replenish the supply. This process is called angiogenesis and can lead to unrestricted tumor growth and ultimately metastasis. One approach to tumor therapy is to try to restrict the growth of blood vessels. You'll have a homework problem on this later in the course.

Another application of our description of diffusion is to the motion of bacteria. If you put a capillary containing a solution of attractant (e.g. aspartate) in a solution of bacteria, the bacteria congregate near the mouth of this capillary. How do the creatures know where to go? In the early 1970's Howard Berg (a former professor at Harvard) analyzed this problem using a microscope that could automatically track the motion of a single bacterium. He found that the bacteria underwent a biased random walk. They swam in a straight line at about 14 μm/s for a bit less than a second, then did a random "twiddle", and started off in a new direction.

[SHOW MOVIE OF BACTERIAL TWIDDLING]

He then found that when the bacteria were swimming up a concentration

gradient, they were less likely to twiddle, and when they were swimming down a concentration gradient they were more likely to twiddle. Thus the distribution of step sizes became biased towards the higher concentration of chemoattractant. This model is in quantitative agreement with the distributions of bacteria around sources of attractant. Berg spent much of his career investigating how bacteria know whether their life is getting better or worse, i.e. how bacteria remember.

6 Continuum description

So far we have been considering the motion of a single particle. Now we will adopt a different approach, and describe the concentration, $c(x,t)$ of a collection of Brownian particles.

What happens if we start with many particles, with some initial density profile $c_0(x)$? If the concentration of the particles is low enough, they don't hinder each other's motion. Each particle diffuses as if there weren't any others around. We can write the total concentration as the superpositions of the concentration profiles originating from each point in the solution. That is:

$$c(x,t) = \int_{-\infty}^{\infty} dx' c_0(x') \frac{1}{\sqrt{4\pi D t}} \exp -\frac{(x-x')^2}{4Dt}. \tag{22}$$

But look! Suppose we define

$$G(x,t) \equiv \frac{1}{\sqrt{4\pi D t}} \exp -\frac{(x)^2}{4Dt}. \tag{23}$$

Then the concentration at time t is just a convolution:

$$c(x,t) = G(x,t) \otimes c_0(x) \tag{24}$$

The quantity $G(x,t)$ is called the Greens Function and is a tool widely used in solving linear partial differential equations. There are a lot of other mathematical tricks for solving diffusion equations, some of which can be more efficient computationally–but the Green's function is the most intuitive approach.

6.1 Diffusion Equation

Now we're going to step back and develop a different language for describing diffusion. Suppose you have some concentration profile $c_0(x)$. Over time, the particles will go from regions of higher concentration to lower concentration, eventually leading to a uniform concentration everywhere. We want to calculate how this happens. Suppose you have a solution with an initial concentration profile $c(x)$, and you divide it up into very thin slices, of width Δx. In a time step Δt, half the particles in each slice go to the left, and half go to the right (imagine that each particle flips a coin). The width of the slice, Δx, and the time to randomize the particles in it, Δt, are related by $\Delta x = \sqrt{(2D\Delta t)}$.

The flux (number of particles per unit area) across the line separating the slice at x from the slice at $x + \Delta x$ is:

$$j = \frac{\frac{1}{2}[c(x) - c(x + \Delta x)]\Delta x}{\Delta t}. \tag{25}$$

We can rearrange this equation (multiplying on the right by $\Delta x/\Delta x$) to get

$$j = -\left(\frac{\Delta x^2}{2\Delta t}\right)\frac{c(x + \Delta x) - c(x)}{\Delta x}. \tag{26}$$

We defined the diffusion coefficient as $\langle \Delta x^2 \rangle / 2\Delta t$. And if this process were happening in 3-d rather than 1-d, the derivative would become a gradient. So the flux of particles due to diffusion is:

$$\mathbf{j} = -D\nabla c. \tag{27}$$

This equation is known as Fick's law. We can write the fact that if more particles flow into a region than flow out of it, the concentration must increase within that region as:

$$\frac{\partial c}{\partial t} = -\nabla \cdot \mathbf{j} \tag{28}$$

Combining these two equations (and assuming that D does not vary with position!) gives:

$$\frac{\partial c}{\partial t} = D\nabla^2 c. \tag{29}$$

This is the diffusion equation. We will come back to this many, many times during the course, and decorate it with other terms to account for electric fields, chemical reactions, and fluid flow.

To solve a PDE, you need to specify the boundary conditions. There are a few possibilities. If you have an impermeable barrier, there can be no flux of particles across it. This means that

$$\hat{n} \cdot \nabla c = 0 \tag{30}$$

at the boundary. Alternatively you can clamp the concentration. Then c is fixed, and the flux at the boundary may be anything.

How does one measure the diffusion coefficient? There are three main techniques: single-molecule tracking, Fluorescence Recovery After Photobleaching (FRAP), and Fluorescence Correlation Spectroscopy (FCS). Single-molecule tracking is conceptually simple. You watch many individual particles diffuse in the microscope, and you track their positions using automated software. Once you have the (x, y) coordinates as a function of time, you can plot $\langle \Delta x^2 \rangle + \langle \Delta y^2 \rangle$ vs. t. If you're tracking in two dimensions, the slope of this line is $4D$. There are some technicalities you need to be careful of when doing this, like how to account for errors in your localization measurements; what do to if the molecules are confined or run into boundaries; how to compensate for flow; and what to

do if your population of molecules is heterogeneous. But these are details for the afficionados.

The second technique is called FRAP. This makes use of the fact (which is usually annoying) that fluorescent molecules photobleach, or go dark, if exposed to too much light. You can use a targeted laser beam to bleach the molecules in a small spot, and then watch the bleached hole fill in over time. By fitting the recovery trace to a numerical or analytical simulation of the diffusion equation, one can get the diffusion coefficient. Here the main trick is to make sure your simulation of diffusion has the correct geometry and boundary conditions. If you bleach molecules in a weird pattern, inside of a weirdly shaped cell, then the dynamics of the recovery might not fit a simple formula. When you do the fitting, you need to pay attention to whether you are fitting 1D, 2D, 3D or mixed diffusion. [SHOW EXAMPLE OF FRAP EXPERIMENT]

The third approach is FCS. FCS lies between single-molecule and bulk techniques: you focus a laser beam to a small spot in a dilute solution (typically $\sim$1 nM). As molecules diffuse in and out of the spot, there are fluctuations in the number of molecules in the spot, which leads to fluctuations in the fluorescence brightness. If the mean number of molecules in the spot is $\langle n \rangle$, the fluctuations are of amplitude $\sqrt{\langle n \rangle}$. [PROOF: Suppose you have a container with a total of N molecules, and your laser beam samples a tiny fraction of total volume $\epsilon = \frac{V_{focus}}{V_{tot}} \ll 1$. The probability that a particular molecule is in the laser focus is just $\langle n_1 \rangle = \epsilon$, so the mean occupancy of the laser focus is $\langle n \rangle = \sum_{i=1}^{N} \epsilon = N\epsilon = CV_{focus}$, where $C = N/V_{tot}$ is the concentration. The variance in the single-particle occupancy is $\sigma_1^2 = \langle n_1^2 \rangle - \langle n_1 \rangle^2 = \epsilon - \epsilon^2 \approx \epsilon$. The variance in the total occupancy is $\sigma_n^2 = N\epsilon = \langle n \rangle$.] The duration of these fluctuations is proportional to the time it takes a molecule to diffuse across the laser spot. So if you know the geometry of your laser spot, and you can measure these fluctuations, then you can calculate the diffusion coefficient. Typically you don't know the laser spot geometry well enough, so instead you first calibrate with a dye of known diffusion coefficient, then measure your unknown molecules. As in the other techniques, there can be complications: if your test molecule fluctuates in intensity (some molecules blink on and off), or if there is any flow, or if the diffusion coefficients of the fluorescent species are heterogeneous, then these FCS measurements get more complex to interpret. We'll talk about FCS in more detail later in the semester.

7 Diffusion-like equations in physics

We've been writing a lot of equations, so now I'd like to relate diffusion to other things you may have studied, maybe in other physics or chemistry classes. Let's start by comparing Fick's law to the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi. \tag{31}$$

If there is no external potential, then we can divide through by $i\hbar$ to get

$$\frac{\partial}{\partial t}\psi = \frac{i\hbar}{2m}\nabla^2\psi, \tag{32}$$

which you see is identical to the diffusion equation if we make the substitution $D \to i\hbar/2m$. That might seem like a strange thing to do, but these are just numbers. All of the mathematical tricks you learned for solving the Schrödinger equation can also be applied to solving the diffusion equation (when both have zero potential–turning on a potential ruins the analogy).

As an interesting aside, if you have particles that are diffusing and replicating or decaying at a constant rate, (e.g. a suspension of bacteria, or a solution of an unstable molecule), then there is an analogy to the full Schrödinger equation, with a potential. You just make the substitution $k \to -iV/\hbar$, where k is the growth or decay rate.

This means that all of the techniques you learned for solving Schroedinger's equation can be used for the diffusion equation: separation of variables, Fourier expansions, perturbation theory, and so on. If analogies to quantum mechanics don't make things more intuitive, don't worry about it.

The diffusion equation shows up in some other places too. Heat propagates through a solid in exactly the same way that particles diffuse. The temperature in a solid obeys:

$$\frac{\partial T}{\partial t} = k\nabla^2 T, \tag{33}$$

where k is called the *thermal diffusivity*. It has the same units as the diffusion coefficient. For water, $k = 0.143mm^2/s = 0.143 \times 10^{-6}m^2/s$.

When you have a fluid flow that happens to lie completely in a plane, you can define the vorticity, ω, by $\omega = \nabla \times \mathbf{v}$, where v is the flow profile. Under certain conditions the vorticity obeys a diffusion equation, where

$$\frac{\partial \omega}{\partial t} = \kappa\nabla^2\omega, \tag{34}$$

and κ is called the kinematic viscosity and has the units of a diffusion coefficient. The kinematic viscosity of water is about $10^{-6}m^2/s$.

If you have a collection of sources or sinks of particles connected to a large reservoir, it's possible to set up concentration profiles that are not everywhere uniform. At long time you reach *steady state* but not equilibrium. That is, you have fluxes of particles, but the concentration isn't changing anywhere. In this case Fick's law becomes:

$$\nabla^2 c = 0. \tag{35}$$

Notice that the diffusion coefficient doesn't appear here. D determines how long it takes to reach steady state, but not what the ultimate state is. Solutions to this equation are called harmonic functions. The height of a soap bubble on a wire frame is given by a harmonic function. There are many useful theorems about harmonic functions. One that is particularly useful is that all the maxima and minima of such a function lie on the boundaries. This means that if you

have a steady state concentration profile, the extrema of the profile cannot be out floating in free space. This should make sense, because a concentration maximum would diffuse away, while a concentration minimum would get filled in.

When you studied electricity and magnetism, you may have come across Poisson's equation in electrostatics:

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}. \tag{36}$$

Well, if the charge density ρ is zero in some region of space, then a solution to the electrostatics problem in that region also gives a solution to the steady state diffusion equation.

So suppose you have a roughly spherical cell of radius R producing a signalling molecule in a stationary, homogeneous medium. If the concentration of the molecule at the surface of the cell is c_0 and very far from the cell the concentration is 0, then I can immediately tell you that the steady state concentration at radius r from the cell is:

$$c(r) = \frac{c_0 R}{r}. \tag{37}$$

So to solve steady state diffusion problems we can use the whole toolbox of electrostatics: Gauss' law, separation of variables, image charges.

For example, here is a challenge problem: consider a spherical source of particles near an absorbing planar boundary. How could you calculate the steady-state concentration of particles everywhere in space? Now consider a spherical source of particles near a reflecting boundary: what's the steady-state particle distribution? (Hint: map these onto electrostatics problems and use the method of images!).

Let's end by considering the example of a non-motile bacterium sitting in a pond, consuming oxygen. Far from the bacterium the concentration of oxygen is a constant C_0. The radius of the bacterium is R. If the bacterium needs a lot of oxygen, the concentration at its surface will be low. Let's say it's zero. So $C(R) = 0$. What's the steady state oxygen concentration everywhere in space? Well we can solve this by analogy to the electrostatics of a charged sphere. We know that $C(r) \sim 1/r$ is a valid spherically symmetric solution to Poisson's equation. So we just have to match boundary conditions to get:

$$C(r) = C_0(1 - R/r). \tag{38}$$

To get the flux of oxygen to the bacterium we need the gradient at its surface. We know that the gradient is only in the $\hat{r}$ direction, so we have

$$\begin{aligned} \nabla C &= \frac{\partial C}{\partial r}\hat{r} \\ &= \frac{C_0 R}{r^2}\hat{r}. \end{aligned} \tag{39}$$

We're interested in the gradient at the surface, so

$$\mathbf{J} = -\frac{C_0 D}{R}\hat{r}. \tag{40}$$

The total flux of oxygen is the gradient times the area, $4\pi R^2$, or

$$\Phi = 4\pi C_0 D R. \tag{41}$$

This is another limit on the size of creatures that live by diffusion. The metabolic needs grow as R^3, but the nutrients only arrive at a rate that grows with R.

We all know from electrostatics that sharp points concentrate electric fields. Electric field is the gradient of the potential, so it plays an analogous role to mass flux in diffusion. So if you have an absorbing object with some complex shape, the diffuse flux density is highest towards sharp points on the object. Most eukaryotic cells have a structure called a primary cilium, which is a little rod that pokes out of the cell surface. There have been many debates about what the primary cilium is for, but one role seems to be in chemical sensing. Many ligand-gated receptors decorate the surface of the primary cilium. If you have a change in the ambient concentration of a signaling molecule, receptors on a lightning rod-shaped cilium will respond faster than receptors on a flat surface. I've posted on the website a nice 2021 paper analyzing how geometry affects the sensing properties of primary cilia.

There's a famous paper by Ed Purcell (Harvard Prof. who discovered NMR) called "Life at low Reynolds Number" which analyzes nutrient and oxygen transport to a bacterium. Many of the same arguments apply to the primary cilium. In both cases, there's a way around the diffusion limit on binding substrates: move through the solution. By either swimming (in the case of bacteria) or beating (in the case of flagella), you can go to the molecules, rather than waiting for the molecules to come to you.

In the case of electrodeposition of copper from $CuSO_4$ solutions this leads to pretty fractal patterns. If diffusion of nutrients limits the growth bacteria in a colony, you can get strikingly similar patterns.

[SEE PICTURE ON PPT]

A Proof of the Central Limit Theorem

We start by proving a surprising and very useful relation between convolutions and Fourier transforms.

The δ-function can be expressed as:

$$\begin{aligned} \delta(x - s_1 - s_2) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{ik(s_1+s_2-x)} \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{iks_1} e^{iks_2} e^{-ikx}. \end{aligned} \tag{42}$$

Substituting this into the expression for $P_2(x)$ gives:

$$P_2(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-ikx} \int_{-\infty}^{\infty} ds_1 w(s_1) e^{iks_1} \int_{-\infty}^{\infty} ds_2 w(s_2) e^{iks_2}. \tag{43}$$

This equation is amazing. The two integrals over s_1 and s_2 are just the expressions for the Fourier transforms of the probability distributions $P(s_1)$ and $P(s_2)$, respectively. And the first integral is just the formula for the inverse Fourier transform. This equation shows that to convolve two functions, all you have to do is take their Fourier transforms, multiply them together, and then take the inverse Fourier transforms. Convolutions are convoluted; but Fourier transforms and multiplication are our friends.

In summary, the mathematical fact we just established is, if:

$$F(x) = G(x) \otimes H(x)$$

then

$$\tilde{F}(k) = \tilde{G}(k) \times \tilde{H}(k), \tag{44}$$

where the $\tilde{}$ indicates the Fourier transform.

Back to the hog. If we now consider a walk composed of N steps, we would have to convolve the distributions from each of the N steps. Alternatively we can work in the Fourier domain. The Fourier transform of the probability distribution after N steps is:

$$\tilde{P}(k) = \prod_{j=1}^{N} \tilde{w}_j(k). \tag{45}$$

We assume that all the $\tilde{w}_j$ are the same, whence the equation becomes

$$\tilde{P}(k) = \tilde{w}(k)^N. \tag{46}$$

Now comes the slightly slippery bit of the argument, so hang on tight. If $w(s)$ has finite variance, then $\tilde{w}(k)$ goes gradually to zero as $|k|$ gets large. For large N this means that $\tilde{w}(k)^N$ decreases very rapidly. So let's expand $\tilde{w}(k)$ about $k = 0$. Recall the definition of the Fourier transform:

$$\tilde{w}(k) = \int_{-\infty}^{\infty} ds w(s) e^{iks}. \tag{47}$$

Taylor expand the exponential to get

$$\begin{aligned} \tilde{w}(k) &= \int_{-\infty}^{\infty} ds w(s)(1 + iks - \frac{1}{2}k^2 s^2 + \ldots). \\ &= 1 + ik\mu - \frac{1}{2}k^2\langle s^2 \rangle. \end{aligned} \tag{48}$$

Rather than expanding $\tilde{w}(k)^N$, we'll expand $\ln(\tilde{w}(k)^N)$, and then take the exponential of the result. This is a very useful trick for sharply varying functions for which the regular Taylor series does not converge. This trick is called the "cumulant expansion". So:

$$\begin{aligned} \ln(\tilde{w}(k)^N) &= N \ln(\tilde{w}(k)) \\ &\approx N \ln(1 + ik\mu - \frac{1}{2}k^2\langle s^2 \rangle) \end{aligned} \tag{49}$$

. Next we'll use the Taylor series of the logarithm, to second order:

$$\ln(1+y) = y - \frac{1}{2}y^2 + \ldots \tag{50}$$

So we get:

$$\begin{aligned} \ln(\tilde{w}(k)^N) &\approx N(ik\mu - \frac{1}{2}k^2\langle s^2\rangle - \frac{1}{2}k^2\mu^2). \\ &= N(ik\mu - \frac{1}{2}k^2\langle \Delta s^2\rangle). \end{aligned} \tag{51}$$

Putting this back in the exponential yields:

$$\tilde{P}(k) = \exp\left[iNk\mu - \frac{1}{2}Nk^2\langle \Delta s^2\rangle\right]. \tag{52}$$

So the probability distribution is:

$$P(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dk \exp\left[-ikx + iNk\mu - \frac{1}{2}Nk^2\langle \Delta s^2\rangle\right]. \tag{53}$$

This is a standard Gaussian integral whose solution you can look up in a book. The result is:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-(x-N\mu)^2/2N\langle \Delta s^2\rangle\right]. \tag{54}$$