

## Lecture 2: Diffusion

### 1 Introduction

If you put a drop of red dye in water, it will slowly diffuse throughout the water. Why does this happen? How fast does it happen? What is going on microscopically?

The microscopic mechanism of diffusion is very simple: the dye molecules start densely concentrated near one point. Then they get bumped by neighboring molecules until they are spread out all over. To model this process, we can suppose that the dye molecule moves a distance  $\ell$  between collisions and after each collision its direction is completely randomized. This approximation is called a **random walk**. Although the distance  $\ell$  between collisions has some variation and the direction of scattering is somewhat correlated with the initial direction, because molecules collide billions of times per second, the law of large numbers applies to their net displacement and random walks provide an excellent approximation to real diffusion.

Random walks are actually quite common. They can be used to model any stochastic process. For another example, say you're playing blackjack with a friend. You are both expert players and evenly matched. Sometimes you win, sometimes she wins. Each time you play, you bet 1 dollar. This is a 1-D random walk. Say you play  $N$  games. Although we can't say who will be winning after  $N$  games, we can predict how by much they would be winning.

The 2-dimensional random walk is sometimes called the **drunkard's walk**. The idea is that a drunkard leaves a party late at night, takes a step in one direction, then gets totally disoriented and takes a step in another direction. How far will she get after  $N$  steps?

### 2 1D random walk

Let's work out the blackjack problem. We'll make it a little more interesting. Say you have a probability  $a$  of winning and your opponent has a probability  $b = 1 - a$  of winning. If you play  $N$  times, the chance of you winning  $m$  of them is

$$B_N(m) = a^m b^{N-m} \binom{N}{m} \quad (1)$$

This is known as the **binomial distribution**. The factor

$$\binom{N}{m} = {}_N C_m = \frac{N!}{m!(N-m)!} \quad (2)$$

is known as the **binomial coefficient**. It is the number of ways of picking  $m$  of the games for you to have won out of the  $N$  total games.<sup>1</sup>

Binomial coefficients come up in the expansion of powers of sums. Namely

$$(a+b)^N = \sum_{m=0}^N a^m b^{N-m} \binom{N}{m} \quad (3)$$

So the binomial distribution is simply the  $m^{\text{th}}$  term in this sum. In fact, this relationship makes it easy to see that the probabilities sum to 1:

$$\sum_{m=0}^N B_N(m) = \sum_{m=0}^N a^m b^{N-m} \binom{N}{m} = (a+b)^N = 1 \quad (4)$$

---

1. To see this, first note that if we want to sort  $N$  objects,  $N$  of them can go first, then  $N-1$  second, and so on, so there are  $N!$  permutations. If we are instead splitting into a set of size  $m$  and a set of size  $N-m$ , we don't care about the order within each set. So we have to divide by the permutations of each set, giving  $\frac{N!}{m!(N-m)!}$ .

since  $b = 1 - a$ .

How much can you expect to be winning after  $N$  games? This is determined by the number of games you win, namely the expected value of  $m$ :

$$\langle m \rangle = \sum_{m=0}^N m B_N(m) = \sum_{m=0}^N m a^m b^{N-m} \binom{N}{m} \quad (5)$$

Although  $b = 1 - a$  we can compute the sum on the right most easily if we allow  $a$  and  $b$  to be unrelated. Then, we note that this sum is the same as the sum in Eq. (4) if we differentiate with respect to  $a$ , then multiply by  $a$ . So we have

$$\sum_{m=0}^N m a^m b^{N-m} \binom{N}{m} = a \partial_a \left[ \sum_{m=0}^N a^m b^{N-m} \binom{N}{m} \right] = a \partial_a (a+b)^N = N a (a+b)^{N-1} \quad (6)$$

Now that we can compute the sum for any  $a$  and  $b$  we can take the case of interest where  $b = 1 - a$  and get

$$\langle m \rangle = N a \quad (7)$$

Similarly, the standard deviation is

$$\sigma_m = \sqrt{\sum_{m=0}^N (m^2 - \langle m \rangle^2) B_N(m)} = \sqrt{N a b} \quad (8)$$

As  $N \rightarrow \infty$  the binomial distribution approaches a Gaussian, by the central limit theorem. Thus knowing the mean and standard deviation, we know the whole answer:

$$B_N(m) \rightarrow \frac{1}{\sqrt{2\pi N a b}} \exp\left[-\frac{(m - N a)^2}{2 N a b}\right] \quad (9)$$

We can derive this by studying the falloff of  $\ln B_N(m)$  at large  $N$ , but using the central limit theorem is easier. Try checking the agreement of Eqs. (1) and (9) for some numerical values.

For the blackjack game,  $\langle m \rangle = N a$  is the expected number of times you win. The expected number of times you lose is  $\langle m \rangle = N$ . If you gain \$1 when you win, or lose \$1 when you lose, then

$$\langle \text{winnings} \rangle = N(a - b) \times \$1 \quad (10)$$

and the standard deviation is (try it to test your understanding):

$$\sigma_{\text{winnings}} = 2\sqrt{N a b} \times \$1 \quad (11)$$

For a fair match  $a = b = \frac{1}{2}$  and so the expected winnings are  $\langle \text{winnings} \rangle = 0$  with standard deviation

$$\sigma = \sqrt{N} \quad (12)$$

That  $\sigma$  grows as  $\sqrt{N}$  is exactly what we expect for the sum of random values by the central limit theorem. The  $a = b = \frac{1}{2}$  case is sometimes called an **unbiased 1D random walk**.

For example, if you play 100 games for \$1 each and are evenly matched, then  $\langle \text{winnings} \rangle = 0$  and  $\sigma_{\text{winnings}} = \sqrt{100} \times \$1 = \$10$ . This means that after 100 games, we don't know who's winning but there is a 32% chance someone is up by at least \$10.

For an unbiased 1D random walk, the mean displacement is 0. In this case, the typical scale for displacement is better described by the RMS fluctuation, which reduces to the standard deviation when the mean is zero. That is, the RMS fluctuation is  $x_{\text{rms}} = \sigma = \sqrt{N}$ . Typically, RMS fluctuations are used for quantities that average to zero (as in a unbiased random walk), but there is no hard and fast rule about when to use the mean displacement and when to use the RMS fluctuation. Just like there are many measures of uncertainty ( $\sigma$ , fraction within 50% of mean, width at half maximum, etc), there are many measures of displacement ( $\bar{x}$ ,  $x_{\text{rms}}$ , etc.). Typically these are all similar (when nonzer) and determined by dimensional analysis, so you can just substitute one for another, often depending on which leads to the simplest looking final expression.

Let us compare the binomial distribution to the Poisson distribution. The binomial distribution  $B_N(m)$  is defined for discrete  $N$  and  $m$ , in contrast to the Poisson distribution  $P_m(t) = \frac{(\lambda t)^m}{m!} e^{-\lambda t}$  which has discrete  $m$  but continuous  $t$ . For a binomial distribution, the smallest interval is one discrete step, with probability of occurrence  $a$ . For Poisson, we can take an arbitrarily small timestep  $\Delta t$  with probability  $dP = \lambda \Delta t$ . If we identify the interval for the binomial distribution with that of the Poisson distribution by setting  $a = \lambda \Delta t = \lambda \frac{t}{N}$ , and then take the limit  $N \rightarrow \infty$  we find

$$\lim_{N \rightarrow \infty} B_N(m) = \lim_{N \rightarrow \infty} \underbrace{\binom{N}{m}}_{\text{pick } m \text{ to decay}} \underbrace{\left(1 - \lambda \frac{t}{N}\right)^{N-m}}_{N-m \text{ do not decay}} \underbrace{\left(\lambda \frac{t}{N}\right)^m}_{m \text{ decays}} = \frac{(\lambda t)^m}{m!} e^{-\lambda t} \quad (13)$$

which formally recovers the Poisson distribution from the binomial distribution. Note that this requires taking  $N \rightarrow \infty$  holding  $\lambda t = Na$  fixed, and therefore  $a \rightarrow 0$  and  $b \rightarrow 1$ . In other words, a Poisson process is like a random walk that always goes in one direction but you don't know when the step will be taken.

Although they can be related, as we have seen, you should really think of binomial and Poisson distributions as being relevant in different contexts: binomial is used when the steps are discrete and incoherent (random directions) and Poisson is used when time is continuous but the steps are coherent (counts always increase). If we are flipping coins, then  $a = \frac{1}{2}$  is fixed, and so the Poisson distribution is not relevant since it needs  $a = 0$ . For a decay process, a decay can happen at any time  $t$  and so the binomial distribution is not appropriate. After a given time  $t$ , a Poisson process can have potentially an infinite number of events. With a binomial process the time is the number of steps  $N$ , so the number of possible events is always bounded.

## 2.1 Random walks in 2D and 3D

For the 2D case, a popular picture of the random walk is a drunkard stumbling around. In each time step she moves a distance  $L$  in some random direction. In 3D you can imagine a dye molecule diffusing in water and in each time step it bumps into something, and then gets buffeted into a different direction. For simplicity, we'll assume in the 2D and 3D cases that the distance is the same each step and the angle totally random. Where will the drunkard or molecule be after  $N$  steps?

Let us say that in the  $j^{\text{th}}$  step she moves by a displacement  $\vec{\ell}_j$ . The vectors  $\vec{\ell}_j$  all have length  $\ell$ . The dot product of two vectors is

$$\vec{\ell}_j \cdot \vec{\ell}_k = \ell^2 \cos \theta_{jk} \quad (14)$$

where  $\theta_{jk}$  is the angle between the two steps. Since we are assuming the angle is random, then the expectation value of this dot product is zero:

$$\langle \vec{\ell}_j \cdot \vec{\ell}_k \rangle = \ell^2 \frac{1}{\pi} \int_0^\pi d\theta \cos \theta = 0 \quad (15)$$

Now let  $\vec{s}_N$  be the total displacement from the origin after  $N$  timesteps.

$$\vec{s}_N = \sum_{j=1}^N \vec{\ell}_j \quad (16)$$

Then

$$\langle \vec{s}_N^2 \rangle = \langle (\vec{s}_{N-1} + \vec{\ell}_N)^2 \rangle = \langle \vec{s}_{N-1}^2 \rangle + 2 \langle \vec{s}_{N-1} \cdot \vec{\ell}_N \rangle + \langle \vec{\ell}_N^2 \rangle \quad (17)$$

Now, the angle between  $\vec{\ell}_N$  and  $\vec{s}_{N-1}$  is totally random, so  $\langle \vec{s}_{N-1} \cdot \vec{\ell}_N \rangle = 0$ , and  $\langle \vec{\ell}_N^2 \rangle = \ell^2$ . So we find

$$\langle \vec{s}_N^2 \rangle = \langle \vec{s}_{N-1}^2 \rangle + \ell^2 \quad (18)$$

Since this is true for any  $N$  we can compute that

$$\langle \vec{s}_N^2 \rangle = \langle \vec{s}_{N-1}^2 \rangle + \ell^2 = \langle \vec{s}_{N-2}^2 \rangle + 2\ell^2 = \dots = N\ell^2 \quad (19)$$

and the RMS distance moved is  $\sqrt{N}\ell$ , just like in the 1D case.

### 3 Diffusion from random walks

Diffusion refers to the net spreading of the distribution of molecules due to random molecular motion. Think about an individual molecule in a gas, say some CO molecule coming out of a car's exhaust. It leaves the exhaust and moves in a straight line until it hits another molecule, in which case it is buffeted essentially randomly in a different direction. As all the CO molecules are doing the same thing, on average, the net effect is a diffusion of the CO gas. We want to compute the probability distribution  $P_t(x)$  for where a CO molecule is after a time  $t$  and then use this to determine the equation of motion of the density of the gas.

#### 3.1 Collisions in a gas

It's helpful to discuss random walks for gases in terms of a set of convenient physical quantities. An important one is

- $\tau$  = the **collision time** is the average time a molecule goes before colliding with another molecule

The number of collisions in a time  $t$  is then

$$N = \frac{t}{\tau} \quad (20)$$

A related quantity is

- $\ell$  = the **mean free path** is the average distance a molecule goes between collisions

The mean free path is related to the collision time by

$$\ell = \bar{v} \tau \quad (21)$$

where

- $\bar{v}$  = the **average molecular speed**,  $\bar{v} = \langle |\vec{v}| \rangle$ .

Sometimes a more useful quantity is the root-mean-square velocity  $v_{\text{rms}} = \sqrt{\langle v^2 \rangle}$ . We can also use the speed of sound  $c_s$  in a gas, which is of course limited by the speed by which the molecules move. All three of these,  $\bar{v}$ ,  $v_{\text{rms}}$  and  $c_s$  are related by coefficients of order one, as we will see once we understand gases in more detail in future lectures. For example, in air at room temperature,  $\bar{v} = 467 \frac{m}{s}$ ,  $v_{\text{rms}} = 507 \frac{m}{s}$  and  $c_s = 346 \frac{m}{s}$ .

The mean free path is related to the density and size of the molecules. Treating molecules as spheres of radius  $R$ , two molecules will hit if their centers are within  $2R$  of each other. Thus you can think of a moving molecule as having an effective cross sectional area of  $\sigma = \pi(2R)^2$ . This effective cross sectional area is also called the **collisional cross section**. After  $N$  collisions a molecule will have swept out a volume  $V = N\ell\sigma$ . The number of molecules it hits during this sweeping is  $N = Vn$  with

- $n$  = the **number density** = number of molecules per unit volume

We will use number density a lot in statistical mechanics. It is interchangeable with the

- $\rho$  = the **mass density**.

as  $\rho = mn$  where  $m$  is the mass of a molecule (or the average mass of a molecule if the gas is mixed).

Thus,

$$\ell = \frac{1}{n\sigma} \quad (22)$$

Bigger molecules have bigger cross sectional areas so they will have smaller mean free paths. Since liquids are more dense than gases, generally they will have smaller mean free paths.

For example, the radius of a typical atom is around the Bohr radius  $a_0 = 0.05 \text{ nm}$ . So an air molecule, such as  $N_2$  or  $O_2$ , has a radius of around  $R \approx 2a_0 \sim 0.1 \text{ nm}$ . Thus  $\sigma \approx \pi(2R)^2 = 0.14 \text{ nm}^2$  in air. Air has a density of  $\rho = 1.3 \frac{\text{kg}}{\text{m}^3}$  and an average mass of  $m = 4.81 \times 10^{-26} \frac{\text{kg}}{\text{molecule}}$ , so its number density is  $n = \frac{\rho}{m} = 2.6 \times 10^{25} \frac{1}{\text{m}^3}$ . Note that  $n^{-1/3} = 3.3 \text{ nm}$  so air molecules are around  $3 \text{ nm}$  apart on average. The mean free path is  $\ell = \frac{1}{n\sigma} = 0.26 \mu\text{m}$ . The collision time is then  $\tau = \frac{\ell}{\bar{v}} = 0.57 \text{ ns}$ .

These are useful numbers to have in your head: in air at room temperature, molecules have velocities around  $v \sim 500 \frac{m}{s}$ , are around  $R \sim 0.1 \text{ nm}$  big and  $n^{-1/3} \sim 1 \text{ nm}$  apart. They collide around once every nanosecond (one billion times per second) after having moved around  $\ell \sim 100 \text{ nm}$  (one thousand molecule lengths).

### 3.2 Diffusion from random walks

Let's now consider the probability distribution  $P_t(x)$  for where a CO molecule is after a time  $t$ . We'll start in one dimension. Treating molecular interactions as a random walk, we take  $a = b = \frac{1}{2}$  since the molecule should be equally likely to be knocked left as right. Such a random walk is unbiased. For an unbiased random walk, the mean displacement is  $\bar{x} = 0$  and therefore does not tell us much about how fast the molecules are diffusing. Instead, the RMS displacement is more useful. The RMS displacement after a time  $t$  is, from Eq. (12)

$$x_{\text{rms}} = \sqrt{N} \ell = \sqrt{\frac{t}{\tau}} \ell = \sqrt{\ell \bar{v} t} \quad (23)$$

This  $x \sim \sqrt{t}$  behavior is the key characteristic of a random walk. Note that this is going to be a much smaller distance than an unhindered molecule would move on average,  $\Delta x \sim vt$ .

Knowing the mean ( $\bar{x} = 0$ ) and the standard deviation ( $\sigma = \sqrt{\ell \bar{v} t}$ ) we can immediately write down the full probability distribution for large times ( $t \gg \tau$ ) using the central limit theorem:

$$P_t(x) = \sqrt{\frac{1}{2\pi \ell \bar{v} t}} \exp\left[-\frac{x^2}{2t \ell \bar{v}}\right] \quad (24)$$

Note that probability distribution satisfies the differential equation

$$\frac{\partial P_t(x)}{\partial t} = D \frac{\partial^2 P_t(x)}{\partial x^2} \quad (25)$$

where  $D = \frac{1}{2} \ell \bar{v}$ . This is the 1D diffusion equation. You can easily check by plugging Eq. (24) into Eq. (25).

We defined  $P_t(x)$  as a probability distribution for one particular CO molecule in a gas. But the same probability distribution holds for any molecule. Since there are usually an enormous number  $N \sim 10^{24}$  of gas molecules, if each one has a probability  $P_t(x)$  of being at the the point  $x$  then number density will be simply

$$n(x, t) = NP_t(x) \quad (26)$$

To be precise, the number density is not exactly the same as the probability distribution since, classically, a particle is either at a particular position or not. So we should think of  $n(x, t) = NP_t(x)$  as referring to the number density averaged over time. (We'll return to this averaging in the next lecture, in the context of ergodicity.)

Thus we find

$$\frac{\partial n(x, t)}{\partial t} = D \frac{\partial^2 n(x, t)}{\partial x^2} \quad (27)$$

In 2 or 3 dimensions, the resulting equation is the rotationally symmetric version of this:

$$\boxed{\frac{\partial n(\vec{x}, t)}{\partial t} = D \vec{\nabla}^2 n(\vec{x}, t)} \quad (28)$$

This is known as the **diffusion equation**. It describes how substances move due to random motion. The coefficient is

$$D = \frac{1}{2} \ell \bar{v} = \frac{1}{2} \frac{\ell^2}{\tau} \quad (29)$$

This coefficient  $D$  is called the **diffusion constant** and the relation  $D = \frac{1}{2} \frac{\ell^2}{\tau}$  is known as the **Einstein-Smoluchowski equation**.

For example, in air  $\ell \approx 0.26 \mu\text{m}$  and  $\tau \approx 0.57 \text{ns}$  so  $D = 5.9 \times 10^{-5} \text{m}^2/\text{s}$ . This means that it takes 1 day for an individual gas molecule to diffuse 1 meter. Then it takes 100 days for it to diffuse 10 meters. These numbers are characteristic of diffusion processes: diffusion in air over macroscopic distances is generally very slow, and almost always dominated by convection (see Section 4.1) and other forms of energy transport. On the other hand, on smaller length scales where convection is irrelevant like a cell, diffusion can be dominant. For example, the diffusion constant for proteins in water is around  $D \approx 10^{-11} \frac{\text{m}^2}{\text{s}}$ . This is tiny compared to  $D$  for gases, but the typical distances proteins diffuse are also tiny, and time scales like distance squared. To diffuse across a cell of size  $10^{-4} \text{m}$  it takes 15 minutes. To diffuse across a cell nucleus of size  $6 \mu\text{m}$  it takes only 3.6 seconds.

Just because the diffusion equation looks simple does not mean it has trivial consequences! For example, it is mathematically identical to the Schrödinger equation, which accounts for a great variety of interesting physics. It also also mathematically identical to the heat equation: heat conduction is a diffusive process.

The diffusion equation is linear, so that if  $n_1(\vec{x}, t)$  and  $n_2(\vec{x}, t)$  are solutions, then so is their sum. In particular if we start with a bunch of particles at some positions  $x_i$ , then they will diffuse independently of each other. This gives us a way to solve the diffusion equation in general. For one particle, starting at  $\vec{x}=0$ , the solution is given by Eq. (24). Note that at  $t=0$ , this solution really does represent a localized source. In fact, the limit as  $t \rightarrow 0$  of this solution is one of the possible definitions of a  $\delta$ -function:

$$\lim_{t \rightarrow 0} \left( \frac{1}{4\pi Dt} \right)^{3/2} \exp\left[ -\frac{\vec{x}^2}{4Dt} \right] = \delta^3(\vec{x}) \quad (30)$$

where  $\delta^3(\vec{x}) = \delta(x)\delta(y)\delta(z)$ . Thus Eq. (24) is a solution to the diffusion equation with boundary condition  $n(\vec{x}, t) = \delta^3(\vec{x})$  at time  $t=0$ : it describes the diffusion away from a point source. Since any function can be described as a set of points, we can construct any solution to the diffusion equation by combining the solutions as in Eq. (24). More precisely, if  $n_0(\vec{x}, 0)$  is the number density at time 0, then the solution for all times is

$$n(\vec{x}, t) = \int d^3y \left( \frac{1}{2\pi\ell\bar{v}t} \right)^{3/2} \exp\left[ -\frac{(\vec{x}-\vec{y})^2}{2t\ell\bar{v}} \right] n_0(\vec{y}, 0) \quad (31)$$

To check this, we note that the right-hand side satisfies the diffusion equation and Eq. (30) verifies the boundary condition at  $t=0$ . Solving differential equations in this way is known as the **Green's function method**.<sup>2</sup> It converts solving a difficult differential equation to doing an integral.

Eq. (31) has a simple physical interpretation: the number of molecules at a point  $\vec{x}$  are those that have walked there randomly from any other point  $\vec{y}$  over the time  $t$ .

## 4 Fick's laws of diffusion

The approach to diffusion we discussed was based on a microscopic picture of random walks of individual molecules. We can also approach diffusion from the continuum perspective. Let us continue to denote the number density by  $n(\vec{x}, t)$  and let us also denote the local velocity as the vector field  $\vec{v}(\vec{x}, t)$ . It can be helpful to think of  $n(\vec{x}, t)$  as the density of a fluid, like water in a stream, and  $\vec{v}(\vec{x}, t)$  as its velocity at the point  $\vec{x}$  at time  $t$ . For simplicity, let's assume that  $n$  and  $\vec{v}$  are constant in the  $y$  and  $z$  directions, so they only depend on  $x$ ,  $n = n(x, t)$ ,  $\vec{v} = (v_x(x, t), 0, 0)$ . Now, the total number of molecules between  $x_1$  and  $x_2$  can only change if particles flow in or out of that region. So

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} dx n(x, t) = \underbrace{n(x_1, t)v_x(x_1, t)}_{\text{number coming in past } x_1 \text{ per unit time}} - \underbrace{n(x_2, t)v_x(x_2, t)}_{\text{number going out past } x_2 \text{ per unit time}} \quad (32)$$

$$= - \int_{x_1}^{x_2} dx \partial_x [n(x, t)v_x(x, t)] \quad (33)$$

<sup>2</sup>. In general, a Green's function satisfies  $\mathcal{O}G(x, t) = \delta(x)\delta(t)$  for some differential operator  $\mathcal{O}$ . In our case, the Green's function is  $G(\vec{x}, t) = P(\vec{x}, t)\theta(t) = \sqrt{\frac{1}{4\pi Dt}} \exp\left[ -\frac{\vec{x}^2}{4Dt} \right] \theta(t)$  which satisfies  $[\partial_t - D\nabla^2]G(\vec{x}, t) = \delta^3(\vec{x})\delta(t)$ .

Pulling the  $\frac{d}{dt}$  on the left into the integral, and using that  $x_1$  and  $x_2$  are arbitrary, we get

$$\frac{\partial}{\partial t}n(x, t) = -\frac{\partial}{\partial x}J_x(x, t) \quad (34)$$

where

$$J_x(x, t) = n(x, t)v_x(x, t) \quad (35)$$

The 3D version of this equation is called the **continuity equation**

$$\frac{\partial n}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \quad (36)$$

and  $\vec{J}(\vec{x}, t)$  is called the **flux**. The flux is the number density times velocity  $\vec{J}(\vec{x}, t) = n(\vec{x}, t) \times \vec{v}(\vec{x}, t)$ . It gives the number of particles passing by a given point per unit area per unit time. Note that the velocity field and the density are in principle independent, like position and velocity are independent in classical mechanics.

How does the flux related to the density during diffusion? Well, if the density is constant in position, then the net diffusion should be zero and the flux should vanish. Conversely, if the density has some spacial gradient, then there should be net flux from high density to low density. Thus, the leading order thing we could imagine is that  $\vec{J}$  is proportional to the concentration gradient:

$$\vec{J} = -D\vec{\nabla}n \quad (37)$$

with  $D$  a proportionality constant. We put in the minus sign so that  $D$  would be a positive number (if  $\partial_x n > 0$ , so the gradient increases to the right, then particles flow to the *left*). This is known as **Fick's first law**. It's a law because we didn't derive it (neither did Fick), it just seems reasonable. Note that Fick's first law is not some general property of fluxes – you can certainly have a nonzero flux at constant density, like current in a river, if there's some potential driving the flow. Fick's first law it is a statement about fluxes in *diffusive systems*, where there is no other source of motion other than random motion.

Once we have Eq. (37) we can plug into the continuity equation, Eq. (36) to get

$$\frac{\partial n(\vec{x}, t)}{\partial t} = D\vec{\nabla}^2 n(\vec{x}, t) \quad (38)$$

This is also known as **Fick's second law**. It is none other than the diffusion equation. Since Fick's second law follows from Fick's first law, in fact, we have justified Fick's first law through our analysis of random walks. Moreover, through our analysis of random walks, we have related  $D$  to properties of the gas,  $D = \frac{1}{2}\ell\bar{v}$  as in Eq. (29).

For example, say we have some lemmings that come out of a hole, walk left or right randomly, and maybe fall off a cliff a distance  $a$  away. Let their density be  $n(z)$ , with  $z = 0$  the hole and  $z = a$  the cliff. At a given  $z$ , over a short time  $\delta t$ , the number of lemmings that leave to the right is proportional to  $n(z)$  and those that come in from the right is proportional to  $n(z + \delta z)$ . Similarly, the number that leave to the left is  $n(z)$  and those coming in from the left proportional to  $n(z - \delta z)$ . So the net change in the number at  $z$  is going to be proportional to  $n(z + \delta t) - n(z) \propto [n(z + \delta z) - n(z)] + [-n(z) + n(z - \delta z)]$ . That is,  $\frac{\partial n}{\partial t} \propto \frac{\partial^2 n}{\partial z^2}$ . This is a physical way to see why diffusion only occurs if the second derivative is nonzero. In a steady state situation,  $\frac{\partial n}{\partial t} = 0$ , the solution to the diffusion equation is  $n(z) = n_0 + \frac{z}{a}(n_1 - n_0)$ . Then  $J_z = -\frac{D}{a}(n_1 - n_0)$ . This is a constant flux of lemmings coming out of the hole and falling off the cliff. Note that since  $J_z$  is constant,  $\partial_z J_z = 0$  and the continuity equation is satisfied.

To get a feel for how fast diffusion is, the diffusion constant in water for nitrogen molecules is  $D = 2 \times 10^{-9} \frac{m^2}{s}$ . Recalling from Eq. (22) that  $\ell = \frac{1}{n\sigma}$ , so  $D = \frac{1}{2} \frac{\bar{v}}{n\sigma}$ , bigger molecules should have smaller diffusion rates. Indeed, benzene molecules  $C_6H_6$  in water have  $D = 1 \times 10^{-9} \frac{m^2}{s}$ . For large molecules like proteins in water the diffusion constant is even smaller  $D \approx 10^{-11} \frac{m^2}{s}$ . In gases, densities  $n$  are smaller so  $\ell$  is larger and the diffusion constants are generally larger. For example, CO molecules in air at room temperature and pressure have  $D = 2 \times 10^{-5} \frac{m^2}{s}$ .

To use the diffusion constant, we can either plug in the exponential solution, Eq. (24), or more simply use Eq. (23):

$$x_{\text{rms}} = \sqrt{2Dt} \quad (39)$$

For example, taking a dye molecule in water with  $D \approx 10^{-9} \frac{m^2}{s}$ , to move  $\Delta x = 1 m$  would take  $\frac{(\Delta x)^2}{2D} = 31$  years. So clearly diffusion is not the main mechanism by which dyes move around in water.

By the way, thermal conduction is very much like diffusion. Instead of the diffusion equation, temperature satisfies the **heat equation**:

$$\frac{\partial T(x, y, z, t)}{\partial t} = \alpha \vec{\nabla}^2 T(x, y, z, t) \quad (40)$$

where  $\alpha$  is called the coefficient of thermal diffusivity. This equation describes diffusion of temperature, rather than diffusion of particle number. The derivation of the heat equation is identical to the derivation of Fick's second law, with conservation of energy replacing conservation of particle number. The analog of Fick's first law for thermal conduction is called **Fourier's law**. Fourier's law is an empirical observation that the rate of heat flow is proportional to the temperature difference. We'll return to thermal conduction when we talk about temperature and heat in future lectures.

## 4.1 Convection (optional)

Diffusion refers to the motion of a molecule through random collisions. Think of a liquid in equilibrium and just try to follow one molecule. **Convection** occurs when the system is not in equilibrium to begin with. In such situations, there can be coherent convective currents, like a hot or cold wind, that move the dye much faster than through a random walk. Or if you dropped the dye into the water with a dropper it hits with some force and has some inertia; then it takes a while for the system to equilibrate and the dye molecules are for a while moving much faster than they would if there were only diffusion.

If there is some external effect causing the medium to flow with velocity  $\vec{v}_{\text{conv}}(\vec{x}, t)$ , then there will be flux even if there is no concentration gradient. We can introduce the **convective flux**

$$\vec{J}_{\text{conv}}(\vec{x}, t) = \vec{v}_{\text{conv}} n(\vec{x}, t) \quad (41)$$

to describe this situation. Adding this convective flux to the diffusive flux, we get a new term:

$$\frac{\partial n(\vec{x}, t)}{\partial t} = D \vec{\nabla}^2 n(\vec{x}, t) + \vec{v}_{\text{conv}} \cdot \vec{\nabla} n(\vec{x}, t) \quad (42)$$

This is called the **generalized diffusion equation** and describes situations where diffusion and convection are both important. Unfortunately, it is usually impossible to determine  $\vec{J}_{\text{conv}}(\vec{x}, t)$ , since when there is convection usually molecules are all moving around in different directions and it is a horribly non-linear process. Think about this next time you pour milk into your coffee – all those little eddy currents and funny shapes are convective. Good luck describing them analytically! Convection is almost always studied with numerical simulations.

So for diffusion to actually be visible, a system has to be very calm – no temperature, density or chemical gradients. A place where diffusion is more important than convection is in biology. In biological systems, temperature is often very constant, convection is small, and molecules do not have to move very far. Diffusion of heat (thermal conduction) is the dominant mechanism of heat transfer in solids, for example as you heat up a pan on the stove. However, when you heat a room, convection dominates and the heat equation, Eq. (40) is not relevant.

## 5 Brownian motion

An important application of the diffusion equation is to study Brownian motion. In 1827, a botanist named Robert Brown collected some pollen one Spring afternoon and put it in some water in a Petri dish in his lab, then went to bed. When he woke up, he found that the pollen grain had moved a significant distance. "It's alive!" he concluded. In fact, the pollen moved not because it is alive, but rather because it underwent a random walk due to the water molecules surrounding it constantly giving it little kicks. This movement is called **Brownian motion**, after Robert Brown.

Brownian motion refers to the random walk of a large particle due to stochastic collisions with smaller particles. Although each hit from a small particle does very little, the hits add up to a macroscopically observable displacement. You can see Brownian motion easily with a microscope, where a dust particle or a bacterium will move a finite distance in a reasonable time. What is fascinating about this migration is that you cannot resolve the small molecules, like water molecules, in the microscope, so it looks like the big particle is moving by magic. Of course, it is not magic, and indeed we can deduce the existence of “invisible” molecules from Brownian motion of something visible. Einstein used this insight to measure Avogadro’s number, as we will now see.

The molecular collisions have another effect too – they slow down a moving particle, through a drag force. Indeed, drag, that you experience running your hand through the air or in water is a collective effect of many small molecules impeding the motion. Drag forces are macroscopic and can be measured without ever talking about molecules. For example we could rub the particle to make it electrically charged, then pull it with the electric force and measure the resistance. Or we could tie a tether to it, add a weight, and pull it with gravity. A drag force, by definition, slows down a particle, so if  $\vec{v}=0$  it should vanish. Thus, the leading effect in an expansion around  $\vec{v}=0$  of drag is that it is linear in the velocity. Calling the mass of the particle  $m$  and the external force  $\vec{F}_{\text{ext}}$ , the drag force we can then write

$$m \frac{d^2 \vec{x}}{dt^2} + \mu \frac{d \vec{x}}{dt} = \vec{F}_{\text{ext}} + \vec{F}_B \quad (43)$$

where  $\vec{x}$  is the position of the particle,  $\mu$  is the **drag coefficient** (also called **mobility**), and  $\vec{F}_B$  is the force due to Brownian motion. That is,  $\vec{F}_B$  is a rapidly varying function of time, representing the small forces that the molecules give from whatever direction they come in at. Generally, when  $\vec{F}_{\text{ext}}$  is nonzero then  $\vec{F}_B$  can be neglected. So we turn  $\vec{F}_{\text{ext}}$  on to measure  $\mu$ , then turn  $\vec{F}_{\text{ext}}$  off.

Once  $\mu$  is measured and  $\vec{F}_{\text{ext}}$  turned off we can look at the distance the dust particle moves due to Brownian motion alone. Of course, the expected value is  $\langle \vec{x} \rangle = 0$ , by symmetry, so we want to look at the RMS displacement  $x_{\text{rms}} = \sqrt{\langle \vec{x}^2 \rangle}$ . We want to know how  $\vec{x}^2$  changes with time:

$$\frac{d}{dt} \vec{x}^2 = 2 \vec{x} \cdot \frac{d \vec{x}}{dt} = 2 \vec{x} \cdot \vec{v} \quad (44)$$

First note that Eq. (43) implies

$$\frac{d}{dt} \vec{v} = -\frac{\mu}{m} \vec{v} + \frac{1}{m} \vec{F}_B \quad (45)$$

So then we have

$$\frac{d}{dt} \vec{x} \cdot \vec{v} = \left( \frac{d}{dt} \vec{x} \right) \cdot \vec{v} + \vec{x} \cdot \left( \frac{d}{dt} \vec{v} \right) = \vec{v} \cdot \vec{v} - \frac{\mu}{m} \vec{x} \cdot \vec{v} + \frac{1}{m} \vec{x} \cdot \vec{F}_B \quad (46)$$

Now let’s take the average of all the quantities in this equation over very many molecular collisions. Since the force  $\vec{F}_B$  due to molecular collisions is in a random direction, on average it is totally uncorrelated to the current position of the particle, so  $\langle \vec{x} \cdot \vec{F}_B \rangle = 0$  as in Eq. (15). Then we have simply

$$\frac{d}{dt} \langle \vec{x} \cdot \vec{v} \rangle = \langle \vec{v}^2 \rangle - \frac{\mu}{m} \langle \vec{x} \cdot \vec{v} \rangle \quad (47)$$

The general solution of this equation starting at  $\vec{x}=0$  is

$$\langle \vec{x} \cdot \vec{v} \rangle = \frac{m}{\mu} \langle \vec{v}^2 \rangle \left( 1 - e^{-\frac{\mu t}{m}} \right) \quad (48)$$

For late times  $t \gg \frac{m}{\mu}$ , the exponential is very small and we see that  $\langle \vec{x} \cdot \vec{v} \rangle = \frac{m}{\mu} \langle \vec{v}^2 \rangle$  is constant in time. Then, from Eq. (44) we get

$$\frac{d}{dt} \langle \vec{x}^2 \rangle = \frac{2m}{\mu} \langle \vec{v}^2 \rangle \quad (49)$$

So that

$$\langle \vec{x}^2 \rangle = \frac{2m}{\mu} \langle \vec{v}^2 \rangle t \quad (50)$$

The RMS displacement is therefore

$$x_{\text{rms}} = \sqrt{\langle \vec{x}^2 \rangle} = v_{\text{rms}} \sqrt{\frac{2m}{\mu}} \sqrt{t} \quad (51)$$

where  $v_{\text{rms}} = \sqrt{\langle \vec{v}^2 \rangle}$ . Comparing to Eq. (39) we have

$$D = \frac{1}{\mu} \langle m \vec{v}^2 \rangle \quad (52)$$

This is known as the Einstein relation.

Since  $\mu$  is the macroscopic drag coefficient and  $x_{\text{rms}}$  and  $t$  are measurable, this equation lets us determine the average kinetic energy of each molecule. Using also the total energy in a system, we can then determine the number of molecules. More precisely, we can derive an explicit formula for  $N_A$  if we borrow a result from Lecture 4, that the average kinetic energy per particle is  $\langle \frac{1}{2} m \vec{v}^2 \rangle = \frac{3R}{2N_A} T$  with  $R$  the ideal gas constant and  $T$  the temperature. (It should not be obvious to you that the average kinetic energy is determined by the temperature, but it's true as we'll show in Lecture 4.) Then Eq. (52) implies

$$N_A = \frac{3RT}{\mu D} \quad (53)$$

So by measuring the temperature (with a thermometer), the drag coefficient (with an external force) and the rate of diffusion (from Brownian motion), the number of water molecules  $N_A$  can be determined. This is how Albert Einstein proposed to measure Avogadro's number  $N_A$  using Brownian motion in 1905.

## 5.1 Viscosity

Consider a ball falling through a fluid. It is accelerated from a downward force due to gravity and encounters resistance, or drag, from the fluid. There are two sources of drag. The first, called **inertial drag**, is due to the ball bonking molecules in front of it and speeding them up. In a time  $\Delta t$  the ball will sweep out a volume  $\Delta V = \pi R^2 |\vec{v}| \Delta t$ , thereby accelerating a mass  $\Delta m = \rho \Delta V$  of fluid to velocity roughly  $\vec{v}$ . The average acceleration is then  $\vec{a} = \frac{\vec{v}}{\Delta t}$ , so that the force is  $\vec{F}_{\text{inertial}} = \Delta m \vec{a} = \pi R^2 \rho |\vec{v}| \vec{v}$ . This goes like the *square* of the velocity. This inertial drag is relevant at large velocities, but for slow objects it is always going to be subdominant to a drag force that is linear in  $v$ . For slow objects (slow compared to typical molecular speeds), the dominant drag force is due to viscosity, called **viscous drag**. It is viscous drag that is important for Brownian motion, and related to random walks.

Viscosity is another physical effect whose microscopic origin is in the stochastic collision of molecules. The more precise name for viscosity is dynamic shear viscosity. It measures how a fluid responds to shear forces: you push the top layer and ask how much the bottom layer moves. Intuitively, viscosity is a measure of how well a fluid flows.

A shear force is applied to an area. Think of floating a block of wood on water and applying an external force  $\vec{F}_{\text{ext}}$  to move it parallel to the surface. The bigger the area of the block, the more force it puts on the water. The water then responds by picking up some velocity  $\vec{v}$ , in the same direction as the force. The deeper you go into the water, the slower it will go. So we might expect

$$\eta \frac{\vec{v}}{z} \stackrel{?}{=} \frac{\vec{F}_{\text{ext}}}{A} \quad (54)$$

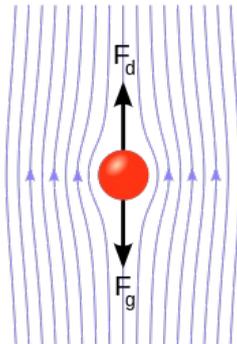
with  $z$  the depth. This isn't quite right, because we don't know that the  $z$  dependence is exactly  $\frac{1}{z}$ .

The right way to think about this shear force is that we apply it only to the top of the water. When the top of the water moves, it pulls along the layer below that (by layer we mean layer of molecules, or just some abstract infinitesimal thickness of the fluid), and so on. So we write

$$\eta \frac{\partial \vec{v}}{\partial z} = \frac{\vec{F}_{\text{ext}}}{A} \quad (55)$$

The parameter  $\eta$  in this equation is called the dynamic shear **viscosity**.

Now let's return to the sphere falling through the fluid. It has a downward force due to gravity. As it moves down, it displaces the molecules, but also imparts velocity to the fluid in the direction transverse to its motion



**Figure 1.** A ball falling through a viscous fluid has a downward force due to gravity and a drag force due to viscosity.

The definition of viscosity tells us that the sphere induces a velocity of the fluid

$$\eta \frac{\partial \vec{v}}{\partial z} = \frac{\vec{F}_g}{4\pi R^2} \quad (56)$$

where  $4\pi R^2$  is the surface area of the sphere. At the surface of the sphere, the fluid velocity is the same as the sphere velocity. Thus we can solve this equation to see how the fluid velocity changes with distance. The faster the sphere is falling, the faster the fluid will go. Eventually, all the energy used by gravity to accelerate the sphere will be taken up by the work done to move the water, and the sphere will stop accelerating. Thus, there will be some effective drag force. Working out all the factors (an annoyingly tedious calculation), the result is

$$\vec{F}_{\text{drag}} = 6\pi\eta R \vec{v} \quad (57)$$

This is known as the **Stokes drag force**. This equation relates a property of the fluid (the viscosity), to the resistance experienced by an external object as it is being forced through the fluid (the viscous drag force). It is linear in the velocity of the object, compared to the inertial force which was quadratic in velocity, so at small velocities, the viscous drag force dominates. Keep in mind that viscosity is a property of the fluid itself, while drag depends on what is being dragged through the fluid. The Stokes force applies when that thing is a sphere of radius  $R$ . For a different object, like a cube, the precise form of the force would be different, but the scaling with  $\eta$  and  $\vec{v}$  and some measure of size  $R$  is universal. Indeed, it is fixed by dimensional analysis.

Recall that we defined mobility as the drag coefficient in Eq. (43). So

$$\mu = 6\pi\eta R \quad (58)$$

This is known as the **Stokes relation**. Plugging it into the Einstein relation in Eq. (52) gives

$$D = \frac{1}{6\pi\eta R} \langle m \vec{v}^2 \rangle \quad (59)$$

This is known as the **Einstein-Stokes relation**. It relates the diffusion constant  $D$  and the viscosity, telling us that viscosity also has its origin in the microscopic random walks of the molecules in the fluid.

## 6 Summary

In this lecture we have studied diffusion. The main concepts to understand are

- **Random walks:** A particle/system has a fixed probability of moving in each direction. We are interested in the net motion after  $N$  steps.
- If we move left with probability  $a$  and right with probability  $b = 1 - a$ , the chance of taking  $m$  steps right is given by the **binomial distribution**:  $B_N(m) = a^m b^{N-m} \binom{N}{m}$  where  $\binom{N}{m} = \frac{N!}{m!(N-m)!}$ .

- The average distance moved in a random walk after  $N$  steps scales as  $\sigma \sim \sqrt{N}\ell$  where  $\ell$  is the step size. This is true in any number of dimensions.
- A bunch of molecules random-walking is called **diffusion**. Then the number of steps  $N$  is proportional to the time  $t$  so the distance moved is  $\Delta x \sim \sqrt{t}\ell$ . This scaling of **distance**  $\propto$  **square-root of time** is the characteristic feature of diffusion.
- In the continuum limit, molecules are described by a **number-density**  $n(\vec{x}, t)$ .
- The number density satisfies the **diffusion equation**  $\frac{\partial n(\vec{x}, t)}{\partial t} = D\vec{\nabla}^2 n(\vec{x}, t)$  when diffusion is dominant.
- The exact solution to the diffusion equation with boundary conditions  $n(\vec{x}, t) = \delta^3(\vec{x})$  is  $n(\vec{x}, t) = \sqrt{\frac{1}{4\pi Dt}} \exp\left[-\frac{\vec{x}^2}{4Dt}\right]$ . This confirms the scaling  $\Delta x \sim \sqrt{t}$ .
- The diffusion equation comes from a modeling of the microscopic system as undergoing random walks. An alternative classical-field approach (no particles) uses **Fick's laws**. These assume the conservation of the amount of stuff (the continuity equation  $\frac{\partial n}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0$ , where  $\vec{J} = n \times \vec{v}$  is the **flux**) and that flux is proportional to the concentration gradient  $\vec{J} = -D\vec{\nabla}n$ , leads to the same diffusion equation.
- The conduction of heat is described by the **heat equation**:  $\frac{\partial T(\vec{x}, t)}{\partial t} = \alpha\vec{\nabla}^2 T(\vec{x}, t)$ . This equation has identical form as the diffusion equation, and so also has  $\Delta x \sim \sqrt{t}$ .
- In liquids and gases, diffusion is rarely dominant unless the system is very calm. **Convection**, coming from non-equilibrium initial conditions, is usually is more important. In solids, diffusion and the heat equation work well.
- A large molecule in a reasonably calm liquid or gas can often be modeled well by a random walk/diffusion. When this applies, we say it undergoes **Brownian motion**. For such molecules, the diffusion constant is given by the **Einstein relation**:  $D = \frac{1}{\mu}\langle m\bar{v}^2 \rangle$  with  $\mu$  the drag coefficient. Measuring  $D$  and  $\mu$  is one way to measure Avogadro's number:  $N_A = \frac{3RT}{\mu D}$ .
- The drag coefficient is also related to **viscosity**  $\eta$  by  $\mu = 6\pi\eta R$ . Then  $D = \frac{1}{6\pi\eta R}\langle m\bar{v}^2 \rangle$ . This is called the **Einstein-Stokes relation**,

The main lesson from the last section is simply that the drag force, mobility, diffusion, viscosity and random walks are all related. I don't expect you to remember all these formulas, and I certainly don't want you to memorize them. Just try to have the basic ideas straight. Viscosity is a macroscopically measurable property of a material. When a *large particle* moves in a viscous material, it undergoes Brownian motion. The bigger the particle, the smaller the diffusion constant, and the slower it moves. The higher the viscosity, the larger the drag force, and the slower the particle moves.

Small molecules moving in a fluid also undergo random walks. For small molecules, of around the same size as the molecules in the fluid, the diffusion constant is  $D = \frac{1}{2}\ell\bar{v}$  with  $\ell$  the mean free path. In this case, it is not useful to think in terms of drag forces and viscosity since the diffusing particle has essentially no inertia.

A tricky point from this lecture is that **diffusion is an equilibrium phenomena**. Equilibrium does not mean there is no time-dependence at all, but that the macroscopic properties of the system are static. We'll discuss equilibrium more next lecture. You should think of  $\frac{\partial n(\vec{x}, t)}{\partial t} = D\vec{\nabla}^2 n(\vec{x}, t)$  as describing the change in the number density of a small subset of the molecules in a bath, for example, dye in water, or of a macromolecule undergoing Brownian motion in a bath. The water in the bath doesn't change, even though individual water molecules might move. When the whole system is out of equilibrium, for example when the boundary conditions are open (gas released into a vacuum) or if you violently drop the dye in, then the diffusion equation doesn't apply.