Physics 115A: Statistical Physics

Prof. Clare Yu
email: cyu@uci.edu
phone: 949-824-6216
Office: 210E RH

Spring 2009

LECTURE 1

Introduction

So far your physics courses have concentrated on what happens to one object or a few objects given an external potential and perhaps the interactions between objects. For example, Newton's second law $F = ma$ refers to the mass of the object and its acceleration. In quantum mechanics, one starts with Schroedinger's equation $H\psi = E\psi$ and solves it to find the wavefunction ψ which describes a particle. But if you look around you, the world has more than a few particles and objects. The air you breathe and the coffee you drink has lots and lots of atoms and molecules. Now you might think that if we can describe each atom or molecule with what we know from classical mechanics, quantum mechanics, and electromagnetism, we can just scale up and describe 10^{23} particles. That's like saying that if you can cook dinner for 3 people, then just scale up the recipe and feed the world. The reason we can't just take our solution for the single particle problem and multiply by the number of particles in a liquid or a gas is that the particles interact with one another, i.e., they apply a force on each other. They see the potential produced by other particles. This makes things really complicated. Suppose we have a system of N interacting particles. Using Newton's equations, we would write:

$$\frac{d\vec{p}_i}{dt} = m\frac{d^2\vec{x}_i}{dt^2} = \sum_{j\neq i}\vec{F}_{ij} = -\nabla V_i(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N) \tag{1}$$

where $\vec{F}_{ij}$ is the force on the ith particle produced by the jth particle. V_i is the total potential energy of the ith particle assuming that all the forces are conservative. We must also specify the initial conditions, i.e., the initial positions and velocities of all the particles. Then we must solve the N coupled second order partial differential equations. This is quite a daunting task. It's true that we now have powerful computers and molecular dynamics simulations carry out such tasks, but they can only handle a few thousand particles and track their motions for perhaps a 10^{-8} seconds. So we throw up our hands. This is not the way to go.

Fortunately, there is a better way. We can use the fact that there are large numbers of particles to apply a statistical analysis to the situation. We are usually not interested in the detailed microscopics of a liquid or a gas. Instead we are usually interested in certain macroscopic quantities like

- pressure
- temperature
- entropy
- specific heat
- electrical conductivity
- magnetic and electric susceptibilities
- etc.

These concepts don't make sense for one atom or even a few particles. For example, what is the temperature of one atom? That doesn't make sense. But if we have a whole bucket of atoms, then temperature makes sense. Since we are not concerned with the detailed behavior of each and every particle, we can use statistical methods to extract information about a system with $N \gg 1$ interacting particles. For example, we may be interested in the average energy of the system, rather than the energy of each particle. Statistical methods work best when there are large numbers of particles. (Think of polling.) Typically we have on the order of a mole's worth of particles. A mole is the number of atoms contained in 12 grams of carbon 12. (A carbon 12 atom has 6 protons and 6 neutrons for a total 12 nucleons. Its atomic weight is 12.) So a mole has 6.022×10^{23} atoms. This is called Avogadro's number.

There are two basic approaches to describing the properties of a large number of particles:

Thermodynamics: This is essentially a postulational approach, or a deductive approach. On the basis of a few general assumptions, general relationships between and among macroscopic parameters are deduced. No reference is made to the microscopic details of the system.

Statistical Mechanics: This is an inductive approach. With only a few assumptions, relations between the macroscopic parameters of a system are induced from a consideration of the microscopic interactions. This approach yields all of the thermodynamic results plus a good deal more information about the system of interacting particles.

Probability

We will talk more about probability in the discussion section. If we flip a coin, the chance of heads is 1/2 and the chance of tails is 1/2. The probability of an outcome is the number of ways of getting that outcome divided by the total number of outcomes. We often write $P(x_i)$ for the probability that an outcome is x_i. So if we flip 2 coins, then P(head,head)=1/4, and P(head,tail)=1/2 because there are a total of 4 possible outcomes and 2 ways of getting one coin heads and one coin tails.

Binary Model

Let's work out a simple example of a binary model. Rather than heads or tails, think of a line of spins which can be up ($\uparrow$) or down ($\downarrow$). The spins correspond to magnetic moments $\pm m$. $+m$ corresponds to $\uparrow$ and $-m$ corresponds to $\downarrow$. Suppose we have N sites, each with a spin. On each site the spin can be up or down. The total number of arrangements is $2 \times 2 \times 2 \times 2...2 = 2^N$. So the total number of states or outcomes is 2^N. We can denote a state by

$$\uparrow\uparrow\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow \ldots \tag{2}$$

Now suppose we want the number of configurations with n up sites, regardless of what order they're in. The number of down sites will be N-n, since the total number of sites is N. The number of such configurations is

$$\frac{N!}{n!(N-n)!} \tag{3}$$

To see where this comes from, suppose we have N spins of which n are up and (N-n) are down. How many ways are there to arrange them? Recall that the number of ways to arrange N objects with one object on each site (or in each box) is

- the first place can be occupied by any one of the N objects
- the second place can be occupied by any one of the N-1 remaining objects

etc. until the last place can be occupied by the last 1 object. So there are

$$N(N-1)(N-2)...1 = N! \tag{4}$$

configurations. In these configurations the n up spins can be arranged in any of $n!$ ways. Also the remaining (N-n) sites have down spins which can be arranged in any of $(N-n)!$ ways. The spins are regarded as distinguishable. So there are 2 ways to have $\uparrow\uparrow$: $\uparrow\uparrow$ and $\uparrow\uparrow$. They look the same, but if one were red and the other blue, you would see the difference (red, blue) and (blue, red). We are dividing out this overcounting; that's what the denominator $n!(N-n)!$ is for.

So the probability of a state with n up spins and (N-n) down spins is

$$\begin{aligned} P(n\uparrow, N) &= \frac{N!}{n!(N-n)!} \div (\text{total number of states}) \\ &= \frac{N!}{n!(N-n)!}\,\frac{1}{2^N} \end{aligned} \tag{5}$$

Now suppose that the probability of a site getting $\uparrow$ is p and the probability of getting $\downarrow$ is $q = 1 - p$. So far we have considered up and down to be equally probable, so $p = q = 1/2$. But what if $p \neq q$? This might be caused by an external magnetic field which biases the spins one way or another. If we have 2 spins, the probability of getting 2 up spins is p^2. (Recall that the probability of flipping a coin twice and getting heads both times is 1/4.) If we want n up spins and (N-n) down spins, then the probability is

$$\begin{aligned} P(n\uparrow, N) &= (\text{number of arrangements}) \times p^n\, q^{N-n} \\ &= \frac{N!}{n!(N-n)!}\, p^n\, q^{N-n} \end{aligned} \tag{6}$$

When $p = q = 1/2$, this reduces to our previous result (5). Equation (6) is called the binomial distribution because the prefactor $N!/n!(N-n)!$ is the coefficient in the binomial expansion:

$$(a+b)^N = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} a^{N-n}\, b^n \tag{7}$$

Normalization, Averages, Second Moment

Let's go back to our general considerations of probability distribution functions. Let $P(x_i)$ be the probability distribution for the probability that an outcome is x_i. If we sum over all possible outcomes, one of them is bound to happen, so

$$\sum_{x_i} P(x_i) = 1 \tag{8}$$

In other words, the probability distribution function $P(x_i)$ is normalized to one. One can also have a probability distribution function $P(x)$ of a continuous variable x. For example, suppose you leave a meter stick outside on the ground. $P(x)dx$ could be the probability that the first raindrop to hit the stick will strike between position x and $x + dx$. $P(x)$ is called the probability density. The normalization condition is

$$\int dx P(x) = 1 \tag{9}$$

Note that $\int dx P(x)$ is the area under the curve $P(x)$. The normalization sets this area equal to 1.

The average of some quantity x is

$$< x >= \overline{x} = \frac{\int dx P(x) x}{\int dx P(x)} \tag{10}$$

Averages can be denoted by $< x >$ or $\overline{x}$. For a discrete quantity

$$\overline{x} = \frac{\sum_i P(x_i) x_i}{\sum_i P(x_i)} \tag{11}$$

If the probability is normalized, then the denominator is 1. For example, suppose $N(A_i)$ is the number of people with age A_i. Then the average age is

$$\begin{aligned}
\overline{A} &= \frac{\sum_{A_i=0}^{\infty} N(A_i)\, A_i}{\sum_{A_i=0}^{\infty} N(A_i)} \\
&\to \frac{\int dA\, N(A)\, A}{\int dA\, N(A)}
\end{aligned} \tag{12}$$

The probability that a person is A old is

$$P(A) = \frac{N(A)}{\int dA\, N(A)} \tag{13}$$

Notice that this satisfies the normalization condition. More generally, if $f(x)$ is a function of x, then the average value of $f(x)$ is

$$\overline{f(x)} = \frac{\int dx P(x) f(x)}{\int dx P(x)} \tag{14}$$

One of the more useful functions concerns the deviation of x from the mean $\overline{x}$:

$$\Delta x = x - \overline{x} \tag{15}$$

Then

$$\begin{aligned}
\overline{\Delta x} &= \overline{x - \overline{x}} \\
&= \overline{x} - \overline{x} \\
&= 0
\end{aligned} \tag{16}$$

A more useful quantity is the square of the deviation from the mean

$$\begin{aligned}
\overline{(\Delta x)^2} &= \overline{(x - \overline{x})^2} \\
&= \overline{(x^2 - 2x\overline{x} + \overline{x}^2)} \\
&= \overline{x^2} - 2\overline{x}\overline{x} + \overline{x}^2 \\
&= \overline{x^2} - \overline{x}^2
\end{aligned} \tag{17}$$

This is known as the second moment of x about its mean. The first moment is just $\overline{x}$.

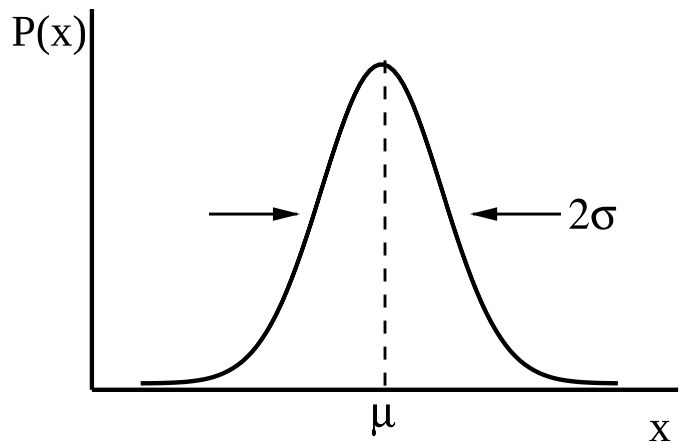
In general the nth moment of x about its mean is given by

$$\overline{(x - \overline{x})^n} = \frac{\int dx (x - \overline{x})^n P(x)}{\int dx P(x)} \tag{18}$$

Two other terms that are often used are the most probable value of x and the median. The most probable value of x is the maximum of $P(x)$. The median x_{med} is the value of x such that half the values of x are greater than x_{med}, i.e., $x > x_{med}$, and half the values of x are less than x_{med}, i.e., $x < x_{med}$. In terms of the area under the curve, the median is the vertical dividing line such that half the area lies to the left of the median and half to the right of the median.

Gaussian Distribution and Central Limit Theorem

One of most useful distributions is the Gaussian distribution. This is sometimes called the bell curve which is well known as the ideal grade distribution.



The formula for a Gaussian distribution is

$$P(x)dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx \tag{19}$$

where $\mu = \overline{x}$ is the mean. The coefficient is set so that the normalization condition is satisfied. $\sigma^2 = \overline{(\Delta x)^2} = \overline{(x - \overline{x})^2}$. 2σ is the width of the distribution. There is a 68% chance that $-\sigma \leq x \leq \sigma$. One obtains this by integrating $P(x)$ from $-\sigma$ to $+\sigma$. σ is sometimes called the root–mean–square (rms) deviation or the standard deviation. You will have the opportunity to check some of these assertions in your homework.

For a Gaussian distribution, all the higher moments can be expressed in terms of the first and second moments. For example $\overline{(x - \overline{x})^4} = 3\overline{(x - \overline{x})^2}$.

There is one other interesting aspect of Gaussian distributions. It's called the central limit theorem. It can be stated in terms of a random walk in one dimension. A drunk starts at a lamp post and staggers back and forth along the one dimensional sidewalk. Let $w(s)ds$ be the probability that a step length lies in the range between s and $s + ds$. No matter what the probability distribution $w(s)$ for each step may be, as long as the steps are statistically independent and $w(s)$ falls off rapidly enough as $|s| \to \infty$, the total displacement x will be distributed according to the Gaussian law if the number of steps N is sufficiently large. This is called the central limit theorem. It is probably the most famous theorem in mathematical probability theory. The generality of the result also accounts for the fact that so many phenomena in nature (e.g., errors in measurement) obey approximately a Gaussian distribution.

In fact, if we go back to the binomial distribution given in eq. (6), then in the limit of large N, this becomes the Gaussian distribution (see Reif 1.5). In this case, x represents the net magnetization, i.e., the difference between the number of up and down spins:

$$x = N_\uparrow - N_\downarrow \tag{20}$$

The average magnetization

$$\mu = (p - q)N \tag{21}$$

and the mean square deviation is

$$\sigma^2 = 4Npq \tag{22}$$

More on the Random Walk

(Reference: Howard C. Berg, *Random Walks in Biology.*)

The random walk is an important concept. Let us go back to the drunk starting at the lamp post and stumbling either to the right or the left. Consider an ensemble of M drunks. Let $x_i(N)$ be the position of the ith drunk after N steps. Suppose everyone starts at the origin so that $x_i(N = 0) = 0$ for all i. Let the step size be of fixed length s. Then

$$x_i(N) = x_i(N-1) \pm s \tag{23}$$

The average displacement from the origin is

$$\begin{aligned} \overline{x}_i(N) &= \frac{1}{M}\sum_{i=1}^{M} x_i(N) \\ &= \frac{1}{M}\sum_{i=1}^{M} [x_i(N) = x_i(N-1) \pm s] \\ &= \frac{1}{M}\sum_{i=1}^{M} x_i(N-1) \\ &= \overline{x}_i(N-1) \\ &= 0 \end{aligned} \tag{24}$$

So the mean displacement is zero because on average, half of the displacements are negative and half are positive. But this does not mean that all the drunks are sitting at the origin after N steps.

It is useful to look at the spread in positions. A convenient measure of spreading is the root-mean-square (rms) displacement $\sqrt{\overline{x^2(N)}}$. Here we average the square of the displacement rather than the displacement itself. Since the square of a negative or a positive number is positive, the result will be positive.

$$x_i^2(N) = x_i^2(N-1) \pm 2sx_i(N-1) + s^2 \tag{25}$$

Then we compute the mean,

$$\begin{aligned}
\overline{x_i^2}(N) &= \frac{1}{M}\sum_{i=1}^{M} x_i^2(N) \\
&= \frac{1}{M}\sum_{i=1}^{M}\left[x_i^2(N-1) \pm 2sx_i(N-1) + s^2\right] \\
&= \overline{x_i^2}(N-1) + s^2 \qquad (26)
\end{aligned}$$

So $\overline{x^2}(0) = 0$, $\overline{x^2}(1) = s^2$, $\overline{x^2}(2) = 2s^2$,..., and $\overline{x^2}(N) = Ns^2$. So the rms displacement

$$\sqrt{\overline{x^2}(N)} = \sqrt{N}s \qquad (27)$$

The rms displacement is a good measure of the typical distance that the drunk reaches. Notice that it goes as the square root of the number of steps. If τ is the time between successive steps, then it takes a time $t = N\tau$ to take N steps. So

$$\sqrt{\overline{x^2}(N)} = \sqrt{\frac{t}{\tau}}s \sim \sqrt{t}s \qquad (28)$$

Notice that the rms displacement goes as square root of the time t.

A random walk is what particles execute when the diffuse. Think of a opening a bottle of perfume. The perfume molecules diffuse through the air. We can define a diffusion coefficient D by

$$D = \frac{s^2}{2\tau} \qquad (29)$$

(The factor of 2 is by convention.) Then in one dimension

$$\begin{aligned}
\overline{x^2}(t) &= 2Dt \\
\sqrt{\overline{x^2}(t)} &= \sqrt{2Dt} \qquad (30)
\end{aligned}$$

For 2 dimensions, if the x and y directions are independent,

$$\overline{r^2}(t) = \overline{x^2}(t) + \overline{y^2}(t) = 4Dt \qquad (31)$$

Similarly for 3 dimensions

$$\overline{r^2}(t) = \overline{x^2}(t) + \overline{y^2}(t) + \overline{z^2}(t) = 6Dt \qquad (32)$$

Notice that in all dimensions, the mean square displacement is linear in the number of steps N and in the time t. The diffusion coefficient D characterizes the migration of particles in a given kind of medium at a given temperature. In general, it depends on the size of the particle, the structure of the medium, and the absolute temperature. For a small molecule in water at room temperature, $D \simeq 10^{-5}$ cm^2/sec.

If we go back to the drunk who is taking steps of size s to the right or to the left, then we can use the binomial distribution. Let p be the probability to take a step to the right and $q = 1 - p$ be the probability to take a step to the left. After taking N steps, the mean number $\overline{n}_1$ of steps to the right is (see Reif Eq. (1.4.4)):

$$\overline{n}_1 = \sum_{n_1=0}^{N} P(n_1, N) n_1 = Np \tag{33}$$

The mean number of steps to the left is

$$\overline{n}_2 = Nq \tag{34}$$

Let the displacement be

$$\Delta x = ms = (n_1 - n_2)s \tag{35}$$

The mean displacement is (Reif Eq. (1.4.6)):

$$\overline{\Delta x} = \overline{m}s = (\overline{n}_1 - \overline{n}_2)s = N(p - q)s \tag{36}$$

So if $p = q = 1/2$, then the mean displacement $\overline{\Delta x} = 0$.

The displacement squared is

$$(\Delta x)^2 = (\Delta m)^2 s^2 = (m - \overline{m})^2 s^2 \tag{37}$$

The dispersion of the net displacement or the mean square displacement is

$$\overline{(\Delta x)^2} = \overline{(\Delta m)^2} s^2 = 4Npqs^2 \tag{38}$$

where we used Reif Eq. (1.4.12). If $p = q = 1/2$, then the mean square displacement $\overline{(\Delta x)^2} = Ns^2$ which is what we got before. The rms displacement is

$$\sqrt{\overline{(\Delta x)^2}} = 2\sqrt{Npq}s \sim \sqrt{N}s \tag{39}$$

If $p = q = 1/2$, then the rms displacement is $\sqrt{\overline{(\Delta x)^2}} = \sqrt{N}s$ which is what we got before. Notice again that the rms displacement goes as the square root of the number of steps. Again the characteristic or typical displacement is given by $\sqrt{\overline{(\Delta x)^2}} \sim \sqrt{N}s$.

Poisson Distribution

(Ref.: C. Kittel, *Thermal Physics*) It is worth mentioning another widely used distribution, namely, the Poisson distribution. This distribution is concerned with the occurrence of small numbers of objects in random sampling processes. For example, if on average there is one bad penny in 1000, what is the probability that no bad pennies will be found in a given sample of 100 pennies? The problem was first consided and solved by Poisson (1837) in a study of the role of luck in criminal and civil trials in France.

So suppose that there are N trials and the probability for a desired outcome is p. The probability that there will be n desired outcomes is

$$P(n) = \frac{\langle n \rangle^n e^{-\langle n \rangle}}{n!} \tag{40}$$

where the mean number of desired events is $\langle n \rangle = Np$. You will derive the Poisson distribution in your homework starting from the binomial distribution.

Quantum Mechanics

Since you are just starting your quantum mechanics course, let's just go over what you need to know for starters. Mostly it's just notation and jargon. Particles or systems of particles are described by a wavefunction ψ. The wavefunction is a function of the coordinates of the particles or of their momenta, but not both. You can't specify both the momentum and position of a particle because of the Heisenberg uncertainty principle

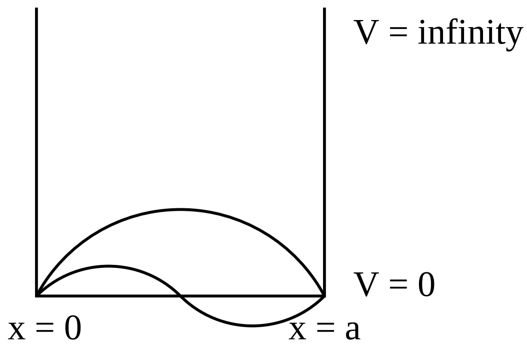
$$\Delta x \Delta p_x \gtrsim \hbar \tag{41}$$

So the wavefunction of a particle is written $\psi(\vec{r})$. The wavefunction is determined by Schroedinger's equation

$$H\psi = E\psi \tag{42}$$

where H is the Hamiltonian and E is the energy. The Hamiltonian is the sum of the kinetic and potential energies. You solve Schroedinger's equation to determine the wavefunction. Schroedinger's equation can be written as a matrix equation or as a second order differential equation. Often the wavefunction solutions are labeled by definite values of energy, momentum, etc. These values can only have discrete values, i.e., they are quantized. They are called quantum numbers. We can label the different wavefunction solutions by their quantum numbers. Quantum numbers correspond to conserved quantities like energy, momentum, angular momentum, spin angular momentum, etc. Perhaps the most familiar example of this is the electronic orbitals in the hydrogen atom. The electron orbitals, like the 1s and 2p states, have definite discrete energies and definite values of the orbital angular momentum. For example, the energy of the 1s state in hydrogen is 1 Rydberg or 13.6 eV. The electron in this state can be described by a wavefunction $\psi_{1s}(r)$. The quantum numbers of this state are $n = 1$ and $L = 0$. $n = 1$ means that it has energy $E_1 = 13.6$ eV and L is the orbital angular momentum. (s is the spectroscopic notation for $L = 0$.) Solving the Schroedinger equation for the hydrogen atom yields many wavefunctions or orbitals (1s, 2s, 2p$_x$, 2p$_y$, 2p$_z$, 3s, etc.).

Another example is a particle in a box.



The energy eigenvalues
$$E_n = \frac{p^2}{2m} = \frac{n^2\pi^2\hbar^2}{2ma^2} \tag{43}$$
Each value of n corresponds to a different eigenvalue of the energy E_n. Notice that the energy levels are not equally spaced; they get farther apart as you go up in energy. Each value of n also corresponds to a different wavefunction
$$\psi_n(x) = A\sin(px/\hbar) = A\sin\left(\frac{n\pi x}{a}\right) \tag{44}$$
Notice that the more nodes there are, the more wiggles there are, and the higher the energy is.

A harmonic oscillator is another example. This is just the quantum mechanical case of a mass attached to a spring. In this case the potential is a parabola rather than being a square well. A particle of mass m in this potential oscillates with frequency ω. The Hamiltonian is
$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \tag{45}$$
You will learn how to solve this in your quantum mechanics course. Let me just write down the result. The energy eigenvalues are
$$E_n = (n + \frac{1}{2})\hbar\omega \tag{46}$$
where $n = 0, 1, 2, 3, ...$ Notice once again that the energy levels are quantized. In this case they are evenly spaced by an amount $\Delta E = \hbar\omega$. The harmonic oscillator has Gaussian wavefunctions (Hermite polynomials).

Wavefunctions can describe more than one particle. You can have a wavefunction associated with a state with many particles. We call this a many body wavefunction. In this case we can describe a state with N particles by $\psi_{\{n\}}(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$. The subscript $\{n\}$ denotes the quantum numbers that label the wavefunction.