Apr. 1

Microcanonical Ensemble

Now we begin the study of statistical mechanics proper. The first thing that we must do is to understand what is meant by *thermodynamic equilibrium.* In principle, we would like to understand exactly when a system is in thermodynamic equilibrium and how fast a system comes to equilibrium. However, I hope you will be happy if I give you a general, practical prescription for computing the coarse-grained properties of states of thermal equilibrium. By "practical", I mean a prescription that does not involve solving the microscopic equations of motion.

To begin our study of thermal equilibrium, I will focus on a particularly simple example, that of an *ideal gas.* This is a collection of N structureless atoms in a volume V. These atoms have strong, energy-conserving interactions with one another, allowing them to exchange energy and momentum. However—the is the *ideal* aspect of an ideal gas—these interactions occupy a negligible amount of time in the flight path of an atom. The atoms are confined to a box with smooth, rigid walls, so that the atoms can bounce off the walls without losing energy. This system of atoms has a total energy E that is conserved. I will assume, for simplicity, that these atoms obey classical mechanics.

If we were to analyze this problem at the microscopic level, we would begin from a complete set of initial conditions. This would be a specification of the positions and momenta of each atom at a specific time. To solve the full mechanics problem, we would integrate the equations of motion for each atom, following each along its linear trajectory, computing the deflection at each scattering, and continuing along the new trajectory. This calculation is basically hopeless to carry out in practice. Even if we could actually know all of the initial conditions, we would rapidly lose control, because the deflection angles depend very sensitively on the impact parameters, which we would never be able to compute exactly. To be able to compute the macroscopic properties of this sytem, we need to look at it in a completely different way.

Here is possible different approach. I claim that the coarse-grained properties of the gas are the same as those of an ensemble of states in which *every state of energy E appears with equal probability.* This collection of states is called the *microcanonical ensemble.* This hypothesis is obviously incorrect at a microscopic level. The gas is in a *particular* state of energy E. The specific positions of the particles reflect that particularity. Nevertheless, it is possible that any *typical* state of energy E will have the same density and distribution of atoms as we find in this particular state, that is, the same properties at the level above that of specific atomic coordinates. In this case, the hypothesis would make sense.

I will postpone further discussion of the validity of the microcanonical ensemble and instead go directly to the calculation of the macroscopic properties of the gas based on this hypothesis.

To do this, we first need a more precise definition of the microcanonical ensemble. Again, I will work here with systems obeying classical mechanics. Such systems are described by phase space variables

$$\{q_i, p_i\}$$

The energy of the system is given in terms of a Hamiltonian function by

$$E = \mathcal{H}(q, p)$$

Define a small quantity δE, and let $\Theta_{\delta E}$ denote a restriction to energies in an interval of size δE about E,

$$\Theta_{\delta E}(q, p, E) = \begin{cases} 1 & \text{if } |\mathcal{H}(q,p) - E| < \frac{\delta E}{2} \\ 0 & \text{otherwise} \end{cases}$$

Then I propose to compute properties of the microcanonical ensemble in the following way: Let $\mathcal{O}(q, p)$ be an observable associated with the system. The expected value of $\mathcal{O}(q, p)$ in the microcanonical ensemble is given by the average of $\mathcal{O}(q, p)$ over states in the energy slice,

$$\langle \mathcal{O} \rangle = \frac{\int dq\, dp \;\; \Theta_{\delta E}(q\, p\, E) \;\; \mathcal{O}(q, p)}{\int dq\, dp \;\; \Theta_{\delta E}(q\, p\, E)}$$

It is useful to define

$$\Omega(E)\,\delta E = \int dq\,dp\ \Theta_{\delta E}(q,p,E)$$

so that $\Omega(E)$ is the *area* of the surface of constant energy in the full phase space.

In analyzing the idea gas, I will ignore the small time intervals in which the atoms are actually scattering from one another. Then the Hamiltonian is given by

$$H(q,p) = \sum_{j=1}^{N} \frac{|\vec{p}_j|^2}{2m}$$

I will assume that the number of particles is very large, of the order of Avogadro's number $N_A = 6 \times 10^{23}$.

The first conclusion from this formalism is that, since $H(q,p)$ is independent of the q_i, each atom is equally likely to be anywhere in the box. This makes intuitive sense for a description of the equilibrium state. It is not that the microcanonical ensemble predicts a fixed constant density of particles. It does allow for the possibility that all of the atoms are, for example, in the right-hand half of the volume. However, the probability of this is very small,

$$\left(\tfrac{1}{2}\right)^N = \exp\left[-N \log 2\right]$$

where $N \sim N_A$. This is a *small* probability indeed.

Still, in the microcanonical ensemble, there is typically a small difference between the number of particles in the left-hand side of the box and the number in the right-hand side of the box. It is instructive to compute this quantity. The probability that there are m extra particles on the right-hand side is that same as that for m extra heads in a series of N coin flips, or m extra steps to the right in a 1-dimensional random walk of N steps,

$$\frac{1}{2^N}\binom{N}{\frac{N+m}{2}} = \frac{1}{2^N}\,\frac{N!}{\left(\frac{N+m}{2}\right)!\,\left(\frac{N-m}{2}\right)!}.$$

Using Stirling's formula

$$\log(n!) = (n+\tfrac{1}{2})\log n - n + \tfrac{1}{2}\log 2\pi + O(\tfrac{1}{n})$$

we worked out the approximation to this formula for $N \gg 1$ in the previous lecture. It is

$$P_N(m) = \frac{1}{\sqrt{2\pi N}} e^{-m^2/2N}$$

From this formula,

$$\langle m \rangle = 0$$

as expected, and

$$\langle m^2 \rangle = N$$

So typically values of the *relative* fluctuation m/N are

$$\frac{m}{N} \sim \frac{1}{\sqrt{N}} \sim 10^{-12}$$

The conclusion is that, while there are configurations on the surface of constant energy that have the atoms all on one side, *typical* configurations have a distributions of atoms that is uniform to a very good approximation.

Next, we can compute the momentum distribution. The total energy is given by

$$E = \sum_{j=1}^{N} \frac{|\vec{p}_j|^2}{2m}$$

Thus, states of a given energy lies on a *sphere* of radius $(2mE)^{1/2}$ in a space of $3N$ dimensions. To compute averages, we first need to know

$$\Omega_N(E)$$

which is proportional to the area of this sphere. There is a nice way to work out the area of the unit sphere in d dimensions. Consider the integral

$$\int d^d x \; e^{-x^2} = (\sqrt{\pi})^d$$

In polar coordinates, this is

$$\int_0^\infty dx \; x^{d-1} \, A_d \; e^{-x^2}$$

where A_d is the area of the unit sphere in d dimensions. Change variables to $y = x^2$, with $dy = 2xdx$. Then

$$\pi^{d/2} = \frac{1}{2} A_d \int_0^\infty dy \; y^{d/2 - 1} \, e^{-y}$$

I hope that you recognize the integral on the right-hand side as the Gamma function

$$\Gamma(z) = \int_0^\infty dy \; y^{z-1} e^{-y}$$

The basic property of $\Gamma(z)$ is its recursion formula

$$\Gamma(z) = (z-1)\,\Gamma(z-1)$$

derived by integration by parts,

$$\Gamma(z) = -y^{z-1}e^{-y}\Big|_0^{\infty} + (z-1)\int_0^{\infty} dy\, y^{z-2}e^{-y}$$

$$= 0 + (z-1)\,\Gamma(z-1)$$

It is easy to see that $\Gamma(1) = 1$, then $\Gamma(2) = 1$, $\Gamma(3) = 2$, and, in general

$$\Gamma(n+1) = n!$$

Also

$$\Gamma(1/2) = \int_0^{\infty} dy\, y^{-1/2}e^{-y} = 2\int_0^{\infty} dx\, e^{-x^2} = \sqrt{\pi}$$

which implies that

$$\Gamma(1/2) = \sqrt{\pi} \qquad \Gamma(3/2) = \tfrac{1}{2}\sqrt{\pi} \qquad \Gamma(n+1/2) = \frac{1}{2}\cdot\frac{3}{2}\cdots\frac{2n-1}{2}\cdot\sqrt{\pi}$$

Then we have a final formula for A_d,

$$A_d = \frac{2\pi^{d/2}}{\Gamma(d/2)} = \frac{2\pi^{d/2}}{(d/2-1)!}$$

You can easily check that the first few cases come out correctly:

$$d=1 \quad A_d = 2 \quad ; \quad d=2 \quad A_d = 2\pi \quad ; \quad d=3 \quad A_d = 4\pi \quad ; \quad d=4 \quad A_d = 2\pi^2$$

For the problem at hand, $\Omega(E)$ requires the area of a sphere of radius $R = (2mE)^{1/2}$ and thickness

$$\Delta = [2m(E + \delta E/2)]^{1/2} - [2m(E - \delta E/2)]^{1/2}$$

$$= \frac{1}{2}\left(2m/E\right)^{1/2} \delta E$$

Then

$$\Omega_N(E)\,\delta E = V^N A_{3N} (2mE)^{\frac{3N-1}{2}} \cdot \frac{1}{2}\left(\frac{2m}{E}\right)^{1/2} \delta E$$

$$= V^N \frac{\pi^{3N/2}}{(\frac{3N}{2}-1)!} (2mE)^{\frac{3N}{2}-1} \cdot 2m \cdot \delta E$$

Here is a relevant quantity to address with this formula: What is the probability that a given atom has momentum $\vec{p}$? We can express this probability as a distribution

$$d^3p\, f(\vec{p}) \qquad \text{such that} \qquad \int d^3p\, f(\vec{p}) = 1$$

To compute $f(\vec{p})$, recognize that, if this one particle has momentum $\vec{p}$ and energy $p^2/2m$, the other $(N-1)$ particles have energy $E - p^2/2m$. Then

$$d^3p\, f(\vec{p}) = \frac{d^3p \; V \; \Omega_{N-1}(E - p^2/2m)\,\delta E}{\Omega_N(E)\,\delta E}$$

First, work out the dependence on $|p|$,

$$\Omega_{N-1}\left(E - p^2/2m\right) \sim \left(E - p^2/2m\right)^{\frac{3N}{2} - 5/2}$$

$$\sim \left(1 - \frac{p^2/2m}{E}\right)^{\frac{3}{2}N - 5/2}$$

Typically, the energy is divided more or less equally among all of the particles, so we expect

$$\frac{\frac{p^2}{2m}}{E} \sim \frac{1}{N} \ll 1$$

For $(p^2/2m)/E \ll 1$, the above formula takes the form

$$\left(1 - \frac{\alpha}{N}\right)^{N\cdot\beta} \xrightarrow[\lim N\to\infty]{} e^{-\alpha\beta}$$

Using this result, we can replace

$$\left(1 - \frac{p^2/2m}{E}\right)^{\frac{3}{2}N - \cdots} \longrightarrow e^{-p^2/2m \cdot \frac{3N}{2E}}$$

The factor

$$\frac{2E}{3N}$$

will permeate the following discussion, so I would like to give it a name:

$$\frac{E}{\left(\frac{3}{2}N\right)} = T \qquad \text{"temperature"}$$

then

$$f(\vec{p}) \sim e^{-\frac{p^2/2m}{T}} = \exp\left[-\frac{\text{Particle energy}}{T}\right]$$

This distribution of momenta is called the *Boltzmann distribution.* We will justify it from a more general point of view in a later lecture.

There are still a few loose ends in our derivation of $f(\vec{p})$ that should be cleaned up. First of all, we should check that the normalization derived from the above theory is correct. Go back to the microcanonical expression for $f(\vec{p})$ and pick up the contributions to the prefactor,

$$f(\vec{p}) = \frac{1}{\pi^{3/2}}\left[\frac{\left(\frac{3}{2}N-\frac{5}{2}\right)!}{\left(\frac{3}{2}N-1\right)!}\right]^{-1}\frac{1}{(2mE)^{3/2}}\; e^{-\frac{p^2/2m}{T}}$$

Now apply Stirling's formula,

$$\log\frac{\left(\frac{3}{2}N-\frac{5}{2}\right)!}{\left(\frac{3}{2}N-1\right)!} = \log\frac{\Gamma\left(\frac{3}{2}N-\frac{3}{2}\right)}{\Gamma\left(\frac{3}{2}N\right)}$$

$$= \left(\tfrac{3}{2}N-2\right)\log\left(\tfrac{3}{2}N-\tfrac{5}{2}\right) - \left(\tfrac{3}{2}N-\tfrac{5}{2}\right) - \left(\tfrac{3}{2}N-\tfrac{1}{2}\right)\log\left(\tfrac{3}{2}N-1\right) + \left(\tfrac{3}{2}N-1\right) + O\left(\tfrac{1}{N}\right)$$

$$= -\frac{3}{2}\log\frac{3}{2}N + \frac{3}{2}N\left(-\frac{5/2}{3/2N}\right) - \frac{3}{2}N\left(-\frac{1}{3/2N}\right) + \frac{3}{2} + O\left(\frac{1}{N}\right)$$

$$= -\frac{3}{2}\log\frac{3N}{2} + O\left(\frac{1}{N}\right)$$

This gives

$$f(\vec{p}) = \frac{1}{\left[\pi(2mE)\cdot\left(\frac{3N}{2}\right)^{-1}\right]^{3/2}}\; e^{-\frac{p^2/2m}{T}}$$

or, finally,

$$f(\vec{p}) = \frac{1}{(2\pi m T)^{3/2}} e^{-p^2/2mT}$$

Using the formulae for Gaussian integrals from the previous lecture, we see that this expression is correctly normalized as a probability,

$$\int d^3p \; f(\vec{p}) = 1$$

From the properties of the Gaussian integral,

$$\langle \vec{p} \rangle = 0$$

The mean square momentum is

$$\langle (p_i)^2 \rangle = \frac{\int dp_i \; p_i^2 \, e^{-p_i^2/2mT}}{\int dp_i \; e^{-p_i^2/2mT}} = mT$$

so that

$$\left\langle \frac{p_i^2}{2m} \right\rangle = \frac{1}{2} T \qquad \left\langle \frac{|\vec{p}|^2}{2m} \right\rangle = \frac{3}{2} T$$

This last result is called the *equipartition theorem.* Again, I will give a more general derivation later in the course.

According to the equipartition theorem, the average energy of each particle is

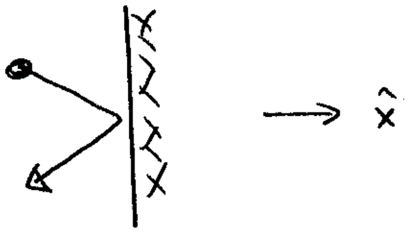
$$\frac{3}{2} T$$

A given particle will have an energy that is different from this value by a factor of order 1. But, if we consider a collection of many particles, the typical fluctuation of the total energy will be smaller relative to the whole. In a subset of M particles, the root mean square fluctuation in energy will be of the order of $\sqrt{M}T$, and so the *relative* fluctuation will be of order

$$1/\sqrt{M}$$

Even for a small number of particles in the collection, the typical energy of the collection is close to the average.

There is one more important property of a gas that we can compute from the microcanonical ensemble, the *pressure* that the gas exerts on the walls of the box. Consider a wall perpendicular to the $\hat{x}$ axis. Atoms reflect off this wall and reverse their $\hat{x}$ motion, without transfering any *energy* to the wall



A particle does transfer *momentum* in the amount

$$\Delta\vec{p} = 2p_x \ \hat{x}$$

In the time interval Δt, all atoms with $p_x > 0$ that are a distance less than

$$v_x \Delta t = \frac{p_x}{m} \Delta t$$

from the wall will transfer momentum to the wall. The total momentum transfer in time Δt is then

$$\Delta P_x = \int_{p_x > 0} d^3p \; f(\vec{p}) \cdot \frac{N}{V} \cdot \left(\frac{p_x}{m} \Delta t\right) \cdot 2p_x \cdot (\text{Area})$$

The rate of momentum transfer is just the force exerted on the wall. The force per unit area, the pressure, is given by

$$\left(\frac{\text{Force}}{\text{Area}}\right) = \frac{N}{V} \int_{p_x > 0} d^3p \; f(\vec{p}) \left(\frac{p_x^2}{2m}\right) \cdot 4$$

$$= \frac{N}{V} \cdot \frac{1}{2} \cdot \frac{1}{2} T \cdot 4$$

Thus,

$$\text{pressure} = P = \frac{NT}{V}$$

We have now derived the basic properties of an ideal gas in terms of the underlying parameter T:

$$E = \frac{3}{2} NT \qquad P = \frac{NT}{V}$$

The analysis was simple. It seems, perhaps, too simple. However, in the next lecture, I will raise these ideas to a higher level of generality.