

4 Kinetic Theory and Statistical Mechanics

May 2, 2007

Introduction

The idea that the familiar states of matter — solids, liquids and gases — could be explained in terms of atoms or molecules closely packed together, more loosely packed together, or moving more or less independently of one another, is a very old one. The simplicity and generality of the gas laws encouraged the advocates of this view to think that the description of gases in terms of molecules might be correspondingly simple.

The picture of a dilute gas that leads to the gas laws can be described as follows. The molecules of the gas are moving round in a container at high speeds (comparable with the speed of sound in the gas). Because the gas is dilute the molecules are rather far apart, and do not interact significantly with one another unless they get unusually close. Such an event we call a *collision*, and in a collision there is exchange of energy and momentum between the molecules, but the total energy is conserved, as is the total momentum of the pair. When the molecules collide with the walls of the container, their momentum is changed, and the average rate at which they exchange momentum with the walls is what we call the pressure force. Boyle's Law follows from this, and Avogadro's Law follows if we identify the Kelvin temperature as something proportional to the average kinetic energy of a molecule.

It is very satisfactory that we can get the relations between pressure, temperature and volume for a dilute gas in terms of this picture, but we can go a lot further, while losing some of the simplicity of the picture. The pressure depends only on the translational motion of molecules, but the argument can be extended to say that the internal modes of motion of a molecule should also have an average kinetic energy proportional to T . This leads to predictions about specific heats of gases and solids, but these predictions are only partially fulfilled. This failure led Kelvin, in 1900, to publish an article called *Nineteenth century clouds over the dynamical theory of heat and light*. This was just before quantum theory came to the rescue.

This theory also allows one to make less well-based predictions about how gases differ from ideal gases when they are not so dilute. The van der Waals equation was the result of one such attempt. Once quantum mechanics was in place, 80 years ago, it was calculated that well-separated molecules should attract one another with a $1/r^6$ potential, which van der Waals assumed in deriving his equation of state for the gas–fluid system. It is named the van der Waals potential in honor of the man who devised the equation of state for simple fluids.

Since molecules can travel a certain distance, known as the *mean free path* λ , between collisions, in an inhomogeneous gas they will move between different regions, smoothing

out the inhomogeneities. The rate at which they transport kinetic energy from high temperature regions to low temperature regions gives the thermal conductivity, the rate at which they transport momentum determines the viscosity, and the rate at which they transport different forms of matter determines the *diffusion constant*. Each of these *transport coefficients* is proportional both to the mean free path of the molecules and to their typical speeds.

In these notes I discuss the basis of the gas laws in section 4.1, and follow this with a discussion of equipartition and specific heats in section 4.2. There is then a brief discussion of collisions, mean free paths, and their influence on transport coefficients in section 4.4. These topics are covered in a similar way in many text-books. I then give brief discussions of how the law of large numbers smooths out fluctuations in large systems in section 4.7, and a discussion of diffusion and Brownian movement in section 4.5.

Then I will finish these notes with a brief account of the necessary parts of probability theory. It is not so important that you understand it at this time, although there are many varied reasons for learning it eventually. I strongly recommend a book by Vinay Ambegaokar called *Reasoning about luck: probability and its uses in physics*, published by the Cambridge University Press in 1996, which was written for nonscientists. For a more technical treatment see Chapter 16 of *Mathematical methods in the physical sciences* by Mary Boas (Wiley 1983), or some equivalent textbook in mathematical methods. I describe the Boltzmann distribution in section 4.6 using a rather different approach from that used in some books, I give the basic ideas of probability theory in section 4.8, and finish off with two useful mathematical results in section 4.9.

4.1 Pressure and the momentum of molecules

The picture that we use of the molecules in a dilute gas is that they are so well separated that most of the time they are moving in straight lines at constant velocity. Occasionally they collide with one another, preserving total energy and momentum, and even less frequently they collide with the walls of the container, transferring momentum to the wall, and keeping the same energy, at least on the average.

If one considers the molecules bouncing off a wall normal to the x direction, each molecule transmits an impulse $2mv_x$ to the wall when it bounces, and the rate at which molecules hit the wall is proportional to $v_x N/V$ per unit area, so that the average force per unit area is equal to N/V times the average of mv_x^2 . Comparison with the ideal gas equation of state $pV = Nk_B T$, where the Boltzmann constant k_B is related to the gas constant R and Avogadro's number N_0 by $R = N_0 k_B$, shows that

$$\frac{1}{2} \langle mv_x^2 \rangle_{\text{ave}} = \frac{pV}{2N} = \frac{1}{2} k_B T, \quad (1)$$

and so each component of the motion of each molecule of an ideal gas contributes an average value $k_B T/2$ to the kinetic energy. This gives us a new interpretation of the

ideal gas scale of temperature (Kelvin), as proportional to the average kinetic energy of each mode of motion of each molecule. We notice that there is a tendency for the kinetic energies of colliding molecules to equalize, so that the typical speed is proportional to the square root of the temperature divided by the mass, and the typical magnitude of momentum is proportional to the square root of the mass times temperature.

We discussed pressure as the force on a wall, but we are interested in it also in the interior of a gas, for example when we discuss the change of pressure with altitude. In the interior of an ideal gas pressure is not a force on any particular unit area of solid, but is the rate at which the momentum of the molecules flows across any unit area in the gas. Consider a slab of gas bounded by two parallel planes of area A , one at the height z , the other at the height $Z + \Delta Z$; these two planes are not solid, but just planes in the interior of the gas. The atoms passing across the lower plane from below are carrying upwardly directed momentum upwards into the slab, while those passing across it from above are carrying downward momentum downwards out of the slab. Both these processes increase the net positive z component of momentum in the slab, so there is an upward flow of momentum into the slab at Z . Similarly, at $Z + \Delta Z$ there is an upward flow of momentum out of the slab. Since the density of molecules decreases with height, more momentum flows in at Z than flows out at $Z + \Delta Z$. This sounds mysterious until one thinks about the fact that all the molecules in the slab are acted on by the gravitational force $-mg\hat{z}$, and this upward flow of momentum is just what is needed to balance the negative acceleration, so that the momentum density remains constant in equilibrium.

4.2 Equipartition and heat capacities

We know that for a gas which is a mixture of different types of molecules, each type of molecule makes an independent contribution to the pressure, which is known as the partial pressure of that molecular species.

Classical statistical mechanics, based on the Boltzmann distribution, shows that the equipartition of kinetic energy between the different modes of motion of a molecule should be much more general than that, applying to internal degrees of freedom, such as molecular rotations and vibrations, as well as to center of mass motion. If that is taken literally it says that the total kinetic energy of a system of N_A atoms at temperature T should have kinetic energy $3N_A k_B T/2$. We also know that in a harmonic oscillator, there is, on average, as much excess potential energy (above the equilibrium potential energy) as there is kinetic energy, so each vibrational mode of motion should contribute a potential energy $k_B T/2$, in addition.

For gases this says that the heat capacity at constant volume should be $3Nk_B/2$ from the translational motion of the molecules, $Nk_B/2$ for each rotational mode of motion, and Nk_B for each vibrational mode (one half from kinetic energy, one half from potential energy). This would give a heat capacity $7Nk_B/2$ for a diatomic molecule (rotational

motion about two different axes, and vibration along the molecular axis), and $6Nk_B$ for a nonlinear triatomic molecule (three rotational axes and three vibrational modes). These would give ratios of the heat capacities $\gamma = C_p/C_v$ as $9/7$ for a diatomic molecule, and $7/6$ for the triatomic molecule.

In reality the heat capacities at room temperature are much less than that for most gases, except for monatomic gases. Typically, in gases like oxygen and nitrogen, C_v is $5Nk_B/2$, and γ is $7/5$, and in triatomic molecules γ may not be much less than $4/3$. These are the values one gets if the rotational motion is contributing fully, but the vibrational motion is frozen out. The explanation came from Planck's quantum theory, which says that a vibrational mode of frequency ν does not contribute significantly to the heat capacity if

$$h\nu > k_B T, \quad (2)$$

where h is Planck's constant. At lower temperatures this condition may be reached even for the rotational frequencies, particularly if the moments of inertia are small, as they are for hydrogen and methane, so C_v drops below $5Nk_B/2$, and γ rises towards $5/3$. The expression that quantum statistical mechanics gives for the average energy of a vibrational mode with frequency ν is

$$\bar{E}_\nu = \frac{1}{2}h\nu \coth \frac{h\nu}{2k_B T},$$

which is close to $k_B T$ for large T , and goes rapidly to the temperature-independent zero-point energy $h\nu/2$ at low temperatures.

For solids, equipartition says that each of the N_a atoms, held in position by its neighbors, should have three vibrational degrees of freedom, giving the solid a total average energy of $3N_A k_B T$ and a heat capacity of $3N_A k_B$. This is the Dulong-Petit law for the heat capacity of solids, obeyed for many solids at room temperature. Bonding between atoms in metals or between ions in ionic crystals is typically weaker than covalent bonding in molecules, so vibrational frequencies are lower. Diamond is both covalently bonded and has low atomic mass, so it has a much lower heat capacity than given by the Dulong-Petit law and graphite also has low atomic specific heat. For any solid the heat capacity drops as the temperature is lowered. Again the quantum condition (2) shows at what temperature the heat capacity drops.

4.3 Black-body radiation

In a cavity, whether it is evacuated or simply transparent, in which the walls are kept at a uniform temperature, equilibrium will be set up with the electromagnetic field in the cavity, so that the emission of radiation by the walls balances the absorption of radiation by the walls. Each different mode of the electromagnetic waves satisfies the equipartition law, provided its frequency is below the limit given by (2). It is fairly easy to count the possible modes of the waves if one uses a simple geometry such as a cavity with rectangular

faces, and the answer for the number of waves with wavenumber (2π over the wavelength) less than K in a cavity of volume V is

$$N(K) = \frac{VK^3}{3\pi^2}. \quad (3)$$

According to equation (2) these will each have an energy $k_B T$ if the condition $hcK/2\pi < k_B T$ is satisfied, where c is the speed of light; it is $k_B T$ because both electric and magnetic fields contribute separately to equipartition. This gives the total electromagnetic energy in the cavity as

$$N(K)k_B T \approx \frac{8\pi(k_B T)^4}{3h^3 c^3} V. \quad (4)$$

This is not the correct formula, as the cut-off at a sharp value of K is not correct, but Planck's quantum theory of cavity radiation gave the correct shape of spectrum, with a peak frequency proportional to $k_B T/h$, total energy proportional to T^4 , and heat capacity proportional to T^3 .

The Planck formula enables the temperature in the interior of a furnace to be measured with a spectrometer, and the surface temperatures of stars are measured by observing the shape of the spectrum. The most spectacular example of a black-body spectrum is in the cosmic background radiation, discussed in Sec. 3.7. This is a close fit to equilibrium at 2.735 K, despite the fact that it has not been in equilibrium for the last 10^{10} years, when it left matter at a temperature of 3000 K in a region receding from us so fast that the Doppler shift reduces the frequency by a factor of more than 10,000.

4.4 Collisions and transport coefficients

We consider a monatomic gas in which the atoms have a radius a , in the sense that $2a$ is the distance at which there is a strong repulsion between the atoms (this can be determined from the density of the solid at low temperatures). Similar arguments can be constructed for roughly spherical molecules like methane or carbon tetrachloride. If we ignore the longer range attractive force there will be a collision whenever the center of one atom gets within a distance $2a$ of another. We can calculate the frequency with which a particular atom has a collision by imagining a cylinder of radius $2a$ and length ut , volume $4\pi a^2 ut$ where u is the speed of the atom and t is the time interval. With a density of atoms N/V , a typical time τ for a collision is given by

$$4\pi a^2 u \tau = \frac{V}{N}, \quad \text{or} \quad \tau = \frac{V/N}{4\pi a^2 u} = \frac{\lambda}{u}.$$

where the mean free path λ is

$$\lambda = \frac{V/N}{4\pi a^2}. \quad (5)$$

There is a better way of putting this argument using probability theory in the form developed in sec. 4.8. The probability that a particular second atom is not in the volume $4\pi a^2 l$ swept out by the first atom in going a distance l is just $1 - 4\pi a^2 l/V$, so the probability that none of the other $n - 1$ atoms are in this volume is

$$\left[1 - 4\pi a^2 l/V\right]^{N-1} \approx e^{-l/\lambda} . \quad (6)$$

where the mean free path is

$$\lambda = \frac{V/N}{4\pi a^2} . \quad (7)$$

This only takes into account the short-range, repulsive, part of the intermolecular interaction, and even for that I have hidden some subtleties. The $1/r^6$ attractive potential produces some additional scattering, but, because the potential is weaker, it tends to produce scattering at smaller angles, and to be less important at higher temperatures.

Once a mean free path is known for a gas, transport coefficients such as diffusion coefficients, viscosities, and thermal conductivities can be estimated. The simplest sort of process to consider is diffusion of one sort of molecule in a gas that is primarily of another sort. This is the process by which, for example, if a small quantity of a heavy gas like ozone is released at the bottom of a flask of still air, it will eventually spread evenly through the flask. The diffusion rate is determined by the mean free path λ of the ozone molecule in air and by the typical speed of the ozone molecule. If the molecule goes a distance λ in a time τ before it is scattered into a direction unrelated to its initial direction, in time $n\tau$ it will have made n steps each of length λ , but with no relation between the directions of the steps. This gives a process known as a *random walk*. If one considers two successive time intervals, the mean square displacement can be written as

$$\begin{aligned} \langle(\delta\mathbf{r}_1 + \delta\mathbf{r}_2)^2\rangle_{\text{ave}} &= \langle(\delta\mathbf{r}_1)^2\rangle_{\text{ave}} + \langle(\delta\mathbf{r}_2)^2\rangle_{\text{ave}} + 2\langle\delta\mathbf{r}_1\delta\mathbf{r}_2\rangle_{\text{ave}} \\ &= 2\langle(\delta\mathbf{r}_1)^2\rangle_{\text{ave}} , \end{aligned} \quad (8)$$

since the crossed term vanishes if the time interval is large enough that there is no correlation between the directions of successive steps. The argument can easily be generalized to show that the root mean square of the time gone in n steps, in time $\tau' = n\tau$, is \sqrt{n} times the root mean square displacement in time τ ; the distance gone is proportional to the square root of the time elapsed.

In time τ about one half of the molecules lying within a distance $\lambda = v\tau$ to the left of the plane $x = x_0$ will cross from $x < x_0$ to $x > x_0$, while one half of the molecules lying within a distance $\lambda = v\tau$ to the right of the plane $x = x_0$ will cross from $x > x_0$ to $x < x_0$. If the density ρ_i of impurity is the same on the two sides of $x = x_0$ the flows in the two directions across $x = x_0$ will cancel, but if there is a density gradient there will be a net

flow across an area A in time τ given by

$$Aj_x\tau = \frac{1}{2}A\lambda[\rho_i(x - \lambda/2, t) - \rho_i(x + \lambda/2, t)] \approx -A\frac{\lambda^2}{2}\frac{\partial\rho_i}{\partial x}, \quad (9)$$

so that the current density due to diffusion is $-\lambda^2/2\tau$ times the density gradient.

In addition, since the total mass of the material is conserved, the rate of increase of the amount of material in a slab of thickness λ , area A , is equal to the difference of the flow through the plane $x = x_0 - \lambda/2$ and the plane $x = x_0 + \lambda/2$. This can be written as

$$A\lambda\frac{\partial\rho_i}{\partial t} = A[j_x(x_0 - \lambda/2) - j_x(x_0 + \lambda/2)] = -A\lambda\frac{\partial j_x}{\partial x},$$

which gives the matter conservation condition in the form

$$\partial\rho_i/\partial t + \partial j_x/\partial x = 0. \quad (10)$$

Combination of Eqs. (9) and (10) gives the *diffusion equation*

$$\frac{\partial\rho_i}{\partial t} = D\frac{\partial^2\rho_i}{\partial x^2}, \quad (11)$$

where the *diffusion constant* D is equal to $\lambda^2/2\tau$. An important solution of such an equation has the form

$$\rho_i \propto \frac{1}{\sqrt{t}}e^{-x^2/4Dt}, \quad (12)$$

which describes a gaussian shape spreading out in such a way that its width is proportional to the square root of the time elapsed.

Because it is hard to avoid convection currents in a gas, it is easier to observe it for liquids. The spreading of a dye in a still liquid, or of coffee on a wet paper towel are good examples. For these examples, however, the mean free path is usually comparable with interatomic spacing.

There is a similar equation for the spread of a temperature change in a constant temperature background. In kinetic theory we describe the molecules whose average kinetic energy is raised (or lowered) as carrying their extra kinetic energy a distance λ_{th} in time τ_{th} , so that the extra kinetic energy does a random walk in space. The diffusion constant for thermal transport, as I showed in sec. 3.4, is the thermal conductivity divided by the specific heat per unit volume. The thermal conductivity is therefore given by

$$k = \rho c D_{th}.$$

If the layers of a gas are moving at different velocities, the momentum characteristic of one layer will be transported a distance λ_{visc} normal to the layer in time τ_{visc} , and give up

its momentum to the new layer. The viscous diffusion coefficient is given by $D_{visc} = \eta/\rho$, so that the coefficient of viscosity is given by

$$\eta = \rho D_{visc} = \frac{mN}{2V} u \lambda_{visc} . \quad (13)$$

Comparison with Eq. (7) shows that this leads to a viscosity coefficient for a gas which is independent of pressure; a surprising result produced by the fact that the decrease in density when pressure is reduced is balanced by the longer distance that momentum is carried.

4.5 Brownian movement

Brownian movement is a visible manifestation of the random motion which occurs for molecules in an ideal gas, and it has most of the same characteristics. The main difference is that one molecule in a gas like air has its direction changed drastically by a single collision with a similar molecule, but a particle large enough to see under a microscope will contain something like 10^{10} molecules, so has 10^5 times the momentum of a typical gas molecule if it has the same kinetic energy. It takes much longer to change its direction than a single molecule. Einstein worked out the theory of this using Stokes' formula $6\pi\eta av$ for the force on a spherical particle of radius a moving with speed v in a fluid with coefficient of viscosity η , and this gives the time for a particle of mass m to lose its initial momentum as $m/6\pi\eta a$, which increases like a^2 .

This time is much too small to resolve by eye, and what can be seen if one takes a set of photographs at equal time intervals is a set of completely random displacements from one picture to the next. The average displacement is zero in equilibrium (in the demonstration in class the smoke particles were not in equilibrium, but were falling in the air), because the displacement is equally likely to be in any direction, but the mean square displacement has a typical value that depends on the time interval. If one considers two successive time intervals, the mean square displacement can be described by Eq. (8), which shows that the root mean square displacement in time $2\tau'$ is $\sqrt{2}$ times the root mean square displacement in time τ' ; the distance gone is proportional to the square root of the time elapsed, and we have a random walk again.

In order to determine the Boltzmann constant, or Avogadro's number, observation of Brownian movement has to be combined with a knowledge of the time τ it takes a particle to lose knowledge of its initial momentum, since it is only for this short time interval that it has a typical speed proportional to $\sqrt{k_B T/m}$. For longer time intervals τ' the ratio of displacement to time is reduced by a factor $(\tau/\tau')^{1/2}$. However, the process known as sedimentation equilibrium, which you looked at in Problem 32 of Chapter 24, direct observation of the particles does give such a measurement, provided the weight W (allowing for buoyancy if the fluid is a liquid) and radius a of the particles is known.

They are continually falling under gravity at the steady speed given by the equality of the weight and the Stokes viscous force, which balance for $-v_z = W/6\pi\eta a$. However, they do not fall all the way, because the random velocity given by molecular collisions with the particles is superposed on this, and the result is a density of particles given by the Boltzmann distribution as

$$n(z) = n_0 \exp(-Wz/k_B T) ; \quad (14)$$

the density, in real space or in velocity space, is proportional to the exponential of the energy over $k_B T$. The normalization constant n_0 is determined by the condition that the integral over space should give the total number of particles.

A more direct method of observing the random motion of small particles is made possible by high precision optics. Since the particles that scatter the light are moving, the light they scatter will have a Doppler shift, and the Doppler broadening of the scattered light gives the root mean square value of the velocity. If one knows the mass of the particles this can lead to a measure of the Boltzmann constant, or, more usefully, since one knows the Boltzmann constant, it gives a way of measuring the mass of the particles. I am not sure how large are the largest particles that can have their speeds measured in this way; the more massive the particle the less is the Doppler broadening of the scattered light.

4.6 The Boltzmann distribution

James Clerk Maxwell and Ludwig Boltzmann argued that energy-conserving collisions between molecules should spread out their momenta in such a way that the probability of finding the one particular component of velocity, say v_x , close to a particular value is given by the frequency curve, or *probability density* equal to

$$f(v_x) \propto \exp(-mv_x^2/2k_B T) . \quad (15)$$

This is the Maxwell–Boltzmann distribution for velocity, and is a special case of the more general Boltzmann distribution, which gives a probability proportional to $\exp[-E(\mathbf{p}, \mathbf{r})/k_B T]$. I am not going to discuss why an exponential function of the energy arises. That question is more suitable for a serious course in statistical mechanics.

The equipartition law follows readily from the Maxwell–Boltzmann distribution. For example, the contribution of v_x to the average kinetic energy per molecule can be written as

$$\left\langle \frac{1}{2}mv_x^2 \right\rangle_{\text{ave}} = \int_{-\infty}^{\infty} \frac{1}{2}mv_x^2 f(v_x) dv_x , \quad (16)$$

and eq. (16) gives this as

$$\left\langle \frac{1}{2}mv_x^2 \right\rangle_{\text{ave}} = \frac{\int_{-\infty}^{\infty} \frac{1}{2}mv_x^2 \exp[-mv_x^2/2k_B T] dv_x}{\int_{-\infty}^{\infty} \exp[-mv_x^2/2k_B T] dv_x} = \frac{1}{2}k_B T .$$

A simple integration by parts leads from the second part to the third part of this equation, and one does not even need to know the value of the infinite integral in the denominator. Its value is derived in sec. 4.9.

Similarly, for a particle moving in a parabolic potential of the form

$$U(x, y, z) = U_0 + \frac{1}{2}K_1(x - x_0)^2 + \frac{1}{2}K_2(y - y_0)^2 + \frac{1}{2}K_3(z - z_0)^2, \quad (17)$$

each part of the potential energy gives an independent factor in the Boltzmann distribution, and each part gives an average value of the form

$$\left\langle \frac{1}{2}K_1x^2 \right\rangle_{\text{ave}} = \frac{\int_{-\infty}^{\infty} \frac{1}{2}K_1x^2 \exp[-K_1x^2/2k_B T] dx}{\int_{-\infty}^{\infty} \exp[-K_1x^2/2k_B T] dx} = \frac{1}{2}k_B T, \quad (18)$$

in addition to the constant term U_0 , where the same integration by parts is used.

The gravitational energy of a molecule or small particle in a constant gravitational field does not have the same average value. The Boltzmann distribution gives the probability density for the height z as

$$f(z) \propto \exp\left(\frac{-mgz}{k_B T}\right), \quad (19)$$

which is just the familiar form for the isothermal atmosphere. For this distribution, the average value of the potential energy is

$$\langle mgz \rangle_{\text{ave}} = \frac{\int_0^{\infty} mgz \exp[-mgz/k_B T] dz}{\int_0^{\infty} \exp[-mgz/k_B T] dz} = k_B T, \quad (20)$$

so that the average potential energy is $k_B T$ in a linear potential, rather than the $k_B T/2$ in a parabolic potential.

The Maxwell distribution of speeds, quoted in many books, follows from Eq. (16). The motions in the three different directions are independent of one another, so the probabilities can be multiplied together, giving the probability of the three-dimensional velocity lying in some range R of the three-dimensional velocity space as

$$C \int \int \int_{\mathcal{R}} \exp\left[-m(v_x^2 + v_y^2 + v_z^2)/2k_B T\right] dv_x dv_y dv_z,$$

where C is a constant. If we write the three components of \mathbf{v} in polar coordinates as $v \sin \theta \cos \phi, v \sin \theta \sin \phi, v \cos \theta$, this becomes

$$C \int v^2 dv \int \sin \theta d\theta \int d\phi \exp\left(-mv^2/2k_B T\right).$$

When the angles are integrated over this gives the probability of finding the speeds in a certain range \mathcal{R}' as

$$4\pi C \int_{\mathcal{R}'} v^2 dv \exp(-mv^2/2k_B T) ,$$

which is the Maxwell distribution of speeds.

Quantum mechanics makes the Boltzmann distribution particularly simple, because in quantum mechanics a system of volume V has a discrete system of energy levels j with energy E_j , so we can deal with probabilities rather than probability densities. The Boltzmann factor for each energy level is proportional to $\exp(-E_j/k_B T)$. The probability P_j is just

$$P_j = \frac{\exp(-E_j/k_B T)}{\sum_i \exp(-E_i/k_B T)} , \quad (21)$$

where the denominator ensures that the probabilities add up to unity

4.7 Fluctuations and the law of large numbers

Since the pressure and the heat capacity are determined by the average values of the kinetic energy, the actual value of the force on a wall or of the kinetic energy in a substance must fluctuate away from that average value. In some cases one can calculate these fluctuations easily, in other cases they are more difficult. Section 4.8 gives a calculation of the fluctuations in the number of particles in a given small volume v inside a big volume $V \gg v$ containing N molecules of an ideal gas. While the average number of particles in v is $\bar{n} = Nv/V$ the root mean square departure from that average (standard deviation) is $\sqrt{\bar{n}}$, so the fluctuations get smaller and smaller relative to \bar{n} as the average number increases. Similarly the actual impulse delivered to a small area of wall in a given time will differ from the average value by an amount proportional to the square root of the number n of molecules hitting the area in that time, so the ratio of the standard deviation to the average will be proportional to $n^{-1/2}$. This is discussed in detail in connection with eq. (30).

These are typical examples of the operation of the law of large numbers in statistical mechanics. The fluctuation of the energy of a molecule from its mean value is quite large, comparable with the mean value. If we were dealing with a collection of N molecules rather than just one the fluctuations would have been increased by a factor of \sqrt{N} , but the mean energy would increase by a factor of N , so the ratio of the fluctuations to the total energy decrease like $1/\sqrt{N}$. In electrical measurements, shot noise and thermal noise are important examples of this kind of fluctuation. Shot noise comes from the counting of individual electrons, while thermal noise comes from the temperature-dependent fluctuations in the current.

4.8 Probability and probability density

Probabilities measure the relative frequencies of discrete events. For a set of mutually exclusive events the probabilities P_n must add up to unity; the equation

$$\sum_{\text{all } n} P_n = 1 \quad (22)$$

represents the certainty that something happens.

Probability theorists are divided into apriorists, who assume that a tossed coin is equally likely to come up heads or tails unless there is information to the contrary, and Bayesians, who take probability to be a measured quantity. The rest of us indiscriminately pick ideas from both sides of the argument. The *a priori* argument says that any possible sequence of cards is equally likely to be dealt, or that any possible arrangement of N molecules in a volume V is equally likely, but we actually modify some of these assumptions as a result of experience, observation, and measurement.

Another important feature of probability theory is that if the probability of an event α is P_α and the probability of a **completely independent** event β is P_β , then the probability of both α and β happening is

$$\text{Prob}\{\alpha \cap \beta\} = P_\alpha P_\beta . \quad (23)$$

A simple counting of possibilities, valid whether or not the two events are uncorrelated, gives the probability of either α or β as

$$\text{Prob}\{\alpha \cup \beta\} = P_\alpha + P_\beta - \text{Prob}\{\alpha \cap \beta\} , \quad (24)$$

since the sum of probabilities counts twice over both of them happening. When we are dealing with independent molecules, as we are in the theory of the ideal gas, eq. (24) is the key result that enables us to write the probabilities for the whole collection of molecules as the product of probabilities for the individual molecules.

I gave a simple example of this in class, when we considered the probability P_n that n out of N molecules are in a subvolume v of the volume V . The probability that any particular molecule is in the volume v is $p = v/V$, and the probability that it is in the rest of the volume $V - v$ is $q = 1 - p$. Since in an ideal gas the position of one molecule does not influence the position of another molecule, the probability that n particular molecules will be in v and $N - n$ particular molecules will be in $V - v$ is equal to $p^n q^{N-n}$, according to eq. (24). To find the probability that any one of the $N!/n!(N - n)!$ choices of n objects out of a collection of N will be in the volume v we have to add the probabilities for each of the different choices together, to get the result

$$P_n = \frac{N!}{n!(N - n)!} p^n q^{N-n} . \quad (25)$$

If you remember the binomial theorem from your high school algebra you should recognize that the sum over all n from 0 to N is just the expansion

$$\sum_0^N \frac{N!}{n!(N-n)!} p^n q^{N-n} = (p+q)^N = 1, \quad (26)$$

in agreement with eq. (23). Further differentiation of this relation gives

$$\bar{n} = \sum_0^N n P_n = p \frac{\partial}{\partial p} (p+q)^N = pN, \quad (27)$$

and

$$\overline{n(n-1)} = \sum_0^N n(n-1) P_n = p^2 \frac{\partial^2}{\partial p^2} (p+q)^N = p^2 N(N-1). \quad (28)$$

Equation (28) gives the entirely obvious result that the average number of molecules in the volume v is Nv/V . It is also easy to show, by using the result $P_n/P_{n-1} = (N-n+1)p/nq$, that the most probable value of n is the integer immediately below $p(N+1)$, very close to \bar{n} . Equation (29) gives the result that the variance of the number of molecules in the volume is

$$\sigma^2 = \overline{(n-\bar{n})^2} = pqN = q\bar{n}, \quad (29)$$

so that the standard deviation σ is, for q close to unity, close to the square root of \bar{n} . This gives a demonstration of the law of large numbers for uncorrelated events as it applies to this *binomial distribution*.

In quantum mechanics the possible states in a volume V are discrete states j with energies E_j , and so the Boltzmann distribution gives a probability given in eq. (22). From this one can write the average energy at temperature T as

$$U = \sum_j E_j \frac{\exp(-E_j/k_B T)}{\sum_i \exp(-E_i/k_B T)}, \quad (30)$$

and the constant volume heat capacity as

$$C_v = \frac{dU}{dT} = k_B \sum_j \left(\frac{E_j}{k_B T} \right)^2 \frac{\exp(-E_j/k_B T)}{\sum_i \exp(-E_i/k_B T)} - k_B \left(\frac{U}{k_B T} \right)^2. \quad (31)$$

This shows that the heat capacity is k_B times the variance of $U/k_B T$ for a system.

In classical statistical mechanics the possible states of a system of molecules are described by the positions and momenta, or positions and velocities, of all the particles. We therefore have to describe the probabilities of continuous variables. There is no way we can assign a probability of a particle being at a particular point in space or having a particular velocity;

instead we work with a *probability density* whose integral over a particular region for the variable or variables in question gives the probability that the variables lie in that region. Just as the sum of probabilities over all possible values of a discrete set of possibilities adds up to unity, so also the integral of a probability density over all possible values of the variables is unity. For noninteracting molecules the probability density for the center of mass motion can be written as can be written as

$$p(\mathbf{r}, \mathbf{v}) \propto \exp \left[-\frac{U(\mathbf{r}) + \frac{1}{2}mv^2}{k_B T} \right]. \quad (32)$$

It is an important feature of the exponential form of the Boltzmann distribution that where different parts of the energy are separate from one another, as the energies of ideal noninteracting molecules are, the joint probability density for the whole system of particles can be written as the product of probability densities for the individual molecules, and the joint probability densities for position and velocity can be written as the product of the probability densities for position and velocity,

$$p_{rv}(\mathbf{r}, \mathbf{v}) = p_r(\mathbf{r})p_v(\mathbf{v}), \quad (33)$$

where the two factors are the probability densities given by

$$p_r(\mathbf{r}) = \frac{\exp[-U(\mathbf{r})/k_B T]}{\int \exp[-U(\mathbf{r})/k_B T] d^3r}, \quad p_v(\mathbf{v}) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp \left[\frac{-m\mathbf{v}^2}{2k_B T} \right], \quad (34)$$

One could use such a joint probability density to calculate the mean square of a component of velocity as

$$\begin{aligned} \overline{v_x^2} &= \int d^3r \int d^3v v_x^2 p_{rv}(\mathbf{r}, \mathbf{v}) = \int p_r(\mathbf{r}) d^3r \int v_x^2 p_v(\mathbf{v}) d^3v \\ &= \int v_x^2 p(v_x) dv_x \int p(v_y) dv_y \int p(v_z) dv_z = \int v_x^2 p(v_x) dv_x, \end{aligned} \quad (35)$$

so that it reduces to the usual expression in terms of the probability density $p(v_x)$ for a single component.

The particle density is just proportional to the probability density, but with a normalization factor such that its integral over all space is equal to the number of particles N rather than to unity, so we have

$$n(\mathbf{r}) = N p_r(\mathbf{r}). \quad (36)$$

4.9 Two useful mathematical results

Because gaussian probabilities of the form

$$p(x) \propto \exp \left(-\frac{x^2}{2\sigma^2} \right) \quad (37)$$

keep occurring, it is useful to know how to normalize this probability. This can be worked out by squaring the integral over all x ,

$$\left[\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \right]^2 = \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy ,$$

and then converting the double integral over x, y to an integral over polar coordinates

$$\left[\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \right]^2 = \int_0^{\infty} r dr \int_0^{2\pi} d\theta \exp\left(-\frac{r^2}{2\sigma^2}\right) .$$

The integral over θ just gives 2π , and the integral over r gives σ^2 , so we have

$$\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx = \sqrt{2\pi}\sigma . \quad (38)$$

The other useful result for handling distributions like the binomial distribution is Stirling's approximation for the factorial of a large number. Since we can write the logarithm of $n!$ as the sum over $\ln m$ for $m \leq n$, so a first guess at an approximation is

$$\ln n! \approx \int_1^n \ln x dx = n(\ln n - 1) .$$

A more careful evaluation gives the Stirling formula

$$\ln n! \approx n(\ln n - 1) + \frac{1}{2} \ln(2\pi n) , \quad n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n} . \quad (39)$$

This result obtained by an eighteenth century mathematician gives very accurate results even for quite small values of n