

2. Physics and probability

Here we take up our study of many interacting molecules. We will mainly be concerned with macroscopic properties, i.e. properties that involve averages over the states of many of the molecules, for example pressure, temperature, or total energy. Further, we will mostly (but not exclusively) study equilibrium states, where the macroscopic properties have settled down to a time-independent average value with small fluctuations. We will try to see how such macroscopic properties can be derived from the fundamental physics of molecules: we will use classical mechanics where appropriate, or, when needed, quantum theory.

Of course, it is not obvious how macroscopic behavior arises from mechanics. There are many features of the everyday world which seem non-mechanical. Notably, our experience abounds with irreversible processes: people get older, not younger, ice melts but does not spontaneously refreeze — but mechanics is reversible. And some things that seem consistent with mechanics, like making a perpetual motion machine by extracting heat from the air (and doing nothing else), somehow never get done.

There is a theory of macroscopic behavior based on a few reasonable ideas (like the impossibility of making a perpetual motion machine) which is very successful: it is called *classical thermodynamics*: its centerpiece is the non-mechanical quantity, entropy. The formulation of thermodynamics was more or less complete in the late nineteenth century. At first, it seemed to be unrelated to mechanics, and even inconsistent with it. In the next chapter we will look at classical thermodynamics.

The question soon arose whether it was possible to derive thermodynamics from mechanics. This problem was called, in nineteenth century language, the *rational foundation of thermodynamics*. It was a huge problem, and many Very Serious People (including Poincaré and Ernst Mach) doubted it could be done at all. However, J. Willard Gibbs, James Clerk Maxwell, and particularly, Ludwig Boltzmann (Figure 2.1) solved the problem at the end of the nineteenth century. This chapter and the next will be concerned with these developments.

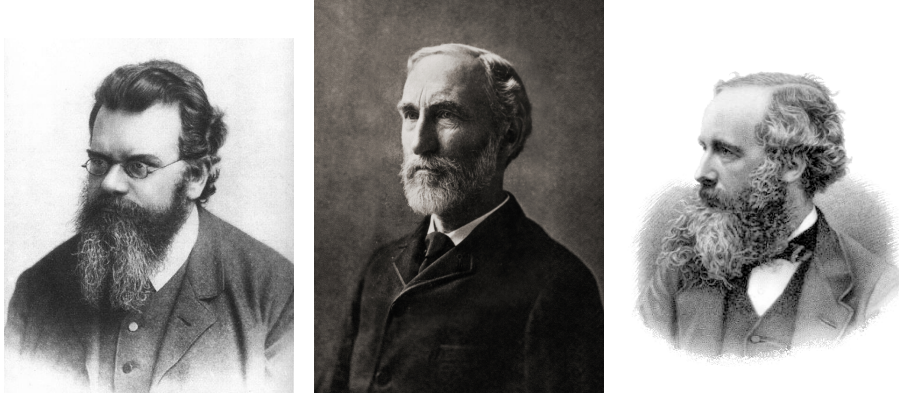


Figure 2.1.: The founders of statistical physics. From left, Ludwig Boltzmann (1844-1906), J. Willard Gibbs (1839-1903), and James Clerk Maxwell (1831-1879). Images in the public domain from Wikimedia Commons.

2.1. Pressure and temperature in a gas

Now consider a gas, and see how far we get by applying mechanics. We can ask why a gas exerts pressure, force per unit area, on the walls of its container. This was answered in the eighteenth century by Daniel Bernoulli. He pointed out that repeated impacts of large numbers of molecules on the walls of a container would look like a steady force.

2.1.1. Kinetic definition of pressure

Here is the Bernoulli argument in modern terms. The force is given by Newton's laws in the form $F = \Delta p / \Delta t$, where Δp is the total momentum change of the particles that hit the wall in Δt . Now consider the molecules in the gas. They will have some distribution of momentum and position at a given time. We call the distribution function, $f(\mathbf{p}, \mathbf{r}, t)$. Its definition is this: the number of molecules with position within a volume element d^3r centered at \mathbf{r} , and similarly for \mathbf{p} is:

$$dN = f(\mathbf{p}, \mathbf{r}) d^3p d^3r. \quad (2.1)$$

Here, and subsequently, we will write d^3r for $dx dy dz$, and similarly for d^3p . Clearly,

$$\int f(\mathbf{p}, \mathbf{r}) d^3p d^3r = N.$$

For the case of uniform density, it is convenient to absorb a factor of $1/V$ into f , and write $\int f(\mathbf{p}) d^3p = N$.

If we want to average some molecular quantity, $g(\mathbf{p}, \mathbf{r})$ over all the molecules, then we should write:

$$\bar{g} = \frac{1}{N} \int d^3p d^3r f(\mathbf{p}, \mathbf{r}) g(\mathbf{p}, \mathbf{r}); \quad \text{or} \quad \frac{1}{N} \int d^3p f(\mathbf{p}) g(\mathbf{p}). \quad (2.2)$$

The function, f , was introduced by Maxwell and Boltzmann.

This function is perfectly well defined in or out of equilibrium. For a nearly-ideal gas it satisfies a kinetic equation, the Boltzmann equation. The solutions to this equation are the content of the kinetic theory of gases; they can give calculations of many dynamic properties such as relaxation to equilibrium and fluxes in driven systems. In this book we look at equilibrium only, and drop the time dependence. We will compute the pressure from f .

Let us assume that the density of the gas, $n = N/V$ is the same everywhere (neglect gravity, for example). We need the number of molecules that will strike area dA of the wall in Δt . Take dA to lie in the y, z plane and consider some value of \mathbf{p} . The molecules need to be closer to the wall than $(p_x/m)\Delta t$, so that they are in volume $\delta V = dA(p_x/m)\Delta t$; see Figure 2.2. This number is

$$dN = \frac{\delta V}{V} f(\mathbf{p}) d^3p.$$

Each collision reverses p_x so the total momentum transfer is:

$$\Delta p_x = \frac{dA \Delta t}{V} \int \frac{2p_x^2}{m} f(\mathbf{p}) d^3p.$$

The integral is taken for $p_x < 0$. Now the pressure is the force per unit area so we have:

$$\begin{aligned} p &= \frac{\Delta p_x}{dA \Delta t} \\ &= \frac{1}{V} \int \frac{p_x^2}{m} f d^3p \\ &= \frac{2}{V} \int \frac{p_x^2}{2m} f d^3p \\ &= \frac{2N}{3V} \overline{\frac{p^2}{2m}}. \end{aligned} \quad (2.3)$$

Here $\overline{p^2} = \overline{p_x^2} + \overline{p_y^2} + \overline{p_z^2}$, and we have used Eq. (2.2). In the second line a factor of 2 disappeared so that we can integrate over all \mathbf{p} , and the

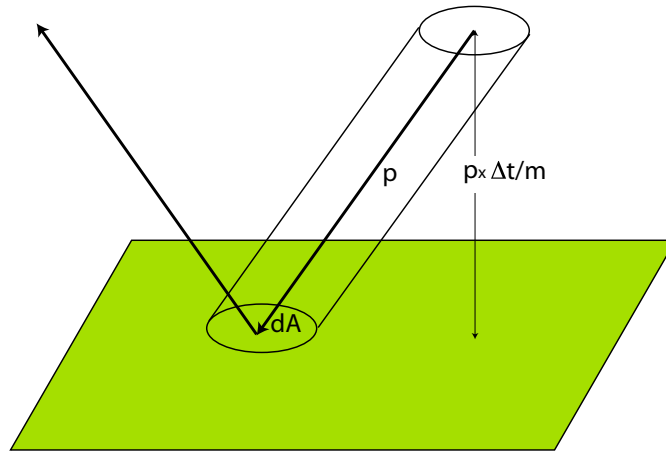


Figure 2.2.: Geometry for computing the momentum transfer to the wall. Any molecule within $p_x \Delta t / m$ of the wall will hit it in time Δt , and the total number to hit with a given angle is the number within a cylinder of volume $dA p_x \Delta t / m$.

factor of $1/3$ in the last line means that we assume that the momentum distribution is isotropic: $\overline{p_x^2} = \overline{p_y^2} = \overline{p_z^2}$. Thus:

$$pV = N \frac{2}{3} \overline{t_E} \quad (2.4)$$

That is, pV is $2N/3$ times the kinetic energy per molecule in the gas, t_E (or $2/3$ of the total kinetic energy). In d dimensions this becomes $pV = (2N/dV) \overline{t_E}$.

We should compare this to the empirical equation of state of an ideal gas which includes Boyle's law, Charles' law, and the law of Guy-Lussac:

$$pV = \nu R(T' + 273.15) = N k_B T. \quad (2.5)$$

Here ν is the number of moles in the gas, and R is the gas constant, 8.314 J/K . The number of molecules is $N = \nu N_o$ where N_o is Avogadro's number, so that $k_B = R/N_o = 1.38 \times 10^{-23} \text{ J/K} = 1.38 \times 10^{-16} \text{ erg/K}$. The temperature is given in two scales, T' , the Celsius temperature, $^{\circ}\text{C}$, and T the Kelvin temperature, K .

2.1.2. Temperature and temperature scales

The equation of state is a bit strange at this point because we have not really given a definition of temperature. As we have pointed out above,

temperature specifies what is meant by hot and cold: heat (energy) flows from hot bodies to cold ones when they are in contact. A higher temperature means hotter. A thermometer measures temperature; if we put a thermometer in contact with a body and if no heat flows they are at the same T . However, what do we use for a thermometer?

It took many centuries to give a more complete definition. Issac Newton measured temperature by thermal expansion, and realized that he needed fixed points to define the scale. He assigned zero as the freezing point of water, and 33 as the boiling point. (The reason why Newton chose the number 33 is discussed by Dan Brown (2009).) Anders Celsius chose the same fixed points, and took the number of degrees between them to be 100. A modified version of his scale is used today almost everywhere but the United States. However, this is not enough, since we have not said what it means for a thermometer to be linear, i.e., how do we measure temperatures between the fixed points.

William Thomson, Lord Kelvin, gave a much more fundamental definition when he realized that by adding a constant to the temperature on the right side of Eq. (2.5) he could make the equation linear in temperature. The advantage is that all ideal gases have the same equation of state, so we have a universal thermometer, namely $T = PV/k_B N$. The constant, k_B , determines the size of the degree.

This also defines the absolute zero of T by $T = T' - 273.15$. What this means will become clearer later, but we can get some insight by comparing Eq. (2.4) and Eq. (2.5):

$$k_B T = (2/3)\overline{t_E}, \quad (2.6)$$

(or $k_B T = (2/d)\overline{t_E}$.) Taken literally, this means that the thermal motion of a gas ceases at $T = 0$. Of course, all gases liquefy long before that, and it is an extrapolation.

For the moment we take Eq. (2.6) as our first definition of temperature. We will do so even for interacting systems. This is something that will be justified later.

2.1.3. Virial theorem

We can look at the previous section another way by considering a quantity defined by Rudolf Clausius, the *virial*; see Goldstein et al. (2002):

$$G = \sum_j \mathbf{r}_j \cdot \mathbf{p}_j.$$

Consider the time derivative of G :

$$\begin{aligned}
 \frac{dG}{dt} &= \sum_j \left[\frac{d\mathbf{r}_j}{dt} \cdot \mathbf{p}_j + \mathbf{r}_j \cdot \frac{d\mathbf{p}_j}{dt} \right] \\
 &= \sum_j \left[m \left(\frac{d\mathbf{r}_j}{dt} \right)^2 + \mathbf{r}_j \cdot \mathbf{F}_j \right] \\
 &= 2T_E + \sum_j \mathbf{r}_j \cdot \mathbf{F}_j,
 \end{aligned} \tag{2.7}$$

where T_E is the total kinetic energy and \mathbf{F}_j is the force on atom j . The time average of any quantity is defined to be the integral of the quantity over $[0, \tau]$ divided by τ . Applying this:

$$\langle dG/dt \rangle_\tau \equiv \frac{1}{\tau} \int_0^\tau dG/dt = G(\tau) - G(0). \tag{2.8}$$

In equilibrium this quantity is zero – that is what we mean by equilibrium. Thus:

$$2 \langle T_E \rangle_\tau = - \left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j \right\rangle_\tau. \tag{2.9}$$

There are two sorts of forces on an atom: there are the the internal forces, \mathbf{F}_j^{int} and the forces on the wall, $\mathbf{F}^{wall} = -p\mathbf{n}dA$, where \mathbf{n} is the normal to the wall. For the latter we can write:

$$\left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{wall} \right\rangle_\tau = -p \int \mathbf{r} \cdot \mathbf{n} dA = -p \int \nabla \cdot \mathbf{r} d^3r = -3pV. \tag{2.10}$$

We have used Gauss' theorem. Thus, by comparing Eq. (2.6) and replacing the average over the distribution by the time average:

$$\begin{aligned}
 \langle T_E \rangle_\tau &= N \overline{t_E} = \frac{3}{2}pV - \frac{1}{2} \left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{int} \right\rangle_\tau \\
 pV &= Nk_B T + \frac{1}{3} \left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{int} \right\rangle_\tau
 \end{aligned} \tag{2.11}$$

In 2d the $1/3$ becomes $1/2$. If the internal forces can be neglected, we have the ideal gas equation, Eq. (2.5). In practice, the ideal gas equation applies for dilute gases. For air at atmospheric pressure, it works quite well.

Eq. (2.11) is very useful in computer simulations. It is possible to compute the pressure by counting collisions on the walls, but Eq.

(2.11) is superior numerically because it uses all the molecules. And, it allows the use of periodic boundary conditions, no walls, and still get the pressure.

We can express the equation of state in terms of the pair distribution function, $g(R)$ of Eq. (1.31) in the following way. First note that \mathbf{F}_j^{int} is the total force on atom j due to all the others:

$$\mathbf{F}_j^{int} = \sum_{k \neq j} \mathbf{F}(\mathbf{r}_j - \mathbf{r}_k).$$

Now note that:

$$\begin{aligned} \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{int} &= \sum_{j \neq k} \mathbf{r}_j \cdot \mathbf{F}(\mathbf{r}_j - \mathbf{r}_k) \\ &= \frac{1}{2} \sum_{j \neq k} (\mathbf{r}_j - \mathbf{r}_k) \cdot \mathbf{F}(\mathbf{r}_j - \mathbf{r}_k), \end{aligned} \quad (2.12)$$

by Newton's third law. Recall the definition;

$$c(\mathbf{r}, \mathbf{s}) = \left\langle \sum_{j \neq k} \delta(\mathbf{r} - \mathbf{r}_j) \delta(\mathbf{r} - \mathbf{r}_k) \right\rangle,$$

Eq. (1.30). Combining this with Eq. (2.12) gives:

$$\begin{aligned} \left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{int} \right\rangle &= \frac{1}{2} \int d\mathbf{r} d\mathbf{s} c(\mathbf{r}, \mathbf{s}) (\mathbf{r} - \mathbf{s}) \cdot \mathbf{F}(\mathbf{r} - \mathbf{s}) \\ &= -\frac{1}{2} \int d\mathbf{r} d\mathbf{s} c(\mathbf{r}, \mathbf{s}) (\mathbf{r} - \mathbf{s}) \cdot \nabla_{\mathbf{r}} \phi(|\mathbf{r} - \mathbf{s}|) \\ &= - \int d\mathbf{r} d\mathbf{s} c(\mathbf{r}, \mathbf{s}) R \left(\frac{\partial \phi}{\partial R} \right), \end{aligned} \quad (2.13)$$

where $R = |\mathbf{r} - \mathbf{s}|$. We have used the identity $\nabla_{\mathbf{r}} \phi(r) = (\mathbf{r}/r) d\phi/dr$. Now change variables in the integral, use translational invariance, and $c = n^2 g(R)$:

$$\left\langle \sum_j \mathbf{r}_j \cdot \mathbf{F}_j^{int} \right\rangle = -\frac{Vn^2}{2} \int d\mathbf{R} R \left(\frac{\partial \phi}{\partial R} \right) g(R). \quad (2.14)$$

This gives another form for the equation of state:

$$\frac{p}{k_B T} = n - \frac{n^2}{6k_B T} \int d\mathbf{R} R \left(\frac{\partial \phi}{\partial R} \right) g(R). \quad (2.15)$$

2.1.4. Dense gases; the van der Waals' equation

We can go a bit further by following the work of Johannes van der Waals and guessing the difference in the equation of state that interactions might cause.

First we note, from the previous chapter, that molecules cannot really explore the whole volume of the gas. When we try to compress a gas so that the hard cores touch, the pressure must rise enormously. So we can, tentatively, replace the volume by the “free volume”:

$$V_f = V - Nb, \quad (2.16)$$

where b is of the order of the molecular volume, i.e., of order σ^3 .

The attractive forces can be expected to change the equation of state by changing the pressure. If a molecule approaches the wall of the container, it is pulled back by the attractive forces. Thus the “kinetic pressure” that Bernoulli considered is not the observed pressure, but bigger. The difference must be proportional to the total attractive force between pairs of molecules within the range of the interaction:

$$p = p_K - an^2, \quad (2.17)$$

where n is the number density. Combining Eq. (2.16) with Eq. (2.17) gives:

$$\begin{aligned} p_K V_f &= \left(p + a \frac{N^2}{V^2}\right)(V - Nb) = Nk_B T ; \\ p &= \frac{nk_B T}{1 - nb} - an^2. \end{aligned} \quad (2.18)$$

The remarkable thing about this simple derivation and relatively simple equation is that it correctly represents a great deal of physics reasonably well. For example let us use the equation to look at *isotherms*, curves of constant temperature on a n, p plot; see Figure 2.3. As we can see, as the temperature decreases the isotherms depart from the ideal gas equation (which is linear in this representation). The prediction is that at a certain temperature, the critical temperature, T_c , there is a place where the curve becomes flat; i.e. where $\partial p / \partial n = 0$. Note that this quantity is proportional to the inverse of the compressibility, $-(1/V)\partial V / \partial p$. What is happening is that the attractive interactions are leading to a liquid-gas transition, and at this point the compressibility is infinite because the gas “wants” to condense. This is an effect which is also observed in nature for any gas: it is called the critical point.

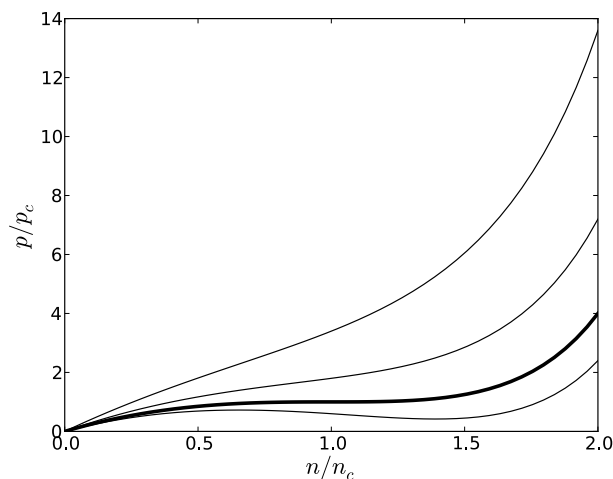


Figure 2.3.: Isotherms of the van der Waals equation. All quantities are scaled to their critical values, i.e., p/p_c as a function of n/n_c (see problems). The isotherms are, from top to bottom: $T/T_c = 1.2, 1.6, 1.0$ (bold line, the critical isotherm) and 0.9 . Note that for the bottom curve an increase in density can cause a decrease in pressure: see text.

Below the critical point the isotherms become non-monotonic. This is unphysical. We cannot have $\partial p/\partial n < 0$ because then the gas would collapse spontaneously. We will see in Chapter 8 how to interpret this behavior, and how to produce the familiar gas-liquid coexistence from the equation.

Van der Waals' constants are available in published tables. For example, for N_2 , $a = 1.370 \text{ bar L}^2/\text{mol}^2$, $b = 0.0387 \text{ L/mol}$.

2.2. Classical systems

What we have done so far is encouraging: simple mechanical methods lead to important results for macroscopic systems. However, we have not gotten very far: for example, we have no way to compute $f(\mathbf{p})$ from the Hamiltonian, or, say, the van der Waals constants a, b . And we have used some ideas rather loosely: we talk about “randomness”, for example, without understanding how it might arise. We will now take a closer look at the fundamental issues in macroscopic systems.

2.2.1. Phase space and phase trajectories

We first define the problem, for the moment using classical mechanics. In a macroscopic system we want to predict the result of experiments to measure physical quantities. Such measurements involve taking data by means of some instrument. Let us suppose that we take data at some sampling rate, so we generate measurements at a series of time points, $t_j, j = 1, \dots, m$.

A classical system obeys Hamilton's equations, so that all the information is in the set $q_i(t), p_i(t)$. Suppose there are N molecules. We then have $6N$ numbers for each time. We imagine that these numbers are the coordinates of a point in a $6N$ dimensional space, *phase space*. We denote a point in phase space by $\gamma(t_j) = (q_i(t_j), p_i(t_j))$. The $\gamma(t)$ form the *phase trajectory*; it is generated by the equations of motion. The measurements correspond to a discrete set of points on the trajectory. A measurement means that we can find whether a phase point is within some volume of phase space, $\delta\Gamma$. If we take $\delta\Gamma = h^{3N}$, where h is Planck's constant, we are measuring each p, q pair as well as is consistent with Heisenberg's principle. (We will show later that the factor should be h , not some other quantity of the same order, see Section 6.1.4.) Of course, in a normal measurement we cannot measure this well. We call each such region of Γ -space a microstate. We will use the same notation later for quantum states.

What do we know about $\gamma(t)$? One thing is clear for a closed system: $\gamma(t)$ lives on a surface of constant energy, $\mathcal{H}(q_i, p_i) = E$. This is called the energy shell: it has dimension $6N - 1$. The area of the energy shell is found by integrating over phase space and restricting the variables to the surface $\mathcal{H} = E$ with a Dirac δ -function:

$$\Omega = \int d\Gamma \delta(\mathcal{H} - E). \quad (2.19)$$

This area depends on energy and real-space constraints, e.g. the size of the box containing the system.

The number of microstates on the energy shell is:

$$W = \Omega\Delta/h^{3N}, \quad (2.20)$$

where we have multiplied by an energy increment, Δ , to make the units come out right. (You could think of Δ as the accuracy with which we know E .) Now this is not quite right: if we have N_k identical particles of type k we have overcounted the number of states: states that differ by interchange of identical particles are the same. We really should

write:

$$W = \frac{\Omega \Delta}{\prod_k N_k! h^{3N}}. \quad (2.21)$$

The significance of the factorials in the denominator, sometimes called the rule of correct Boltzmann counting, will be discussed later.

2.2.2. Time averages and phase-space averages

Boltzmann and Maxwell made the bold and remarkable proposition that when a system is in equilibrium $\gamma(t)$ visits *all* of the energy shell, and does so uniformly, so that the trajectory spends equal time in every microstate. This means that γ will pass arbitrarily close to every point on the energy shell. Such a trajectory is called ergodic.

What we measure in an experiment is some quantity which depends on the mechanical coordinates, $Q(q_i, p_i)$. Macroscopic experiments always involve averages over time, $\langle Q \rangle_\tau$, (c.f. Eq. (2.8)). Recall, for example, the time average in the kinetic pressure. This is what we want to calculate. But, if the phase point uniformly explores the energy shell, it follows that, as $\tau \rightarrow \infty$:

$$\begin{aligned} \frac{1}{\tau} \int_0^\tau Q(\gamma(t)) dt &= \frac{\int Q(q_i, p_i) \delta(\mathcal{H} - E) d\Gamma}{\Omega} \\ \langle Q \rangle_\tau &= \langle Q \rangle, \end{aligned} \quad (2.22)$$

since both integrals explore the same points in a different order. This equation defines the phase (or ensemble) average: $\langle Q \rangle$. Thus the physical quantity on the left is the same as a phase space average.

Gibbs gave a colorful interpretation of this equation: he imagined that we have a large number, or ensemble¹, of identical equilibrated systems with different initial conditions so that, at some time t , they densely and uniformly cover the energy shell. Then we average over this set of systems *at a fixed instant of time*. He called this the ensemble average, and the proposition is that the time average is equal to the ensemble average. Gibbs gave the ensemble with fixed energy (the one we are studying) the odd name *microcanonical*.

We can say this another way: define the *probability* for a system point to be in a region of phase space to be $P \equiv \rho(q_i, p_i) d\Gamma$, where ρ

¹The word ensemble means set or group in this context.

is called the ensemble density. We can write, for our particular case:

$$\begin{aligned}\langle Q \rangle &= \int Q \rho \, d\Gamma; \\ \rho_{mc} &= \frac{\delta(\mathcal{H} - E)}{\int \delta(\mathcal{H} - E) \, d\Gamma}.\end{aligned}\tag{2.23}$$

The subscript *mc* refers to the fact that this particular ρ is for the microcanonical ensemble. We will meet other ensembles later. The first line of the equation is general.

When we call ρ a probability density we mean nothing more than this: if we make a large number, \mathcal{N} , of observations of Q and record the number giving a certain value, q then $P = \mathcal{N}(q)/\mathcal{N}$. This is called the *frequentist* interpretation of probability. It is also the same as Eq. (2.22). It does not imply any “real” randomness in the system such as the randomness observed in a radioactive decay. It just looks similar.

There are several remarks to make here. It is not obvious, but it will become so, that computing the ensemble average is much easier than computing the time average. In practice, the results from using Eq. (2.22) are remarkably good: this is how everyone does statistical mechanics, and when you can do it, the results agree with experiments *perfectly* in almost all cases. Note that we do not need to know how equilibrium is approached, but only the states available: to find $\langle Q \rangle$ we do not solve equations of motion.

Also, as we will see below, this approach gives us a neat definition of entropy:

$$S = k_B \ln W\tag{2.24}$$

This famous discovery of Boltzmann completely solves the problem of the rational foundation of thermodynamics. We will discuss this at length in the next chapter.

Thus it is clear that Boltzmann’s assumption of ergodicity is useful. There is a much more difficult question: is it correct? We could imagine that the time average and the ensemble average are the same for some other reason, for example, or that the agreement with experiment is accidental. There has been more than a century of very interesting work on this question, which will be treated in the next sections.

2.2.3. Ergodicity and mixing

This subject has a large literature, much of it highly mathematical. For introductions, see Lebowitz & Penrose (1973), Penrose (1979),

Uhlenbeck, Ford & Montroll (1974), and Chapter 26 of Ma (1985). We will sketch a few major developments.

Liouville theorem

The first thing to notice is that the way trajectories cover phase space is special to Hamiltonian systems. In other kinds of systems things are quite different; for example, in the presence of friction trajectories can converge into an attractor or a fixed point. However, we deal with closed systems here, so nothing like this happens. A demonstration of this is given by Liouville's theorem, based on the work of J. Liouville, but first published by Gibbs. It gives the equation of motion for the ensemble density, ρ .

To get an equation of motion, we start with a bunch of initial conditions (perhaps corresponding to initial experimental uncertainty); we represent them by $\rho(q_i, p_i, t = 0)$. The fraction of systems in $d\Gamma$ is $\rho d\Gamma$. Now we follow the systems in time to find $\rho(t)$.

Since systems cannot be created or destroyed ρ acts like the density of a "fluid" in phase space, namely it obeys a continuity equation of the form:

$$\partial\rho/\partial t + \nabla \cdot (\rho\mathbf{v}) = 0.$$

We need to interpret \mathbf{v} in the $6N$ dimensional phase space as (\dot{q}_i, \dot{p}_i) . Similarly the divergence is $(\partial/\partial q_i, \partial/\partial p_i)$. That is:

$$\begin{aligned} \frac{\partial\rho}{\partial t} &= -\sum_i \left[\frac{\partial(\rho\dot{q}_i)}{\partial q_i} + \frac{\partial(\rho\dot{p}_i)}{\partial p_i} \right] \\ &= -\sum_i \left[\frac{\partial\rho}{\partial q_i} \dot{q}_i + \frac{\partial\rho}{\partial p_i} \dot{p}_i \right] - \rho \sum_i \left[\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i} \right] \\ &= -\sum_i \left[\frac{\partial\rho}{\partial q_i} \dot{q}_i + \frac{\partial\rho}{\partial p_i} \dot{p}_i \right] \\ &\quad - \rho \sum_i \left[\frac{\partial^2\mathcal{H}}{\partial q_i \partial p_i} - \frac{\partial^2\mathcal{H}}{\partial p_i \partial q_i} \right]. \end{aligned} \tag{2.25}$$

The last term is zero, and thus:

$$\frac{\partial\rho}{\partial t} + \sum_i \left[\frac{\partial\rho}{\partial q_i} \dot{q}_i + \frac{\partial\rho}{\partial p_i} \dot{p}_i \right] = \frac{d\rho}{dt} = 0. \tag{2.26}$$

Here $d\rho/dt$ is the Lagrangian derivative, the change of ρ as it is swept along in the Hamiltonian flow. Thus $\rho(\gamma(0)) = \rho(\gamma(t))$, i.e., the density

around a phase point is constant as it moves around on the energy surface. The fluid that we deal with is incompressible. In the language of chaos theory, there are no attractors for Hamiltonian systems².

Another consequence is that if ρ is constant over a volume, $\Delta\Gamma$, and zero elsewhere, then the volume is conserved as it moves around under the dynamics, though it will, in general, change shape. This follows from the fact that $\rho = \mathcal{N}/\Delta\Gamma$, where \mathcal{N} is the number of systems represented.

Ergodic theorems

Boltzmann reasoned that if you consider a whole trajectory, the density near one point is the same as the density near its image. Thus if the point goes nearly everywhere on the energy shell, the density is constant nearly everywhere. Thus we should average over the whole shell with $\rho = \text{constant}$. That is exactly Eq. (2.22). However, we do not know that a *single* trajectory goes everywhere.

Formal theorems by George D. Birkhoff and John von Neumann made this idea rigorous, see Uhlenbeck et al. (1974). Birkhoff proved that if the energy shell could not be divided into pieces such that trajectories never cross the boundary (this is called “metrically indecomposable”) then Eq. (2.22) holds. The trick, then, is to show what class of Hamiltonians have this property. Physically reasonable examples were hard to produce for a long time. However, there has been quite a lot of progress since Birkhoff and Neumann.

Mixing

It is necessary to point out that ergodicity is really not enough. A real macroscopic system, or even the few atom system we have simulated, has a stronger property called *mixing*. To see this consider a one-dimensional harmonic oscillator. Its phase space is two-dimensional, and since the conserved Hamiltonian is $p^2 + q^2$ (in the proper units) the energy shell is a circle. Consider a group of initial conditions: they will all travel around the circle so the system is certainly ergodic. However, it is really different from an equilibrium macroscopic system since it is periodic with period 2π ; it never “settles” down to a steady state.

We need to make a further assumption: we require that the states in any small group of initial conditions spreads out in such a way that

²For the chaos groupies: the only generic fixed points are saddle points and foci.

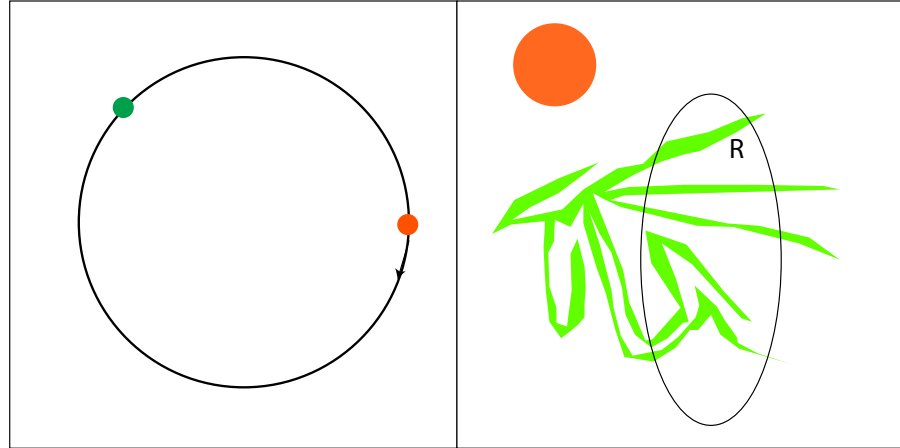


Figure 2.4.: Mixing. Left: the energy shell of a harmonic oscillator and a group of initial conditions. They flow around the circle unchanged in shape. In order to have mixing we need a positive Lyapunov exponent (see Eq. (2.28).) In this case (and even for a non-linear oscillator) the Lyapunov exponent is zero. Right: In a mixing system the initial conditions in the small disk spread out into Gibbs tendrils while preserving volume. In the limit of large times the fraction of the image that overlaps any region R is simply the fraction of R in the whole shell. This means that the tendrils cover the shell uniformly.

they cover the shell *uniformly*. That is, take ρ to be non-zero for some small region initially, and consider the $\rho(t)$, i.e. the result of letting all the points in the region develop for time t . Suppose t is a long time. A system is mixing if:

$$\int \rho(t) R d\Gamma = \frac{\int R d\Gamma}{\Omega} \quad (2.27)$$

for any phase space function R . For example, if R is constant over some region and zero elsewhere (a so-called indicator function), this says that the fraction of points that end up in that region is just the fraction of the shell that R occupies. See Figure 2.4.

Why is this important for our problem? It means, for example, that if we have some initial condition that stays in a periodic orbit, as in the harmonic oscillator, neighboring points will wander away, because any small initial region will spread out to cover the shell. Almost all initial conditions will settle down to equilibrium.

Gibbs (1902) already had this idea. He compared the spread of initial conditions to stirring ink into water. The volume of ink is

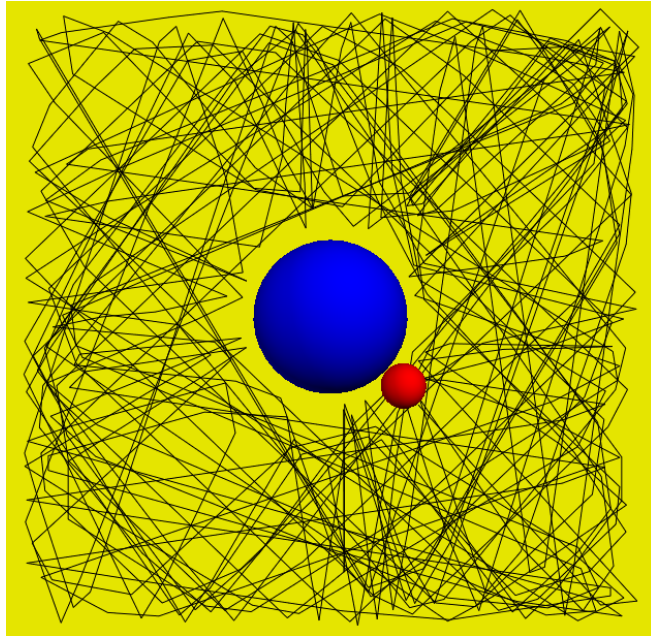


Figure 2.5.: The track of the Sinai billiard in the x, y plane. On the cover is a similar track for a rectangular billiard table.

constant, but after stirring it is uniformly dispersed. As in Figure 2.4, there will be tendrils of the initial blob that get finer and finer in time (to conserve volume as required by the Liouville theorem). Any measurement with finite precision will see the tendrils dispersed over the volume.

Instability of orbits: playing billiards

It is clear that some systems of many atoms are not mixing — the ideal gas with no interactions at all is an example. From the work of Yakov Sinai we can give an example of a simple Hamiltonian system which is more-or-less like interesting physical systems, and which is mixing and ergodic. This is called the Sinai billiard; see Penrose (1979).

For the Sinai billiard a hard disk bounces against the walls of a square box, and also against a circular obstacle in the center. We show, in Figure 2.5 the trajectory of the disk. In this case the phase space is four dimensional, x, y, p_x, p_y but the energy shell is three-dimensional. It is also very simple because for elastic collisions $|\mathbf{p}|$ is conserved, the momentum can only change its direction, θ . Thus the

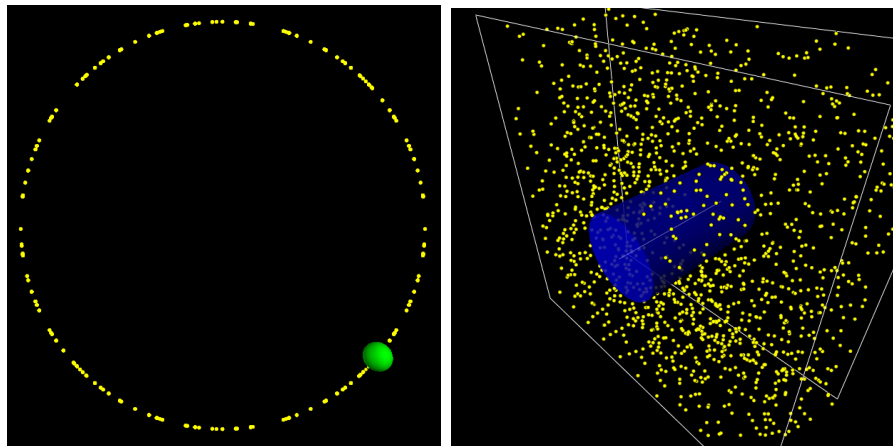


Figure 2.6.: Left: the subspace of phase space in the p_x, p_y plane. The points are samples taken from the track. The momentum changes in direction only. Right: the energy shell. The axes in the front are x, y and the axis parallel to the excluded blue cylinder is θ , where $0 \leq \theta < 2\pi$ is the angle from the x -axis of \mathbf{p} .

energy shell is a circle in \mathbf{p} -space. We can plot discrete sampling times in the three-dimensional space x, y, θ . The points fill up the space, as we see in the Figure 2.6.

We can trace the source of mixing here: the ball bounces against a convex surface, so two trajectories that start from a point with directions differing by $\delta\theta_o$ will have the angular difference magnified. The result, after a number of collisions will be that

$$\delta\theta_t \approx \delta\theta_o \exp(\alpha t). \quad (2.28)$$

The number, α , is called a Lyapunov exponent, and is positive for diverging trajectories.

Lorentz gas and hard sphere gas

Research in this area has extended to deal with more realistic systems, notably the Lorentz gas (hard disks bouncing on fixed obstacles) and interacting hard spheres. The Lorentz model is used for electrons in solids: the obstacles are impurities in a crystal. It is close to the standard way to treat electrical conduction; see Sander (2009).

As of this writing, there are some solid results in this area, e.g. Simányi (2009). Oddly enough, it is easier to prove mixing for systems

of a few hard spheres than for many, and for small numbers like 2 and 3 in dimensions 2 or greater the theorem is proved.

Numerical “proofs”

The mathematicians are hard at work in this area, but a physicist may wonder if their results are really important. For a physicist (in any case, for this physicist) it is sufficient to know that numerical investigations of realistic models of microscopic behavior give the results that we expect from ergodic/mixing theory. This is certainly the case, and the rigorous mathematical proofs of Boltzmann’s proposition in semi-realistic cases go to support this position. In fact, most textbooks in statistical physics simply skip the discussion altogether, and start by assuming that Eq. (2.22) is correct.

2.2.4. Objections to the theory

When Boltzmann and Maxwell advanced the idea of ergodicity it met with fierce opposition. The opponents claimed that it was not consistent with experiment. These objections are all incorrect. It is enlightening to see why.

We will not discuss further the small but vocal class of scientists such as Mach who had not yet (at the end of the nineteenth century!) accepted the idea that matter is made of atoms and molecules. This is of only historical interest. There were more substantive and interesting objections raised. These are known in the older literature as the *Wiederkehrerwand* (objection about recurrences) and the *Umkehrerwand* (objection about reversibility).

Recurrence and reversal paradoxes

In his famous work that founded modern chaos theory, Poincaré proved a theorem that seemed contrary to the ideas of Boltzmann. For a closed dynamical system of the type we are studying he showed that the system will return infinitely often to an arbitrarily small region near to its original state. This is troubling: it seems to say that if we confine a gas to a box of volume V , and then break a wall so that the atoms can wander in a bigger box, say of volume $2V$, and then wait, the atoms will go back into the smaller box. This objection was raised by Ernst Zermelo.

This is not really an objection. In fact, ergodic theory says the same thing: we average over all phase space, including parts where all the atoms are in the smaller volume. The problem is how long we have to wait. Ergodic theory gives us a hint: consider an ideal gas, for which \mathcal{H} only depends on p_i . Now the fraction of phase space occupied in the initial state is:

$$\frac{\int \delta(\mathcal{H}(p_i) - E) d^{3N}p \int_V d^{3N}q}{\int \delta(\mathcal{H}(p_i) - E) d^{3N}p \int_{2V} d^{3N}q} = \frac{V^N}{(2V)^N} = \frac{1}{2^N}. \quad (2.29)$$

For a mixing system, if we sample the observations \mathcal{N} times, we will find the original state $\mathcal{N}/2^N$ times. So for two particles 1/4 of the observations will be like this, but for a macroscopic system the fraction will be of order $2^{10^{20}}$. Put another way, we will have to wait $2^N \tau$, where τ is the shortest time for an independent measurement (e.g. the time for particle to traverse the system). This time, called the Poincaré cycle time, is longer than the age of the universe.

Boltzmann's answer to Zermelo (translated in Brush & Hall (2003)) was marked by his characteristic caustic wit:

Thus when Zermelo concludes, from the theoretical fact that the initial states in a gas must recur — without having calculated how long a time this will take — that the hypotheses of gas theory must be rejected or else fundamentally changed, he is just like a dice player who has calculated that the probability of a sequence of 1000 one's is not zero, and then concludes that his dice must be loaded since he has not yet observed such a sequence!

Johann Loschmidt objected to Boltzmann's ideas on the ground that mechanics has time-reversal invariance. Thus you cannot deduce irreversible behavior, like approach to equilibrium, from mechanics. Boltzmann's response was, in effect, that if you *prepare* a system out of equilibrium, the boundary conditions set a direction of time, not the equation of motion. In fact, if the experimenter is working on the system, the Hamiltonian is different for $t < 0$, so we should not expect reversibility. An interesting discussion of this point is given by Ambegaokar & Clerk (1999) in terms of the Ehrenfest “dog-flea” model (equivalent to the Ising model with $J = h = 0$).

There are deep philosophical questions about the “arrow of time” connected with this point. Our interest here is physics, not philosophy; we will go no further.

2.2.5. Relaxation times

Mathematical approaches to ergodic theory are silent on the question of relaxation times. This is as it should be: relaxation depends crucially on the system studied. A system such as the Sinai billiard can approach equilibrium very quickly. Other systems do not.

To chose a random example, diamonds are not the ground state of carbon at room temperature: the stable structure is graphite — diamonds are not forever. However, the time to convert your diamond ring to an ugly bit of pencil lead is very long, as witnessed by the presence of diamonds in old geological formations. We may confidently expect that if we are willing to wait many geological eras for conversion, and then do our time averaging, we would get a correct average. We could even wait a Poincaré cycle time and hope our ring comes back.

Does this mean that statistical physics is useless for diamond? Not at all: we can use a constrained ensemble, namely assume that the very long conversion time is infinite, and get good results for diamonds in the laboratory.

One more point needs to be made: for a system that, empirically, comes to equilibrium quickly, we do not need to wait a Poincaré cycle time to do our averages. The Gibbs tendrils (see Figure 2.4) cover the whole shell coarsely at first, and then more and more finely. The time to wait for averages over R to settle down depends on the size of R . If R is large, in some sense, we have a coarse-grained measurement, a typical macroscopic experiment. For finer details, we have to wait longer.

This explains why for experiments (and numerical experiments) we can use phase space averages for macroscopic purposes as long as we don't insist on *very* fine details.

2.3. Quantum systems

We have confined our discussion to classical mechanics. For quantum systems the situation is more involved because of the discrete spectrum and the structure of eigenstates.

In fact, the situation is quite confusing. We could imagine that, for an isolated system, we prepare a quantum state in wavefunction, $\Psi(\{\mathbf{r}\}, t)$. Here $\{\mathbf{r}\}$ is the set of all the coordinates of all the particles.

Now we can, as is usual in quantum mechanics, expand in energy eigenfunctions:

$$\Psi = \sum_k c_k e^{-iE_k t/\hbar} \psi_k. \quad (2.30)$$

The problem here is that if the ψ_k are really energy eigenfunctions, this is an exact solution of the problem. The analogue of visiting the whole energy shell doesn't happen: we are stuck with a particular combination of eigenfunctions, the c_k , for all time. And, there will be quantum interference effects between the various eigenfunctions. This is exactly the situation discussed in quantum mechanics textbooks for two level systems, say in magnetic resonance.

2.3.1. Random phases

However, in a big quantum system this is never observed. The lovely interference effects of small, isolated systems disappear because of interactions with the environment. The reason, roughly speaking, is that statistical systems have many closely spaced eigenvalues so that transitions between them are impossible to avoid. What seems to happen is that these interactions don't have much effect on the energy, but scramble the phases causing decoherence of the wavefunctions.

The practical effect of this is as follows: suppose we take Ψ and try to compute an expectation value so that we can observe something. But, we need to average over the “stray” interactions that scramble the phases. We will denote this average as $\bar{\cdot}$. Then, for some operator, \hat{R} :

$$\begin{aligned} \overline{\langle \Psi | \hat{R} | \Psi \rangle} &= \sum_{k,l} \overline{c_k^* c_l} e^{i(E_k - E_l)t/\hbar} \langle \psi_k | \hat{R} | \psi_l \rangle \\ &\rightarrow \sum_k \overline{|c_k|^2} \langle \psi_k | \hat{R} | \psi_k \rangle. \end{aligned} \quad (2.31)$$

The off-diagonal terms “average out”, but the diagonal terms do not, because the phases cancel. Further, in our situation of nearly conserved energy, we must assume that the average means that each of the diagonal terms average to be the same:

$$\overline{|c_k|^2} = 1/\mathcal{N},$$

where \mathcal{N} is the degeneracy, the number of states with nearly the same energy. For example, for the Ising model in zero field, $\mathcal{N} = 2^N$.

2.3.2. Density matrix

Another way to put it is to write down the thermal equilibrium *density matrix*, which is the quantum analogue of the phase space density. For all the degenerate states:

$$\rho_{mc}^{k,l} \equiv \overline{c_k^* c_l} = \delta_{k,l} / \mathcal{N}. \quad (2.32)$$

The analogue to Eq. (2.23) is:

$$\langle \hat{R} \rangle = \sum_{k,l} \rho_{mc}^{k,l} \langle \psi_k | \hat{R} | \psi_l \rangle = \sum_k \frac{\langle \psi_k | \hat{R} | \psi_k \rangle}{\mathcal{N}}. \quad (2.33)$$

Sometimes it is useful (though we will never use this in this book) to define a density operator, in this case the microcanonical version.

$$\hat{\rho}_{mc} = \frac{\delta(E - \hat{\mathcal{H}})}{\sum_k \langle k | \delta(E - \hat{\mathcal{H}}) | k \rangle} = \frac{\delta(E - \hat{\mathcal{H}})}{\text{Tr}(\delta(E - \hat{\mathcal{H}}))} \quad (2.34)$$

Here, Tr means the trace. The average can be written:

$$\langle R \rangle = \text{Tr}(\hat{\rho} \hat{R}) \quad (2.35)$$

This all sounds complicated, and justifying it (at the same level that we did classical systems) would be. However, the recipe for calculation is simple. For a nearly closed quantum system average over all the degenerate states with equal weight. As we will see, this is not terribly hard — if the model is tractable. For the simple case of the Ising model, we can get useful explicit answers this way, as we will see.

Later we will see how to deal with systems at fixed temperature. In this case we average over states with a probability distribution (the Boltzmann factor). This changes the form of the density matrix. There are several examples of this in Chapter 4.

2.4. Method of the most probable distribution

It is instructive to follow Boltzmann and use the ideas above to derive the velocity distribution function, $f(\mathbf{p})$, for atoms in an ideal gas.

2.4.1. Maxwell-Boltzmann distribution

In a classical ideal gas we can consider molecules separately, and each one lives in its own phase space, called μ -space, which is 6 dimensional.

We divide μ -space into many cells of the same size $d\mu = d\mathbf{r}d\mathbf{p}$ located at different $\mathbf{r}_i, \mathbf{p}_i$.

We want to know how many molecules are in each cell; we call this set $n_i, i = 1, \dots, M$ where $\sum_1^M n_i = N$. Molecules in cell i have energy $\epsilon_i = p_i^2/2m$. The total energy, for non-interacting particles, is $E = \sum \epsilon_i n_i$.

To get to Γ -space we note that a given set, $\{n_i\}$ will live in a region of volume:

$$d\Gamma = d\mu_1^{n_1} d\mu_2^{n_2} \dots d\mu_M^{n_M}.$$

However, there are many different places in Γ -space that correspond to the same set of n_i 's, namely the number of ways to permute N molecules among M cells given $\{n_i\}$. This number is:

$$\frac{N!}{n_1! n_2! \dots n_M!}.$$

The probability to have $\{n_i\}$ is proportional to the total volume in phase space occupied:

$$\Gamma(\{n_i\}) = N! \prod_i \frac{d\mu_i^{n_i}}{n_i!}. \quad (2.36)$$

We need to maximize the volume with respect to each of the occupation numbers.

We take the logarithm and use two Lagrange multipliers to preserve energy and number conservation. We are led to the following equation for the maximum with respect to each n_j :

$$\frac{\partial}{\partial n_j} \left[\ln \Gamma(\{n_i\}) - \alpha \sum_i n_i - \beta \sum_i \epsilon_i n_i \right] = 0$$

We can use Stirling's approximation in the form $\ln n_i! \approx n_i \ln n_i - n_i$. We must solve:

$$\begin{aligned} 0 &= \frac{\partial}{\partial n_j} \sum_i [-(n_i \ln n_i - n_i) + n_i \ln d\mu_i - \alpha n_i - \beta \epsilon_i n_i], \\ 0 &= -\ln n_j - \alpha - \beta \epsilon_j + \ln d\mu_j \\ n_j &= d\mu_j e^{-\alpha} e^{-\beta \epsilon_j}. \end{aligned} \quad (2.37)$$

We set $e^{-\alpha} = A$. We can determine the constants, A, β by applying the constraints. For the ideal gas:

$$E = \sum_i \epsilon_i n_i = \int d\mathbf{p}d\mathbf{r} (p^2/2m) A e^{-\beta p^2/2m};$$

$$N = \sum_i n_i = \int d\mathbf{p}d\mathbf{r} A e^{-\beta p^2/2m}.$$

In particular, carrying out some Gaussian integrals we get:

$$\frac{E}{N} \equiv \overline{t_E} = \frac{\int d\mathbf{p} (p^2/2m) e^{-\beta p^2/2m}}{\int d\mathbf{p} e^{-\beta p^2/2m}} = \frac{3}{2\beta}. \quad (2.38)$$

If we compare this with Eq. (2.6) we have:

$$\begin{aligned} \beta &= 1/k_B T \\ n_i &\propto f(\mathbf{p}_i) \propto \exp(-p_i^2/2mk_B T). \end{aligned} \quad (2.39)$$

This is the famous Maxwell-Boltzmann distribution. The normalization, and some variants on the distribution are in the problems.

It is possible to show that, not only is this the most probable distribution, but most of phase space is occupied by distributions that differ very little from this one. The method is to show that most of the probability is contained in the region where the fractional difference of the average of n_i differs from the Maxwell-Boltzmann value by less than $\mathcal{O}(1/\sqrt{N}) = 10^{-10}$; see Huang (1987). We will get the distribution and the fluctuations in a different way later.

2.4.2. Fermi distribution

We can use the same method for a system of non-interacting fermions, the ideal Fermi gas. For non-interacting particles the antisymmetry of the wavefunction means that we cannot have more than one fermion in each single particle state. The derivation of Boltzmann makes no such restriction, and thus breaks down at high density because each cell has a degeneracy that is too large.

We proceed in the same way, but the volume elements on phase space, $d\mu_i$ need to be replaced by cells that group a number of quantum states, g_i . Now, as before, we have to figure out the number of ways to put the fermions into the cells. For classical particles, each cell gave a factor $w_i = d\mu^{n_i}/n_i!$. Now we need to make sure that there is no double occupancy. Thus the number of interchanges is different: we must distribute n_i particles in g_i states with single occupancy:

$$w_i = \frac{g_i!}{n_i!(g_i - n_i)!}.$$

Now we proceed as before:

$$\Gamma(\{n_i\}) = \prod_i \frac{g_i!}{n_i!(g_i - n_i)!}. \quad (2.40)$$

Now take the log, use Stirling's approximation, and set the variation to zero subject to two Lagrange multipliers. The result is:

$$n_i/g_i = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1}, \quad (2.41)$$

We have set $\alpha = -\beta\mu$ for reasons that will become evident in the next chapter. We will show that $\beta = 1/k_B T$ below. This is the Fermi-Dirac distribution. We will use it below for the physics of electrons in metals.

2.4.3. Bose distribution

For the Bose gas we need to do the counting differently because now there is no restriction on the number in a state, but we still need to make sure that we count different quantum states for indistinguishable particles. We can see how to count by picturing the cell, g_i , as a line with n_i particles and $g_i - 1$ boundaries between the different states in the cell. For example:

$$\circ \circ || \circ | \circ \circ || \dots,$$

means that the first state has 2 particles, the second is empty (two adjacent partitions) the third has one, the fourth has two, and the fifth and sixth states have no particles. The number of ways to realize this is the number of distinct ways to permute $n_i + g_i - 1$ partitions and particles:

$$w_i = \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}.$$

Proceeding as before gives:

$$\Gamma(\{n_i\}) = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}. \quad (2.42)$$

And therefore:

$$n_i/g_i = \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1}. \quad (2.43)$$

This is the Bose-Einstein distribution. We will use it below for the physics of liquid He, and for thermal properties of phonons and photons.

2.4.4. Classical limit

We can take the classical limit for both distributions by noting that if $n_i/g_i \ll 1$ we should get back to classical physics where particle

identity plays no real role. This can occur, in both cases, if $\beta(\epsilon_i - \mu) \gg 1$. Then we have:

$$\frac{1}{e^{\beta(\epsilon_i - \mu)} \pm 1} \rightarrow e^{\beta\mu} e^{-\beta\epsilon_i}. \quad (2.44)$$

We have the Maxwell-Boltzmann distribution, as we should expect, if we identify $\beta = 1/k_B T$. Later we will show that μ is the thermodynamic chemical potential.

Suggested reading

There are many excellent references and textbooks for this subject that the student can explore.

The classic undergraduate texts are:

Kittel & Kroemer (1980).

Reif (1965)

There is an undergraduate text which emphasizes numerics:

Gould & Tobochnik (2010)

At the graduate level there are many texts. Here are a few choices:

Huang (1987)

Pathria & Beale (2011)

Landau & Lifshitz (1980)

Chaikin & Lubensky (1995)

Ma (1985)

Peliti (2011)

Problems

1. The Maxwell-Boltzmann distribution is easy to observe in molecular dynamics. But we need to do some preliminary work first.

a.) Write down the full expression for $f(\mathbf{p})$ including the constants in front in 2 and 3 dimensions. Do this by requiring that $\int d\mathbf{p} f(\mathbf{p}) = N$.

b.) Find the distribution of speeds, $f(v)$, $v = p/m$, in 2 and 3d. Figure out the mean speed, i.e.:

$$\int_0^\infty v f(v) dv / \int_0^\infty f(v) dv.$$

c.) In a molecular dynamics code display a histogram of the speeds, and compare to the results above. You should average the histogram over time after equilibration. Use the total kinetic energy to get the temperature.

2. a.) Generalize your MD program to include an attractive force. Show that for small enough initial energy you get something resembling condensation.

b.) Use Eq. (2.11) to get the pressure in molecular dynamics. Show that for high enough T you approach the ideal gas law. Plot some isotherms in the non-ideal region.

3. a.) The derivation of the Maxwell-Boltzmann distribution can be generalized to the case when each molecule is subject to gravity. Work out $f(z, \mathbf{p})$ for this case. Show that the dependence on z gives rise to the *barometric distribution*, $n(z) = n(0)e^{-\beta mgz}$.

b.) Derive the result of a.) by macroscopic reasoning as follows:

- Argue that $p(z) - p(z + dz) = n(z)mgdz$.
- Use this to make a first-order differential equation relating $p(z)$ to $n(z)$.
- Use the ideal gas law, and solve the equation.

4. Does the reasoning leading to the Maxwell-Boltzmann distribution change if some particles are heavier than others? Suppose that there is just one particle with mass M and the rest have mass $m < M$. What is the mean kinetic energy of the heavy particle? What is its mean speed? Verify this roughly with molecular dynamics.

5. Figure out p_c, T_c, n_c for the van der Waals equation in terms of a, b . You need to set $dp/dn = d^2p/dn^2 = 0$. Explain why.

6. Fill in the steps in Eq. (2.38).

7. Scuba divers use a compressed air cylinder called an Aluminum-80 which means that 80 cubic ft of air at atmospheric pressure, room temperature, is jammed into a cylinder that you can carry on your back. The pressure is 3000 psi. (One atmosphere is 14.7 psi.) What is the internal volume of the cylinder? Work this out

for an ideal gas and a real gas (using the van der Waals constants for Nitrogen).

8. Consider a hot gas in a furnace with a hole through which a spectral line is observed. Show that the line is Doppler broadened so that the wavelength distribution of light intensity is given by:

$$I(\lambda) \propto \exp \left[-\frac{mc^2(\lambda - \lambda_0)^2}{2\lambda_0^2 k_B T} \right]$$

Here T is the temperature of the furnace, m the mass of the molecule, and λ_0 the wavelength when the molecule is at rest.

Hint: The Doppler effect works in the following way: the observed wavelength is:

$$\lambda \approx \lambda_0(1 + v_x/c),$$

where v_x is the velocity of the molecule emitting the light along the line of sight.

9. Compute the probability of having more than 0.001% difference in the number of molecules of ideal gas in two sides of a room. Suppose there are $N = 10^{20}$ total. We want $P(|R - L|/N > 10^{-5})$. Here $R + L = N$ and $R, L =$ number on right, left.

Hint: Assume that each molecule is on each side of the room with equal probability. Argue that the probability for a given value of R is $(N!/R!L!)2^{-N}$. Use Stirling's approximation, and express the result in terms of $m = (R - L)/N$ for small m . You should show that

$$P(m) \propto e^{-Nm^2/2}.$$

Find the normalization by setting

$$\int_{-\infty}^{\infty} P(m) dm = 1.$$

Argue that the probability, $P(|m| > r)$ is given by $2 \int_r^{\infty} P(m) dm$.

10. Here is an example of using Liouville's theorem. Show that the Verlet algorithm in one dimension is *symplectic*. This means that it conserves the phase space area $dx \times dp$. This is a property of the exact dynamics as well, as Liouville's theorem shows, so it is a desirable property for the approximate, numerical dynamics. It also implies that energy is conserved very well by Verlet – but this is more complicated, and we will not discuss it.

Hint: Start by showing that the algorithm in the problem in the previous chapter can be rewritten as follows. Set $m = 1$ so that $v = p$. Suppose you start with x_n, v_n at the n^{th} step. Then:

$$(a) \quad v_{n+1/2} = v_n + dt * a(x_n)/2,$$

$$(b) \quad x_{n+1} = x_n + dt * v_{n+1/2},$$

$$(c) \quad v_n = v_{n+1/2} + dt * a(x_{n+1})/2.$$

Now consider this as three transformations,

$$(a) \quad A : (x_n, v_n) \rightarrow (x_n, v_{n+1/2});$$

$$(b) \quad B : (x_n, v_{n+1/2}) \rightarrow (x_{n+1}, v_{n+1/2})$$

$$(c) \quad C : (x_{n+1}, v_{n+1/2}) \rightarrow (x_{n+1}, v_{n+1}).$$

If the area is to be preserved, the Jacobian of the transformation must be unity (make sure you remember what a Jacobian is). Show that $J = J(C)J(B)J(A) = 1$.