

## 2. Classical Gases

Our goal in this section is to use the techniques of statistical mechanics to describe the dynamics of the simplest system: a gas. This means a bunch of particles, flying around in a box. Although much of the last section was formulated in the language of quantum mechanics, here we will revert back to classical mechanics. Nonetheless, a recurrent theme will be that the quantum world is never far behind: we'll see several puzzles, both theoretical and experimental, which can only truly be resolved by turning on  $\hbar$ .

### 2.1 The Classical Partition Function

For most of this section we will work in the canonical ensemble. We start by reformulating the idea of a partition function in classical mechanics. We'll consider a simple system – a single particle of mass  $m$  moving in three dimensions in a potential  $V(\vec{q})$ . The classical Hamiltonian of the system<sup>3</sup> is the sum of kinetic and potential energy,

$$H = \frac{\vec{p}^2}{2m} + V(\vec{q})$$

We earlier defined the partition function (1.21) to be the sum over all quantum states of the system. Here we want to do something similar. In classical mechanics, the state of a system is determined by a point in *phase space*. We must specify both the position and momentum of each of the particles — only then do we have enough information to figure out what the system will do for all times in the future. This motivates the definition of the partition function for a single classical particle as the integration over phase space,

$$Z_1 = \frac{1}{h^3} \int d^3q d^3p e^{-\beta H(p,q)} \quad (2.1)$$

The only slightly odd thing is the factor of  $1/h^3$  that sits out front. It is a quantity that needs to be there simply on dimensional grounds:  $Z$  should be dimensionless so  $h$  must have dimension (length  $\times$  momentum) or, equivalently, Joules-seconds ( $J_s$ ). The actual value of  $h$  won't matter for any physical observable, like heat capacity, because we always take  $\log Z$  and then differentiate. Despite this, there is actually a correct value for  $h$ : it is Planck's constant,  $h = 2\pi\hbar \approx 6.6 \times 10^{-34} J_s$ .

It is very strange to see Planck's constant in a formula that is supposed to be classical. What's it doing there? In fact, it is a vestigial object, like the male nipple. It is redundant, serving only as a reminder of where we came from. And the classical world came from the quantum.

---

<sup>3</sup>If you haven't taken the [Classical Dynamics](#) course, you should think of the Hamiltonian as the energy of the system expressed in terms of the position and momentum of the particle.

### 2.1.1 From Quantum to Classical

It is possible to derive the classical partition function (2.1) directly from the quantum partition function (1.21) without resorting to hand-waving. It will also show us why the factor of  $1/h$  sits outside the partition function. The derivation is a little tedious, but worth seeing. (Similar techniques are useful in later courses when you first meet the path integral.) To make life easier, let's consider a single particle moving in one spatial dimension. It has position operator  $\hat{q}$ , momentum operator  $\hat{p}$  and Hamiltonian,

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q})$$

If  $|n\rangle$  is the energy eigenstate with energy  $E_n$ , the quantum partition function is

$$Z_1 = \sum_n e^{-\beta E_n} = \sum_n \langle n | e^{-\beta \hat{H}} | n \rangle \quad (2.2)$$

In what follows, we'll make liberal use of the fact that we can insert the identity operator anywhere in this expression. Identity operators can be constructed by summing over any complete basis of states. We'll need two such constructions, using the position eigenvectors  $|q\rangle$  and the momentum eigenvectors  $|p\rangle$ ,

$$\mathbf{1} = \int dq |q\rangle \langle q| \quad , \quad \mathbf{1} = \int dp |p\rangle \langle p|$$

We start by inserting two copies of the identity built from position eigenstates,

$$\begin{aligned} Z_1 &= \sum_n \langle n | \int dq |q\rangle \langle q| e^{-\beta \hat{H}} \int dq' |q'\rangle \langle q'| n \rangle \\ &= \int dq dq' \langle q | e^{-\beta \hat{H}} | q' \rangle \sum_n \langle q' | n \rangle \langle n | q \rangle \end{aligned}$$

But now we can replace  $\sum_n |n\rangle \langle n|$  with the identity matrix and use the fact that  $\langle q' | q \rangle = \delta(q' - q)$ , to get

$$Z_1 = \int dq \langle q | e^{-\beta \hat{H}} | q \rangle \quad (2.3)$$

We see that the result is to replace the sum over energy eigenstates in (2.2) with a sum (or integral) over position eigenstates in (2.3). If you wanted, you could play the same game and get the sum over any complete basis of eigenstates of your choosing. As an aside, this means that we can write the partition function in a basis independent fashion as

$$Z_1 = \text{Tr} e^{-\beta \hat{H}}$$

So far, our manipulations could have been done for any quantum system. Now we want to use the fact that we are taking the classical limit. This comes about when we try to factorize  $e^{-\beta\hat{H}}$  into a momentum term and a position term. The trouble is that this isn't always possible when there are matrices (or operators) in the exponent. Recall that,

$$e^{\hat{A}}e^{\hat{B}} = e^{\hat{A}+\hat{B}+\frac{1}{2}[\hat{A},\hat{B}]+\dots}$$

For us  $[\hat{q}, \hat{p}] = i\hbar$ . This means that if we're willing to neglect terms of order  $\hbar$  — which is the meaning of taking the classical limit — then we can write

$$e^{-\beta\hat{H}} = e^{-\beta\hat{p}^2/2m} e^{-\beta V(\hat{q})} + \mathcal{O}(\hbar)$$

We can now start to replace some of the operators in the exponent, like  $V(\hat{q})$ , with functions  $V(q)$ . (The notational difference is subtle, but important, in the expressions below!)

$$\begin{aligned} Z_1 &= \int dq \langle q | e^{-\beta\hat{p}^2/2m} e^{-\beta V(\hat{q})} | q \rangle \\ &= \int dq e^{-\beta V(q)} \langle q | e^{-\beta\hat{p}^2/2m} | q \rangle \\ &= \int dq dp dp' e^{-\beta V(q)} \langle q | p \rangle \langle p | e^{-\beta\hat{p}^2/2m} | p' \rangle \langle p' | q \rangle \\ &= \frac{1}{2\pi\hbar} \int dq dp e^{-\beta H(p,q)} \end{aligned}$$

where, in the final line, we've used the identity

$$\langle q | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{ipq/\hbar}$$

This completes the derivation.

## 2.2 Ideal Gas

The first classical gas that we'll consider consists of  $N$  particles trapped inside a box of volume  $V$ . The gas is “ideal”. This simply means that the particles do not interact with each other. For now, we'll also assume that the particles have no internal structure, so no rotational or vibrational degrees of freedom. This situation is usually referred to as the monatomic ideal gas. The Hamiltonian for each particle is simply the kinetic energy,

$$H = \frac{\vec{p}^2}{2m}$$

And the partition function for a single particle is

$$Z_1(V, T) = \frac{1}{(2\pi\hbar)^3} \int d^3q d^3p e^{-\beta\vec{p}^2/2m} \quad (2.4)$$

The integral over position is now trivial and gives  $\int d^3q = V$ , the volume of the box. The integral over momentum is also straightforward since it factorizes into separate integrals over  $p_x$ ,  $p_y$  and  $p_z$ , each of which is a Gaussian of the form,

$$\int dx e^{-ax^2} = \sqrt{\frac{\pi}{a}}$$

So we have

$$Z_1 = V \left( \frac{mk_B T}{2\pi\hbar^2} \right)^{3/2}$$

We'll meet the combination of factors in the brackets a lot in what follows, so it is useful to give it a name. We'll write

$$Z_1 = \frac{V}{\lambda^3} \quad (2.5)$$

The quantity  $\lambda$  goes by the name of the *thermal de Broglie wavelength*,

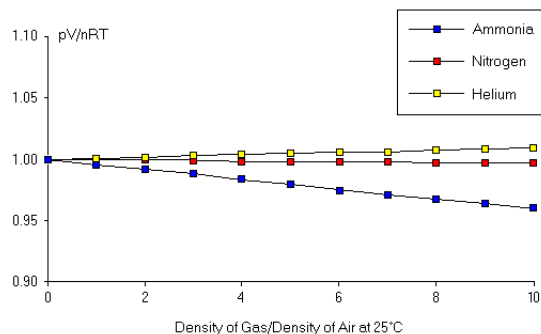
$$\lambda = \sqrt{\frac{2\pi\hbar^2}{mk_B T}} \quad (2.6)$$

$\lambda$  has the dimensions of length. We will see later that you can think of  $\lambda$  as something like the average de Broglie wavelength of a particle at temperature  $T$ . Notice that it is a quantum object – it has an  $\hbar$  sitting in it – so we expect that it will drop out of any genuinely classical quantity that we compute. The partition function itself (2.5) is counting the number of these thermal wavelengths that we can fit into volume  $V$ .

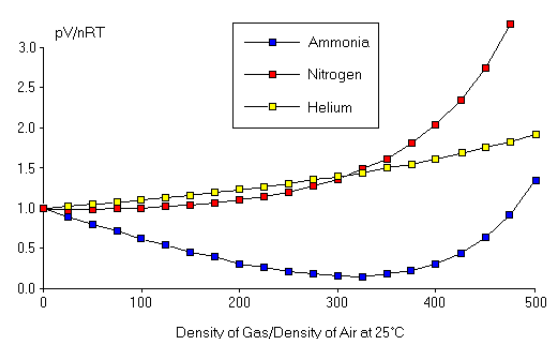
$Z_1$  is the partition function for a single particle. We have  $N$ , non-interacting, particles in the box so the partition function of the whole system is

$$Z(N, V, T) = Z_1^N = \frac{V^N}{\lambda^{3N}} \quad (2.7)$$

(Full disclosure: there's a slightly subtle point that we're brushing under the carpet here and this equation isn't quite right. This won't affect our immediate discussion and we'll explain the issue in more detail in Section 2.2.3.)



**Figure 8:** Deviations from ideal gas law at sensible densities



**Figure 9:** Deviations from ideal gas law at extreme densities

Armed with the partition function  $Z$ , we can happily calculate anything that we like. Let's start with the pressure, which can be extracted from the partition function by first computing the free energy (1.36) and then using (1.35). We have

$$\begin{aligned}
 p &= -\frac{\partial F}{\partial V} \\
 &= \frac{\partial}{\partial V}(k_B T \log Z) \\
 &= \frac{Nk_B T}{V}
 \end{aligned} \tag{2.8}$$

This equation is an old friend – it is the ideal gas law,  $pV = Nk_B T$ , that we all met in kindergarten. Notice that the thermal wavelength  $\lambda$  has indeed disappeared from the discussion as expected. Equations of this form, which link pressure, volume and temperature, are called *equations of state*. We will meet many throughout this course.

As the plots above show<sup>4</sup>, the ideal gas law is an extremely good description of gases at low densities. Gases deviate from this ideal behaviour as the densities increase and the interactions between atoms becomes important. We will see how this comes about from the viewpoint of microscopic forces in Section 2.5.

It is worth pointing out that this derivation should calm any lingering fears that you had about the definition of temperature given in (1.7). The object that we call  $T$  really does coincide with the familiar notion of temperature applied to gases. But the key property of the temperature is that if two systems are in equilibrium then they have the same  $T$ . That's enough to ensure that equation (1.7) is the right definition of temperature for *all* systems because we can always put any system in equilibrium with an ideal gas.

<sup>4</sup>Both figures are taken from the web textbook “*General Chemistry*” and credited to John Hutchinson.

### 2.2.1 Equipartition of Energy

The partition function (2.7) has more in store for us. We can compute the average energy of the ideal gas,

$$E = -\frac{\partial}{\partial\beta} \log Z = \frac{3}{2}Nk_B T \quad (2.9)$$

There's an important, general lesson lurking in this formula. To highlight this, it is worth repeating our analysis for an ideal gas in arbitrary number of spatial dimensions,  $D$ . A simple generalization of the calculations above shows that

$$Z = \frac{V^N}{\lambda^{DN}} \quad \Rightarrow \quad E = \frac{D}{2}Nk_B T$$

Each particle has  $D$  degrees of freedom (because it can move in one of  $D$  spatial directions). And each particle contributes  $\frac{1}{2}Dk_B T$  towards the average energy. This is a general rule of thumb, which holds for all classical systems: the average energy of each free degree of freedom in a system at temperature  $T$  is  $\frac{1}{2}k_B T$ . This is called the *equipartition of energy*. As stated, it holds only for degrees of freedom in the absence of a potential. (There is a modified version if you include a potential.) Moreover, it holds only for classical systems or quantum systems at suitably high temperatures.

We can use the result above to see why the thermal de Broglie wavelength (2.6) can be thought of as roughly equal to the average de Broglie wavelength of a particle. Equating the average energy (2.9) to the kinetic energy  $E = p^2/2m$  tells us that the average (root mean square) momentum carried by each particle is  $p \sim \sqrt{mk_B T}$ . In quantum mechanics, the de Broglie wavelength of a particle is  $\lambda_{dB} = h/p$ , which (up to numerical factors of 2 and  $\pi$ ) agrees with our formula (2.6).

Finally, returning to the reality of  $d = 3$  dimensions, we can compute the heat capacity for a monatomic ideal gas. It is

$$C_V = \left. \frac{\partial E}{\partial T} \right|_V = \frac{3}{2}Nk_B \quad (2.10)$$

### 2.2.2 The Sociological Meaning of Boltzmann's Constant

We introduced Boltzmann's constant  $k_B$  in our original definition of entropy (1.2). It has the value,

$$k_B = 1.381 \times 10^{-23} \text{ JK}^{-1}$$

In some sense, there is no deep physical meaning to Boltzmann's constant. It is merely a conversion factor that allows us to go between temperature and energy, as reflected

in (1.7). It is necessary to include it in the equations only for historical reasons: our ancestors didn't realise that temperature and energy were closely related and measured them in different units.

Nonetheless, we could ask why does  $k_B$  have the value above? It doesn't seem a particularly natural number. The reason is that both the units of temperature (Kelvin) and energy (Joule) are picked to reflect the conditions of human life. In the everyday world around us, measurements of temperature and energy involve fairly ordinary numbers: room temperature is roughly  $300\text{ K}$ ; the energy required to lift an apple back up to the top of the tree is a few Joules. Similarly, in an everyday setting, all the measurable quantities —  $p$ ,  $V$  and  $T$  — in the ideal gas equation are fairly normal numbers when measured in SI units. The only way this can be true is if the combination  $Nk_B$  is a fairly ordinary number, of order one. In other words the number of atoms must be huge,

$$N \sim 10^{23} \tag{2.11}$$

This then is the real meaning of the value of Boltzmann's constant: *atoms are small*.

It's worth stressing this point. Atoms aren't just small: they're really really small.  $10^{23}$  is an astonishingly large number. The number of grains of sand in all the beaches in the world is around  $10^{18}$ . The number of stars in our galaxy is about  $10^{11}$ . The number of stars in the entire visible Universe is probably around  $10^{22}$ . And yet the number of water molecules in a cup of tea is more than  $10^{23}$ .

### Chemist Notation

While we're talking about the size of atoms, it is probably worth reminding you of the notation used by chemists. They too want to work with numbers of order one. For this reason, they define a *mole* to be the number of atoms in one gram of Hydrogen. (Actually, it is the number of atoms in 12 grams of Carbon-12, but this is roughly the same thing.) The mass of Hydrogen is  $1.6 \times 10^{-27}\text{ kg}$ , so the number of atoms in a mole is Avogadro's number,

$$N_A \approx 6 \times 10^{23}$$

The number of moles in our gas is then  $n = N/N_A$  and the ideal gas law can be written as

$$pV = nRT$$

where  $R = N_A k_B$  is called the Universal gas constant. Its value is a nice sensible number with no silly power in the exponent:  $R \approx 8\text{ JK}^{-1}\text{mol}^{-1}$ .

### 2.2.3 Entropy and Gibbs's Paradox

“It has always been believed that Gibbs's paradox embodied profound thought. That it was intimately linked up with something so important and entirely new could hardly have been foreseen.”

*Erwin Schrödinger*

We said earlier that the formula for the partition function (2.7) isn't quite right. What did we miss? We actually missed a subtle point from quantum mechanics: quantum particles are indistinguishable. If we take two identical atoms and swap their positions, this doesn't give us a new state of the system – it is the same state that we had before. (Up to a sign that depends on whether the atoms are bosons or fermions. We'll discuss this aspect in more detail in Sections 3.5 and 3.6.) However, we haven't taken this into account. We wrote the expression  $Z = Z_1^N$  which would be true if all the  $N$  particles were distinguishable, for example, if each of the particles were of a different type. But this naive partition function overcounts the number of states in the system when we're dealing with indistinguishable particles.

It is a simple matter to write down the partition function for  $N$  indistinguishable particles. We simply need to divide by the number of ways to permute the particles. In other words, for the ideal gas the partition function is

$$Z_{\text{ideal}}(N, V, T) = \frac{1}{N!} Z_1^N = \frac{V^N}{N! \lambda^{3N}} \quad (2.12)$$

The extra factor of  $N!$  doesn't change the calculations of pressure or energy since, for each, we had to differentiate  $\log Z$  and any overall factor drops out. However, it does change the entropy since this is given by,

$$S = \frac{\partial}{\partial T} (k_B T \log Z_{\text{ideal}})$$

which includes a factor of  $\log Z$  without any derivative. Of course, since the entropy is counting the number of underlying microstates, we would expect it to know about whether particles are distinguishable or indistinguishable. Using the correct partition function (2.12) and Stirling's formula, the entropy of an ideal gas is given by,

$$S = Nk_B \left[ \log \left( \frac{V}{N\lambda^3} \right) + \frac{5}{2} \right] \quad (2.13)$$

This result is known as the *Sackur-Tetrode* equation. Notice that not only is the entropy sensitive to the indistinguishability of the particles, but it also depends on  $\lambda$ . However, the entropy is not directly measurable classically. We can only measure entropy differences by integrating the heat capacity as in (1.10).

The benefit of adding an extra factor of  $N!$  was noticed before the advent of quantum mechanics by Gibbs. He was motivated by the change in entropy of mixing between two gases. Suppose that we have two different gases, say red and blue. Each has the same number of particles  $N$  and sits in a volume  $V$ , separated by a partition. When the partition is removed the gases mix and we expect the entropy to increase. But if the gases are of the same type, removing the partition shouldn't change the macroscopic state of the gas. So why should the entropy increase? This is referred to as the *Gibbs' paradox*. Including the factor of  $N!$  in the partition function ensures that the entropy does not increase when identical atoms are mixed<sup>5</sup>.

### 2.2.4 The Ideal Gas in the Grand Canonical Ensemble

It is worth briefly looking at the ideal gas in the grand canonical ensemble. Recall that in such an ensemble, the gas is free to exchange both energy and particles with the outside reservoir. You could think of the system as some fixed subvolume inside a much larger gas. If there are no walls to define this subvolume then particles, and hence energy, can happily move in and out. We can ask how many particles will, on average, be inside this volume and what fluctuations in particle number will occur. More importantly, we can also start to gain some intuition for this strange quantity called the chemical potential,  $\mu$ .

The grand partition function (1.39) for the ideal gas is

$$\mathcal{Z}_{\text{ideal}}(\mu, V, T) = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_{\text{ideal}}(N, V, T) = \exp\left(\frac{e^{\beta\mu} V}{\lambda^3}\right)$$

From this we can determine the average particle number,

$$N = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z} = \frac{e^{\beta\mu} V}{\lambda^3}$$

Which, rearranging, gives

$$\mu = k_B T \log\left(\frac{\lambda^3 N}{V}\right) \quad (2.14)$$

If  $\lambda^3 < V/N$  then the chemical potential is negative. Recall that  $\lambda$  is roughly the average de Broglie wavelength of each particle, while  $V/N$  is the average volume taken

---

<sup>5</sup>Be warned however: a closer look shows that the Gibbs paradox is rather toothless and, in the classical world, there is no real necessity to add the  $N!$ . A clear discussion of these issues can be found in E.T. Jaynes' article "*The Gibbs Paradox*" which you can download from the [course website](#).

up by each particle. But whenever the de Broglie wavelength of particles becomes comparable to the inter-particle separation, then quantum effects become important. In other words, to trust our classical calculation of the ideal gas, we must have  $\lambda^3 \ll V/N$  and, correspondingly,  $\mu < 0$ .

At first sight, it is slightly strange that  $\mu$  is negative. When we introduced  $\mu$  in Section 1.4.1, we said that it should be thought of as the energy cost of adding an extra particle to the system. Surely that energy should be positive! To see why this isn't the case, we should look more closely at the definition. From the energy variation (1.38), we have

$$\mu = \left. \frac{\partial E}{\partial N} \right|_{S,V}$$

So the chemical potential should be thought of as the energy cost of adding an extra particle *at fixed entropy and volume*. But adding a particle will give more ways to share the energy around and so increase the entropy. If we insist on keeping the entropy fixed, then we will need to reduce the energy when we add an extra particle. This is why we have  $\mu < 0$  for the classical ideal gas.

There are situations where  $\mu > 0$ . This can occur if we have a suitably strong repulsive interaction between particles so that there's a large energy cost associated to throwing in one extra. We also have  $\mu > 0$  for fermion systems at low temperatures as we will see in Section 3.6.

We can also compute the fluctuation in the particle number,

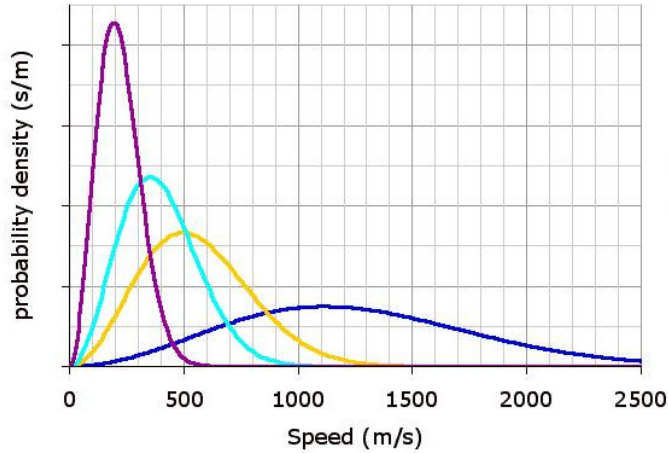
$$\Delta N^2 = \frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} \log \mathcal{Z}_{\text{ideal}} = N$$

As promised in Section 1.4.1, the relative fluctuations  $\Delta N / \langle N \rangle = 1/\sqrt{N}$  are vanishingly small in the thermodynamic  $N \rightarrow \infty$  limit.

Finally, it is very easy to compute the equation of state in the grand canonical ensemble because (1.45) and (1.48) tell us that

$$pV = k_B T \log \mathcal{Z} = k_B T \frac{e^{\beta\mu} V}{\lambda^3} = k_B T N \quad (2.15)$$

which gives us back the ideal gas law.



**Figure 10:** Maxwell distribution for Noble gases: *He*, *Ne*, *Ar* and *Xe*.

### 2.3 Maxwell Distribution

Our discussion above focusses on understanding macroscopic properties of the gas such as pressure or heat capacity. But we can also use the methods of statistical mechanics to get a better handle on the microscopic properties of the gas. Like everything else, the information is hidden in the partition function. Let's return to the form of the single particle partition function (2.4) before we do the integrals. We'll still do the trivial spatial integral  $\int d^3q = V$ , but we'll hold off on the momentum integral and instead change variables from momentum to velocity,  $\vec{p} = m\vec{v}$ . Then the single particle partition function is

$$Z_1 = \frac{m^3 V}{(2\pi\hbar)^3} \int d^3v e^{-\beta m \vec{v}^2 / 2} = \frac{4\pi m^3 V}{(2\pi\hbar)^3} \int dv v^2 e^{-\beta m v^2 / 2}$$

We can compare this to the original definition of the partition function: the sum over states of the probability of that state. But here too, the partition function is written as a sum, now over speeds. The integrand must therefore have the interpretation as the probability distribution over speeds. The probability that the atom has speed between  $v$  and  $v + dv$  is

$$f(v)dv = \mathcal{N} v^2 e^{-mv^2/2k_B T} dv \quad (2.16)$$

where the normalization factor  $\mathcal{N}$  can be determined by insisting that probabilities sum to one,  $\int_0^\infty f(v) dv = 1$ , which gives

$$\mathcal{N} = 4\pi \left( \frac{m}{2\pi k_B T} \right)^{3/2}$$

This is the *Maxwell distribution*. It is sometimes called the Maxwell-Boltzmann distribution. Figure 10 shows this distribution for a variety of gases with different masses at the same temperature, from the slow heavy Xenon (purple) to light, fast Helium (blue). We can use it to determine various average properties of the speeds of atoms in a gas. For example, the mean square speed is

$$\langle v^2 \rangle = \int_0^\infty dv v^2 f(v) = \frac{3k_B T}{m}$$

This is in agreement with the equipartition of energy: the average kinetic energy of the gas is  $E = \frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k_B T$ .

### Maxwell's Argument

The above derivation tells us the distribution of velocities in a non-interacting gas of particles. Remarkably, the Maxwell distribution also holds in the presence of any interactions. In fact, Maxwell's original derivation of the distribution makes no reference to any properties of the gas. It is very slick!

Let's first think about the distribution of velocities in the  $x$  direction; we'll call this distribution  $\phi(v_x)$ . Rotational symmetry means that we must have the same distribution of velocities in both the  $y$  and  $z$  directions. However, rotational invariance also requires that the full distribution can't depend on the direction of the velocity; it can only depend on the speed  $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$ . This means that we need to find functions  $F(v)$  and  $\phi(v_x)$  such that

$$F(v) dv_x dv_y dv_z = \phi(v_x)\phi(v_y)\phi(v_z) dv_x dv_y dv_z$$

It doesn't look as if we possibly have enough information to solve this equation for both  $F$  and  $\phi$ . But, remarkably, there is only one solution. The only function which satisfies this equation is

$$\phi(v_x) = A e^{-Bv_x^2}$$

for some constants  $A$  and  $B$ . Thus the distribution over speeds must be

$$F(v) dv_x dv_y dv_z = 4\pi v^2 F(v) dv = 4\pi A^3 v^2 e^{-Bv^2} dv$$

We see that the functional form of the distribution arises from rotational invariance alone. To determine the coefficient  $B = m/2k_B T$  we need the more elaborate techniques of statistical mechanics that we saw above. (In fact, one can derive it just from equipartition of energy.)

### 2.3.1 A History of Kinetic Theory

The term *kinetic theory* refers to the study of gases through their underlying atomic constituents. The discussion given above barely scratches the surface of this important subject.

Kinetic theory traces its origin to the work of Daniel Bernoulli in 1738. He was the first to argue that the phenomenon that we call pressure is due to the constant bombardment of tiny atoms. His calculation is straightforward. Consider a cubic box with sides of length  $L$ . Suppose that an atom travelling with momentum  $v_x$  in the  $x$  direction bounces elastically off a wall so that it returns with velocity  $-v_x$ . The particle experiences a change in momentum is  $\Delta p_x = 2mv_x$ . Since the particle is trapped in a box, it will next hit the wall at a time  $\Delta t = 2L/v_x$  later. This means that the force on the wall due to this atom is

$$F = \frac{\Delta p_x}{\Delta t} = \frac{mv_x^2}{L}$$

Summing over all the atoms which hit the wall, the force is

$$F = \frac{Nm\langle v_x^2 \rangle}{L}$$

where  $\langle v_x^2 \rangle$  is the average velocity in the  $x$ -direction. Using the same argument as we gave in Maxwell's derivation above, we must have  $\langle v_x^2 \rangle = \langle v^2 \rangle/3$ . Thus  $F = Nm\langle v \rangle^2/3L$  and the pressure, which is force per area, is given by

$$p = \frac{Nm\langle v^2 \rangle}{3L^3} = \frac{Nm\langle v^2 \rangle}{3V}$$

If this equation is compared to the ideal gas law (which, at the time, had only experimental basis) one concludes that the phenomenon of temperature must arise from the kinetic energy of the gas. Or, more precisely, one finds the equipartition result that we derived previously:  $\frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k_B T$ .

After Bernoulli's pioneering work, kinetic theory languished. No one really knew what to do with his observation nor how to test the underlying atomic hypothesis. Over the next century, Bernoulli's result was independently rediscovered by a number of people, all of whom were ignored by the scientific community. One of the more interesting attempts was by John Waterson, a Scottish engineer and naval instructor working for the East India Company in Bombay. Waterson was considered a crackpot. His 1843 paper was rejected by the Royal Society as "nothing but nonsense" and he wrote up his results in a self-published book with the wonderfully crackpot title "Thoughts on Mental Functions".

The results of Bernouilli and Waterson finally became accepted only after they were re-discovered by more established scientists, most notably Rudolph Clausius who, in 1857, extended these ideas to rotating and vibrating molecules. Soon afterwards, in 1859, Maxwell gave the derivation of the distribution of velocities that we saw above. This is often cited as the first statistical law of physics. But Maxwell was able to take things further. He used kinetic theory to derive the first genuinely new prediction of the atomic hypothesis: that the viscosity of a gas is independent of its density. Maxwell himself wrote,

*”Such a consequence of the mathematical theory is very startling and the only experiment I have met with on the subject does not seem to confirm it.”*

Maxwell decided to rectify the situation. With help from his wife, he spent several years constructing an experimental apparatus in his attic which was capable of providing the first accurate measurements of viscosity of gases<sup>6</sup>. His surprising theoretical prediction was confirmed by his own experiment.

There are many further developments in kinetic theory which we will not cover in this course. Perhaps the most important is the Boltzmann equation. This describes the evolution of a particle’s probability distribution in position and momentum space as it collides with other particles. Stationary, unchanging, solutions bring you back to the Maxwell-Boltzmann distribution, but the equation also provides a framework to go beyond the equilibrium description of a gas. You can read about this in the lecture notes on [Kinetic Theory](#).

## 2.4 Diatomic Gas

“I must now say something about these internal motions, because the greatest difficulty which the kinetic theory of gases has yet encountered belongs to this part of the subject”.

*James Clerk Maxwell, 1875*

Consider a molecule that consists of two atoms in a bound state. We’ll construct a very simple physicist’s model of this molecule: two masses attached to a spring. As well as the translational degrees of freedom, there are two further ways in which the molecule can move

---

<sup>6</sup>If you’re in Cambridge, you can see the original apparatus down the road in the corridor of the Cavendish lab.

- Rotation: the molecule can rotate rigidly about the two axes perpendicular to the axis of symmetry, with moment of inertia  $I$ . (For now, we will neglect the rotation about the axis of symmetry. It has very low moment of inertia which will ultimately mean that it is unimportant.)
- Vibration: the molecule can oscillate along the axis of symmetry

We'll work under the assumption that the rotation and vibration modes are independent. In this case, the partition function for a single molecule factorises into the product of the translation partition function  $Z_{\text{trans}}$  that we have already calculated (2.5) and the rotational and vibrational contributions,

$$Z_1 = Z_{\text{trans}} Z_{\text{rot}} Z_{\text{vib}}$$

We will now deal with  $Z_{\text{rot}}$  and  $Z_{\text{vib}}$  in turn.

### Rotation

The Lagrangian for the rotational degrees of freedom is<sup>7</sup>

$$L_{\text{rot}} = \frac{1}{2} I (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \quad (2.17)$$

The conjugate momenta are therefore

$$p_{\theta} = \frac{\partial L_{\text{rot}}}{\partial \dot{\theta}} = I \dot{\theta} \quad , \quad p_{\phi} = \frac{\partial L_{\text{rot}}}{\partial \dot{\phi}} = I \sin^2 \theta \dot{\phi}$$

from which we get the Hamiltonian for the rotating diatomic molecule,

$$H_{\text{rot}} = \dot{\theta} p_{\theta} + \dot{\phi} p_{\phi} - L = \frac{p_{\theta}^2}{2I} + \frac{p_{\phi}^2}{2I \sin^2 \theta} \quad (2.18)$$

The rotational contribution to the partition function is then

$$\begin{aligned} Z_{\text{rot}} &= \frac{1}{(2\pi\hbar)^2} \int d\theta d\phi dp_{\theta} dp_{\phi} e^{-\beta H_{\text{rot}}} \\ &= \frac{1}{(2\pi\hbar)^2} \sqrt{\frac{2\pi I}{\beta}} \int_0^{\pi} d\theta \sqrt{\frac{2\pi I \sin^2 \theta}{\beta}} \int_0^{2\pi} d\phi \\ &= \frac{2Ik_B T}{\hbar^2} \end{aligned} \quad (2.19)$$

---

<sup>7</sup>See, for example, Section 3.6 of the lecture notes on [Classical Dynamics](#).

From this we can compute the average rotational energy of each molecule,

$$E_{\text{rot}} = k_B T$$

If we now include the translational contribution (2.5), the partition function for a diatomic molecule that can spin and move, but can't vibrate, is given by  $Z_1 = Z_{\text{trans}} Z_{\text{rot}} \sim (k_B T)^{5/2}$ , and the partition function for a gas of these object  $Z = Z_1^N / N!$ , from which we compute the energy  $E = \frac{5}{2} N k_B T$  and the heat capacity,

$$C_V = \frac{5}{2} k_B N$$

In fact we can derive this result simply from equipartition of energy: there are 3 translational modes and 2 rotational modes, giving a contribution of  $5N \times \frac{1}{2} k_B T$  to the energy.

### Vibrations

The Hamiltonian for the vibrating mode is simply a harmonic oscillator. We'll denote the displacement away from the equilibrium position by  $\zeta$ . The molecule vibrates with some frequency  $\omega$  which is determined by the strength of the atomic bond. The Hamiltonian is then

$$H_{\text{vib}} = \frac{p_\zeta^2}{2m} + \frac{1}{2} m \omega^2 \zeta^2$$

from which we can compute the partition function

$$Z_{\text{vib}} = \frac{1}{2\pi\hbar} \int d\zeta dp_\zeta e^{-\beta H_{\text{vib}}} = \frac{k_B T}{\hbar\omega} \quad (2.20)$$

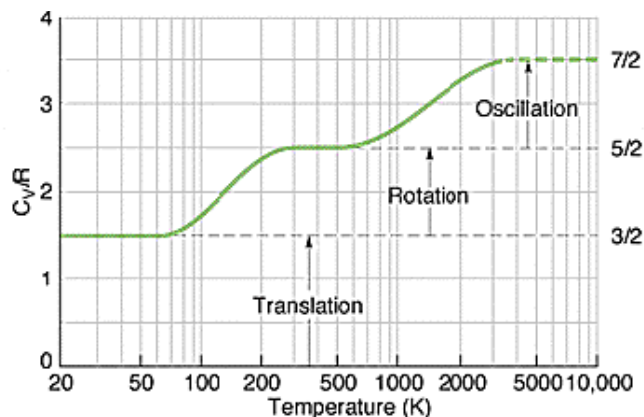
The average vibrational energy of each molecule is now

$$E_{\text{vib}} = k_B T$$

(You may have anticipated  $\frac{1}{2} k_B T$  since the harmonic oscillator has just a single degree of freedom, but equipartition works slightly differently when there is a potential energy. You will see another example on the problem sheet from which it is simple to deduce the general form.)

Putting together all the ingredients, the contributions from translational motion, rotation and vibration give the heat capacity

$$C_V = \frac{7}{2} N k_B$$



**Figure 11:** The heat capacity of Hydrogen gas  $H_2$ . The graph was created by P. Eyland.

This result depends on neither the moment of inertia,  $I$ , nor the stiffness of the molecular bond,  $\omega$ . A molecule with large  $I$  will simply spin more slowly so that the average rotational kinetic energy is  $k_B T$ ; a molecule attached by a stiff spring with high  $\omega$  will vibrate with smaller amplitude so that the average vibrational energy is  $k_B T$ . This ensures that the heat capacity is constant.

Great! So the heat capacity of a diatomic gas is  $\frac{7}{2} N k_B$ . Except it's not! An idealised graph of the heat capacity for  $H_2$ , the simplest diatomic gas, is shown in Figure 11. At suitably high temperatures, around  $5000 K$ , we do see the full heat capacity that we expect. But at low temperatures, the heat capacity is that of monatomic gas. And, in the middle, it seems to rotate, but not vibrate. What's going on? Towards the end of the nineteenth century, scientists were increasingly bewildered about this behaviour.

What's missing in the discussion above is something *very* important:  $\hbar$ . The successive freezing out of vibrational and rotational modes as the temperature is lowered is a quantum effect. In fact, this behaviour of the heat capacities of gases was the first time that quantum mechanics revealed itself in experiment. We're used to thinking of quantum mechanics as being relevant on small scales, yet here we see that it affects the physics of gases at temperatures of  $2000 K$ . But then, that is the theme of this course: how the microscopic determines the macroscopic. We will return to the diatomic gas in Section 3.4 and understand its heat capacity including the relevant quantum effects.

## 2.5 Interacting Gas

Until now, we've only discussed free systems; particles moving around unaware of each other. Now we're going to turn on interactions. Here things get much more interesting.

And much more difficult. Many of the most important unsolved problems in physics are to do with the interactions between large number of particles. Here we'll be gentle. We'll describe a simple approximation scheme that will allow us to begin to understand the effects of interactions between particles.

We'll focus once more on the monatomic gas. The ideal gas law is exact in the limit of no interactions between atoms. This is a good approximation when the density of atoms  $N/V$  is small. Corrections to the ideal gas law are often expressed in terms of a density expansion, known as the *virial expansion*. The most general equation of state is,

$$\frac{p}{k_B T} = \frac{N}{V} + B_2(T) \frac{N^2}{V^2} + B_3(T) \frac{N^3}{V^3} + \dots \quad (2.21)$$

where the functions  $B_j(T)$  are known as virial coefficients.

Our goal is to compute the virial coefficients from first principles, starting from a knowledge of the underlying potential energy  $U(r)$  between two neutral atoms separated by a distance  $r$ . This potential has two important features:

- An attractive  $1/r^6$  force. This arises from fluctuating dipoles of the neutral atoms. Recall that two permanent dipole moments,  $p_1$  and  $p_2$ , have a potential energy which scales as  $p_1 p_2 / r^3$ . Neutral atoms don't have permanent dipoles, but they can acquire a temporary dipole due to quantum fluctuations. Suppose that the first atom has an instantaneous dipole  $p_1$ . This will induce an electric field which is proportional to  $E \sim p_1 / r^3$  which, in turn, will induce a dipole of the second atom  $p_2 \sim E \sim p_1 / r^3$ . The resulting potential energy between the atoms scales as  $p_1 p_2 / r^3 \sim 1/r^6$ . This is sometimes called the *van der Waals* interaction.
- A rapidly rising repulsive interaction at short distances, arising from the Pauli exclusion principle that prevents two atoms from occupying the same space. For our purposes, the exact form of this repulsion is not so relevant: just as long as it's big. (The Pauli exclusion principle is a quantum effect. If the exact form of the potential is important then we really need to be dealing with quantum mechanics all along. We will do this in the next section.)

One very common potential that is often used to model the force between atoms is the *Lennard-Jones potential*,

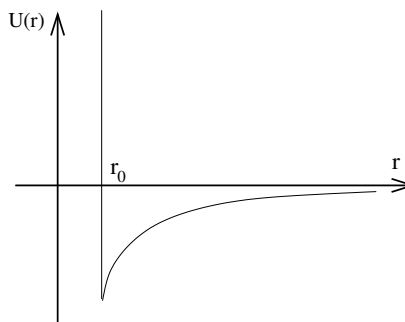
$$U(r) \sim \left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^6 \quad (2.22)$$

The exponent 12 is chosen only for convenience: it simplifies certain calculations because  $12 = 2 \times 6$ .

An even simpler form of the potential incorporates a *hard core repulsion*, in which the particles are simply forbidden from closer than a fixed distance by imposing an infinite potential,

$$U(r) = \begin{cases} \infty & r < r_0 \\ -U_0 \left(\frac{r_0}{r}\right)^6 & r \geq r_0 \end{cases} \quad (2.23)$$

The hard-core potential with van der Waals attraction is sketched to the right. We will see shortly that the virial coefficients are determined by increasingly difficult integrals involving the potential  $U(r)$ . For this reason, it's best to work with a potential that's as simple as possible. When we come to do some actual calculations we will use the form (2.23).



**Figure 12:**

### 2.5.1 The Mayer f Function and the Second Virial Coefficient

We're going to change notation and call the positions of the particles  $\vec{r}$  instead of  $\vec{q}$ . (The latter notation was useful to stress the connection to quantum mechanics at the beginning of this Section, but we've now left that behind!) If we restrict to two-particle interactions only, the Hamiltonian of the gas is

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i>j} U(r_{ij})$$

where  $r_{ij} = |\vec{r}_i - \vec{r}_j|$  is the separation between particles. The restriction  $i > j$  on the final sum ensures that we sum over each pair of particles exactly once. The partition function is then

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \int \prod_{i=1}^N d^3p_i d^3r_i e^{-\beta H} \\ &= \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \left[ \int \prod_i d^3p_i e^{-\beta \sum_j p_j^2/2m} \right] \times \left[ \int \prod_i d^3r_i e^{-\beta \sum_{j<k} U(r_{jk})} \right] \\ &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3r_i e^{-\beta \sum_{j<k} U(r_{jk})} \end{aligned}$$

where  $\lambda$  is the thermal wavelength that we met in (2.6). We still need to do the integral over positions. And that looks hard! The interactions mean that the integrals don't

factor in any obvious way. What to do? One obvious thing to try is to Taylor expand (which is closely related to the so-called *cumulant expansion* in this context)

$$e^{-\beta \sum_{j < k} U(r_{jk})} = 1 - \beta \sum_{j < k} U(r_{jk}) + \frac{\beta^2}{2} \sum_{j < k, l < m} U(r_{jk})U(r_{lm}) + \dots$$

Unfortunately, this isn't so useful. We want each term to be smaller than the preceding one. But as  $r_{ij} \rightarrow 0$ , the potential  $U(r_{ij}) \rightarrow \infty$ , which doesn't look promising for an expansion parameter.

Instead of proceeding with the naive Taylor expansion, we will instead choose to work with the following quantity, usually called the *Mayer f function*,

$$f(r) = e^{-\beta U(r)} - 1 \quad (2.24)$$

This is a nicer expansion parameter. When the particles are far separated at  $r \rightarrow \infty$ ,  $f(r) \rightarrow 0$ . However, as the particles come close and  $r \rightarrow 0$ , the Mayer function approaches  $f(r) \rightarrow -1$ . We'll proceed by trying to construct a suitable expansion in terms of  $f$ . We define

$$f_{ij} = f(r_{ij})$$

Then we can write the partition function as

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3 r_i \prod_{j < k} (1 + f_{jk}) \\ &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3 r_i \left( 1 + \sum_{j < k} f_{jk} + \sum_{j < k, l < m} f_{jk} f_{lm} + \dots \right) \end{aligned} \quad (2.25)$$

The first term simply gives a factor of the volume  $V$  for each integral, so we get  $V^N$ . The second term has a sum, each element of which is the same. They all look like

$$\int \prod_{i=1}^N d^3 r_i f_{12} = V^{N-2} \int d^3 r_1 d^3 r_2 f(r_{12}) = V^{N-1} \int d^3 r f(r)$$

where, in the last equality, we've simply changed integration variables from  $\vec{r}_1$  and  $\vec{r}_2$  to the centre of mass  $\vec{R} = \frac{1}{2}(\vec{r}_1 + \vec{r}_2)$  and the separation  $\vec{r} = \vec{r}_1 - \vec{r}_2$ . (You might worry that the limits of integration change in the integral over  $\vec{r}$ , but the integral over  $f(r)$  only picks up contributions from atomic size distances and this is only actually a problem close to the boundaries of the system where it is negligible.) There is a term

like this for each pair of particles – that is  $\frac{1}{2}N(N-1)$  such terms. For  $N \sim 10^{23}$ , we can just call this a round  $\frac{1}{2}N^2$ . Then, ignoring terms quadratic in  $f$  and higher, the partition function is approximately

$$\begin{aligned} Z(N, V, T) &= \frac{V^N}{N! \lambda^{3N}} \left( 1 + \frac{N^2}{2V} \int d^3r f(r) + \dots \right) \\ &= Z_{\text{ideal}} \left( 1 + \frac{N}{2V} \int d^3r f(r) + \dots \right)^N \end{aligned}$$

where we've used our previous result that  $Z_{\text{ideal}} = V^N / N! \lambda^{3N}$ . We've also engaged in something of a sleight of hand in this last line, promoting one power of  $N$  from in front of the integral to an overall exponent. Massaging the expression in this way ensures that the free energy is proportional to the number of particles as one would expect:

$$F = -k_B T \log Z = F_{\text{ideal}} - N k_B T \log \left( 1 + \frac{N}{2V} \int d^3r f(r) \right) \quad (2.26)$$

However, if you're uncomfortable with this little trick, it's not hard to convince yourself that the result (2.27) below for the equation of state doesn't depend on it. We will also look at the expansion more closely in the following section and see how all the higher order terms work out.

From the expression (2.26) for the free energy, it is clear that we are indeed performing an expansion in density of the gas since the correction term is proportional to  $N/V$ . This form of the free energy will give us the second virial coefficient  $B_2(T)$ .

We can be somewhat more precise about what it means to be at low density. The exact form of the integral  $\int d^3r f(r)$  depends on the potential, but for both the Lennard-Jones potential (2.22) and the hard-core repulsion (2.23), the integral is approximately  $\int d^3r f(r) \sim r_0^3$ , where  $r_0$  is roughly the minimum of the potential. (We'll compute the integral exactly below for the hard-core potential.) For the expansion to be valid, we want each term with an extra power of  $f$  to be smaller than the preceding one. (This statement is actually only approximately true. We'll be more precise below when we develop the cluster expansion.) That means that the second term in the argument of the log should be smaller than 1. In other words,

$$\frac{N}{V} \ll \frac{1}{r_0^3}$$

The left-hand side is the density of the gas. The right-hand side is atomic density. Or, equivalently, the density of a substance in which the atoms are packed closely together. But we have a name for such substances – we call them liquids! Our expansion is valid for densities of the gas that are much lower than that of the liquid state.

### 2.5.2 van der Waals Equation of State

We can use the free energy (2.26) to compute the pressure of the gas. Expanding the logarithm as  $\log(1+x) \approx x$  we get

$$p = -\frac{\partial F}{\partial V} = \frac{Nk_B T}{V} \left( 1 - \frac{N}{2V} \int d^3r f(r) + \dots \right)$$

As expected, the pressure deviates from that of an ideal gas. We can characterize this by writing

$$\frac{pV}{Nk_B T} = 1 - \frac{N}{2V} \int d^3r f(r) \quad (2.27)$$

To understand what this is telling us, we need to compute  $\int d^3r f(r)$ . Firstly let's look at two trivial examples:

**Repulsion:** Suppose that  $U(r) > 0$  for all separations  $r$  with  $U(r = \infty) = 0$ . Then  $f = e^{-\beta U} - 1 < 0$  and the pressure increases, as we'd expect for a repulsive interaction.

**Attraction:** If  $U(r) < 0$ , we have  $f > 0$  and the pressure decreases, as we'd expect for an attractive interaction.

What about a more realistic interaction that is attractive at long distances and repulsive at short? We will compute the equation of state of a gas using the hard-core potential with van der Waals attraction (2.23). The integral of the Mayer  $f$  function is

$$\int d^3r f(r) = \int_0^{r_0} d^3r (-1) + \int_{r_0}^{\infty} d^3r (e^{+\beta U_0 (r_0/r)^6} - 1) \quad (2.28)$$

We'll approximate the second integral in the high temperature limit,  $\beta U_0 \ll 1$ , where  $e^{+\beta U_0 (r_0/r)^6} \approx 1 + \beta U_0 (r_0/r)^6$ . Then

$$\begin{aligned} \int d^3r f(r) &= -4\pi \int_0^{r_0} dr r^2 + \frac{4\pi U_0}{k_B T} \int_{r_0}^{\infty} dr \frac{r_0^6}{r^4} \\ &= \frac{4\pi r_0^3}{3} \left( \frac{U_0}{k_B T} - 1 \right) \end{aligned} \quad (2.29)$$

Inserting this into (2.27) gives us an expression for the equation of state,

$$\frac{pV}{Nk_B T} = 1 - \frac{N}{V} \left( \frac{a}{k_B T} - b \right)$$

We recognise this expansion as capturing the second virial coefficient in (2.21) as promised. The constants  $a$  and  $b$  are defined by

$$a = \frac{2\pi r_0^3 U_0}{3} \quad , \quad b = \frac{2\pi r_0^3}{3}$$

It is actually slightly more useful to write this in the form  $k_B T = \dots$ . We can multiply through by  $k_B T$  then, rearranging we have

$$k_B T = \frac{V}{N} \left( p + \frac{N^2}{V^2} a \right) \left( 1 + \frac{N}{V} b \right)^{-1}$$

Since we're working in an expansion in density,  $N/V$ , we're at liberty to Taylor expand the last bracket, keeping only the first two terms. We get

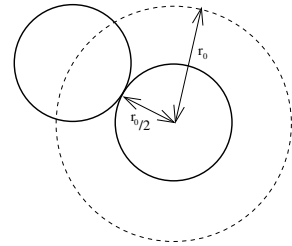
$$k_B T = \left( p + \frac{N^2}{V^2} a \right) \left( \frac{V}{N} - b \right) \tag{2.30}$$

This is the famous *van der Waals equation of state* for a gas. We stress again the limitations of our analysis: it is valid only at low densities and (because of our approximation when performing the integral (2.28)) at high temperatures.

We will return to the van der Waals equation in Section 5 where we'll explore many of its interesting features. For now, we can get a feeling for the physics behind this equation of state by rewriting it in yet another way,

$$p = \frac{N k_B T}{V - bN} - a \frac{N^2}{V^2} \tag{2.31}$$

The constant  $a$  contains a factor of  $U_0$  and so captures the effect of the attractive interaction at large distances. We see that its role is to reduce the pressure of the gas. The reduction in pressure is proportional to the density squared because this is, in turn, proportional to the number of *pairs* of particles which feel the attractive force. In contrast,  $b$  only contains  $r_0$  and arises due to the hard-core repulsion in the potential. Its effect is to reduce the effective volume of the gas because of the space taken up by the particles.



**Figure 13:**

It is worth pointing out where some quizzical factors of two come from in  $b = 2\pi r_0^3/3$ . Recall that  $r_0$  is the minimum distance that two atoms can approach. If we think of each atom as a hard sphere, then they have radius  $r_0/2$  and volume  $4\pi(r_0/2)^3/3$ . Which isn't equal to  $b$ . However, as illustrated in the figure, the excluded volume around each

atom is actually  $\Omega = 4\pi r_0^3/3 = 2b$ . So why don't we have  $\Omega$  sitting in the denominator of the van der Waals equation rather than  $b = \Omega/2$ ? Think about adding the atoms one at a time. The first guy can move in volume  $V$ ; the second in volume  $V - \Omega$ ; the third in volume  $V - 2\Omega$  and so on. For  $\Omega \ll V$ , the total configuration space available to the atoms is

$$\frac{1}{N!} \prod_{m=1}^N (V - m\Omega) \approx \frac{V^N}{N!} \left( 1 - \frac{N^2 \Omega}{2V} + \dots \right) \approx \frac{1}{N!} \left( V - \frac{N\Omega}{2} \right)^N$$

And there's that tricky factor of  $1/2$ .

Above we computed the equation of state for the dipole van der Waals interaction with hard core potential. But our expression (2.27) can seemingly be used to compute the equation of state for any potential between atoms. However, there are limitations. Looking back to the integral (2.29), we see that a long-range force of the form  $1/r^n$  will only give rise to a convergent integral for  $n \geq 4$ . This means that the techniques described above do not work for *long-range* potentials with fall-off  $1/r^3$  or slower. This includes the important case of  $1/r$  Coulomb interactions.

### 2.5.3 The Cluster Expansion

Above we computed the leading order correction to the ideal gas law. In terms of the virial expansion (2.21) this corresponds to the second virial coefficient  $B_2$ . We will now develop the full expansion and explain how to compute the higher virial coefficients.

Let's go back to equation (2.25) where we first expressed the partition function in terms of  $f$ ,

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3 r_i \prod_{j>k} (1 + f_{jk}) \\ &= \frac{1}{N! \lambda^{3N}} \int \prod_i d^3 r_i \left( 1 + \sum_{j>k} f_{jk} + \sum_{j>k, l>m} f_{jk} f_{lm} + \dots \right) \end{aligned} \quad (2.32)$$

Above we effectively related the second virial coefficient to the term linear in  $f$ : this is the essence of the equation of state (2.27). One might think that terms quadratic in  $f$  give rise to the third virial coefficient and so on. But, as we'll now see, the expansion is somewhat more subtle than that.

The expansion in (2.32) includes terms of the form  $f_{ij}f_{kl}f_{mn}\dots$  where the indices denote pairs of atoms,  $(i, j)$  and  $(k, l)$  and so on. These pairs may have atoms in common or they may all be different. However, the same pair never appears twice in a given term as you may check by going back to the first line in (2.32). We'll introduce a diagrammatic method to keep track of all the terms in the sum. To each term of the form  $f_{ij}f_{kl}f_{mn}\dots$  we associate a picture using the following rules

- Draw  $N$  atoms. (This gets tedious for  $N \sim 10^{23}$  but, as we'll soon see, we will actually only need pictures with small subset of atoms.)
- Draw a line between each pair of atoms that appear as indices. So for  $f_{ij}f_{kl}f_{mn}\dots$ , we draw a line between atom  $i$  and atom  $j$ ; a line between atom  $k$  and atom  $l$ ; and so on.

For example, if we have just  $N = 4$ , we have the following pictures for different terms in the expansion,

$$f_{12} = \begin{array}{c} \bullet \quad \bullet \\ 3 \quad 4 \\ \hline \bullet \quad \bullet \\ 1 \quad 2 \end{array} \quad f_{12}f_{34} = \begin{array}{c} \bullet \quad \bullet \\ 3 \quad 4 \\ \hline \bullet \quad \bullet \\ 1 \quad 2 \end{array} \quad f_{12}f_{23} = \begin{array}{c} \bullet \quad \bullet \\ 3 \quad 4 \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \\ 1 \quad 2 \end{array} \quad f_{21}f_{23}f_{31} = \begin{array}{c} \bullet \quad \bullet \\ 3 \quad 4 \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \\ 1 \quad 2 \end{array}$$

We call these diagrams *graphs*. Each possible graph appears exactly once in the partition function (2.32). In other words, the partition function is a sum over all graphs. We still have to do the integrals over all positions  $\vec{r}_i$ . We will denote the integral over graph  $G$  to be  $W[G]$ . Then the partition function is

$$Z(N, V, T) = \frac{1}{N! \lambda^{3N}} \sum_G W[G]$$

Nearly all the graphs that we can draw will have disconnected components. For example, those graphs that correspond to just a single  $f_{ij}$  will have two atoms connected and the remaining  $N - 2$  sitting alone. Those graphs that correspond to  $f_{ij}f_{kl}$  fall into two categories: either they consist of two pairs of atoms (like the second example above) or, if  $(i, j)$  shares an atom with  $(k, l)$ , there are three linked atoms (like the third example above). Importantly, the integral over positions  $\vec{r}_i$  then factorises into a product of integrals over the positions of atoms in disconnected components. This is illustrated by an example with  $N = 5$  atoms,

$$W \left[ \begin{array}{c} \bullet \quad \bullet \\ 3 \quad 4 \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \\ 1 \quad 2 \end{array} \quad \begin{array}{c} \bullet \\ 4 \\ \hline \bullet \\ 5 \end{array} \right] = \left( \int d^3r_1 d^3r_2 d^3r_3 f_{12} f_{23} f_{31} \right) \left( \int d^3r_4 d^3r_5 f_{45} \right)$$



The product  $N!/\prod_l m_l!(l!)^{m_l}$  counts the number of ways to split the particles into  $m_l$   $l$ -clusters, while ignoring the different ways to internally connect each cluster. This is the right thing to do since the different internal connections are taken into account in the integral  $U_l$ .

Combinatoric arguments are not always transparent. Let's do a couple of checks to make sure that this is indeed the right answer. Firstly, consider  $N = 4$  atoms split into two 2-clusters (i.e.  $m_2 = 2$ ). There are three such diagrams,  $f_{12}f_{34} = \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array}$ ,  $f_{13}f_{24} = \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array}$ , and  $f_{14}f_{23} = \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array}$ . Each of these gives the same answer when integrated, namely  $U_2^2$  so the final result should be  $3U_2^2$ . We can check this against the relevant terms in (2.35) which are  $4!U_2^2/2!^22! = 3U_2^2$  as expected.

Another check:  $N = 5$  atoms with  $m_2 = m_3 = 1$ . All diagrams come in the combinations

$$U_3U_2 = \int \prod_{i=1}^5 d^3r_i \left( \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array} + \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array} + \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array} + \begin{array}{c} \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \end{array} \right)$$

together with graphs that are related by permutations. The permutations are fully determined by the choice of the two atoms that sit in the pair: there are 10 such choices. The answer should therefore be  $10U_3U_2$ . Comparing to (2.35), we have  $5!U_3U_2/3!2! = 10U_3U_2$  as required.

Hopefully you are now convinced that (2.35) counts the graphs correctly. The end result for the partition function is therefore

$$Z(N, V, T) = \frac{1}{\lambda^{3N}} \sum_{\{m_l\}} \prod_l \frac{U_l^{m_l}}{(l!)^{m_l} m_l!}$$

The problem with computing this sum is that we still have to work out the different ways that we can split  $N$  atoms into different clusters. In other words, we still have to obey the constraint (2.33). Life would be very much easier if we didn't have to worry about this. Then we could just sum over any  $m_l$ , regardless. Thankfully, this is exactly what we can do if we work in the grand canonical ensemble where  $N$  is not fixed! The grand canonical ensemble is

$$\mathcal{Z}(\mu, V, T) = \sum_N e^{\beta\mu N} Z(N, V, T)$$

We define the *fugacity* as  $z = e^{\beta\mu}$ . Then we can write

$$\mathcal{Z}(\mu, V, T) = \sum_N z^N Z(N, V, T) = \prod_{l=1}^{\infty} \sum_{m_l=0}^{\infty} \left(\frac{z}{\lambda^3}\right)^{m_l} \frac{1}{m_l!} \left(\frac{U_l}{l!}\right)^{m_l} = \prod_{l=1}^{\infty} \exp\left(\frac{U_l z^l}{\lambda^{3l} l!}\right)$$

One usually defines

$$b_l = \frac{\lambda^3}{V} \frac{U_l}{l! \lambda^{3l}} \quad (2.36)$$

Notice in particular that  $U_1 = V$  so this definition gives  $b_1 = 1$ . Then we can write the grand partition function as

$$\mathcal{Z}(\mu, V, T) = \prod_{l=1}^{\infty} \exp\left(\frac{V}{\lambda^3} b_l z^l\right) = \exp\left(\frac{V}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l\right) \quad (2.37)$$

Something rather cute happened here. The sum over all diagrams got rewritten as the exponential over the sum of all *connected* diagrams, meaning all clusters. This is a general lesson which also carries over to quantum field theory where the diagrams in question are Feynman diagrams.

Back to the main plot of our story, we can now compute the pressure

$$\frac{pV}{k_B T} = \log \mathcal{Z} = \frac{V}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l$$

and the number of particles

$$\frac{N}{V} = \frac{z}{V} \frac{\partial}{\partial z} \log \mathcal{Z} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l b_l z^l \quad (2.38)$$

Dividing the two gives us the equation of state,

$$\frac{pV}{Nk_B T} = \frac{\sum_l b_l z^l}{\sum_l l b_l z^l} \quad (2.39)$$

The only downside is that the equation of state is expressed in terms of  $z$ . To massage it into the form of the virial expansion (2.21), we need to invert (2.38) to get  $z$  in terms of the particle density  $N/V$ . Equating (2.39) with (2.21) (and defining  $B_1 = 1$ ), we have

$$\begin{aligned} \sum_{l=1}^{\infty} b_l z^l &= \sum_{l=1}^{\infty} B_l \left(\frac{N}{V}\right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \\ &= \sum_{l=1}^{\infty} \frac{B_l}{\lambda^{3(l-1)}} \left(\sum_{n=1}^{\infty} n b_n z^n\right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \\ &= \left[ 1 + \frac{B_2}{\lambda^3} (z + 2b_2 z^2 + 3b_3 z^3 + \dots) + \frac{B_3}{\lambda^6} (z + 2b_2 z^2 + 3b_3 z^3 + \dots)^2 + \dots \right] \\ &\quad \times [z + 2b_2 z^2 + 3b_3 z^3 + \dots] \end{aligned}$$

where we've used both  $B_1 = 1$  and  $b_1 = 1$ . Expanding out the left- and right-hand sides to order  $z^3$  gives

$$z + b_2 z^2 + b_3 z^3 + \dots = z + \left( \frac{B_2}{\lambda^3} + 2b_2 \right) z^2 + \left( 3b_3 + \frac{4b_2 B_2}{\lambda^3} + \frac{B_3}{\lambda^6} \right) z^3 + \dots$$

Comparing terms, and recollecting the definitions of  $b_l$  (2.36) in terms of  $U_l$  (2.34) in terms of graphs, we find the second virial coefficient is given by

$$B_2 = -\lambda^3 b_2 = -\frac{U_2}{2V} = -\frac{1}{2V} \int d^3 r_1 d^3 r_2 f(\vec{r}_1 - \vec{r}_2) = -\frac{1}{2} \int d^3 r f(r)$$

which reproduces the result (2.27) that we found earlier using slightly simpler methods. We now also have an expression for the third coefficient,

$$B_3 = \lambda^6 (4b_2^2 - 2b_3)$$

although admittedly we still have a nasty integral to do before we have a concrete result. More importantly, the cluster expansion gives us the technology to perform a systematic perturbation expansion to any order we wish.

## 2.6 Screening and the Debye-Hückel Model of a Plasma

There are many other applications of the classical statistical methods that we saw in this chapter. Here we use them to derive the important phenomenon of *screening*. The problem we will consider, which sometimes goes by the name of a “one-component plasma”, is the following: a gas of electrons, each with charge  $-q$ , moves in a fixed background of uniform positive charge density  $+q\rho$ . The charge density is such that the overall system is neutral which means that  $\rho$  is also the average charge density of the electrons. This is the Debye-Hückel model.

In the absence of the background charge density, the interaction between electrons is given by the Coulomb potential

$$U(r) = \frac{q^2}{r}$$

where we're using units in which  $4\pi\epsilon_0 = 1$ . How does the fixed background charge affect the potential between electrons? The clever trick of the Debye-Hückel model is to use statistical methods to figure out the answer to this question. Consider placing one electron at the origin. Let's try to work out the electrostatic potential  $\phi(\vec{r})$  due to this electron. It is not obvious how to do this because  $\phi$  will also depend on the positions of all the other electrons. In general we can write,

$$\nabla^2 \phi(\vec{r}) = -4\pi (-q\delta(\vec{r}) + q\rho - q\rho g(\vec{r})) \quad (2.40)$$

where the first term on the right-hand side is due to the electron at the origin; the second term is due to the background positive charge density; and the third term is due to the other electrons whose average charge density close to the first electron is  $\rho g(\vec{r})$ . The trouble is that we don't know the function  $g$ . If we were sitting at zero temperature, the electrons would try to move apart as much as possible. But at non-zero temperatures, their thermal energy will allow them to approach each other. This is the clue that we need. The energy cost for an electron to approach the origin is, of course,  $E(\vec{r}) = -q\phi(\vec{r})$ . We will therefore assume that the charge density near the origin is given by the Boltzmann factor,

$$g(\vec{r}) \approx e^{\beta q\phi(\vec{r})}$$

For high temperatures,  $\beta q\phi \ll 1$ , we can write  $e^{\beta q\phi} \approx 1 + \beta q\phi$  and the Poisson equation (2.40) becomes

$$\left(\nabla^2 + \frac{1}{\lambda_D^2}\right)\phi(\vec{r}) = 4\pi q\delta(\vec{r})$$

where  $\lambda_D^2 = 1/4\pi\beta\rho q^2$ . This equation has the solution,

$$\phi(\vec{r}) = -\frac{qe^{-r/\lambda_D}}{r} \tag{2.41}$$

which immediately translates into an effective potential energy between electrons,

$$U_{\text{eff}}(r) = \frac{q^2 e^{-r/\lambda_D}}{r}$$

We now see that the effect of the plasma is to introduce the exponential factor in the numerator, causing the potential to decay very quickly at distances  $r > \lambda_D$ . This effect is called *screening* and  $\lambda_D$  is known as the *Debye screening length*. The derivation of (2.41) is self-consistent if we have a large number of electrons within a distance  $\lambda_D$  of the origin so that we can happily talk about average charge density. This means that we need  $\rho\lambda_D^3 \gg 1$ .