Lecture Notes on Statistical Mechanics

Gordon W. Semenoff
Department of Physics and Astronomy
University of British Columbia
6224 Agricultural Road
Vancouver, British Columbia, Canada V6T 1K8

Statistical mechanics is a theoretical framework that applies statistical methods and probability theory to physical systems which are composed of large assemblies of microscopic entities. It does not assume or postulate any natural laws, but it sets out to explain the macroscopic behaviour of a physical system from the behaviour of its microscopic constituents. In particular, its goal is to explain the success of classical thermodynamics in describing the macroscopic thermal properties and behaviour of matter in terms of microscopic properties of the physical system that is involved.

1 Basics of Statistical Mechanics: The Canonical Ensemble

The canonical ensemble is a way to describe a physical system which is in contact with a heat reservoir and where this contact has been allowed to come to equilibrium. Equilibrium is characterized by the temperature of the system and the heat reservoir. To be in equilibrium, the system and the heat reservoir should have the same temperatures.

Let us consider a mechanical system which can take up a number of energy states which we will assume, for the time being, are discrete,

$$E_1, E_2, E_3, \ldots$$

For the next while, it does not matter whether the system is quantum mechanical or classical mechanical, nor do any details of the internal dynamical

workings of the system matter. We only will need the assumption that the dynamics of the system lead to physical states with energies which can be listed as above. We also assume that there are no other conservation laws besides conservation of energy that would prevent the system from transitioning between any of its energy states.

The energy states need not be finite in number. We will assume that there is a lowest energy state, but perhaps not a highest energy state. The states could also be degenerate – meaning that there is more than one state with the same energy. In that case, the number E_a corresponding to the energy of the degenerate state would appear D_a times in the above list where D_a is the order of the degeneracy. [1]

What we will do is construct the "canonical ensemble" which is composed of a very large number, $\mathcal{N}$, of identical copies of our physical system. We then imagine that we put all of these copies into weak contact with each other. Here, the term "weak contact" means that there is some mechanism for energy to flow back and forth between different systems in the ensemble but the interaction energies due to this contact are a negligibly small contribution to the total energy so that the latter can, to as good an approximation as we will ever need, be written as a sum of the energies of each system. For any given individual system in this ensemble, the remainder of the systems, being very large in number, emulate a "heat reservoir".

A "state" of the ensemble is given by specifying the energy E_a of each system in the ensemble. These energies must add up to a total energy $\mathcal{E}$ of the entire ensemble. We shall then assume that we do not know which state the ensemble is in. We therefore can only describe it probabilistically. What we will assume is that the ensemble has equal probability of being in any state that is compatible with the fact that the energies of the system must add up to the energy of the ensemble. This is the hypothesis of "equal a priori probability". It is justified in physics by the idea that the weak interactions between the systems drives transitions between different energy states so that when we average over time, the ensemble spends the same amount of time in each possible state.

[1] The list of energies is discrete. If the energy spectrum is continuous, rather than discrete, this is not a big problem, we could simply (and admittedly somewhat arbitrarily) split the spectrum into energy bins, which we would label and list as above. The criterion for the size of a bin would perhaps be that the properties of the system do not vary discernibly as the energy varies over the interval of the bin. The degeneracy of a bin would be some estimate of the number of states in the bin.

Besides the concept of state of the ensemble, we shall also need the concept of a distribution. For a given state of the ensemble, let us define a distribution by counting the number of systems in the ensemble which are in each of the energy states. A distribution is therefore specified by a sequence of non-negative integers

$$\{n_1, n_2, n_3, \ldots\}$$

where n_1 is the number of systems in the ensemble which have energy E_1, n_2 is the number of systems in the ensemble with energy E_2, and so on.

Each state implies a distribution. If we know what the state is, we simply count how many systems in the ensemble have each value of the energies E_a and we find the distribution.

Conversly, if we know the distribution, we do not have a unique state, there are generally many states that give the same distribution.

A distribution is constrained by the fact that the total energy of the ensemble, $\mathcal{E}$, and the total number of copies of the system in the ensemble, $\mathcal{N}$, are given by

$$\mathcal{E} = n_1 E_1 + n_2 E_2 + n_3 E_3 + \ldots \tag{1}$$

$$\mathcal{N} = n_1 + n_2 + n_3 + \ldots \tag{2}$$

As we have already said, many different states can correspond to a given distribution. The fact that the ensemble is equally likely to be in any given state means that the probability that the system will have a given distribution is proportional to the number of states that correspond to that distribution. Therefore, to determine the likelihood of a distribution, we need to know how many states lead to that distribution.

Amongst all of the possible distributions, there is generally one which is most probable. This is the distribution that is most likely to describe the ensemble and we call it the "most probable distribution". We would like to find this most probable distribution. It is the distribution that the ensemble is most likely to be in at any given time.

In order to proceed, we must learn how to count the number of states of the ensemble that correspond to a given distribution. The systems in the ensemble are identical but they are distinguishable. We could imagine numbering them. Then, we can count the number of states in the distribution as the number of ways of placing $\mathcal{N}$ distinguishable objects into distinct bins in such a way that we have n_1 objects in bin number 1, n_2 objects in bin

number 2 and so on. This will turn out to be the combinatorial number

$$\frac{\mathcal{N}!}{n_1!n_2!\ldots}$$

To see how this comes about, we imagine that we begin by filling the first bin, which we call bin number 1. There are $\mathcal{N}$ distinguishable objects to choose the first object from. This contributes a factor of $\mathcal{N}$ to the total number of ways of achieving the desired filling. Afterward, we must choose the second object for bin number 1. For the second object there are $\mathcal{N}-1$ choices and this contributes a factor of $\mathcal{N}-1$ to the total number and we continue until the last object, for bin 1, that is the n_1'th object for which there are $\mathcal{N}-n_1+1$ choices. The total number of choices made in filling the first bin have now been

$$\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1)$$

But the order in which we placed objects in the bin does not matter. All that matters to us is that there are n_1 objects in total occupying that bin. This means that we have over-counted the ways. We should divide by the number of permutations of the objects in the bin, $n_1!$. Thus, upon filling the first bin, we see that it can be done in

$$\frac{\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1)}{n_1!}$$

different ways.

Then, we begin filling the second bin, which we call bin number 2. The procedure is identical to the first bin but it begins with $\mathcal{N}-n_1$ systems. Choosing the first system has $\mathcal{N}-n_1$ possibilities and contributes a factor of $\mathcal{N}-n_1$ to the number of ways of making the system. We continue with the next choice and so on in exact analogy with what we did for the first bin. What we find is the additional factor

$$\frac{(\mathcal{N}-n_1)(\mathcal{N}-n_1-1)...(\mathcal{N}-n_1-n_2+1)}{n_2!}$$

where the denominator again compensates for over-counting.

Thus, the total number of ways to get the first two bins is

$$\frac{\mathcal{N}(\mathcal{N}-1)\ldots(\mathcal{N}-n_1+1)}{n_1!}\cdot\frac{(\mathcal{N}-n_1)(\mathcal{N}-n_1-1)...(\mathcal{N}-n_1-n_2+1)}{n_2!}$$

Then we continue filling the third bin and so on. At the end of the process, the total number of ways that we could get the bins filled is given by

$$W[n_1, n_2, \ldots] = \frac{\mathcal{N}!}{n_1! n_2! \ldots} = \frac{\mathcal{N}!}{\prod_a n_a!} \tag{3}$$

Now, we can ask the question as to what is the most probable distribution. This is the distribution $\{n_1, n_2, \ldots\}$ with the largest value of $W[n_1, n_2, \ldots]$ in equation (3). We must find the $n_1, n_2, ...$ for which W is a maximum, subject to the constraint that the total energy of the ensemble, $\mathcal{E}$, is fixed and, of course, that the total number of systems in the ensemble is $\mathcal{N}$. This is equivalent to finding a maximum of the function

$$\Gamma[n_1, n_2, \ldots, \beta, \gamma] = \ln W[n_1, n_2, \ldots] + \beta\left[\mathcal{E} - \sum_a E_a n_a\right] + \gamma\left[\sum_a n_a - \mathcal{N}\right] \tag{4}$$

Here, we are using the fact that the logarithm is monotonic so, for example, the existence of a maximum of a real-valued positive function $f(x)$ implies a maximum of the functions $\ln f(x)$ and the maximum must occur at the same value of x. We have also added two Lagrange multipliers, β which enforces the constraint that the total energy is fixed and γ which requires the n_a's to add up to $\mathcal{N}$. This allows us to find an extremum of $\Gamma[n_1, n_2, \ldots, \beta, \gamma]$ without explicit constraints on the n_a's, aside from the fact that they should be integers. Setting the derivatives of $\Gamma[n_1, n_2, \ldots, \beta, \gamma]$ by β and γ to zero give the constraints (1) and (2).

Now, we make the assumption that the non-negative integers n_a are very large. Of course one might expect that this cannot be the case for all of them. In fact if the energy levels are unbounded from above, there must be some very high energy states (for example, all states with $E_a > \mathcal{E}$) for which $n_a = 0$ and our assumption certainly does not hold. A way to control this problem is to impose a high energy cutoff, that is to assume an upper bound on the energies, so that we have a finite number of energy bins. Then we can always make $\mathcal{N}$ large enough that a typical n_a's are as large as we want. Then, we will find a local maximum of Γ amongst those typical values. One can argue (but we will not prove) that the distributions where some of the n_a's are small are atypical and their number cannot compete with the number of states in the distribution which maximized Γ.

Stirling's approximation is the leading orders of the asymptotic expansion

$$\ln[\mathcal{N}!] = \mathcal{N}\ln\mathcal{N} - \mathcal{N} + \frac{1}{2}\ln[2\pi\mathcal{N}] + \frac{1}{12\mathcal{N}} + \ldots$$

where the corrections go to zero at large $\mathcal{N}$ at least as fast as $1/\mathcal{N}^2$.

So, we proceed with the assumption that all of the n_a's are large integers, large enough so that Stirling's formula is sufficiently accurate, whence

$$\Gamma[n_1, n_2, \ldots] = \mathcal{N}\ln\mathcal{N} - \mathcal{N} - \sum_a (n_a \ln n_a - n_a)$$
$$+ \beta\left[\mathcal{E} - \sum_a E_a n_a\right] + \gamma\left[\sum_a n_a - \mathcal{N}\right] \tag{5}$$

$$\Gamma[n_1, n_2, \ldots] =$$
$$\mathcal{N}\left\{-\sum_a (\rho_a \ln \rho_a - \rho_a) + \beta\left[\frac{\mathcal{E}}{\mathcal{N}} - \sum_a E_a \rho_a\right] + \gamma\left[\sum_a \rho_a - 1\right]\right\} \tag{6}$$

$$\rho_a \equiv \frac{n_a}{\mathcal{N}} \tag{7}$$

We have used some straightforward algebra to write the function that we want to maximize to be a function of the density variables $\rho_a = n_a/\mathcal{N}$. Even though n_a is integer valued, if n_a and $\mathcal{N}$ are large enough we can treat their ratio ρ_a as if it is a continuous variable. And, of course, ρ_a is positive and $\sum_a \rho_a = 1$.

To find the maximum of $\Gamma[n_1, n_2, \ldots]$, we take partial derivatives by all of the variables ρ_a, β, γ and we look for solutions of the equations that result when we set each partial derivative to zero,

$$-\ln\rho_a - \beta E_a + \gamma = 0 \tag{8}$$

$$U \equiv \frac{\mathcal{E}}{\mathcal{N}} = \sum_a E_a \rho_a \tag{9}$$

$$\sum_a \rho_a = 1 \tag{10}$$

Upon solving equations (8) and (10) we obtain the Boltzmann distribution

$$\rho_a = \frac{e^{-\beta E_a}}{\sum_a e^{-\beta E_a}} = \frac{e^{-\beta E_a}}{Z} \tag{11}$$

$$Z = \sum_a e^{-\beta E_a} \tag{12}$$

Here, the normalization Z is called the partition function. We have not determined the Lagrange multiplier β. Indeed, we did not solve all of the equations for the minimum. There is one left over, equation (9), which determines the average energy if β where known,

$$U \equiv \frac{\mathcal{E}}{\mathcal{N}} = \sum_a \rho_a E_a = \frac{\sum_a e^{-\beta E_a} E_a}{\sum_a e^{-\beta E_a}} \tag{13}$$

Here, we have introduced the thermodynamic symbol U, which describes the "internal energy" in classical thermodynamics as denoting the average energy of a system in the ensemble. This equation should be solved to determine β as a function of the internal energy, U. However, for most applications of this formalism it is easier, and in fact necessary to keep and interpret the parameter β as a parameter which characterizes the system. It turns out that it has a beautiful thermodynamic interpretation. It is proportional to the inverse of the temperature,

$$\beta = \frac{1}{k_B T} \tag{14}$$

where

$$\begin{aligned} k_B &= 1.380649 \times 10^{-23} \ \frac{m^2 kg}{s^2} \frac{1}{\text{deg K}} \\ &= 8.617333262 \frac{eV}{\text{deg K}} \end{aligned}$$

is Boltzmann's constant. The role of Boltzmann's constant is to fix the temperature units for us (in the above equations it is degrees Kelvin) and it converts temperature units to energy units.

To see that this is the correct interpretation of β, let us examine the logarithm of the partition function

$$F \equiv -\frac{1}{\beta} \ln Z = -\frac{1}{\beta} \ln \left[\sum_a e^{-\beta E_a} \right] \tag{15}$$

The quantity F is related to energy, in fact, if there were only one possible energy level,

$$F = -\frac{1}{\beta} \ln e^{-\beta E_1} = E_1 \tag{16}$$

and F would simply be the expectation value of the energy. However, even when there are only two energy levels, it would differ from the expectation value of the energy. To understand it better, it is suggestive to write its expression, as given in equation (15), as the internal energy U as given in equation (13) plus a correction,

$$\begin{aligned} F &= U - \frac{1}{\beta}\left\{ \ln\left[\sum_a e^{-\beta E_a}\right] + \frac{\sum_a \beta E_a e^{-\beta E_a}}{\sum e^{-\beta E_a}} \right\} \\ &= U - \frac{1}{\beta}\left\{ -\sum_a \rho_a \ln \rho_a \right\} \end{aligned} \tag{17}$$

where the second line in the above formulae follows from some elementary algebra and the definition of the probability ρ_a in equation (11). The bracketed term is a well-known information theoretic quantity, the von Neumann entropy of the probability distribution,

$$S_{\rm vN} = -\sum_a \rho_a \ln \rho_a \tag{18}$$

If we are willing to accept that the thermodynamic entropy of the physical system that we are describing is simply proportional to the von Neumann entropy of its probability distribution, the right-hand-side of equation (15) matches the thermodynamic formula

$$F = U - TS \tag{19}$$

where $\beta = 1/k_B T$, T is the temperature and the entropy is given by

$$S = k_B S_{\rm vN} \tag{20}$$

Then F is identified with the Helmholtz free energy, which one might recall has the thermodynamic interpretation as the maximum amount of work that can be extracted from the system while keeping the temperature fixed. Since S is necessarily positive, F is always less than the internal energy U.

1.1 Classical perfect gas

Let us consider one of the simplest possible examples, a physical system consisting of an assembly of weakly-interacting non-relativistic particles whose dynamics are governed by classical mechanics.

We shall assume that the interactions between the particles are so weak that, to a sufficiently good approximation, the total energy is given by the Hamiltonian which is simply a sum of the kinetic energies of the individual paricles.

$$H = \sum_{b=1}^{N} \frac{\vec{p}_b^2}{2m} \tag{21}$$

We have assumed that all of the particles have the same mass. In fact, for simplicity, we will assume that all of the physical attributes of the particles are identical and we then call them "identical particles". However, in classical mechanics they are not indistinguishable.

A state of a classical system such as this one is gotten by specifying the positions and momenta of all of the particles, that is the set of vectors $\{\vec{q}_1, \vec{p}_1, \ldots, \vec{q}_N, \vec{p}_N\}$. To calculate the partition function, we would take the Boltzmann weight, $e^{-H/k)BT}$ and sum (in this case integrate) over the states,

$$Z[T, V, N] = \int d\vec{q}_1 d\vec{p}_1 \ldots d\vec{q}_n d\vec{p}_n \frac{1}{(2\pi\hbar)^{3N}} \exp\left(-H(\vec{p}_1, \vec{p}_2, ...)/k_B T\right) \tag{22}$$

Here, we are doing a volume integral over the space of all positions and the space of all momenta of each of the particles. We note that, in equation (22) we have introduced a factor $\frac{1}{(2\pi\hbar)^{3N}}$ in order to make the partition function dimensionless. This is needed when we take its logarithm to find the free energy. The argument of any transcendental function must be a dimensionless number, independent of units of measure. The choice of Planck's constant here is entirely arbitrary. We could have chosen any number that has the dimensions of distance times momentum, or angular momentum. We use Planck's constant in this spot since it is known that $2\pi\hbar$ is the size of a unit cell in phase space in semi-classical physics. It plays an important role in Bohr-Sommerfeld quantization and other semi-classical approximations to quantum mechanical behaviour. Anticipating that our classical system should properly be the classical limit of a quantum system, this is a natural choice for this factor.

We can do the integral in equation (22) to get the partition function

$$Z[T, V, N] = V^N \left(\frac{m k_B T}{2\pi\hbar^2}\right)^{\frac{3N}{2}} \tag{23}$$

where the factors of V come from the integrals $V = \int d^3q$. The Helmholtz free energy is then given by

$$F[T, V, N] = -Nk_BT \ln Z[T, V, N]$$

which, using equation (23), results in

$$F[T, V, N] = -Nk_BT \ln\left[V \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] \tag{24}$$

The Helmholtz free energy is a function of the temperature, volume and the number of particles and it is related to the internal energy U and the entropy S by the equation

$$F = U - TS \tag{25}$$

The differential of the Helmholtz free energy is given by the equation

$$dF[T, V, N] \;=\; -\; SdT \;-\; PdV \;+\; \mu dN \tag{26}$$

and, consequently the entropy, pressure and chemical potential can be found from partial derivatives of the free energy as

$$S = -\left.\frac{\partial F}{\partial T}\right|_{V,N} \;,\; P = -\left.\frac{\partial F}{\partial V}\right|_{T,N} \;,\; \mu = \left.\frac{\partial F}{\partial N}\right|_{T,V} \tag{27}$$

Using these equations we get

$$S = k_BN \ln\left[V \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] + \frac{3}{2}k_BN \tag{28}$$

$$P = \frac{k_BNT}{V} \tag{29}$$

$$\mu = -k_BT \ln\left[V \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] \tag{30}$$

We can recognize equation (29) as the ideal gas law – the equation of state of a perfect gas. This tells us that the classical system that we are studying behaves like an ideal gas.

However, equations (28) and (30) are highly problematic. The variables S, V, N are supposed to be extensive thermodynamic variables. T, P, μ are

supposed to be intensive variables. If we double the volume of the system, while keeping the intensive variables fixed, the values of the extensive variables should be doubled. Equations (28) and (30) do not cooperate with what we expect here. In fact, the logic can be sharpend in the form of the Gibbs paradox which tell us that these cannot be the correct equations to describe an ideal gas.

This would seem to be a failure of statistical mechanics applied to a weakly interacting gas of classical particles – their behaviour does not agree with classical thermodynamics. However, it is also known to have a simple resolution. This resolution assumes that when we computed the partition function, we over-counted the states of the system. Of course, there is no basis for this in the strict realm of classical mechanics. In fact, having to do this is a hint directing us toward a quantum version of this problem where the particles would be indistinguishable and their counting indeed ends up with fewer states. And, to be clear, the statistics that we are introducing here and which are called "Maxwell-Boltzmann statistics" are not the correct quantum statistics. But they are a step in that direction. To impose Maxwell-Boltzmann statistics we assume that, in our computation of the partition function, our over-counting of the states can be fixed by dividing by a factor of the number of permutations of the N particles, $N!$, so that we take the partition function to be

$$Z[T,V,N] = \frac{1}{N!}\, V^N \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3N}{2}} \tag{31}$$

Then, assuming that N is large enough that Stirling's approximation

$$\ln N! \sim N\ln N - N = \ln\left[(N/e)^N\right] \quad , \quad \ln e = 1$$

is accurate, we can recalculate the Helmholtz free energy and the other thermodynamic functions that follow from it as

Mawell-Boltzmann Statistics:

$$F[T,V,N] = -Nk_BT\ln\left[\frac{eV}{N}\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] \tag{32}$$

$$S = k_BN\ln\left[\frac{eV}{N}\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] + \frac{3}{2}k_BN \tag{33}$$

$$P = \frac{k_BNT}{V} \tag{34}$$

$$\mu = k_BT - k_BT\ln\left[\frac{eV}{N}\left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] \tag{35}$$

$$U = F + TS = \frac{3}{2}Nk_BT \tag{36}$$

These give a beautiful description of an ideal gas. Our conclusion is that the ideal gas that is studied as a simple example of a thermodynamic system is matched perfectly by the gas of weakly interacting classical particles with Maxwell-Boltzmann statistics.

2 Basics of Statistical Mechanics The Grand Canonical ensemble

The grand canonical ensemble is very similar to the canonical ensemble where we consider a large collection of $\mathcal{N}$ copies of the physical system of interest, put into weak contact, in such a way that the systems are able to exchange both energy and particles. The state of a given system is specified by specifying the number of particles and the total energy of that system, (E_a, N_a). In the ensemble there are n_a systems which take up the state (E_a, N_a). The non-negative integers n_a are then subject to the constraints

$$\sum_a n_1 = \mathcal{N},\ \sum_a n_aE_a = U\mathcal{N},\ \sum_a n_aN_a = N\mathcal{N} \tag{37}$$

where U is the average energy and N is the average number of particles. Again, the number of ways of assembling this state of the ensemble is given by the combinatorial number

$$\frac{N!}{\prod_a n_a!}$$

and we want to find the state of the system which can be assembled in the maximum number of ways subject to the constraints, that is, the maximum of the function

$$\ln\left[\frac{N!}{\prod_a n_a!}\right] + \beta\left[U\mathcal{N} - \sum_a n_a E_a\right] + \alpha\left[N\mathcal{N} - \sum_a n_a N_a\right] + \gamma\left[\mathcal{N} - \sum_a n_a\right]$$

where N is the average number of particles per system. Assuming that the numbers n_a are large enough that Stirling's formula holds, we can write the above function as

$$\mathcal{N}\Bigg[-\sum_a[\rho_a \ln \rho_a - \rho_a] + \beta[U - \sum_a \rho_a E_a] + \alpha[N - \sum_a n_a N_a] + \gamma[1 - \sum_a \rho_a]\Bigg]$$

where $\rho_a = \frac{n_a}{N}$. Now we assume that n_a and $\mathcal{N}$ are both large enough that we can treat ρ_a as a continuously varying variable and we maximize the above function by taking its partial derivatives and setting them to zero. This results in the set of equations

$$-\rho_a - \beta E_a - \alpha N_a - \gamma = 0 \tag{38}$$

$$U = \sum_a \rho_a E_a \tag{39}$$

$$N = \sum_a \rho_a N_a \tag{40}$$

$$\sum_a \rho_a = 1 \tag{41}$$

We solve equations (38) and (41) to obtain

$$\rho_a = \frac{e^{-\frac{E_a}{k_B T} + \frac{\mu}{k_B T} N_a}}{\mathcal{Z}} \tag{42}$$

$$\mathcal{Z}[T, \mu, V] = \sum_a e^{-\frac{E_a}{k_B T} + \frac{\mu}{k_B T} N_a} \tag{43}$$

where $\mathcal{Z}$ is the grand canonical partition function and we have inserted the interpretation of the two Lagrange multipliers α and β in terms of the chemical potential and the temperature, respectively. Indeed, to see that this identification is the appropriate one is a procedure identical to that for the canonical

ensemble, with the difference that the logarithm of the grand canonical partition function is the grand canonical free energy (sometimes called the Landau free energy)

$$\Phi[T, \mu, V] \;=\; F \;-\; \mu N \tag{44}$$

The differential of the grand canonical free energy has the form

$$d\Phi \;=\; -\; SdT \;-\; Nd\mu - PdV \tag{45}$$

and that S and N can therefore be computed by taking the appropriate partial derivatives of the grand canonical free energy,

$$\left.\frac{\partial \Phi}{\partial T}\right|_{\mu,V} = -S \tag{46}$$

$$\left.\frac{\partial \Phi}{\partial \mu}\right|_{T,V} = -N \tag{47}$$

$$\left.\frac{\partial \Phi}{\partial V}\right|_{T,\mu} = -P \tag{48}$$

2.1 Grand canonical perfect gas

Starting with the partition function for the canonical ensemble that we computed for the classical assembly of N weakly interacting particles in equation (31), we can find the grand canonical partition function by multiplying by $\exp\left(\frac{\mu}{k_B T} N\right)$ and summing over N,

$$\mathcal{Z}[T, V, \mu] = \sum_{N=0}^{\infty} e^{\frac{\mu}{k_B T} N} \frac{1}{N!} Z[T, V, N] \tag{49}$$

$$= \sum_{N=0}^{\infty} e^{\frac{\mu}{k_B T} N} \frac{1}{N!}\; V^N \left(\frac{m k_B T}{2\pi\hbar^2}\right)^{\frac{3N}{2}} \tag{50}$$

$$= \exp\left[e^{\frac{\mu}{k_B T}}\; V \left(\frac{m k_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right] \tag{51}$$

and the grand canonical free energy is

$$\Phi[T, \mu, V] = -k_B T V e^{\frac{\mu}{k_B T}} \left(\frac{m k_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}} \tag{52}$$

We can see from this expression that

$$\Phi[T, \mu, V] = -PV = -Nk_BT$$

which is in perfect agreement with what we would obtain from the canonical ensemble computation of F and the Legendre transform

$$\Phi = F - \nu N$$

which relates the Helmholtz free energy F and the grand canonical potential Φ.

3 Basics of Statistical Mechanics The Micro-canonical Ensemble

The expectation value of the energy in the canonical ensemble is the internal energy

$$U = \frac{\sum_a E_a e^{-E_a/k_BT}}{\sum_a e^{-E_a/k_BT}}$$

This is the expectation value in the most probable distribution. Moreover, even in the most probably distribution, there is still a distribution of energy levels. One might wonder how certain one can be that the system actually has energy close to U, even in the most probable distribution. The probabilisic way to address this question is to compute the variance of the energy

$$\Delta U = \frac{1}{U}\sqrt{\frac{\sum_a (E_a - U)^2 e^{-E_a/k_BT}}{\sum_a e^{-E_a/k_BT}}}$$

If we take the heat capacity

$$C_V = \frac{\partial U}{\partial T} = \frac{1}{k_BT^2}\Delta U^2$$

In a wide range of substances, $C_V \sim V$ or $C_V \sim N$ which tells us that

$$\frac{\Delta U}{U} \sim \sqrt{\frac{1}{V}} \to 0 \text{ as } V \to \infty$$

The variance of the energy compared to the energy itself relaxes to zero in the limit where the volume is very large. We interpret this as indicating

that, in such a system, the fluctuations of the energy are very small. The distribution of states in the canonical ensemble is dominated by one state, the one with energy equal to U. We would say that

$$\rho_a = \begin{cases} 0 & E_a \neq U \\ \frac{1}{W} & E_a = U \end{cases}$$

where $W = D_a$ is the degeneracy of the energy state. We might expect that this will only work if W is a large number. In this case, the internal energy of the system is given by U and a direct route to the thermodynamics is the entropy which, if we assume the distribution that is given above, and that the thermodynamic entropy is proportional to the von Neumann entropy of this distribution, gives us Boltzmann's formula[2] for the entropy of an isolated system,

$$S[U] = k_B \ln W[U] \tag{53}$$

where W is the degeneracy of the state of the system with energy U. We recover thermodynamics by noting that the temperature is given by

$$\frac{1}{T} = \frac{\partial S}{\partial U} \tag{54}$$

where other parameters of the system (like the volume) are held fixed when the partial derivative is taken. Once we have the temperature, we can perform the Legendre transform

$$F = U - TS$$

to find the Helmholtz free energy.

3.1 Perfect Gas

We can apply Boltzmann's formula to the perfect gas. What we need to do is to compute $W[U]$. In classical mechanics this is an estimate of how many states have energy U. We can take this as given by the formula

$$W[U, V, N] = \frac{1}{N!} \int d^3q_1 d^3p_1 \ldots d^3q_N d^3p_N \frac{1}{(2\pi\hbar)^{3N}} \delta U \delta(U - \sum_i \vec{p}_i^2/2m) \tag{55}$$

[2]This is not the same as the Boltzmann equation which is used in non-equilibrium situations. We distinguish the two by calling this one "Boltzmann's formula".

where δU is an energy resolution and $\delta(U - \sum_i \vec{p}_i^2/2m)$ is a Dirac delta function. We take this integral as estimating the number of cells is the phase space of a gas of N particles which have energies in the interval $[U - \delta U/2, U + \delta U/2]$ which we take as the "energy bin". Following our discussions of the same factor in the canonical and grand canonical ensembles, we have included the factor $\frac{1}{(2\pi\hbar)^{3N}}$ which renders W dimensionless. We also impose Maxwell-Boltzmann statistics by dividing by the factor $N!$.

We can easily to the integrals in equation (55) to get [3]

$$W[U, V, N] = \frac{V^N}{N!\Gamma[3N/2]} \left(\frac{mU}{2\pi\hbar^2}\right)^{\frac{3}{2}N} \frac{2\delta U}{U} \tag{56}$$

Now, we assume N is very large and use Stirling's formula to get

$$W[U, V, N] = \left(\frac{eV}{N} \left(\frac{\frac{2e}{3} mU/N}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right)^N \frac{2\delta U}{U} \tag{57}$$

Taking the logarithm of this expression gives the formula for the entropy,

$$S[U, V, N] = k_B N \ln\left(\frac{V}{N} \left(\frac{m\frac{2}{3}U/N}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right) + \frac{5}{2} k_B N \tag{58}$$

and we can compute the temperature using equation (54) which gives us the expression

$$\frac{1}{T} = \frac{3}{2} k_B N \frac{1}{U}$$

which is equivalent to the formula for the equipartition of energy that we found for a perfect gas in the canonical ensemble. In fact, if we solve this equation for U as a function of N and T, we can form the entropy

$$S[T, N, V] = k_B N \ln\left(\frac{V}{N} \left(\frac{mk_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}}\right) + \frac{5}{2} k_B N \tag{59}$$

which matches equation (34) that we found using the canonical ensemble. This tells us that all of the termodynamic formulae that we derive from our

[3] We use the fact that the volume of the unit sphere embedded in D dimensions is $2\pi^{D/2}/\Gamma[D/2]$ where $\Gamma[x]$ is Euler's gamma function.

current micro-cannonical approach to the perfect gas will match those of the canonical ensemble as well as the same formulae in classical thermodynamics. The key to the equivalence of the micro-canonical, canonical and grand canonical formalisms has been the fact that N is large (which we needed, for example, to use Stirling's formula in the discussion above).

4 Quantum Perfect Gas

Let us now take our example of a gas of weakly interacting particle and assume that they have quantum mechanical behaviour. The energy of such a system is gotten by finding the eigenvalues of the Hamiltonian

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

and, in addition to this, we must assume that the particles are identical and, in fact indistinguishable so that they behave like fermions or bosons. In these cases, the grand canonical free energies are given by

$$\Phi[T,\mu,V] = k_B T V \int \frac{d^3p}{(2\pi)^3} \ln\left(1 + e^{-\frac{\hbar^2\vec{p}^2/2m-\mu}{k_B T}}\right) \qquad \text{fermions} \qquad (60)$$

$$\Phi[T,\mu,V] = -k_B T V \int \frac{d^3p}{(2\pi)^3} \ln\left(1 - e^{-\frac{\hbar^2\vec{p}^2/2m-\mu}{k_B T}}\right) \qquad \text{bosons} \qquad (61)$$

These expressions can be used to derive the thermodynamics of the quantum ideal gas. These can differ markedly from the classical ideal gas that we have been studying so far, particularly in the low temperature regime where the counting of states using Fermi or Bose statistics if very different from the counting of states using Maxwell-Boltzmann statistics. We might expect that quantum and classical behaviours are similar for highly excited states which should dominate the physical behaviour in the high temperature regime. This is indeed the case and it is possible to show that the behaviour of fermions and bosons are similar there and they are also similar to the classical perfect gas.