

Statistical Mechanics
and
its applications

Dr A. Alavi

Part II Chemistry

2007-2008

I. INTRODUCTION

Statistical Mechanics is concerned with how to describe the behaviour of systems containing large numbers of particles. Its remit is to provide the theoretical apparatus to analyse the equilibrium properties of such systems in terms of the underlying microscopic interactions between the particles. It is common experience (and indeed experimentally verifiable) that the properties of systems containing very large numbers of particles depend only on the specification of a few macroscopic parameters such as total (internal) energy E , temperature T , volume V , pressure P , density ρ , etc. On the other hand, a detailed description of the microscopic state of a system requires knowledge of a huge number (on the order of Avogadro's number) of rapidly fluctuating variables, such as the positions and momenta of each particle. The measurable macroscopic properties should be obtainable via an appropriate *averaging* over the microstates, and how this is done, as well as its theoretical justification, is the subject of Statistical Mechanics.

A. Why do we need to go beyond 'MELT'?

In Dr Keeler's IB course on 'Molecular Energy Levels and Thermodynamics', a great deal of attention was paid to the *single-particle (or molecular) partition function* q . This is defined as follows. Given a particle with available molecular energy states ϵ_s , a partition function was defined as 'a sum over all states':

$$q = \sum_s \exp(-\epsilon_s/kT) \quad (1)$$

You will recall that q measures the *number of thermally accessible states* available to our particle. If we set the zero of energy to be equal to the lowest available energy level (i.e. the ground-state) you will see that $q \geq 1$. When the first excited state is much higher than kT , then $q \sim 1$, whereas when many excited states are thermally accessible, then $q \gg 1$.

For a system of N identical (but non-interacting) particles, we found that the par-

tition function of the whole assembly was given as:

$$Q_N = \frac{q^N}{N!} \quad (2)$$

How does this arise? If the particles were *distinguishable*, then we would obviously obtain

$$Q_{distinguishable} = q^N \quad (3)$$

But for indistinguishable particles this expression contains an over-counting of states. For example, let particle 1 be in state ϵ_{n_1} , particle 2 in state ϵ_{n_2} , and so on. You can see that as long as the N particles occupy N different states, then there are obviously $N!$ permutations of the particles among the chosen states n_1, n_2, \dots, n_N , which are all equivalent. These states should be counted only once in the partition function Q . Therefore we divide the result by $N!$ to arrive at Eq.(2). A moment's thought will show that difficulties arise with this simple $N!$ correction as soon as we have *multiple occupancy* of the states.

For example, consider a two level system with energy levels at $0, \epsilon$, containing two particles. Then, the molecular partition function is:

$$q = 1 + \exp(-\epsilon/kT) \quad (4)$$

and therefore for two indistinguishable particles:

$$\begin{aligned} Q_2 &= \frac{q^2}{2} \\ &= \frac{1}{2} + \exp(-\epsilon/kT) + \frac{1}{2} \exp(-2\epsilon/kT) \end{aligned} \quad (5)$$

However, it is clear that if our particles are indistinguishable, there are only three non-equivalent states, with energies $0, \epsilon, 2\epsilon$, each with degeneracy 1:

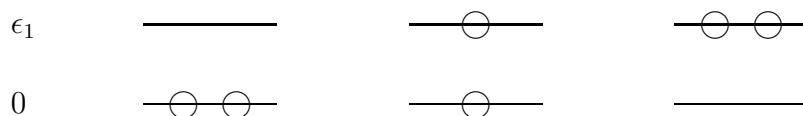


Fig. 1. The three states of an $N = 2$ two-level system

Evaluating the total partition function Q directly we have:

$$Q_2 = 1 + \exp(-\epsilon/kT) + \exp(-2\epsilon/kT) \quad (6)$$

which is clearly different from Eq.(5). The second expression is the correct one, whereas the first is in error for those states in which the energy levels are multiply occupied. It is intuitively clear that when $q \gg N$, so that there are many more levels available than there are particles, then the $N!$ correction to q^N is indeed a very good approximation to the true partition function. In this case we say that the particles obey ‘Boltzon’ statistics.

The usefulness of Q arises from the fact that it determines the Helmholtz free energy via the bridge relation:

$$\boxed{A = -kT \ln Q} \quad (7)$$

Once the Helmholtz Free Energy is known, many thermodynamic properties become computable. Some relations which you should be familiar with are given below. Thus

$$S = -(\partial A/\partial T)_{N,V} \quad \text{Entropy} \quad (8)$$

$$P = -(\partial A/\partial V)_{N,T} \quad \text{Pressure} \quad (9)$$

$$\mu = (\partial A/\partial N)_{V,T} \quad \text{Chemical Potential} \quad (10)$$

We later show the key relation, Eq.(7), remains valid under all circumstances for systems in equilibrium. Nevertheless, the above apparatus for computing Q is useful only when the following conditions are met: (a) The particles are independent (i.e. non-interacting); (b) the number of accessible states greatly exceeds the number of particles (i.e. $q \gg N$, so that the average occupancy of *any* given level is much less than 1).

These two conditions are rather restrictive. First, it hardly needs stating that real particles do interact with each other; real matter is not an ideal gas! Thus the first

condition is already highly restrictive. It does not provide a reasonable description even of moderately dense gases, let alone liquids, solids, and other interesting and common forms of matter. To be able to treat these, even at an approximate level, will require an extension of the formalism to interacting systems. However, even if the particles are non-interacting, the formalism still fails under conditions of low temperatures, high pressures, or for relatively small-mass particles such as electrons and light nuclei.

A few examples will illustrate under which conditions this becomes necessary. Consider for example the translational partition function. The translational energy levels for a particle of mass m in a box of $L^3 = V$ are:

$$\epsilon_{trans}(n_x, n_y, n_z) = \frac{(n_x^2 + n_y^2 + n_z^2)h^2}{8mL^2} \quad (11)$$

In Dr Keeler's course, the translational partition function q_{trans} was evaluated to give [make sure you can derive this expression!]:

$$q_{trans} = \left(\frac{2\pi mkT}{h^2} \right)^{3/2} V \quad (12)$$

Substituting numbers corresponding to helium-4 we obtain $m = 4 \times 1.66 \times 10^{-27}$ kg, $T = 300\text{K}$, $V = L^3 = 1 \text{ cm}^3$, we find

$$q_{trans} = 7.9 \times 10^{24}$$

i.e. a He atom moving unhindered in a cm^3 at 300K has about 10^{25} translational states effectively accessible to it. This enormous number arises because the gap between the translational energy levels are on the order of $h^2/8mL^2 \sim 10^{-37}\text{J}$, whereas $kT \sim 10^{-21}\text{J}$. At room temperature, kT greatly exceeds the spacings between the translational energy levels. On the other hand, at atmospheric pressure ($P = 10^5 \text{ Pa}$) a cm^3 holds $N = PV/kT \sim 2 \times 10^{19}$ particles. Thus the average occupancy of any given translational energy level is on the order of 10^{-5} , and it is very unlikely that any two states have an occupancy exceeding 1. However, if we compress the system by decreasing L by a factor of 100, increasing P to 1 Mbar, or cooling the system by a factor of 100, then we would indeed approach conditions in which $q/N \sim 1$. Quantum effects become relevant at high pressures or low temperatures. Indeed, a more illuminating form

to express q_{trans} which gives a good criterion for when ‘quantum statistics’ becomes necessary is in terms of the thermal de Broglie wavelength, which is defined as:

$$\Lambda = \frac{h}{(2\pi mkT)^{1/2}} \quad [\text{thermal de Broglie wavelength}] \quad (13)$$

from which it follows:

$$q_{trans} = \frac{V}{\Lambda^3} \quad (14)$$

$$Q = \frac{q^N}{N!} \quad (15)$$

$$\approx \left(\frac{Ve}{N\Lambda^3} \right)^N \quad (16)$$

It can be seen that strange things begin to happen when $\Lambda^3 \sim V/N$, i.e. the volume occupied per particle becomes comparable to the volume associated with the quantum mechanical uncertainty of its position as measured by the thermal de Broglie wavelength. For example, according to this formula $Q \rightarrow 0$ as $\Lambda^3 \rightarrow \infty$ (i.e. as the temperature goes to zero, or the mass of the particles go to zero!). What is one to make of $Q = 0$?!

A second case considered in Dr Keeler’s course was *vibrational* partition function for a harmonic oscillator of frequency ν . We know that:

$$\epsilon_{vib}(n) = \left(n + \frac{1}{2}\right)h\nu \quad (17)$$

Substituting in Eq.(1) and performing the sum we obtain [Show!]:

$$q_{vib} = \frac{e^{-h\nu/2kT}}{1 - e^{-h\nu/kT}} \quad (18)$$

Taking now the limit $T \rightarrow 0$ we see that $q \rightarrow 0$. Wait a minute! How can $q = 0$ you ask? This is arising because we didn’t measure our zero of energy from the lowest available state (i.e. $h\nu/2$). If we do, we obtain instead:

$$q_{vib} = \frac{1}{1 - e^{-h\nu/kT}} \quad (\text{removing the ZPE}) \quad (19)$$

and you think all is saved. But no! You see that as soon as $T \ll h\nu/k$, we have $q \rightarrow 1$, and we do not satisfy the condition $q > N$. If we use the formula:

$$A_{vib} = -kT \ln(q_{vib}^N/N!)$$

$$S_{vib} = - \left(\frac{\partial A_{vib}}{\partial T} \right)_{N,V}$$

we obtain:

$$S_{vib} = -Nk \ln(1 - e^{-h\nu/kT}) + U_{vib}/T - Nk(\ln N - 1)$$

$$U_{vib} = \frac{Nh\nu}{e^{h\nu/kT} - 1}$$

and in the limit $T \rightarrow 0$, we have:

$$S_{vib} = -Nk(\ln N - 1) \tag{20}$$

which is clearly nonsense (a negative number!). You can easily see that the $N!$ in the expression for A is what gives rise to the trouble. In Dr Keeler's course, this issue was side-stepped by saying that we can treat the molecular vibrations as distinguishable particles, and thus neglect the $N!$ term. This can be justified in the case of molecular vibrations, in which each normal mode is localised on an individual molecule (not over the whole ensemble of molecules). However, it is not justified in the case of lattice vibrational modes of crystals, where each normal mode is delocalised over the whole crystal; a second very important case is that of photons of light (blackbody radiation), where once again the modes are delocalised over a whole system. For these, we need to treat the proper quantum statistics (Bose-Einstein).

The aim of this course is three-fold: (a) Generalisation of the above formalism to *interacting* particles, (b) the proper description of quantum statistics, (c) establishing a basis to do statistical mechanics with *classical* systems in which energy levels are *not* quantized. You can ask, will this latter be useful? The answer is that, for many problems of interacting particles (such as dense gases, liquids, and solids), classical statistical mechanics suffices to give an excellent treatment of the equilibrium behaviour of systems. We have already seen that the translational partition function is quasi-classical. It remains so when in the limit of strong interactions. The quantum-interacting problem which is necessary to treat correlated electron systems, and some exotic forms of matter, is outside the scope of the present course.

II. GENERAL DESCRIPTION OF EQUILIBRIUM STATISTICAL MECHANICS

The key principle of statistical mechanics can be stated as follows:

If a system can exist in one of M states with energy E_i , $i = 1, \dots, M$, then at equilibrium at temperature T , the probability that the system will be found in a particular state i is:

$$P_i = \frac{\exp(-E_i/kT)}{Q} \quad (21)$$

where

$$Q = \sum_{i=1}^M \exp(-E_i/kT), \quad (22)$$

Furthermore, the expectation value of a quantum-mechanical operator \hat{A} , corresponding to an observable A , is at equilibrium given by:

$$\langle A \rangle = \frac{\sum_i A_i \exp(-E_i/kT)}{Q} \quad (23)$$

where

$$A_i = \langle i | \hat{A} | i \rangle \quad (24)$$

is the expectation value of \hat{A} in the (normalised) state $|i\rangle$. The states $|i\rangle$ are solutions to the time-independent Schrödinger equation:

$$\hat{H}|i\rangle = E_i|i\rangle \quad (25)$$

where \hat{H} is the quantum-mechanical Hamiltonian operator which defines the system of interest. It follows that:

$$E_i = \langle i | \hat{H} | i \rangle \quad (26)$$

We will explore the many consequences of this hypothesis in the coming lectures, but we start by deriving this law by considering a ‘toy’ model.

A. Harmonic system+bath

Consider a system (S) and bath (B), consisting of $N^{(S)}$ and $N^{(B)}$ particles respectively. S and B can exchange energy ('heat'). Let $E^{(S)}$ ($E^{(B)}$) be the instantaneous energy of the system (bath), so that

$$E^{(S)} + E^{(B)} = E \quad (27)$$

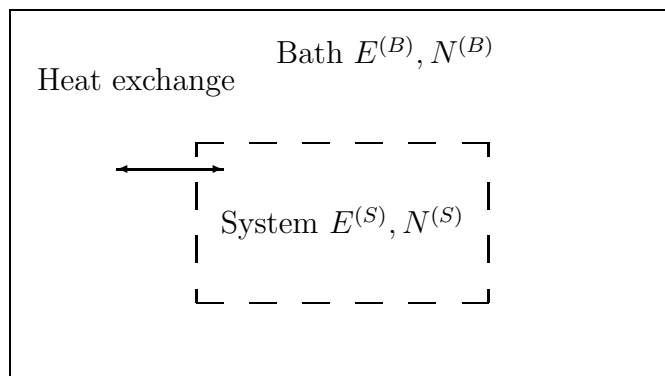
is fixed. (S and B together form an isolated system which does not exchange heat with the rest of the Universe). In our toy model, we will assume that the particles of S and B are separately immersed in a harmonic potential, of characteristic frequency ω . The one-particle energy levels are given by:

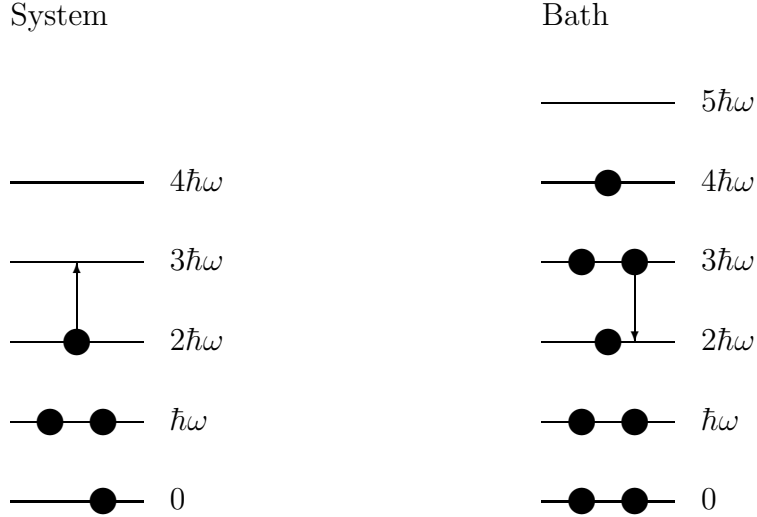
$$\epsilon_n = n\hbar\omega, \quad n = 0, 1, 2, \dots \quad (28)$$

where we are measuring the energy from the ground state $\hbar\omega/2$. Furthermore, we will assume that:

$$N^{(B)} \gg N^{(S)}, \quad E^{(B)} \gg E^{(S)}, \quad (29)$$

i.e. that the bath is much larger (both in terms of particle number and energy) than our system.





$$N^{(S)} = 4, E^{(S)} = 4\hbar\omega$$

$$N^{(B)} = 8, E^{(B)} = 14\hbar\omega$$

Fig. 2. System and Bath are loosely coupled so that they exchange heat. A particle in S gaining one quantum of energy is compensated by a corresponding drop in B .

We can specify the state of a system with N particles by giving a list of N integers (n_1, n_2, \dots, n_N) specifying the level which each particle occupies. For example, for the situation depicted in the Figure, the states of the system (and bath) could be described with:

$$\text{System: } (0, 1, 1, 2) \quad \text{Bath: } (0, 0, 1, 1, 2, 3, 3, 4).$$

In the analysis that follows, we are going to assume that the particles are distinguishable. (This means, for example, that the state $(0, 1, 1, 2)$ is distinguishable from $(1, 0, 1, 2)$. The case for indistinguishable particles will be dealt with later on in the course).

With this notation, the energy of a system is given by:

$$E(n_1, n_2, \dots, n_N) = \sum_{i=1}^N n_i \hbar\omega \tag{30}$$

The question we would like to ask is: *what is the probability that the system is in a particular microstate*, which for shorthand we shall call j , remembering that it is defined by the set of N integers (n_1, n_2, \dots, n_N) .

Let $\Omega(E, N)$ be the number of microstates of a system consisting of N particles with total energy E . For the present harmonic system, which consists of equally spaced single-particle levels, this is fairly easy to compute. First, note that what determines the total energy is the *total number of quanta*, M . i.e.

$$E/(\hbar\omega) = \sum_{i=1}^N n_i = M \quad (31)$$

not the particular distribution of quanta among the particles. Therefore, what we are essentially asking is: in how many ways can we sum up N integers to give a value M ? (in other words how many ways are there to partition M into N integers including zero?)

A moment's thought shows this to be:

$$\Omega(E, N) = \frac{(M + N - 1)!}{M!(N - 1)!} \quad (32)$$

This can be seen as follows. We have M identical quanta to be distributed among N particles. The particles can be imagined to form $N - 1$ partitions, which divide the M quanta. For example for the system in the Figure, we have $M = 4$ and $N = 4$. Therefore we imagine $N - 1 = 3$ walls (designated by a |) which partition the 4 quanta:

particle number: 1 2 3 4
 |q|q|qq

The total number of 'objects' (i.e. q's and |'s) is $M + N - 1$, which can be ordered in $(M + N - 1)!$ ways. However, the permutations of the quanta are all identical, as are those of the walls. Thus we divide by $M!$ and $(N - 1)!$ to obtain the result already stated, Eq.(32). Take a moment to verify explicitly that for the situation shown in the Figure:

$$\Omega(4, 4) = \frac{7!}{4!3!} = 35 \quad (33)$$

Returning to our system+bath example, let us assume that S is in a particular microstate (j) with energy $E^{(S)} = E_j$. Then $E^{(B)} = E - E_j$, and the total number of ways B can exist in is given by $\Omega(E - E_j, N^{(B)}) \equiv \Omega_B(E - E_j)$. Since we are assuming

the system to be in a particular state j , it follows that the number of ways the the complex 'S+B' can exist in, given this condition, is:

$$\Omega_{B+S}(E|j) = \Omega_B(E - E_j) \times 1 \quad (34)$$

Furthermore, the total number of ways S+B can exist, irrespective of the state of the system, is:

$$\Omega_{B+S}(E) = \sum_j \Omega_{B+S}(E|j) \quad (35)$$

$$= \sum_j \Omega_B(E - E_j) \quad (36)$$

where the sum over j is over all microstates of S, and is of course independent of E_j .

The *probability* P_j of observing S in microstate j is given by the ratio:

$$P_j = \frac{\Omega_{B+S}(E|j)}{\Omega_{B+S}(E)} \quad (37)$$

in other words, because of Eq.(34), the probability to observe the system in a particular microstate with energy E_j is proportional to $\Omega_B(E - E_j)$:

$$P(E_j) \propto \Omega_B(E - E_j) \quad (38)$$

Pause for a moment to digest this, since it is the key point of statistical mechanics. It is saying that the probability to observe a system, in equilibrium with a heat bath, in a given state, depends on the *entropy of the heat bath!* If energy flows in to the system, thereby increasing its net energy content, it does so at the expense of the energy of the heat bath. This is not forbidden by physical law. However, the decrease in energy of the heat bath, results in a decrease in the entropy of the heat bath: $\Omega_B(E^{(B)})$ very rapidly decreases (indeed exponentially) as $E^{(B)}$ decreases. Conversely, if the system gives up energy to the bath, the number of microstates of the bath goes up, and hence its entropy goes up. Thus, on purely statistical grounds, the higher the energy of a particular state of the system, the less likely the system will be found in it; and *vice versa*, the lower the energy of a given state, the more likely the system is to be found in it.

Developing Eq. (38) further, we use Stirling's approximation

$$\ln N! \approx N \ln N - N \quad \text{for large } N \quad (39)$$

to expand $\ln \Omega_B$ to obtain [show!]

$$\begin{aligned} \ln[\Omega_B(E)] &= (M + N^{(B)} - 1) \ln(M + N^{(B)} - 1) \\ &\quad - M \ln M - (N^{(B)} - 1) \ln(N^{(B)} - 1) \\ &\quad [M = E/\hbar\omega] \end{aligned}$$

In order to perform a Taylor expansion:

$$\ln[\Omega_B(E - E_j)] = \ln \Omega_B(E) - E_j \frac{\partial \ln \Omega_B}{\partial E} + \dots \quad (40)$$

we need [show!]:

$$\begin{aligned} \frac{\partial \ln \Omega_B}{\partial E} &= \frac{1}{\hbar\omega} \left(\ln(E/\hbar\omega + N^{(B)} - 1) - \ln(E/\hbar\omega) \right) \\ &= \frac{1}{\hbar\omega} \ln \left(1 + \frac{N^{(B)}\hbar\omega}{E} - \frac{\hbar\omega}{E} \right) \\ &= \frac{1}{\hbar\omega} \ln \left(1 + \frac{\hbar\omega}{\bar{\epsilon}} \right) \end{aligned} \quad (41)$$

where in the last line $\bar{\epsilon} = E/N^{(B)}$ is the energy per particle in the bath, and we have taken the thermodynamic limit:

$$E \rightarrow \infty, N^{(B)} \rightarrow \infty, \frac{E}{N^{(B)}} = \bar{\epsilon} \quad (42)$$

Thus, in the thermodynamic limit of the bath, the Taylor expansion can be rigorously truncated at the first-order term, i.e.:

$$\ln[\Omega_B(E - E_j)] = \ln[\Omega_B(E)] - E_j \frac{\partial \ln \Omega_B}{\partial E} \quad (43)$$

$$= \ln[\Omega_B(E)] - E_j \times \beta \quad (44)$$

where

$$\beta = \frac{1}{\hbar\omega} \ln \left(1 + \frac{\hbar\omega}{\bar{\epsilon}} \right) \quad (45)$$

is independent of E . Taking exponentials, we have:

$$\Omega_B(E - E_j) \propto \Omega_B(E) \times \exp(-\beta E_j) \quad (46)$$

The first factor in the product above is constant in relation to E_j . It can be absorbed into the constant of proportionality. Thus, using Eq.(38) we arrive at:

$$P(E_j) \propto \exp(-\beta E_j) \quad (47)$$

where the constant of proportionality can be determined by the condition:

$$\sum_j P(E_j) = 1 \quad (48)$$

i.e.

$$P(E_j) = \frac{\exp(-\beta E_j)}{\sum_j \exp(-\beta E_j)} \quad (49)$$

In the specific example above, β has an explicit meaning in terms of the oscillator energy, viz. Eq.(45), which is close to $1/(\hbar\omega)$. More generally, the value of β is determined via:

$$\beta = \left(\frac{\partial \ln \Omega_B}{\partial E} \right)_{N(B)} \quad (50)$$

Now, you will recall the Boltzmann definition of entropy:

$$S = k \ln \Omega \quad [\text{Boltzmann entropy}] \quad (51)$$

i.e.

$$\beta = \left(\frac{\partial(S/k)}{\partial E} \right)_{N(B)} \quad (52)$$

But from thermodynamics, we know that

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N,V} \quad (53)$$

In the present system we have assumed the volume to be fixed. Therefore we arrive at the general identification:

$$\beta = \frac{1}{kT} \quad (54)$$

A point to note is that in deriving the Boltzmann distribution, Eq.(46), the assumption was made that the Taylor expansion could be truncated at the first order term. This implies that β should be at best only slowly varying with E . In practice, this behaviour is true for systems which are sufficiently large.

As a corollary to the above, if all microstates of the system are equal in energy, it follows that the probability to observe the system in a particular state j is

$$P_j = \frac{1}{\Omega(E)} \quad (55)$$

where E is the energy of the system. Systems that are isolated (i.e. thermally insulated from the rest of the Universe), have their energy fixed. It therefore follows that at equilibrium, all states of an isolated system occur with equal probability.

<p>Isolated System</p> $E, N, V, \rightarrow \Omega(E, N, V)$ $S(E, N, V) = k \ln \Omega(E, N, V)$ $P_j = 1/\Omega(E, N, V)$
--

III. THE CANONICAL ENSEMBLE

We now discuss a more general and formal way to construct probability distributions which take into account a variety of macroscopic constraints which we, the experimenter, can impose on the system. None of the physics changes in essence from the previous discussion; maximizing entropy remains a fundamental postulate.

A commonly encountered situation is a system consisting of a fixed number of particles N , fixed volume V , and in contact with a heat-bath maintained at a temperature T . The macroscopic constraints, therefore, are N, V, T .

The system exchanges energy with the heat bath. Therefore, its energy is not fixed,

but fluctuates about an expected value, $\langle E \rangle$. We would like to know what this expected value is and what its fluctuations are.

Consistent with the three macroscopic constraints, there are, in general, a vast number of compatible microstates. The system, as it evolves in time, will change its microstate from one to another. Let us denote by $|s(t)\rangle$ the microstate in which the system is in at time t . In general $|s(t)\rangle$ will be given by the solution of microscopic equation such as the time-dependent Schrödinger equation:

$$-i\hbar \frac{\partial}{\partial t} |s(t)\rangle = \hat{H} |s(t)\rangle \quad (56)$$

In the equation above, \hat{H} is the Hamiltonian operator for the system of N particles.

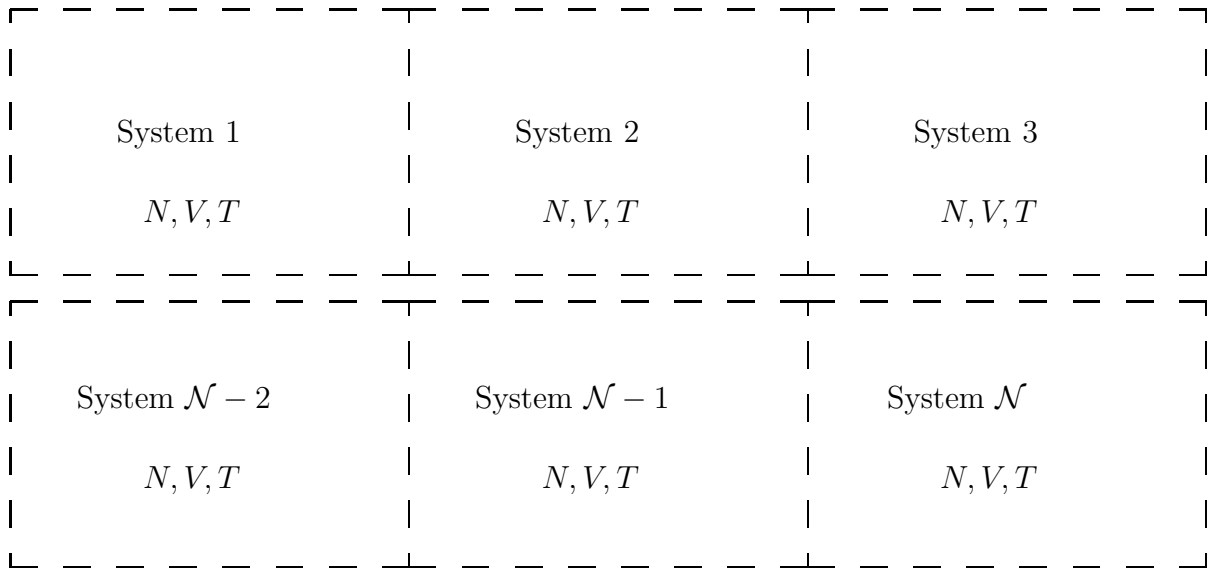
The value of any observable \hat{A} will therefore change with time, as $A(t) = \langle s(t) | \hat{A} | s(t) \rangle$. The measured, time-averaged, value of any observable is then:

$$\bar{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt A(t) \quad [\text{time-average}] \quad (57)$$

where T is the total time over which we observe the system, which in principle should be extended to infinity. Performing such a time-average is in general not feasible for quantum systems; in classical systems, which we will later discuss, it can be done for systems consisting of several hundreds or thousands of particles, but not for macroscopic sizes; and even then only for finite periods of time, extending to a few nano-seconds for systems containing atom-scale particles.

Fortunately, there is an alternative way to compute equilibrium expectation values through a process called ‘ensemble averaging’. We imagine the following mental construction: suppose we replicate the system a vast number, \mathcal{N} times, each replica having the same values of N, V, T , but differing in the precise microstate. This ‘ensemble’ of systems is assumed to be so large in number that any given microstate will occur a number of times taken over the ensemble. The question ‘what is the probability that a given system chosen at random is in a particular microstate’, now can be given a definite meaning: if the microstate j occurs \mathcal{N}_j times in the ensemble, then

$$P_j = \frac{\mathcal{N}_j}{\mathcal{N}} \quad (58)$$



A canonical ensemble of \mathcal{N} systems.

An ensemble average of an observable A then has the form:

$$\begin{aligned}
 \langle A \rangle &= \frac{1}{\mathcal{N}} \sum_j \mathcal{N}_j A_j \\
 &= \sum_j P_j A_j
 \end{aligned}
 \tag{59}$$

It is assumed (the ‘ergodic hypothesis’) that the time-averaged value of A and its ensemble averaged value:

$$\bar{A} = \langle A \rangle \quad [\text{Ergodic Hypothesis}]
 \tag{60}$$

In general, it is very difficult to prove rigorously (mathematically) that the two types of averaging produce the same result. The difficulty lies in the fact that the time evolution of a system may not explore all of the microstates available to the system, but could get trapped in a relatively small number of them. Such ‘broken ergodicity’ systems are known to occur in nature (e.g. glasses); we can assume them to be out-of-equilibrium. They present special cases that are difficult to treat. For our purposes we will simply assume that the ergodic hypothesis holds true.

Returning to the discussion of the P_j , we wish to find out what the set of $\{P_j\}$ is for all states of the system.

Let us define an *entropy function*:

$$S[P_j] = -k \sum_j P_j \ln P_j \quad (61)$$

and make the hypothesis that the best set of $\{P_j\}$ is the one which maximizes S . In attempting to maximize S defined above, we must remember that there are two constraints to be satisfied:

$$\sum_j P_j = 1 \quad [\text{Normalisation}] \quad (62)$$

$$\sum_j E_j P_j = \langle E \rangle \quad (63)$$

These constraints can be satisfied if we introduce the Lagrange multipliers $k\lambda$ and $k\beta$ (the Boltzmann constant k is something of a nuisance here, but we include it to keep with the convention in the field), and maximize an unconstrained function:i.e.

$$\frac{\partial}{\partial P_n} \left(S[P_j] - k\lambda \left(\sum_j P_j - 1 \right) - k\beta \left(\sum_j E_j P_j - \langle E \rangle \right) \right) = 0 \quad (64)$$

for all states n . We obtain:

$$-\ln P_n - 1 - \lambda - \beta E_n = 0 \quad (65)$$

i.e.

$$P_n = \exp(-\beta E_n) \times \exp(-1 - \lambda) \quad (66)$$

But this is precisely the Boltzmann distribution we (laboriously) obtained in Section 2! The normalization condition determines the value of λ :

$$\lambda + 1 = \ln \left[\sum_n \exp(-\beta E_n) \right] \quad (67)$$

$$= \ln Q_N \quad (68)$$

where

$$Q_N(V, T) = \sum_n \exp(-\beta E_n). \quad (69)$$

It is left as an exercise to show that $k\beta$ can be identified as $1/T$. Hint: consider the derivative

$$\left(\frac{\partial S}{\partial \langle E \rangle} \right)_{N, V} \quad (70)$$

which can be shown to equal to $k\beta$ by direct substitution into Eq.(61) [see problem sheet]. But from thermodynamics, we know that

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N, V} \quad (71)$$

Therefore, if we identify $\langle E \rangle$, i.e.the expected value of the total energy of the system, with E , the thermodynamic internal energy, then

$$\beta = \frac{1}{kT}. \quad (72)$$

Thus, the assumption that the entropy is given by Eq.(61) leads to the Boltzmann distribution for P_j :

$$P_j = \frac{\exp(-\beta E_j)}{Q_N} \quad (73)$$

Substituting this back into Eq.(61) we obtain:

$$S/k = - \sum_j P_j \ln P_j \quad (74)$$

$$= - \sum_j \frac{\exp(-\beta E_j)}{Q_N} \times (-\beta E_j - \ln Q_N) \quad (75)$$

$$= \beta \sum_j E_j \frac{\exp(-\beta E_j)}{Q_N} + \ln Q_N \quad (76)$$

$$= \beta \langle E \rangle + \ln Q_N \quad (77)$$

Rearranging this we obtain:

$$-kT \ln Q_N = \langle E \rangle - TS \quad (78)$$

From thermodynamics we know that $E - TS$ is the Helmholtz Free energy, we thus obtain the key bridge relation relating the partition function and the macroscopic world:

$$A(N, V, T) = -kT \ln Q_N. \quad (79)$$

The internal energy $\langle E \rangle$ can be obtained from:

$$\langle E \rangle = \sum_j P_j E_j \quad (80)$$

$$= \frac{1}{Q_N} \sum_j E_j \exp(-\beta E_j) \quad (81)$$

Note that:

$$\langle E \rangle = \sum_j P_j E_j \quad (82)$$

$$= -\frac{1}{Q_N} \frac{\partial \sum_j \exp(-\beta E_j)}{\partial \beta} \quad (83)$$

$$= -\frac{1}{Q_N} \frac{\partial Q_N}{\partial \beta} \quad (84)$$

$$= -\frac{\partial \ln Q_N}{\partial \beta} \quad (85)$$

$$= \frac{\partial(\beta A)}{\partial \beta}. \quad (86)$$

where it is understood that N, V are kept constant in the partial derivative above.

We can similarly compute the fluctuations in energy,

$$(\Delta E)^2 = \langle E^2 \rangle - \langle E \rangle^2 \quad (87)$$

It is left as an exercise to show that

$$-\left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_{N,V} = \langle E^2 \rangle - \langle E \rangle^2. \quad (88)$$

But we know that:

$$\left(\frac{\partial \langle E \rangle}{\partial \beta}\right)_{N,V} = -kT^2 \left(\frac{\partial \langle E \rangle}{\partial T}\right)_{N,V} \quad (89)$$

$$= -kT^2 C_V \quad (90)$$

where we have used the definition of C_V from thermodynamics. This proves the stability criterion that, for a system in thermal equilibrium, $C_V > 0$. The relative fluctuation is given by:

$$\frac{\Delta E}{\langle E \rangle} = \frac{\sqrt{kT^2 C_V}}{E} \quad (91)$$

Since both C_V and E are extensive quantities (i.e. proportional to N), the relative fluctuation is $O(1/\sqrt{N})$, hence negligible when $N \sim 6 \times 10^{23}$. The result Eq.(91) shows that the energy distribution is very narrow for macroscopic systems. The internal energy of a system in a canonical ensemble has negligible probability to deviate significantly from the mean value $E = \langle E \rangle$.

A. An example: spin systems

Para/ferromagnet crystals are often modeled as spin systems. In the simplest case, the magnetic moment μ_B on an atom in a crystalline lattice is represented by a spin variable s_i , which can take values ± 1 , being aligned in parallel or antiparallel to an external magnetic field B . A system of N such spins has 2^N possible configurations.

Let us assume that the spins are non-interacting (i.e. do not interact with each other; they do however interact with the external field). The energy of any spin, therefore, is $\pm\mu_B B$, depending on whether it is parallel ($-\mu_B B$) or antiparallel ($\mu_B B$) to the external field. Therefore, the total energy of an array of N spins of configuration $(s_1, s_2, \dots, s_N) \equiv (s^N)$ is:

$$E(s^N) = \sum_{i=1}^N -B\mu_B \times s_i \quad (92)$$

The canonical partition function is:

$$Q_N = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} \exp(-\beta E(s^N)) \quad (93)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} \prod_i \exp(+\beta\mu_B B s_i) \quad (94)$$

Note that we can factorize each sum (this is a consequence of the spins being non-interacting), with the result

$$Q_N = [\exp(-\beta\mu_B B) + \exp(+\beta\mu_B B)]^N \quad (95)$$

$$= [2 \cosh(\beta\mu_B B)]^N \quad (96)$$

Therefore the Helmholtz energy is:

$$A = -kT \ln Q_N \quad (97)$$

$$= -NkT \ln[2 \cosh(\beta\mu_B B)] \quad (98)$$

Thermodynamic quantities such as E follow easily:

$$E = \left(\frac{\partial(\beta A)}{\partial\beta} \right)_{N,V} \quad (99)$$

$$= \frac{-N\mu_B B}{\cosh \beta\mu_B B} \times \sinh(\beta\mu_B B) \quad (100)$$

$$= -N\mu_B B \tanh(\beta\mu_B B) \quad (101)$$

The specific heat C_V is

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{N,V} = -\frac{1}{kT^2} \times \left(\frac{\partial E}{\partial\beta} \right)_{N,V} \quad (102)$$

$$= -N\mu_B B \operatorname{sech}^2(\beta\mu_B B) \times \frac{-\mu_B B}{kT^2} \quad (103)$$

$$= \frac{N(\mu_B B)^2}{kT^2} \frac{1}{\cosh^2(\beta\mu_B B)} \quad (104)$$

It is instructive to make a plot of E and C_V .

The net magnetization for a given configuration of spins is:

$$M(s^N) = \mu_B \sum_i s_i \quad (105)$$

The expected value can be computed directly by noting $E = -BM$

$$\langle M \rangle = \frac{1}{Q_N} \sum_{s_1} \dots \sum_{s_N} M \exp(+\beta BM) \quad (106)$$

$$= \frac{1}{Q_N} \frac{\partial}{\partial B} \sum_{s_1} \dots \sum_{s_N} \exp(-\beta BM) \times \frac{1}{\beta} \quad (107)$$

$$= \frac{1}{Q_N} \frac{\partial Q_N}{\partial B} \times \frac{1}{\beta} \quad (108)$$

$$= -\frac{\partial A}{\partial B} \quad (109)$$

$$= N\mu_B \tanh(\beta\mu_B B) \quad (110)$$

In other words, at any non-zero T , the net magnetization is zero if $B = 0$. This model cannot be ferromagnet (i.e. acquire a spontaneous magnetization in the absence of an external B field). For that, we will need to include spin-spin interactions (via mean-field theory. More on that later in the course).

The susceptibility can be computed via:

$$\chi = \frac{\partial \langle M \rangle}{\partial B} \tag{111}$$

$$= \frac{N\mu_B^2\beta}{\cosh^2(\beta\mu_B B)} \tag{112}$$

In the limit that $B \rightarrow 0$, we obtain $\chi_0 = N\beta\mu_B^2 \sim \frac{1}{T}$ which is Curie's law. For very large fields one obtains:

$$\cosh(\beta\mu_B B) \approx \frac{1}{2} \exp(\beta\mu_B B) \quad \text{for } \beta\mu_B B \gg 1 \tag{113}$$

i.e.

$$\lim_{B \rightarrow \infty} \chi = N\mu_B^2\beta \times 2^2 \times \exp(-2\beta\mu_B B) \tag{114}$$

$$= 0 \tag{115}$$

IV. CLASSICAL STATISTICAL MECHANICS

A. The Microcanonical System

The emphasis on systems which can exist as discrete states $|i\rangle$, with discrete energy levels E_i which are the solutions to an underlying Schrödinger equation

$$\hat{H}|i\rangle = E_i|i\rangle \tag{116}$$

is convenient when developing the theoretical basis of Statistical Mechanics, but in practice is of fairly limited use. Instead, many systems of real interest can be described to high accuracy using classical mechanics, and it is very important to see how the machinery we have so far developed can be adapted to the classical systems, in which the states and energy levels are no longer discrete.

In classical mechanics, the ‘state’ of a system of N particles is described by specifying the instantaneous positions and momenta of each particle. This is a $6N$ -dimensional space:

$$\Gamma : (\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N) \equiv (\mathbf{r}^N, \mathbf{p}^N) \quad (117)$$

usually called ‘phase space’. The position vectors of the particles, \mathbf{r}_i , will be restricted to the volume of the system, whereas the momentum variables, \mathbf{p}_i , will be unrestricted, and each component of which can range from $-\infty$ to ∞ .

The energy of the system at the point $(\mathbf{r}^N, \mathbf{p}^N)$ is given by the Hamiltonian function:

$$H_N(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^N \frac{p_i^2}{2m_i} + V(\mathbf{r}^N) \quad (118)$$

where m_i is the mass of particle i , and the potential energy function V is assumed to depend only the positions (not momenta) of the particles.

An isolated classical system evolves in time via Newton’s equations of motion:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad (119)$$

This is a second-order differential equation whose solution would give the trajectory followed by each particle as a function of time $\mathbf{r}_i(t)$. The instantaneous momenta can be obtained from

$$\mathbf{p}_i = m_i \mathbf{v}_i = m_i \dot{\mathbf{r}}_i \quad (120)$$

It can be easily shown that for such a system, the energy is conserved:

$$\frac{dH_N}{dt} = 0 \quad (121)$$

which expresses the fact that the total energy of an isolated system is constant (time invariant):

$$H_N(\mathbf{r}^N, \mathbf{p}^N) = E \quad (122)$$

This implies that during the course of its evolution, an isolated system can, at best, visit only those points in phase space which have the same energy. In other words, the

accessible phase space is a constant energy hyper-surface, of dimension $6N - 1$. Such isolated systems, i.e. (N, V, E) fixed are called *microcanonical*.

In just the same way as in quantum mechanics we perform a ‘sum over states’, the corresponding operation, classically, is a integration over phase-space. A volume element in phase space is given by $d\mathbf{r}^N d\mathbf{p}^N$, and integrations performed over phase space are $6N$ dimensional integrals:

$$\sum_j \rightarrow \int_{\Gamma} d\mathbf{r}^N d\mathbf{p}^N$$

Note, however, that the dimension of the integral is (position \times momentum) 3N , (i.e. action 3N) whereas the quantum summation gives just a number. The two can be brought into resolution by recalling that Quantum Mechanics, through the Heisenberg Uncertainty Principle limits, the accuracy to which we can measure simultaneously the position and momentum of a particle. There is always an uncertainty on the order of

$$\Delta x \Delta p_x \sim h \quad (123)$$

for each Cartesian degree of freedom. Therefore, there exists a ‘graininess’ to phase space, a fundamental volume element, which is on the order h^{3N} .

$$\sum_j \rightarrow \frac{1}{h^{3N}} \int_{\Gamma} d\mathbf{r}^N d\mathbf{p}^N \quad [\text{Integration over phase space}] \quad (124)$$

Similarly, in just the same way that, for a quantum system, we defined $\Omega(E)$ to be the number of (discrete) states of an N particle system which have the same energy, we shall define the ‘volume’ of the accessible phase space to be its corresponding classical counterpart:

$$W(E) = \frac{1}{h^{3N}} \int_{\Gamma} d\mathbf{r}^N d\mathbf{p}^N \delta(H_N - E) \quad (125)$$

In the above, the function $\delta(x)$ is a Dirac Delta function, which is defined to have the properties:

$$\int_{\Gamma} dx \delta(x) = 1 \quad (126)$$

$$\int_{\Gamma} dx f(x) \delta(x - x_0) = f(x_0) \quad (127)$$

The job of the Dirac Delta function is to select out of all points in phase space those for which $H_N = E$. Unfortunately, integrals of the type Eq.(125) are extremely difficult to handle, and we will not elaborate much about it.

A further point to note is that phase-points which differ only by a permutation of particle labels must be considered identical. Since there are $N!$ permutations for a system with N particles, we arrive at

$$W(E) = \frac{1}{N!h^{3N}} \int_{\Gamma} d\mathbf{r}^N d\mathbf{p}^N \delta(H_N - E) \quad [\text{Indistinguishable}] \quad (128)$$

The entropy of the microcanonical system is given by the Boltzmann equation:

$$S(E, V, N) = k \ln W \quad (129)$$

B. The Classical Canonical Ensemble

It is much more convenient to deal with classical systems in the Canonical Ensemble, i.e. those which are not isolated, but in contact with a heat bath (albeit N, V are still fixed). The Boltzmann distribution continues to hold for classical systems. Except that it now refers to a probability density. Specifically, the probability to find the system in a volume element $d\mathbf{r}^N d\mathbf{p}^N$ around the phase point $(\mathbf{r}^N, \mathbf{p}^N)$ is given by:

$$P(\mathbf{r}^N, \mathbf{p}^N) d\mathbf{r}^N d\mathbf{p}^N = \frac{1}{N!h^{3N}} \frac{\exp(-\beta H_N)}{Q_N} d\mathbf{r}^N d\mathbf{p}^N \quad (130)$$

where the partition function is now an integral over the $6N$ dimensional phase-space of the system:

$$Q_N(V, T) = \frac{1}{N!h^{3N}} \int_{\Gamma} d\mathbf{r}^N d\mathbf{p}^N \exp(-\beta H_N) \quad (131)$$

As before, the $N!$ has been included in the denominator of Eq.(130) to account for the indistinguishability of the particles: phase points which differ only in a permutation of particle labels are indistinguishable, and must be counted only once in the partition sum.

The Helmholtz energy can be computed from the bridge relation, Eq.(7), from which the thermodynamic properties can be determined. E.g

$$S = - \left(\frac{\partial A}{\partial T} \right)_{N,V} \quad (132)$$

$$P = - \left(\frac{\partial A}{\partial V} \right)_{N,T} \quad (133)$$

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{V,T} \quad (134)$$

$$(135)$$

The expectation value of any function of the phase-space coordinates, $A(\mathbf{r}^N, \mathbf{p}^N)$ can be computed analogously to the quantum systems:

$$\langle A \rangle = \int_{\Gamma} P(\mathbf{r}^N, \mathbf{p}^N) A(\mathbf{r}^N, \mathbf{p}^N) d\mathbf{r}^N d\mathbf{p}^N \quad (136)$$

Of course, performing such integrals can only be performed for a few model systems. Most of the effort in the subject goes into developing (approximate) methods to obtain estimates of $Q_N(V, T)$ and expectation values as given by the above integral.

C. The ideal gas

The simplest system with which we can make analytic progress is the classical ideal gas, for which the particle-particle interactions $V(\mathbf{r}^N) = 0$. In this case, the Hamiltonian is simply:

$$H_N(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^N \frac{p_i^2}{2m} \quad [\text{Ideal gas Hamiltonian}] \quad (137)$$

where we have assumed the particles are of equal mass m . Treating mixtures is not difficult, although care needs to be paid to the $N!$ factors which arise due to ‘indistinguishability’.

We can actually evaluate $Q_N(V, T)$ for the ideal gas. First, note that the integration over the momenta variables can be separated (factorised) from the integration over the spatial coordinates. (This can always be done for Hamiltonians of the form Eq.(118) -

more on this later.)

$$Q_N(V, T) = \frac{1}{N!h^{3N}} \int \exp(-\beta \sum_i p_i^2/2m) d\mathbf{p}^N \times \int d\mathbf{r}^N \quad (138)$$

Second, note that the integration over the positions simply gives volume of the system, raised to the power N :

$$\int d\mathbf{r}^N = V^N \quad \text{Don't confuse the volume } V \text{ with potential energy } V(\mathbf{r}^N)! \quad (139)$$

Third, let us have a look at the integration over the momenta:

$$\int \exp(-\beta \sum_i^N p_i^2/2m) d\mathbf{p}^N \quad (140)$$

We can resolve each p_i^2 into its Cartesian components:

$$p_i^2 = p_{xi}^2 + p_{yi}^2 + p_{zi}^2 \quad (141)$$

Also, the exponential factorizes:

$$\exp(-\beta \sum_i^N p_i^2/2m) = \prod_{i=1}^N \prod_{\alpha=x,y,z} \exp(-\beta p_{\alpha i}^2/2m) \quad (142)$$

Therefore, the integral factorizes into $3N$ identical terms (Gaussian integrals):

$$\left(\int_{-\infty}^{\infty} \exp(-\beta p^2/2m) dp \right)^{3N} = \left(\frac{2m\pi}{\beta} \right)^{3N/2} \quad (143)$$

Combining the results above we have:

$$Q_N(V, T) = \frac{1}{N!h^{3N}} \times \left(\frac{2m\pi}{\beta} \right)^{3N/2} \times V^N \quad (144)$$

In terms of the De Broglie wavelength,

$$\Lambda = \frac{h}{\sqrt{2\pi mkT}}, \quad (145)$$

we have:

$$Q_N(V, T) = \frac{1}{N!} \left(\frac{V}{\Lambda^3} \right)^N \quad [\text{Ideal gas}] \quad (146)$$

The Helmholtz free energy of the ideal gas follows on from Eq.(7)

$$A = -kT \times N \times [\ln(V/N\Lambda^3) + 1] \quad (147)$$

whilst the internal energy, entropy and pressure become [show!]:

$$E = \left(\frac{\partial(\beta A)}{\partial\beta} \right)_{N,V} = \frac{3NkT}{2} \quad (148)$$

$$S = \frac{E - A}{T} = Nk \left(\ln(V/N\Lambda^3) + \frac{5}{2} \right) \quad (149)$$

$$P = - \left(\frac{\partial A}{\partial V} \right)_{NT} = \frac{NkT}{V} \quad (150)$$

D. The configurational integral

It should be noticed that the canonical integral over the momentum degrees of freedom can always be performed for Hamiltonians of the type:

$$H_N(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^N \frac{p_i^2}{2m} + V(\mathbf{r}^N) \quad (151)$$

where the potential energy $V(\mathbf{r}^N)$ depends only the coordinates of the particles. Thus:

$$Q_N(V, T) = \frac{1}{N!h^{3N}} \int \exp(-\beta \sum_i p_i^2/2m) d\mathbf{p}^N \times \int \exp(-\beta V(\mathbf{r}^N)) d\mathbf{r}^N \quad (152)$$

$$= \frac{1}{N!\Lambda^{3N}} Z_N(V, T) \quad (153)$$

where

$$Z_N(V, T) = \int \exp(-\beta V(\mathbf{r}^N)) d\mathbf{r}^N \quad [\text{Configurational Integral}] \quad (154)$$

contains all of the information relating to the particle-particle interactions. Nearly all of the difficulties of statistical mechanics originate from this $3N$ dimensional integral, and indeed most of the research in this field concerns with methods to approximate it.

V. GRAND CANONICAL ENSEMBLE

The fixed macroscopic variables in the canonical ensemble are (N, V, T) , which assumes, therefore, that the system of interest is in thermal contact (and equilibrium)

with a heat bath whose temperature is fixed at T . Besides exchanging energy, however, many real systems also exchange matter, so that N , the number of particles in the system, also fluctuates. This situation can be handled by the grand canonical ensemble. Systems in this ensemble are in thermal and chemical equilibrium with a thermostat which acts as a reservoir of energy and molecules, fixing the temperature T and chemical potential μ . For simplicity we will assume that the molecules in the system are all the same species - extension to multicomponent systems is straightforward. The total volume V of the system must remain fixed. The grand-canonical ensemble, therefore, is a (μ, V, T) ensemble.

We proceed as we did for the Canonical Ensemble, and ask: what is the probability P_{Nj} of finding the system in a given microstate with energy $E_{Nj} \equiv E_j(N, V)$ and number of particles N , compatible with fixed μ, V, T ? Note that the set of compatible microstates includes, in principle, systems with N varying from 0 to ∞ ; for each N (at fixed V, T) is itself a member of a canonical ensemble. The form of the entropy stays the same as before:

$$S/k = - \sum_{Nj} P_{Nj} \ln P_{Nj} \quad (155)$$

$$= - \sum_{N=0}^{\infty} \sum_j P_{Nj} \ln P_{Nj} \quad (156)$$

only now we have an additional constraints on the $\{P_j\}$:

$$\sum_{Nj} P_{Nj} = 1 \quad (157)$$

$$\sum_{Nj} E_{Nj} P_{Nj} = \langle E \rangle \quad (158)$$

$$\sum_{Nj} N P_{Nj} = \langle N \rangle \quad (159)$$

The implication of the final constraint is that we require the expected number of particles in the system to be $\langle N \rangle$. Let us introduce the Lagrange multipliers λ, β, γ to impose the three constraints, and perform an unconstrained maximization:

$$\frac{\partial}{\partial P_{Nj}} \left(S - \lambda \left(\sum_{Nj} P_{Nj} - 1 \right) - \beta \left(\sum_{Nj} E_{Nj} P_{Nj} - \langle E \rangle \right) - \gamma \left(\sum_{Nj} N P_{Nj} - \langle N \rangle \right) \right) = 0 \quad (160)$$

obtaining:

$$-\ln P_{Nj} - 1 - \lambda - \beta E_{Nj} - \gamma N = 0 \quad (161)$$

i.e.

$$P_{Nj} = \exp(-\lambda - 1) \exp(-\beta E_{Nj}) \exp(-\gamma N) \quad (162)$$

The requirement for normalization implies:

$$\sum_{Nj} \exp(-\lambda - 1) \exp(-\beta E_{Nj}) \exp(-\gamma N) = 1, \text{ i.e.} \quad (163)$$

$$\exp(\lambda + 1) = \sum_{Nj} \exp(-\beta E_{Nj}) \exp(-\gamma N) \equiv \Xi \quad (164)$$

where we have defined Ξ to be the above sum over (N, j) . Note that this sum can be regarded as a sum over all N and, for a given N , as a sum over all states j which have that particular number of particles:

$$\Xi = \sum_{Nj} \exp(-\beta E_{Nj}) \exp(-\gamma N) \quad (165)$$

$$= \sum_{N=0}^{\infty} \exp(-\gamma N) \sum_j \exp(-\beta E_j(N, V)) \quad (166)$$

$$= \sum_{N=0}^{\infty} \exp(-\gamma N) Q_N(V, \beta) \quad (167)$$

Therefore:

$$P_{Nj} = \frac{\exp(-\beta E_j) \exp(-\gamma N)}{\Xi} \quad (168)$$

But what is the meaning of γ and Ξ ? To this end, let us substitute P_{Nj} back into the expression for the entropy:

$$S/k = - \sum_{Nj} P_{Nj} \ln P_{Nj} \quad (169)$$

$$= - \sum_{Nj} \left(\frac{\exp(-\beta E_{Nj}) \exp(-\gamma N)}{\Xi} \right) \times (-\beta E_{Nj} - \gamma N - \ln \Xi) \quad (170)$$

$$= \beta \langle E \rangle + \gamma \langle N \rangle + \ln \Xi \quad (171)$$

Now, from Thermodynamics, we know that:

$$G = E - TS + PV \quad (172)$$

$$= \mu N \quad (173)$$

where μ is the chemical potential (for a one-component system). So re-arranging it in terms of S :

$$S = \frac{E + PV - \mu N}{T} \quad (174)$$

Therefore comparing Eq.(174) with Eq.(171), and using $\beta = 1/kT$ which we know from the Canonical ensemble, identifying $E = \langle E \rangle$, $N = \langle N \rangle$, we discover:

$$\gamma = -\frac{\mu}{kT} \quad (175)$$

$$\ln \Xi = \frac{PV}{kT} \quad (176)$$

Thus we have:

$$\Xi(\mu, V, T) = \sum_N \exp(\beta\mu N) Q_N(V, T) \quad (177)$$

$$PV = kT \ln \Xi(\mu, V, T) \quad (178)$$

which are expressions for the grand canonical partition function, and the corresponding bridge relation relating Ξ to the macroscopic thermodynamic quantity PV . This latter is often referred to as the *Grand potential*, $-\Omega$:

$$\Omega \equiv -PV \quad [\text{Definition of the Grand Potential}] \quad (179)$$

$$\Omega(\mu, V, T) = -kT \ln \Xi(\mu, V, T) \quad (180)$$

The relation (178) gives the *equation of state* PV of a system in terms of (μ, V, T) (by contrast to the more usual forms of the EOS, which is given in terms of (N, V, T) . More on this below.)

It is sometimes convenient when manipulating the GCE partition function to define the *fugacity*:

$$z \equiv \exp(\beta\mu) \quad [\text{Definition of the fugacity}] \quad (181)$$

and express Ξ as:

$$\Xi(z, V, T) = \sum_N z^N Q_N(V, T) \quad (182)$$

One can ask the question: what is the probability that a system in a GCE at fixed (μ, V, T) has N particles in it, irrespective of the energy? The answer is to sum over all states of the system which have the given value of N , i.e.

$$P_N = \frac{z^N Q_N(V, T)}{\Xi} \quad (183)$$

Similarly, the expected number of particles, $\langle N \rangle$ is given by:

$$\langle N \rangle = \sum_{N=0}^{\infty} N \times P_N \quad (184)$$

$$= \frac{1}{\Xi} \sum_{N=0}^{\infty} N z^N Q_N(V, T) \quad (185)$$

$$= z \times \left(\frac{\partial}{\partial z} \ln \Xi(z, V, T) \right)_{V, T} \quad (186)$$

$$= -\beta z \left(\frac{\partial \Omega}{\partial z} \right)_{V, T} \quad (187)$$

$$= - \left(\frac{\partial \Omega}{\partial \mu} \right)_{V, T}. \quad (188)$$

This gives $N = \langle N \rangle(z, V, T)$, or equivalently $N = \langle N \rangle(\mu, V, T)$. In principle, one can solve one or either equation to obtain $z = z(N, V, T)$ or $\mu = \mu(N, V, T)$. With this in hand, one can re-substitute back into the expression for PV to obtain:

$$PV = f(z(N, V, T), V, T) \quad (189)$$

in effect PV as a function of N, V, T , which is the usual way to express an equation of state.

By this stage, I expect you are totally confused, so let us consider an example (which is, no surprises), the ideal gas(!):

A. The ideal gas in the GCE

Recall that for an ideal gas, $Q_N(V, T) = (1/N!) \times (V/\Lambda^3)^N$. Then:

$$\Xi^{id}(z, V, T) = \sum_N \frac{z^N}{N!} \left(\frac{V}{\Lambda^3}\right)^N \quad (190)$$

$$= \exp(zV/\Lambda^3) \quad [\text{a geometric series}] \quad (191)$$

so we have

$$\Omega^{id} = -kTz \frac{V}{\Lambda^3} \quad (192)$$

and hence:

$$P = kT \frac{z}{\Lambda^3} \quad (193)$$

Also, we can compute $\langle N \rangle$:

$$\langle N \rangle = -\beta z \left(\frac{\partial \Omega}{\partial z} \right)_{V, T} \quad (194)$$

$$= \frac{zV}{\Lambda^3} \quad (195)$$

i.e.

$$z = \frac{\langle N \rangle \Lambda^3}{V} \quad (196)$$

and therefore:

$$\Omega^{id} = -PV = -kT \langle N \rangle \quad (197)$$

which is a familiar result for the EOS of an ideal gas. One can go further. Noting that

$$P_N = \frac{z^N Q_N}{\Xi} \quad (198)$$

$$= z^N \frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N \exp(-\beta PV) \quad (199)$$

$$= z^N \frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N \exp(-\langle N \rangle) \quad (200)$$

$$= \frac{\langle N \rangle^N}{N!} \exp(-\langle N \rangle) \quad (201)$$

which is a Poisson distribution. For large $\langle N \rangle$, this is highly peaked around the most probable value $N^* = \langle N \rangle$. Indeed a second order expansion of

$$\ln P_N = \ln P_{\langle N \rangle + \delta N} \quad (202)$$

in powers of δN leads to

$$P_N \approx \exp(-(\delta N)^2/2\langle N \rangle) \quad (203)$$

which decreases rapidly when $|\delta N| > \sqrt{\langle N \rangle}$.

This result can be generally extended to all systems in a GCE by returning to

$$P_N = z^N Q_N \exp(-\beta PV) \quad (204)$$

$$= z^N \exp(-\beta A) \exp(-\beta PV) \quad (205)$$

and expanding $A = A(N, V, T)$ around the most probable value N^* , (which maximizes P_N):

$$A(N, V, T) = A(N^*, V, T) + (N - N^*) \left(\frac{\partial A}{\partial N} \right)_{V,T} + \frac{1}{2} (N - N^*)^2 \left(\frac{\partial^2 A}{\partial N^2} \right)_{V,T} + \dots \quad (206)$$

where the first derivative term of the Taylor expansion is zero. The second derivative of A is evaluated by noting that the Helmholtz energy is an extensive quantity:

$$A(N, V, T) = N \times a(V/N, T) = N \times a(v, T) \quad (207)$$

where $v = V/N$ is the volume per particle, and a is the Helmholtz energy per particle. Now for fixed V :

$$N \frac{\partial A}{\partial N} = -v \frac{\partial A}{\partial v} \quad (208)$$

i.e.

$$\left(\frac{\partial A}{\partial N} \right)_{V,T} = a(v, T) - v \frac{\partial a}{\partial v} \quad (209)$$

$$= a + P(v, T)v \quad (210)$$

$$\left(\frac{\partial^2 A}{\partial N^2}\right)_{v,T} = -\frac{v}{N} \left(\frac{\partial}{\partial v} \frac{\partial A}{\partial N}\right)_{v,T} \quad (211)$$

$$= -\frac{v}{N} \left(\frac{\partial(a + Pv)}{\partial v}\right)_{v,T} \quad (212)$$

$$= -\frac{v^2}{N} \left(\frac{\partial P}{\partial v}\right)_{v,T} \quad (213)$$

Now the isothermal compressibility is defined to be:

$$\chi_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_{NT} \quad (214)$$

$$= -\left[v \left(\frac{\partial P}{\partial v}\right)_T\right]^{-1} \quad (215)$$

$$(216)$$

from which we get:

$$\left(\frac{\partial^2 A}{\partial N^2}\right)_{v,T} = \frac{v}{N\chi_T} \quad (217)$$

Substituting this into Eq.(206) we obtain:

$$P_N = P_{N^*} \exp(-\beta v(N - N^*)^2/2N^*\chi_T) \quad (218)$$

which is once again a Gaussian distribution with $N^* = \langle N \rangle$; the variance in the number of molecules is:

$$\langle N^2 \rangle - \langle N \rangle^2 = \langle N \rangle \frac{\chi_T}{\beta v} \quad (219)$$

which implies the stability condition $\chi_T > 0$. The relative fluctuation is:

$$\frac{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}{\langle N \rangle} = \frac{1}{\sqrt{\langle N \rangle}} \times \sqrt{\frac{\chi_T}{\beta v}} \quad (220)$$

since $\sqrt{\chi_T/\beta v}$ is an intensive quantity, this shows that the relative fluctuation of N is negligible for large systems, so that open systems behave practically like closed systems with $\langle N \rangle$ molecules: the grand-canonical and canonical ensembles are equivalent for macroscopic systems.

Fluctuations in N become significant only in the vicinity of the liquid-gas critical point, since the compressibility diverges at that point, giving rise, in particular, to critical opalescence.

B. Fluctuations away from equilibrium

So far, we have only considered systems in thermodynamic equilibrium. The second law of thermodynamics tells us that an isolated system that is initially in a non-equilibrium state will spontaneously relax towards thermodynamic equilibrium, increasing the total entropy, so that the final state of equilibrium is characterized by the maximum entropy.

A fruitful way to examine fluctuations away from equilibrium is to consider an isolated system subject to an internal constraint x . (The extension to several constraints is straightforward). For example, such a constraint is a partitioning of the N particles of the system into two sub-systems by introducing a dividing wall. The constraint maintains the system in a non-equilibrium state characterized by the variable x , in addition to the variables such as E, V which specify a state of thermodynamic equilibrium. For each value of x , there corresponds a number $\Omega(x)$ of microstates, and (non-equilibrium) entropy $S(x) = k \ln \Omega(x)$. When the constraint is relaxed, more microstates become available, so that Ω and hence S , increase, while the system evolves towards the macroscopic state characterized by the most probable value, x^* , of the internal variable in the absence of the constraint. In other words, the state of thermodynamic equilibrium is characterized by:

$$\left(\frac{\partial S}{\partial x}\right)_{x=x^*} = 0 \quad (221)$$

We also have, by the fundamental postulate that all available microstates of an isolated system occur with equal probability, that:

$$P(x) = \frac{\Omega(x)}{\Omega} \quad (222)$$

Now, using the maximum term method (see problem sheet):

$$S = k \ln \Omega \approx k \ln \Omega(x^*) \quad (223)$$

i.e. the most probable value of x provides the overwhelming contribution to Ω . Thus we are led to:

$$P(x) = \exp((S(x) - S(x^*))/k) \quad [\text{Microcanonical fluctuation}] \quad (224)$$

Expanding $S(x)$ about $S(x^*)$ in a Taylor series, and noting that the first derivative term is zero, we have:

$$S(x) = S(x^*) + \frac{1}{2}(x - x^*)^2 \times S''(x^*) + \dots \quad (225)$$

leading to a Gaussian distribution:

$$P(x) = \exp(-(x - x^*)^2/2\sigma^2) \quad (226)$$

where

$$\sigma^2 = -\frac{k}{S''} > 0 \quad (227)$$

(since S is a maximum at equilibrium).

A similar argument can be made for systems in contact with a heat bath, i.e. a canonical ensemble. In that case we obtain (show):

$$P(x) = \exp(-(A(x) - A(x^*))/kT) \quad [\text{Canonical fluctuation}] \quad (228)$$

and

$$P(x) = \exp(-(x - x^*)^2/2\sigma^2) \quad (229)$$

with

$$\sigma^2 = \frac{kT}{A''} > 0 \quad (230)$$

since A is a *minimum* at equilibrium (at fixed N, V, T).

VI. STATISTICAL MECHANICS OF INTERACTING PARTICLES AND PHASE TRANSITIONS

The general principles of Statistical Mechanics so far introduced have been illustrated by examples taken from ideal systems in which the particles do not interact. In this and coming sections, we move towards the goal of doing Statistical Mechanics with interacting particles (be they atoms, ions or spins). This is, in general, a very

challenging problem, but one which by necessity we have to face: real particles interact with each other! One limit for which a systematic theory can be developed is that of weak coupling (interactions), for which dilute gases provide a good example; this is the **virial expansion** which will be initially treated. For strong coupling, one must resort to approximation schemes such as **mean-field theory**, to be introduced later on.

A. The virial expansion for classical imperfect gases

Consider a gas of atoms or nearly spherical molecules (e.g. CH₄). Since their charge distribution is (nearly) spherical, they carry no permanent electric moment, and they interact via a spherically symmetric pair potential, $v(r)$, where r denotes the distance between the centres of the two atoms or quasi-spherical molecules. The potential is steeply repulsive for distances $r \leq \sigma$ (where σ is the molecular diameter), due to the **excluded volume effect**, which is a consequence of the Pauli exclusion principle between the electrons. At large distances, $r \gg \sigma$, the van-der-Waals-London dispersion forces take over. The dominant attractive induced dipole-dipole interaction decreases like r^{-6} . A convenient and widely used interpolation between the short-range repulsion and long-range attraction is provided by the Lennard-Jones potential:

$$v(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] \quad (231)$$

For argon, the atomic diameter is $\sigma \approx 0.3$ nm, while the well-depth $\epsilon/k \approx 120$ K ($\epsilon \approx 0.01$ eV), while for methane $\sigma \approx 0.38$ nm, and $\epsilon/k \approx 150$ K.

In a dilute gas, the **reduced density** $n^* = n\sigma^3 \ll 1$, so that the mean distance between nearest neighbour atoms, $d \approx n^{-1/3} \gg \sigma$, i.e. greatly exceeds the range of the interatomic potential $v(r)$. Under these conditions, the particles do not interact with each other except during very rare binary collisions, so that the ideal gas law holds:

$$\frac{P}{kT} = n \quad \text{when } n^* \ll 1. \quad (232)$$

The probability of finding two atoms within a volume of the order of the atomic volume σ^3 is of the order of $(n^*)^2$, and hence deviations from Eq.(232) may be expected as n^*

increases. It therefore seems natural to fit the experimental equation of state data for compressed gases by an expansion in powers of n :

$$\frac{P}{kT} = n + B_2(T)n^2 + B_3(T)n^3 + \dots \quad (233)$$

This is the (originally) empirical **virial expansion**, and the temperature dependent coefficients $B_j(T)$ are the virial coefficients. Statistical Mechanics yields explicit expressions for the $B_j(T)$ in terms of $v(r)$.

We consider here the lowest order correction to the ideal gas law, determined by $B_2(T)$.

The total potential energy of N atoms may be written, to an excellent approximation, as a sum of pair interactions:

$$V(\mathbf{r}^N) = \frac{1}{2} \sum_i \sum_{j \neq i} v(r_{ij}) = \sum_i \sum_{j > i} v(r_{ij}), \quad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \quad (234)$$

The canonical partition function is given by Eq.(153), where Z_N is the configurational integral defined in Eq.(154). The grand partition function may be correspondingly re-expressed as:

$$\Xi = \sum_{N=0}^{\infty} \frac{\zeta^N}{N!} Z_N \quad (235)$$

where $\zeta = z/\Lambda^3$ is the fugacity z (defined in Eq.(181)) divided by the cube of the thermal wavelength Eq.(13). In the ideal gas limit the density n and the fugacity z go to zero. We shall therefore seek successive approximations of Ξ by retaining an increasing number of terms in the infinite series Eq.(235). To first order in ζ :

$$\Xi^{(1)} = 1 + \zeta Z_1 = 1 + \zeta V \quad (236)$$

$$\frac{PV}{kT} = \ln \Xi^{(1)} = \ln(1 + \zeta V) \approx \zeta V \quad (237)$$

$$n = \zeta \frac{\partial(P/kT)}{\partial \zeta} = \zeta \quad (238)$$

Elimination of ζ from the last two equations above leads to the ideal gas law.

To second order in ζ :

$$\Xi^{(2)} = 1 + \zeta Z_1 + \frac{\zeta^2}{2} Z_2 \quad (239)$$

where

$$Z_2 = \int \int \exp(-\beta v(r_{12})) d\mathbf{r}_1 d\mathbf{r}_2 \quad (240)$$

$$= \int \int [1 + f(r_{12})] d\mathbf{r}_1 d\mathbf{r}_2 \quad (241)$$

$$= V^2 + \int \int f(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2 \quad (242)$$

where the **Mayer f-function** has been introduced:

$$f(r) = \exp(-\beta v(r)) - 1 \quad (243)$$

This is a short-ranged function which goes rapidly to zero for $r \gg \sigma$. The double integral can be reduced by a change of variables:

$$(\mathbf{r}_1, \mathbf{r}_2) \rightarrow \left(\mathbf{R}_{12} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2), \mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2 \right) \quad (244)$$

Since the integrand depends only on r_{12} , integration over \mathbf{R}_{12} yields the total volume V , and Z_2 simplifies to:

$$Z_2 = V^2 + V \times a \quad (245)$$

where

$$a = \int f(r) d\mathbf{r} \quad (246)$$

Since $f(r)$ is short ranged, the integration over r can be taken over all space. Substituting the above into Eq.(239), and taking logs, we arrive at:

$$\frac{PV}{kT} = \ln \Xi^{(2)} = \ln \left[1 + \zeta V + \frac{\zeta^2}{2} (V^2 + aV) \right] \quad (247)$$

$$\approx \zeta V + \frac{\zeta^2}{2} \times aV + O(\zeta^3) \quad (248)$$

$$n = \zeta \frac{\partial(P/kT)}{\partial \zeta} = \zeta + a\zeta^2 \quad (249)$$

Elimination of ζ between the last two equations yields the second order result:

$$\frac{P}{kT} = n - \frac{a}{2}n^2 \quad (250)$$

Comparison with the virial expansion leads to the following expression for the second virial coefficient:

$$B_2(T) = -\frac{1}{2} \int f(r) d\mathbf{r} = -2\pi \int_0^\infty f(r) r^2 dr \quad (251)$$

The temperature dependence of $B_2(T)$ comes from the β in the Boltzmann factor in the Mayer function. Similar, more complicated, multi-dimensional cluster integrals may be derived for the higher order virial coefficients, using diagrammatic techniques.

$B_2(T)$ will now be calculated for two simple models of the inter-atomic potential $v(r)$.

(a) The **hard-sphere** model is an extreme simplification of $v(r)$, which retains only its most fundamental feature, i.e. the excluded volume effect, which prevents molecules from overlapping, by taking:

$$v(r) = \begin{cases} \infty & r < \sigma \\ 0 & r > \sigma \end{cases} \quad (252)$$

leading to

$$f(r) = \begin{cases} -1 & r < \sigma \\ 0 & r > \sigma \end{cases} \quad (253)$$

With this Mayer function, $B_2(T)$ is easily calculated:

$$B_2(T) = +2\pi \int_0^\sigma r^2 dr = \frac{2\pi\sigma^3}{3} \quad (254)$$

which is half the excluded volume around one sphere. In this particular case B_2 is independent of temperature (“athermal system”), which reflects the fact that there is no energy scale in the hard-sphere potential. The corresponding correction to the ideal gas law is positive, as one might expect from the purely repulsive nature of the hard-sphere potential.

(b) More generally, consider conformal potentials of the generic form

$$v(r) = \epsilon \times \phi(r/\sigma) \quad (255)$$

where ϵ is an energy scale, σ is the range parameter, and $\phi(x)$ is some “universal” function, assumed to be the same for all atoms and quasi-spherical molecules. The Lennard-Jones potential is an example of a conformal potential. $B_2(T)$ can be cast in the form:

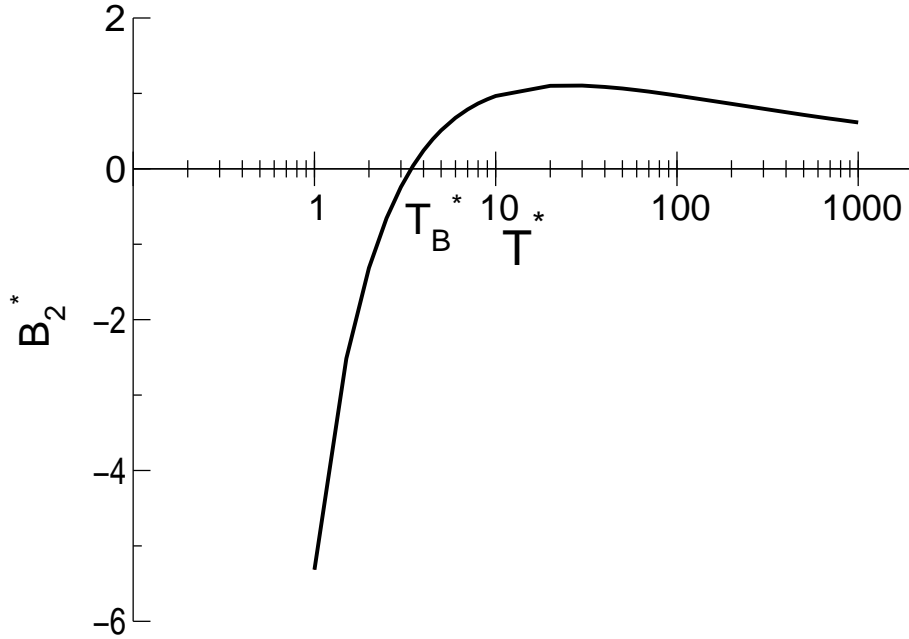
$$B_2(T) = -2\pi \int_0^\infty [\exp(-\beta\epsilon\phi(r/\sigma)) - 1] r^2 dr \quad (256)$$

$$= -2\pi\sigma^3 \int_0^\infty [\exp(-\phi(x)/T^*) - 1] x^2 dx \quad (257)$$

where $x = r/\sigma$ and $T^* = kT/\epsilon$ is the reduced temperature. The reduced (i.e. dimensionless) second virial coefficient, $B_2^* = B_2/\sigma^3$, is a function of T^* only:

$$B_2^*(T^*) = -2\pi \int_0^\infty [\exp(-\phi(x)/T^*) - 1] x^2 dx \quad (258)$$

(258) defines a universal curve. If the atoms or molecules are assumed to interact via a conformal pair potential, where $\phi(x)$ is the same for all species, then the properly reduced values of measured second virial coefficients must all fall on this curve. For the LJ potential where $\phi(x) = 4(x^{-12} - x^{-6})$, the universal curve has the following shape:



The universal curve of B_2^* for the Lennard-Jones potential.

The reduced Boyle temperature, at which B_2 changes sign, $T_B^* \approx 3$.

Experimental data for the second virial coefficient of the rare gases, and of simple quasi-spherical molecules (H_2 , N_2 , CH_4) can be reasonably well fitted to this universal curve over a rather broad range of temperatures, by an appropriate least-squares adjustment of the two parameters ϵ and σ . Some typical values obtained by this procedure are given in the Table I together with corresponding values of the critical temperature, pressure and density, to be later discussed.

For the lightest species, He and H_2 , quantum corrections must be included in the calculation of $B_2(T)$ at the lower temperatures.

B. Law of corresponding states

The “universality” of the 2nd virial coefficient for any functional form of $\phi(x)$ of the pair potential may be generalized to fluids of arbitrary density (including liquids) as the “law of corresponding states”. Given a pair potential of the form (255), which is characterized by a single energy scale ϵ and a single range parameter σ , the configurational

Table I. Critical and potential parameters for some simple molecules

Gas	$\epsilon/k(K)$	σ/nm	T_c^*	n_c^*	$P_c/n_c k T_c$
He	10.2	0.256	0.52	0.174	0.300
Ne	35.8	0.275	1.24	0.300	0.296
Ar	119.8	0.340	1.26	0.316	0.292
Kr	116.7	0.368	1.26	0.326	0.289
Xe	225.3	0.407	1.29	0.338	0.289
H ₂	36.7	0.296	0.91	0.240	0.298
N ₂	95.5	0.375	1.32	0.352	0.292
CH ₄	148.2	0.382	1.29	0.338	0.292

integral $Z_N(V, T)$ can be cast in the form:

$$\frac{Z_N}{V^N} = \frac{1}{V^N} \int_V \dots \int_V \exp(-\epsilon \sum_{i<j} \phi(r_{ij}/\sigma)/kT) d\mathbf{r}_1 \dots d\mathbf{r}_N \quad (259)$$

$$= \frac{1}{V^{*N}} \int_{V^*} \dots \int_{V^*} \exp(-\sum_{i<j} \phi(x_{ij})/T^*) d\mathbf{x}_1 \dots d\mathbf{x}_N \quad (260)$$

$$= \Phi_N(V^*, T^*) \quad (261)$$

where $\mathbf{x} = \mathbf{r}/\sigma$ and $V^* = V/\sigma^3$ are dimensionless quantities. The dimensionless configurational integral Φ_N is a function of the reduced variables V^* and T^* . The partition function then reads:

$$Q_N = \frac{V^N}{N! \Lambda^{3N}} \Phi_N(V^*, T^*) = Q_N^{id} \times \Phi_N(V^*, T^*) \quad (262)$$

where Q_N^{id} is the partition function of an ideal gas. The Helmholtz energy accordingly splits into two parts, an ‘‘ideal’’ and an ‘‘excess’’ part, the latter resulting from the interaction between molecules:

$$A = -kT \ln Q_N^{id} - kT \ln \Phi_N \quad (263)$$

$$= +NkT [\ln(n\Lambda^3) - 1] - kT \ln \Phi_N \quad (264)$$

$$= A^{id} + A^{ex} \quad (265)$$

The reduced excess free energy is a universal function of V^* and T^* for all fluids:

$$\frac{A^{ex}}{kT} = -\ln \Phi_N(V^*, T^*) = N a^{ex}(v^*, T^*) \quad (266)$$

where a^{ex} is the reduced excess free energy per particle, which is a function of the intensive variables $v^* = V^*/N$ and T^* . The resulting equation of state and internal energy similarly split into ideal and excess terms:

$$\frac{PV}{NkT} = -V \times \frac{\partial}{\partial V} \left[\frac{A}{NkT} \right] = 1 - v^* \frac{\partial}{\partial v^*} a^{ex}(v^*, T^*) \quad (267)$$

$$\frac{E}{NkT} = \beta \times \frac{\partial}{\partial \beta} \left[\frac{A}{NkT} \right] = \frac{3}{2} + e^{ex}(v^*, T^*) \quad (268)$$

Eq.(267) shows that the equation of state of systems of particles interacting with a conformal potential is a universal function of the reduced variables v^* and T^* . This is reasonably well borne out by experiment for a large family of quasi-spherical molecules. In particular, the reduced critical temperature, volume and pressure have very similar values for many systems (see Table I). Conversely, measurements of the absolute values of these thermodynamic parameters allows for a quick determination of the molecular interaction parameters ϵ and σ (except for He and H₂, for which quantum effects which violate corresponding states, are important). Thus we find:

$$T_c^* \approx 1.3 \rightarrow \epsilon/k \approx 0.77 \times T_c \quad (269)$$

$$v_c^* = \frac{1}{n_c^*} \approx 3 \rightarrow \sigma \approx (v_c/3)^{1/3} \quad (270)$$

C. Perturbation Theory and the Variational Principle

For a given potential energy function $V(\mathbf{r}^N)$, the explicit calculation of the configurational integral Z_N (and hence the partition function Q_N) is, in general, an impossible task. Suppose that the problem can be solved for a simpler model potential: $V^{(0)}(\mathbf{r}^N)$. This is the case, for example, when $V^{(0)}$ is the potential energy in an external field, acting separately on each atom:

$$V^{(0)}(\mathbf{r}^N) = \sum_{i=1}^N \phi(\mathbf{r}_i) \quad (271)$$

because the configurational integral then factorizes into N independent integrals [Show!]. Another example arises when the pair potential can be split into a short-range repulsive part and a long-range interaction:

$$v(r) = v_0(r) + w(r) \quad (272)$$

The hard-core repulsion may be well represented by the hard-sphere interaction, while $w(r)$ can be treated as a **perturbation**. More generally, let us write:

$$V(\mathbf{r}^N) = V^{(0)}(\mathbf{r}^N) + W(\mathbf{r}^N) \quad (273)$$

where $V^{(0)}$ denotes the **reference** system and W the perturbation. Then, we can rewrite the configurational integral of the full system as:

$$Z_N = \int \exp\left(-\beta [V^{(0)}(\mathbf{r}^N) + W(\mathbf{r}^N)]\right) d\mathbf{r}^N \quad (274)$$

$$= Z_N^{(0)} \times \frac{\int \exp\left(-\beta V^{(0)}\right) \exp\left(-\beta W\right) d\mathbf{r}^N}{Z_N^{(0)}} \quad (275)$$

$$= Z_N^{(0)} \times \langle \exp(-\beta W) \rangle^{(0)} \quad (276)$$

where $\langle \rangle^{(0)}$ denotes a configurational average over the reference system (i.e. weighted by the Boltzmann factor $\exp(-\beta V^{(0)})$). and $Z_N^{(0)}$ is the corresponding configuration integral.

Now, we shall use the fact that the exponential function is *convex*:

$$\exp(x) \geq 1 + x \quad (277)$$

which applied to $\exp(-\beta [W - \langle W \rangle])$ gives:

$$\exp(-\beta [W - \langle W \rangle]) \geq 1 - \beta [W - \langle W \rangle] \quad (278)$$

Taking the statistical average of both sides of this inequality, we find:

$$\langle \exp(-\beta [W - \langle W \rangle]) \rangle = \langle \exp(-\beta W) \rangle \times \exp(\beta \langle W \rangle) \quad (279)$$

$$\geq 1 - \beta [\langle W \rangle - \langle W \rangle] \quad (280)$$

$$= 1 \quad (281)$$

or:

$$\langle \exp(-\beta W) \rangle \geq \exp(-\beta \langle W \rangle) \quad (282)$$

Applying this very general inequality to (276) we arrive at:

$$Z_N = Z_N^{(0)} \times \langle \exp(-\beta W) \rangle^{(0)} \geq Z_N^{(0)} \times \exp(-\beta \langle W \rangle^{(0)}) \quad (283)$$

and hence:

$$Q_N \geq Q_N^{(0)} \exp(-\beta \langle W \rangle^{(0)}) \quad (284)$$

Taking logarithms, we obtain the **Gibbs-Bogoliubov** inequality:

$$A \leq A^{(0)} + \langle W \rangle^{(0)} \quad (285)$$

which states the free energy of the system of interest (i.e. with interaction $V(\mathbf{r}^N)$) is bounded from above by the sum of the free energy of the reference system and the mean value of the perturbation calculated within the reference system ensemble.

The inequality (285) forms the basis of a variational principle in Statistical Mechanics, in much the same way the Rayleigh-Ritz variational principle plays a role in Quantum Mechanics. If $V^{(0)}$ is the potential energy of a **trial** reference system, depending on one or more parameters, say λ , then $A^{(0)}$ and $\langle W \rangle^{(0)}$ depend on λ , and the rhs of (285) can be minimized wrt to λ to provide the lowest upper bound to the free energy A of the system of interest. If $V^{(0)}$ is taken to be of the form (271) the particles are considered as effectively being independent of each other, and the variational principle leads to mean-field theory (MFT). Variants of the mean-field theory will be discussed in the next few sections.

D. van der Waals theory of condensation

Below a substance dependent critical temperature T_c , any gas will liquefy when sufficiently compressed or cooled. To study this important phenomenon, of condensation and liquid-vapour co-existence, we shall follow the classic work of van der Waals

(1873), and split the pair potential between molecules $v(r)$ into a sharply repulsive part at short range, and a smoothly varying attractive part at long range:

$$v(r) = v_0(r) + w(r) \quad (286)$$

The resulting potential energy of the molecules has the form (273). Within the spirit of the perturbation approach, we will approximate the free energy of the fluid by its upper bound:

$$A \approx A^{(0)} + \langle W \rangle^{(0)} \quad (287)$$

Two further approximations are needed in order to get a tractable expression for A .

(a) **Reference system:** When the hard-sphere diameter shrinks to zero, the reference system reduces to an ideal gas, and the configurational integral $Z_N^{(0)} = V^N$. For finite σ , the volume accessible to the centre of any one atom is reduced, due to the presence of the other $N - 1$ particles. The accessible volume is the difference between the total volume V and the excluded volume V' . We shall make the (strong) assumption that the excluded volume is proportional to the number of molecules:

$$V' = N \times b \quad (288)$$

where b is a microscopic volume, of the order of the molecular volume, $b \sim \sigma^3$. This assumption breaks down at higher densities, because there is considerable overlap between the excluded volumes around different molecules (hint: consider the excluded volume of two touching spheres). Nevertheless, using (288), the reference system behaves as an ideal gas occupying the accessible volume $V - Nb$, rather than the total volume V . Under these conditions:

$$Z_N^{(0)} \approx (V - Nb)^N \rightarrow Q_N^{(0)} = \frac{V^N}{N! \Lambda^{3N}} (1 - nb)^N \quad (289)$$

and

$$A^{(0)} = A^{id} + A^{ex} = NkT \left[\ln(n\Lambda^3) - 1 - \ln(1 - nb) \right] \quad (290)$$

The equation of state follows:

$$\frac{P^{(0)}}{kT} = -\frac{\partial}{\partial V} \left(\frac{A^{(0)}}{kT} \right) = \frac{n}{1 - nb} \quad (291)$$

The volume b may now be determined by requiring that to second order in the density n , the expression (291) agrees with the first two terms in the virial expansion, i.e.:

$$n(1 + nb) = n + B_2 n^2 \quad (292)$$

implying that:

$$b = B_2 = \frac{2\pi\sigma^3}{3} \quad (293)$$

where the hard-sphere result for B_2 has been used. However, in view of the approximations used, it is preferable to consider b as being an adjustable parameter.

(b) **Perturbation:** To calculate $\langle W \rangle^{(0)}$ we make a mean-field approximation, by replacing the Boltzmann weight $\exp(-\beta V^{(0)})$ by 1 in the calculation of the statistical average; the latter then becomes a simple volume average of $W = \sum_{i<j} w(r_{ij})$, which amounts to neglecting **correlations** between the positions of different particles:

$$\langle W \rangle^{(0)} \approx \frac{1}{V^N} \int_V \dots \int_V \sum_{i<j} w(r_{ij}) d\mathbf{r}^N \quad (294)$$

$$= \frac{N(N-1)}{2} \frac{V^{N-2}}{V^N} \int_V \int_V w(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2 \quad (295)$$

$$= -Nna \quad (296)$$

where we have approximated $N-1$ by N and

$$a = -\frac{1}{2V} \int_V \int_V w(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2 \quad (297)$$

$$= -\frac{1}{2} \int_V w(r) d\mathbf{r} > 0 \quad (298)$$

since $w(r)$ is negative. Within this mean-field approximation, the total free energy is:

$$A = A^{(0)} - Nna \quad (299)$$

with $A^{(0)}$ given by (290).

The resulting equation of state has the form:

$$\frac{P}{nkT} = \frac{1}{1-nb} - \frac{an}{kT} \quad (300)$$

which can be rearranged into the form proposed by van der Waals in 1873 on purely intuitive grounds:

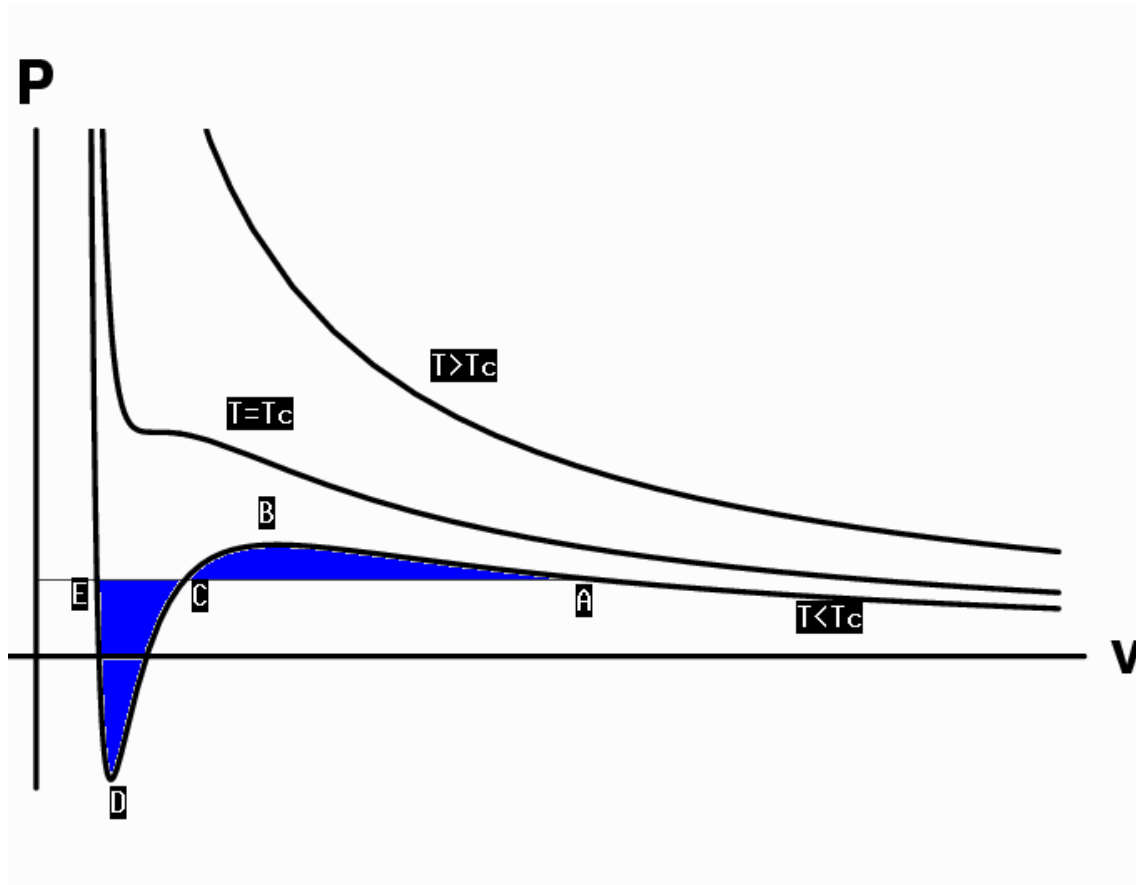
$$\left[P + a \left(\frac{N}{V} \right)^2 \right] [V - Nb] = NkT \quad (301)$$

a may be computed from first principles from a knowledge of the intermolecular pair potential, but in practice it is treated, like b , as an adjustable parameter by fitting (300) to the measured EOS of compressed gases. Thus for N_2 , the optimum parameters turn out to be $a = 0.1408 \text{ Pa m}^6 \text{ mol}^{-2}$, $b = 0.3913 \times 10^{-4} \text{ m}^3 \text{ mol}^{-1}$.

Condensation (gas-liquid transition): Eq(300) may be rewritten as a cubic equation for the volume per molecule $v = 1/n$:

$$v^3 - \left[b + \frac{kT}{P} \right] v^2 + \frac{a}{P} v - \frac{ab}{P} = 0 \quad (302)$$

Along an isotherm (fixed T), this equation has either three real roots, or one (plus two unphysical complex roots), i.e. either 3 or 1 values of v will correspond to any given pressure P . Typical isotherms are given in the figure below.



$P(v)$ isotherms for the Lennard-Jones system for $T < T_c$, $T = T_c$, $T > T_c$. The intersections of the equal-area tie-line with the $P(v)$ curve give the values of the molar volumes at phase-coexistence.

For $T < T_c$, the isotherms $P(v)$ exhibit, between the maximum and the minimum, a region of negative compressibility $\chi_T = -[v(\partial P/\partial v)_T]^{-1}$, which corresponds to unstable thermodynamic states, since they violate the stability condition $\chi_T > 0$. These states are eliminated by replacing the van der Waals loop by a horizontal tie-line reflecting the fact that the condensation of a liquid to a gas occurs at constant P . The location of the horizontal tie-line is determined by the condition of equal chemical potentials in the two coexisting phases (gas g and liquid l). From the thermodynamic

relation:

$$d\mu = d(G/N) = -\frac{S}{N}dT + vdP \quad (303)$$

we conclude that along an isotherm:

$$\mu_l - \mu_g = \int_g^l d\mu = \int_g^l vdP = 0 \quad (304)$$

The last integral may be rewritten as:

$$\int_{P_A}^{P_B} vdP + \int_{P_B}^{P_C} vdP + \int_{P_C}^{P_D} vdP + \int_{P_D}^{P_E} vdP = 0 \quad (305)$$

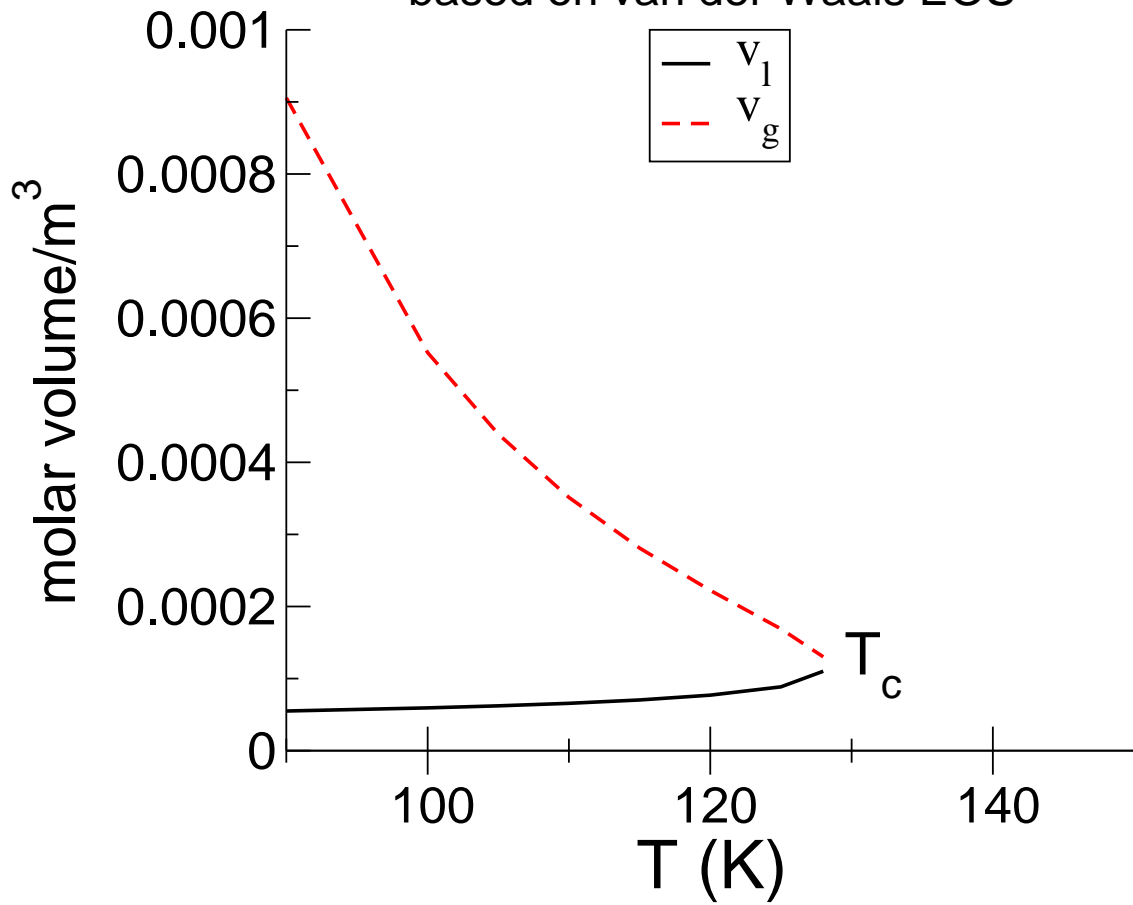
or, rearranging terms:

$$\int_{P_A}^{P_B} vdP - \int_{P_C}^{P_B} vdP = \int_{P_D}^{P_C} vdP - \int_{P_D}^{P_E} vdP \quad (306)$$

which expresses the equal area condition: Area(ABC)=Area(CDE). This condition fixes unambiguously the position of the gas-liquid tie-line AE. The volume change upon condensation $\Delta v = v_l - v_g$ is negative. A non-zero volume change is the signature of a first-order phase transition.

Below is the phase co-existence curve for the N₂ system, based on the parameters given in the text. It has been obtained by numerically implementing the equal-area construction.

Phase coexistence curve for N₂ based on van der Waals EOS



The critical point. The temperature beyond which liquid-gas coexistence ceases corresponds to the $P(v)$ isotherm with an inflexion point; this inflexion point is the critical point, above which there is only a single “supercritical” fluid phase. Thus, the critical point is determined by the conditions:

$$\left(\frac{\partial P}{\partial v}\right)_{T=T_c} = \left(\frac{\partial^2 P}{\partial v^2}\right)_{T=T_c} = 0 \quad (307)$$

Equivalently, and more elegantly, the critical point coordinates (T_c, v_c, P_c) can be obtained by noting that at that T and P , the three roots of the cubic equation (302) coalesce into a triple root for v_c :

$$(v - v_c)^3 = v^3 - 3v_c v^2 + 3v_c^2 v - v_c^3 = 0 \quad (308)$$

By identifying the coefficients of the powers in v in (302) and (308) we obtain the result:

$$v_c = 3b, P_c = \frac{a}{27b^2}, T_c = \frac{8a}{27b} \quad (309)$$

while the dimensionless critical ratio is:

$$\frac{P_c v_c}{kT_c} = \frac{3}{8} = 0.375 \quad (310)$$

which is in rough agreement with the experimental data in Table I. The critical parameters above are expressed in terms of the molecular parameter a and b . The latter may be eliminated from the van der Waals EOS by introducing the reduced variables:

$$t = \frac{T}{T_c}, p = \frac{P}{P_c}, \phi = \frac{V}{V_c} \quad (311)$$

so that (300) can be cast into the form:

$$\left(p + \frac{3}{\phi^2}\right)(3\phi - 1) = 8t \quad (312)$$

which is a special case of the law of corresponding states.

Critical point exponents. At the critical point $(\partial P/\partial v)_{T=T_c} = 0$, the isothermal compressibility χ_T diverges. According to the GCE fluctuation formula Eq.(220), this implies very large fluctuations in the density. As we noted earlier, below T_c the phase transition from gas to liquid is first order, and is characterized by a finite discontinuity in the molar volume Δv , and by a large latent heat $\Delta H = T\Delta S$ (where H is the enthalpy function). Both Δv and ΔH vanish at the critical point, and the phase transition becomes **second** order, signaled by the divergence of the compressibility χ_T . The divergence of χ_T at $v = v_c$ is characterized by a power (or scaling) law:

$$\chi_T \sim (T - T_c)^{-\gamma} \text{ at } v = v_c, T > T_c \quad (313)$$

where γ is a **critical point** exponent.

A second critical exponent, β , governs the way in which the density discontinuity vanishes at T_c :

$$n_l - n_g \sim |T - T_c|^\beta \quad \text{for } T < T_c \quad (314)$$

while yet another exponent, δ , characterizes the variation of the pressure with density n along the critical isotherm:

$$P - P_c \sim |n - n_c|^\delta \times \text{sign}(n - n_c), T = T_c \quad (315)$$

The key experimental finding is that the exponents β, γ, δ (and others) are **universal**, i.e. have the same value for all substances:

$$\beta \approx 0.35, \gamma \approx 1.22, \delta \approx 4.35 \quad (\text{Experiment}) \quad (316)$$

The mean-field EOS (300) does lead to universal critical exponents, but with values which do not agree with experiment:

$$\beta = \frac{1}{2}, \gamma = 1, \delta = 3 \quad (\text{Mean-field theory}) \quad (317)$$

This disagreement is due to the neglect of fluctuations within mean-field theory.

E. The Ising model of Magnetism.

Earlier in this course (section III.A), we encountered a simple model of magnetism, characterized by spin variables $s_i = \pm 1$ located at sites of a lattice. In the case we considered, the spins were non-interacting, and we found that the resulting behaviour described (at best) that of a paramagnet; the system did not show a spontaneous magnetization in the absence of an externally applied field Eq.(110,112).

The Ising model introduces interactions between nearest-neighbour spins, so that contribution to the total energy of adjacent spins s_i and s_j is $-Js_i s_j$. If the **coupling constant** $J > 0$, then this energy is negative for parallel spins, and positive for antiparallel spins. It follows that for such a coupling, alignment of spins is energetically favourable.

The Ising model Hamiltonian $E(s^N) \equiv H(s^N)$ is:

$$H = -J \sum_{\{i,j\}} s_i s_j - \mu_B B \sum_i s_i \quad (318)$$

where the double sum is restricted to nearest-neighbour pairs of spins only, μ_B is the magnetic moment of each spin, and B is the external magnetic field aligned in the $+z$

direction. The coordination number of each spin depends on the lattice type. We shall assume each spin has ν nearest neighbours (4 for a square lattice, 6 for a cubic lattice, etc).

We propose to calculate the partition function Q_N of the Ising model and derive from it the magnetic and thermodynamic properties. To this end, in the spirit of the variational theory, we split the Hamiltonian into two:

$$H = H_0 + H_1 \quad (319)$$

$$H_0 = -\mu_B B \sum_i s_i \quad (320)$$

$$H_1 = -J \sum_{\{i,j\}} s_i s_j \quad (321)$$

Then, the partition function is given, as in Eq.(93)

$$Q_N = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \dots \sum_{s_N=\pm 1} \exp(-\beta H(s^N)) \quad (322)$$

If the coupling between the spins is neglected $J = 0$ (or alternatively, if $kT \gg J$), then we already know the partition function and free energy, from Eq.(96) and Eq.(98):

$$Q_N^{(0)} = [2 \cosh(\beta \mu_B B)]^N \quad (323)$$

$$A^{(0)} = -kT \ln Q_N^{(0)} = -NkT \ln [2 \cosh(\beta \mu_B B)] \quad (324)$$

As the temperature is lowered, the coupling between spins can no longer be neglected. In the first instance, H_1 may be regarded by a perturbation. According to Eq.(276):

$$Q_N = Q_N^{(0)} \times \langle \exp(-\beta H_1) \rangle^{(0)} \quad (325)$$

which to first order in the dimensionless ratio J/kT is:

$$Q_N \approx Q_N^{(0)} \times [1 - \langle \beta H_1 \rangle^{(0)}] \quad (326)$$

which implies:

$$A = A^{(0)} + \langle H_1 \rangle^{(0)} \quad (327)$$

i.e. to first order, the free energy is replaced by its Gibbs-Bogoliubov upper bound. Since the N spins are independent in the reference system,

$$\langle H_1 \rangle^{(0)} = -J \sum_{\{i,j\}} \langle s_i s_j \rangle^{(0)} \quad (328)$$

$$= -J \sum_{\{i,j\}} \langle s_i \rangle^{(0)} \langle s_j \rangle^{(0)} \quad (329)$$

$$= -J \frac{N\nu}{2} \left(\langle s \rangle^{(0)} \right)^2 \quad (330)$$

where in the last step we have used the fact that there are $N\nu/2$ terms in the double sum $\sum_{\{i,j\}}$, and each term must contribute the same amount.

Now, we know that $\langle s \rangle^{(0)}$ is given by Eq.(110):

$$\langle s \rangle^{(0)} = \tanh(\beta\mu_B B) \quad (331)$$

and hence

$$\langle H_1 \rangle^{(0)} = -J \frac{N\nu}{2} (\tanh(\beta\mu_B B))^2 \quad (332)$$

The magnetization is given by:

$$M = -\frac{\partial (A^{(0)} + \langle H_1 \rangle^{(0)})}{\partial B} \quad (333)$$

$$= N\mu \left[\tanh(\beta\mu_B B) + \nu\beta J \frac{\tanh(\beta\mu_B B)}{\cosh^2(\beta\mu_B B)} \right] \quad (334)$$

which in the weak field $B \rightarrow 0$ limit reduces to:

$$M \approx N\beta\mu_B^2 (1 + \nu\beta J) \times B \quad (335)$$

so the zero-field susceptibility is (c.f. Eq.(111)):

$$\chi = \left(\frac{\partial M}{\partial B} \right)_T \Big|_{B=0} \quad (336)$$

$$= N\beta\mu_B^2 (1 + \nu\beta J) \quad (337)$$

$$= \frac{C}{T} \times \left(1 + \frac{\theta}{T} \right) \quad (338)$$

where $C = N\mu_B^2/k$, and $\theta = \nu J/k$ is a characteristic temperature of the spin-spin coupling. The 2nd term is a correction to Curie's law for paramagnetic materials due to this coupling; it is the first term in a high temperature expansion in powers of θ/T .

The above theory still does not predict the existence of a ferromagnetic phase, i.e. one in which $M \neq 0$ when $B = 0$. To proceed further, we need a more accurate treatment of the effect of spin-spin coupling. To this end, we shall suppose that if a spontaneous magnetization does occur in the absence of an external field, this hints at the existence of an “internal” (or “molecular”) field, B_1 , inside the ferromagnetic material, which tends to align the individual spins while being self-consistently generated by their mutual interactions. The existence of such a self-consistent field was postulated by Pierre Weiss in 1907, in his phenomenological theory of ferromagnetism. This was one of the first examples of a mean-field theory. A modern derivation makes use of the Gibbs-Bogoliubov inequality and the variational principle. To this end, we split up the Ising Hamiltonian into a reference term and a perturbation, but in a cunning way:

$$H_0 = -\mu_B(B + B_1) \sum_i s_i \quad (339)$$

$$H_1 = -J \sum_{\{i,j\}} s_i s_j + \mu_B B_1 \sum_i s_i \quad (340)$$

The molecular field B_1 will be determined by the variational principle: it is a variational parameter used to minimize the RHS of the Gibbs-Bogoliubov inequality (285). Note that the reference Hamiltonian H_0 is still a non-interacting Hamiltonian, and therefore the statistical averages can be easily carried out as before, just by replacing B with $B + B_1$:

$$A^{(0)} = -NkT \ln [2 \cosh (\beta\mu_B(B + B_1))] \quad (341)$$

Similarly:

$$\langle H_1 \rangle^{(0)} = -J \frac{N\nu}{2} (\tanh(\beta\mu_B(B + B_1)))^2 + \mu_B B_1 N \langle s \rangle^{(0)} \quad (342)$$

$$= -J \frac{N\nu}{2} (\tanh(\beta\mu_B(B + B_1)))^2 + \mu_B B_1 N \tanh(\beta\mu_B(B + B_1)) \quad (343)$$

The upper bound $A^{(0)} + \langle H_1 \rangle^{(0)}$ must now be minimized with respect to B_1 , i.e. we seek the solution of:

$$\frac{\partial (A^{(0)} + \langle H_1 \rangle^{(0)})}{\partial B_1} = 0 \quad (344)$$

Carrying out the derivative, we obtain the following implicit equation for B_1 [show!]:

$$B_1 = \frac{\nu J}{\mu_B} \tanh[\beta\mu_B(B + B_1)] \quad (345)$$

Recognizing that B_1 is a function of B , the net magnetization M may be calculated from

$$M = - \left(\frac{\partial A}{\partial B} \right)_{T, B_1} - \left(\frac{\partial A}{\partial B_1} \right)_{T, B} \times \left(\frac{\partial B_1}{\partial B} \right)_T \quad (346)$$

with, as usual, $A = A^{(0)} + \langle H_1 \rangle^{(0)}$. Carrying out this derivative, we obtain:

$$M = N\mu_B \tanh[\beta\mu_B(B + B_1)] \quad (347)$$

$$= \frac{N\mu_B^2}{\nu J} B_1 \quad \text{using Eq.(345)}. \quad (348)$$

The magnetization at saturation is $M_s = N\mu_B$; define the **order parameter** $\zeta = M/M_s$. In terms of ζ , Eq.(347) is:

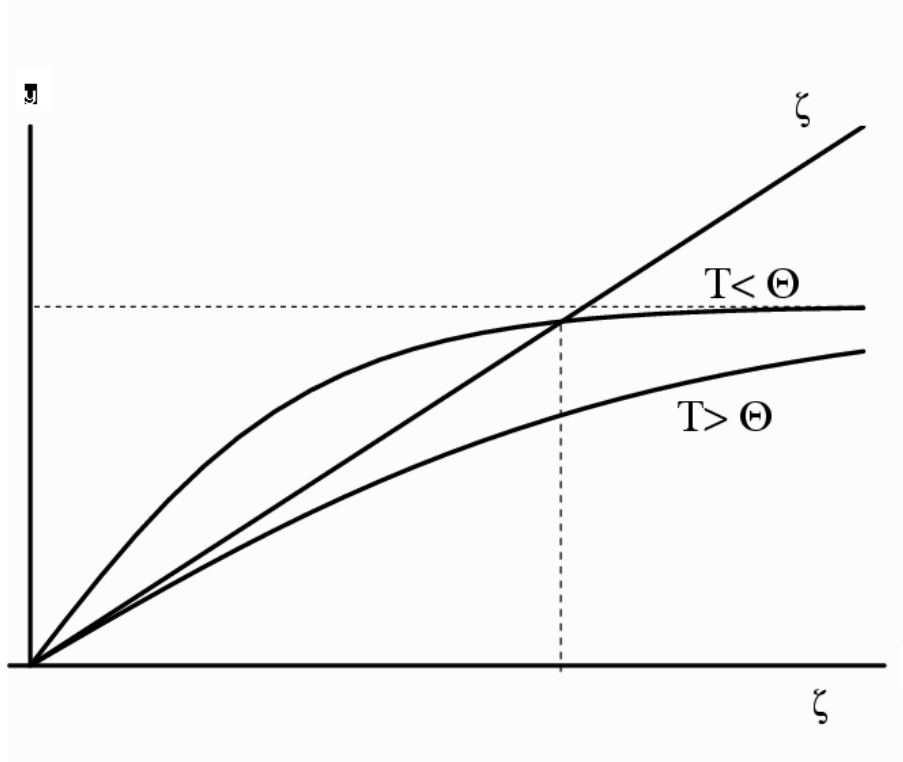
$$\zeta = \tanh[\beta\mu_B B + \beta\nu J\zeta] \quad (349)$$

We are interested in the possibility of spontaneous magnetization (i.e. $\zeta \neq 0$) in the absence of an applied field $B = 0$. Then Eq.(349) reduces to

$$\zeta = \tanh[\theta\zeta/T] \quad (350)$$

where $\theta = \nu J/k$ is the characteristic spin-spin coupling temperature introduced earlier. Eq.(350) is a non-linear equation which must be solved numerically. The maximum slope of $\tanh(\theta\zeta/T)$ as a function of ζ is at the origin, where it is θ/T . If $T < \theta$, then the initial slope will be greater than 1, the equation Eq.(350) has a solution at $\zeta > 0$; otherwise, the only solution is at $\zeta = 0$. Thus a spontaneous magnetization (ferromagnetic phase) occurs at all temperatures below:

$$T_c = \frac{\nu J}{k} \equiv \theta \quad (351)$$



Graph of the equation $\zeta = \tanh [\theta\zeta/T]$.

In the ferromagnetic phase, ζ increases towards its maximum value of 1 at T is lowered towards 0. Just below T_c , ζ is small, and therefore $\tanh(T_c\zeta/T)$ may be expanded in a Taylor series; keeping the first two terms, we have for Eq.(349):

$$\zeta = \tanh(T_c\zeta/T) \approx \frac{T_c}{T}\zeta - \frac{1}{3}\left(\frac{T_c}{T}\right)^3 \zeta^3 \quad (352)$$

which can be solved to give:

$$\zeta^2 = 3\left(\frac{T}{T_c}\right)^2 \left(\frac{T_c - T}{T_c}\right) \approx 3\left(\frac{T_c - T}{T_c}\right) \quad (353)$$

i.e.

$$\zeta \approx \pm\sqrt{3}\left(\frac{T_c - T}{T_c}\right)^{1/2} \quad (354)$$

This in turn implies that the critical exponent β which characterizes the order parameter for the ferromagnetic-paramagnetic transition is $\beta = 1/2$, somewhat larger than the experimentally observed value of $\beta \approx 0.35$. This disagreement is ascribed to the neglect of the effect of fluctuations in the mean-field theory.

Note that this exponent is independent of material properties such as ν, J , i.e. is the same for all ferromagnetic materials (universality), whereas the critical temperature depends on the material properties.

The susceptibility of the paramagnetic phase $T > T_c$ may be calculated in the weak-field limit:

$$\zeta = \beta\mu_B B + \frac{\theta\zeta}{T} \quad (355)$$

yielding:

$$\zeta = \frac{\beta\mu_B}{1 - \theta/T} B \quad (356)$$

The magnetization is:

$$M = N\mu_B\zeta = \frac{N\mu_B^2}{k} \times \frac{1}{T - T_c} \times B \quad (357)$$

giving the susceptibility:

$$\chi = \left(\frac{\partial M}{\partial B} \right)_T = \frac{C}{T - T_c} \quad (358)$$

which diverges at the Curie temperature T_c , with a critical exponent $\gamma = 1$, somewhat smaller than the experimental value of $\gamma \approx 1.22$.

VII. QUANTUM STATISTICS (NON-INTERACTING PARTICLES)

In the last part of this course, we will return to quantum systems. You will recall that classical (or “Boltzon”) statistics applies whenever the molecular partition function $q \gg N$, so that the average occupation number of any given molecular energy-level is much less than unity. This approximation breaks down whenever $\Lambda \approx \sigma$, i.e. the thermal de Broglie wavelength becomes comparable to the interatomic distance, as it does for low-mass particles (electrons, protons etc), and at low temperatures or high pressures. Here we shall give a systematic account of statistical mechanics as it applies to non-interacting quantum systems; for **bosons** the particles obey Bose-Einstein statistics, and for **fermions** they obey Fermi-Dirac statistics. The fundamental difference between the two is that the N -particle wavefunction which describes a microstate of the whole system must have a definite parity under particle exchange; it must be symmetric for bosons and antisymmetric for fermions, i.e.:

$$\begin{aligned} \Psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N) &= \pm \Psi(\mathbf{r}_1, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N) & (359) \\ &+ \quad \text{for bosons} \\ &- \quad \text{for fermions} \end{aligned}$$

Bosons are particles with integer spin (e.g. photons, hydrogen atom); fermions are particles with half-integer spin (electrons, protons). For fermion systems, the antisymmetry of the wavefunction implies the Pauli exclusion principle for non-interacting particles.

A. Occupation number representation

Consider a system with molecular (i.e. single particle) states labeled α , $\alpha = 0, \dots, \infty$, with energies $\epsilon_\alpha \geq 0$ (we can take the zero of energy so that the ground-state energy $\epsilon_0 \geq 0$). In general, the energy levels will depend on the external potential, and on the volume V of the system. The state of a system consisting of N indistinguishable particles (which must occupy the available single-particle states) can be uniquely specified by specifying the occupation numbers of each single-particle state, n_α , which can

a priori taken on the values $0, 1, 2, \dots, N$. For an isolated system with total energy E , we must have:

$$\sum_{\alpha} n_{\alpha} = N, \quad (360)$$

$$\sum_{\alpha} \epsilon_{\alpha} n_{\alpha} = E. \quad (361)$$

For bosons, there is no restriction on the value n_{α} can take:

$$0 \leq n_{\alpha} \leq N, \quad \text{all } \alpha, \text{ (bosons)} \quad (362)$$

For fermions, the Pauli exclusion principle restricts the occupation numbers to a maximum of one:

$$0 \leq n_{\alpha} \leq 1, \quad \text{all } \alpha, \text{ (fermions)} \quad (363)$$

The canonical partition function can be written as:

$$\begin{aligned} Q_N &= \sum'_{\{n_{\alpha}\}} \exp(-\beta E) \\ &= \sum'_{\{n_{\alpha}\}} \exp\left(-\beta \sum_{\alpha} \epsilon_{\alpha} n_{\alpha}\right) \end{aligned} \quad (364)$$

where the sum $\sum'_{\{n_{\alpha}\}}$ is shorthand for a sum over all allowable distributions of the N particles over the one-particle quantum states. The prime ' over the summation indicates that the restriction (360) must be strictly adhered to.

Writing out the sums explicitly, and factorizing the exponential, we have:

$$\begin{aligned} Q_N &= \sum_{n_0} \sum_{n_1} \dots' \prod_{\alpha} \exp(-\beta n_{\alpha} \epsilon_{\alpha}) \\ &= \sum_{n_0} \sum_{n_1} \dots' \prod_{\alpha} [\exp(-\beta \epsilon_{\alpha})]^{n_{\alpha}} \end{aligned}$$

The constraint (360) prevents the summation to be carried out analytically. However, this inconvenient constraint can be lifted by going to the grand canonical ensemble.

The grand partition function is:

$$\Xi = \sum_{N=0}^{\infty} z^N Q_N \quad (365)$$

$$= \sum_{N=0}^{\infty} \left(\sum_{n_0} \sum_{n_1} \dots [z \exp(-\beta\epsilon_0)]^{n_0} \times [z \exp(-\beta\epsilon_1)]^{n_1} \times \dots \right) \quad (366)$$

$$= \sum_{n_0} \sum_{n_1} \dots [z \exp(-\beta\epsilon_0)]^{n_0} \times [z \exp(-\beta\epsilon_1)]^{n_1} \times \dots \quad (367)$$

$$= \prod_{\alpha} \sum_{n_{\alpha}} [z \exp(-\beta\epsilon_{\alpha})]^{n_{\alpha}} \quad (368)$$

where in going from (366) to (367) we allow the n_{α} to vary over their full range, thereby covering all possible values of N . Of course, the allowable range of n_{α} depends on the type of particle present, boson or fermion.

B. Bose-Einstein and Fermi-Dirac statistics

(a) **Bosons.** In this case $0 \leq n_{\alpha} \leq \infty$. Therefore:

$$\sum_{n_{\alpha}=0}^{\infty} [z \exp(-\beta\epsilon_{\alpha})]^{n_{\alpha}} = \frac{1}{1 - z \exp(-\beta\epsilon_{\alpha})} \quad (369)$$

and hence the grand partition function for BE statistics reduces to

$$\Xi^{BE}(z, V, T) = \prod_{\alpha} \frac{1}{1 - z \exp(-\beta\epsilon_{\alpha})}. \quad (370)$$

The EOS follows from Eq. (178):

$$\frac{PV}{kT} = \ln \Xi^{BE} = - \sum_{\alpha} \ln [1 - z \exp(-\beta\epsilon_{\alpha})] \quad (371)$$

while the mean number of particles in V is:

$$\langle N \rangle = \sum_{\alpha} \langle n_{\alpha} \rangle = z \frac{\partial}{\partial z} \ln \Xi^{BE} \quad (372)$$

$$= \sum_{\alpha} \frac{1}{z^{-1} \exp(\beta\epsilon_{\alpha}) - 1} \quad (373)$$

The mean number of particles in state α defines the distribution function:

$$f(\epsilon_{\alpha}) = \langle n_{\alpha} \rangle \quad (374)$$

Comparison of the second and last sum above shows:

$$f^{BE}(\epsilon_{\alpha}) = \frac{1}{\exp(\beta(\epsilon_{\alpha} - \mu)) - 1} \quad (375)$$

where we have used $z = e^{\beta\mu}$. This is the Bose-Einstein distribution.

(b) **Fermions.** In this case $n_\alpha = 0$ or 1.

$$\sum_{n_\alpha=0}^1 [z \exp(-\beta\epsilon_\alpha)]^{n_\alpha} = 1 + z \exp(-\beta\epsilon_\alpha) \quad (376)$$

The remaining formulae follow easily:

$$\Xi^{FD}(z, V, T) = \prod_{\alpha} [1 + z \exp(-\beta\epsilon_\alpha)] \quad (377)$$

$$\frac{PV}{kT} = \ln \Xi^{FD} = \sum_{\alpha} \ln [1 + z \exp(-\beta\epsilon_\alpha)] \quad (378)$$

$$\langle N \rangle = \sum_{\alpha} \langle n_{\alpha} \rangle = z \frac{\partial}{\partial z} \ln \Xi^{FD} \quad (379)$$

$$= \sum_{\alpha} \frac{1}{z^{-1} \exp(\beta\epsilon_\alpha) + 1} \quad (380)$$

$$f^{FD}(\epsilon_\alpha) = \frac{1}{\exp(\beta(\epsilon_\alpha - \mu)) + 1} \quad (381)$$

Note that the BE and FD distributions differ only in a sign in the denominator. This apparently minor difference leads to completely different behaviour of ideal fermion and boson gases at low temperatures.

C. The weak-degeneracy limit

To make contact with the classical Boltzmann distribution, consider the limit $z \ll 1$ (i.e. highly negative chemical potential). Since $\epsilon_\alpha \geq 0$ for all α , $\exp(\beta\epsilon_\alpha) \geq 1$, so that:

$$f(\epsilon_\alpha) = \frac{1}{z^{-1} \exp(\beta\epsilon_\alpha) \mp 1} \approx z \exp(-\beta\epsilon_\alpha) \quad (382)$$

which is the classical Maxwell-Boltzmann distribution:

$$f^{MB}(\epsilon_\alpha) = \exp(-\beta(\epsilon_\alpha - \mu)) \quad (383)$$

To see the significance of the limit $z \ll 1$, compute:

$$\langle N \rangle = \sum_{\alpha} \langle n_{\alpha} \rangle = z \sum_{\alpha} \exp(-\beta\epsilon_\alpha) = z \times q \quad (384)$$

where q is the molecular partition function. If we assume our particles are structure-less (and spin-less), then q is just the translational partition function, which we know from Eq.(12):

$$q = \frac{V}{\Lambda^3} \quad (385)$$

Therefore:

$$z = \frac{\langle N \rangle}{V} \times \Lambda^3 = \frac{\Lambda^3}{v} \quad (386)$$

where v is the mean volume per atom. Thus the requirement $z \ll 1$ is equivalent to:

$$\Lambda = \frac{h}{(2\pi mkT)^{1/2}} \ll v^{1/3} \approx \sigma \quad (387)$$

where σ is roughly the mean distance between nearest-neighbour particles, which is the criterion we had intuitively derived earlier. Thus classical statistical mechanics is valid whenever the thermal de Broglie wavelength (which increases as T decreases) is much less than the mean distance between neighbouring atoms. Λ is roughly the extent of the wavepacket associated with a quantum particle at temperature T (or, in other words, the smearing in position implied by the Heisenberg Uncertainty principle). Thus the classical limit is reached when the wavepackets of neighbouring atoms cease to overlap. The **degeneracy** temperature T_d above which classical statistics holds is:

$$T_d = \frac{h^2}{2\pi m k v^{2/3}} \quad (388)$$

Quantum statistics becomes important for $T < T_d$, and the gas is said to be **degenerate**.

Consider the case of a weakly degenerate (or nearly classical) gas. Since $z \ll 1$, we can expand $\ln \Xi$ in powers of z . According to Eqs.(370) and (377) we have:

$$\ln \Xi = \mp \sum_{\alpha} \ln [1 \mp z e^{-\beta \epsilon_{\alpha}}] \quad (389)$$

$$= \sum_{l=1}^{\infty} (\pm)^{l+1} \frac{z^l}{l} \sum_{\alpha} e^{-l\beta \epsilon_{\alpha}} \quad (390)$$

The sum over states can be evaluated by noting that the translational energy levels in a periodic box of length L have the form:

$$\epsilon_{\mathbf{p}} \equiv \epsilon_{n_x, n_y, n_z} = \frac{|\mathbf{p}|^2}{2m} = \frac{h^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2), \quad \mathbf{p} = \hbar \times \frac{2\pi}{L}(n_x, n_y, n_z) \quad (391)$$

and replacing the sum by an integral:

$$\sum_{\alpha} e^{-l\beta\epsilon_{\alpha}} \rightarrow \frac{V}{h^3} \int e^{-l\beta p^2/2m} d\mathbf{p} \quad (392)$$

$$= \frac{V}{\Lambda^3 l^{3/2}} \quad (393)$$

It follows that (+1 for bosons, -1 for fermions):

$$\frac{PV}{kT} = \ln \Xi = \frac{V}{\Lambda^3} \sum_{l=1}^{\infty} (\pm)^{l+1} \frac{z^l}{l^{5/2}} \quad (394)$$

Similarly,

$$\langle N \rangle = z \frac{\partial}{\partial z} \ln \Xi = \frac{V}{\Lambda^3} \sum_{l=1}^{\infty} (\pm)^{l+1} \frac{z^l}{l^{3/2}} \quad (395)$$

To lowest order in z we find:

$$\frac{P}{kT} = \frac{z}{\Lambda^3} \quad (396)$$

$$n = \frac{\langle N \rangle}{V} = \frac{z}{\Lambda^3} \quad (397)$$

leading to the familiar classical ideal gas EOS:

$$\frac{P}{kT} = n \quad (398)$$

D. Photons and Blackbody radiation

An important application of BE statistics is to the equilibrium properties of electromagnetic radiation trapped in a cavity of volume V . The radiation is in thermal equilibrium, at temperature T , with the atoms of the surface of the cavity. These atoms can absorb and emit quanta of radiations, called photons. It was precisely to interpret the radiation spectrum emitted by such a cavity (called a “black body”) that

Max Planck, in 1900, introduced the concept of quantization of the energy of electromagnetic waves. The plane wave modes are characterized by a wavevector \mathbf{k} and an angular frequency ω , related by:

$$\omega = c|\mathbf{k}| \quad (399)$$

where c is the speed of light in vacuum ($c \approx 3 \times 10^8 \text{ m s}^{-1}$).

Planck's quantization hypothesis states that the energy of the electromagnetic waves of frequency ω can take on values:

$$E = n\hbar\omega, \quad n = 0, 1, 2, \dots \quad (400)$$

With each quantum of energy, one may associate a "particle" of energy $\hbar\omega$ and a momentum $\hbar\mathbf{k}$, the photon. Photons are bosons of zero mass and spin 1, which however, have only 2 (rather than 3) projections along the momentum, associated with the 2 polarization states ϵ . Due to the absorption and emission of photons by the cavity walls, their number is not conserved. Since the electromagnetic modes (\mathbf{k}, ϵ) are independent (non-interacting), we may calculate the partition function of each mode separately. Labeling each mode (\mathbf{k}, ϵ) by the index α , we find:

$$q_\alpha = q_{\mathbf{k},\epsilon} = \sum_{n=0}^{\infty} \exp(-n\beta\hbar\omega_\alpha) \quad (401)$$

$$= \frac{1}{1 - \exp(-\beta\hbar\omega_\alpha)} \quad (402)$$

Let n_α be the number of photons in the mode α . The total energy of the electromagnetic radiation inside the cavity is:

$$E(\{n_\alpha\}) = \sum_{\alpha} n_\alpha \hbar\omega_\alpha \quad (403)$$

and the total partition function is given by the product over the partition function of individual modes:

$$Q = \prod_{\alpha} q_\alpha = \prod_{\alpha} \frac{1}{1 - \exp(-\beta\hbar\omega_\alpha)}. \quad (404)$$

The mean number of photons in mode α is:

$$\langle n_\alpha \rangle = \frac{1}{q_\alpha} \sum_{n=0}^{\infty} n \exp(-\beta n \hbar \omega_\alpha) \quad (405)$$

$$= -kT \frac{\partial}{\partial \hbar \omega_\alpha} \ln q_\alpha \quad (406)$$

$$= \frac{1}{\exp(\beta \hbar \omega_\alpha) - 1}. \quad (407)$$

Comparison with Eq.(375) show that photons obey BE statistics with zero chemical potential $\mu = 0$. Since $\mu = (\partial A / \partial N)_{V,T}$, this implies that the free energy is a minimum with respect to the particle number: the number of photons adjusts itself to minimize the free energy. The total number of photons is not conserved.

Planck's radiation law. The mean value of the total energy of the radiation inside the cavity:

$$\langle E \rangle = \sum_{\alpha} \hbar \omega_{\alpha} \langle n_{\alpha} \rangle \quad (408)$$

For a macroscopic volume V , the sum over modes may be replaced by an integral, according to the prescription:

$$\sum_{\mathbf{p}, \epsilon} \rightarrow 2 \times \frac{V}{h^3} \int d\mathbf{p} \quad (409)$$

(Recall that the volume of an elementary cell in phase space, $d\mathbf{r} \times d\mathbf{p} \sim h^3$ [see (123)] so that in a interval $d\mathbf{p}$ of momentum space, there are $\sim (V/h^3) \times d\mathbf{p}$ states for a system of volume V . The factor of two in (409) is arises due to the two states of polarization of light). Therefore, we have:

$$\langle E \rangle = \frac{V}{4\pi^3} \int d\mathbf{k} \frac{\hbar \omega}{\exp(\beta \hbar \omega) - 1} \quad (410)$$

$$= \frac{V}{4\pi^3} \int_0^{\infty} \frac{\hbar \omega}{\exp(\beta \hbar \omega) - 1} 4\pi k^2 dk \quad (411)$$

$$= \frac{V}{c^3 \pi^2} \int_0^{\infty} \frac{\hbar \omega^3}{\exp(\beta \hbar \omega) - 1} d\omega \quad (412)$$

The energy density per unit volume may be written as an integral over all frequencies:

$$\frac{\langle E \rangle}{V} = \int_0^{\infty} u(\omega, T) d\omega \quad (413)$$

where $u(\omega, T)$ is given by:

$$u(\omega, T) = \frac{\hbar}{c^3 \pi^2} \frac{\omega^3}{\exp(\beta \hbar \omega) - 1} \quad (414)$$

As a function of ω , $u(\omega, T)$ is peaked at a frequency ω_{max} which is easily shown to scale linearly with T as:

$$\omega_{max} \approx 2.822 \times \frac{kT}{\hbar} \quad (415)$$

(Hint: differentiate $u(\omega, T)$ wrt ω , and make substitution $x = \beta \hbar \omega$ to obtain an equation in x , which can be numerically solved). At room temperature (300 K), $\omega_{max} \approx 10^{14}$ Hz, corresponding to infra-red radiation. At the surface of the sun, $T = 6000$ K, corresponding to visible light.

The integral (412) can be further re-written in terms of x as:

$$\frac{\langle E \rangle}{V} = \frac{k^4 T^4}{\pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^3}{e^x - 1} dx \quad (416)$$

$$= \frac{k^4 T^4}{\pi^2 c^3 \hbar^3} \times \frac{\pi^4}{15} \quad (417)$$

$$= \frac{\pi^2 k^4}{15 c^3 \hbar^3} \times T^4 \quad (418)$$

i.e. the energy of the electromagnetic radiation inside the cavity increases like T^4 , and the specific heat varies like T^3 .

As an exercise, you can show (see problems) that the Helmholtz free energy and pressure go as:

$$\frac{A}{V} = -\frac{1}{3} \frac{\langle E \rangle}{V} = -\frac{\pi^2 k^4}{45 c^3 \hbar^3} \times T^4 \quad (419)$$

$$P = +\frac{1}{3} \frac{\langle E \rangle}{V} = \frac{\pi^2 k^4}{45 c^3 \hbar^3} \times T^4 \quad (420)$$

E. The ideal Fermi-Dirac gas

The basic relations for the FD distribution have been given above c.f. Eq.(381). Due to the Pauli exclusion principle, the energy levels can be at most occupied by $g_s = (2s + 1)$ fermions, for fermions with spin s . For electrons and protons $s = 1/2$, and $g_s = 2$.

At $T = 0$, the N non-interacting fermions will fill the lowest (N/g_s) energy levels, starting with the ground-state $\epsilon_{\mathbf{p}} = p^2/(2m) = 0$. The energy of the highest occupied orbital is the **Fermi energy**, ϵ_F , determined by:

$$\langle N \rangle = \int_0^{\epsilon_F} g(\epsilon) d\epsilon \quad (421)$$

where $g(\epsilon)$ is the density of states (the number of states in the energy interval $[\epsilon, \epsilon + d\epsilon]$). You can show that:

$$g(\epsilon) = g_s \times \frac{2\pi V}{h^3} \times (2m)^{3/2} \times \epsilon^{1/2} \quad (422)$$

Substitution of (422) into (421), yields:

$$\epsilon_F = \frac{h^2}{2m} \left(\frac{3}{4\pi g_s} \right)^{2/3} \times n^{2/3} \quad (423)$$

where n is the number density of particles. One may associate a Fermi momentum for ϵ_F , via:

$$\epsilon_F = \frac{p_F^2}{2m} \rightarrow p_F = h \left(\frac{3}{4\pi g_s} \right)^{1/3} n^{1/3} \quad (424)$$

At $T = 0$, all occupied energy levels have $\epsilon \leq \epsilon_F$, so that $|\mathbf{p}| \leq p_F$: in momentum space all quantum states are occupied inside the Fermi sphere of radius p_F . The internal energy at $T = 0$ is:

$$\langle E \rangle = \int_0^{\epsilon_F} \epsilon g(\epsilon) d\epsilon = \frac{3}{5} N \epsilon_F \quad (425)$$

where you should make sure you can derive the second equality. The $T = 0$ pressure and chemical potentials similarly follow:

$$P_0 = - \left(\frac{\partial \langle E \rangle}{\partial V} \right)_{T=0} = \frac{2}{3} \frac{\langle E \rangle}{V} \quad (426)$$

$$\mu_0 = \left(\frac{\partial \langle E \rangle}{\partial N} \right)_{T=0, V} = \epsilon_F \quad (427)$$

The $T = 0$ Fermi-Dirac distribution becomes a step-function:

$$f_{T=0}(\epsilon) = \begin{cases} 1 & \epsilon < \epsilon_F \\ 0 & \epsilon > \epsilon_F \end{cases} \quad (428)$$

One may associate a Fermi temperature T_F with the Fermi energy:

$$kT_F = \epsilon_F \quad (429)$$

Comparison of (423) with (388) shows that T_F equals the degeneracy temperature T_d to within a numerical factor of order 1. Thus the ideal Fermi gas is degenerate at temperatures below T_F . Fermi temperatures are typically very high. Consider the case of the conduction electrons of Cu. Here $n \approx 8.5 \times 10^{22} \text{ cm}^{-3}$, for which we find $T_F \approx 8.2 \times 10^4 \text{ K}$, far exceeding room temperature. The mean energy per electron, $3\epsilon_F/5 = 3kT_F/5$ far exceeds the thermal energies.

The finite temperature corrections to the thermodynamic quantities can be obtained via an expansion of the FD distribution around the Fermi energy. It can be shown that the second order correction to μ is:

$$\mu = kT_F \left[1 - \frac{\pi^2}{12} \left(\frac{T}{T_F} \right)^2 + \dots \right] \quad (430)$$

whilst the internal energy goes as:

$$\langle E \rangle = \frac{3N}{5} kT_F \left[1 + \frac{5\pi^2}{12} \left(\frac{T}{T_F} \right)^2 + \dots \right] \quad (431)$$

leading to specific heat which is *linear* in T :

$$C_V = \frac{\pi^2}{2} Nk \times \frac{T}{T_F} \quad (432)$$

A linear behaviour of C_V at low temperatures is well obeyed experimentally. It should be contrasted with the contribution from lattice vibrations, which vary as $(T/\Theta)^3$, where Θ is a vibrational temperature. As a result, at low temperatures $T < \Theta$, the specific heat of metals is dominated by the electronic contribution. The measured experimental slope is, however, systematically larger than the theoretically predicted $\pi^2 Nk/(2T_F)$. This can be accounted for phenomenologically by associating an effective mass m^* with the electrons, which enters the Fermi energy ϵ_F . Typically $1 < m^*/m \leq 2$, and is a consequence of the interactions between electrons.

Problems

1. Verify expression Eq.(12) for the translational partition function of a particle of mass m in a box of dimension L^3 at a temperature T . Evaluate (a) the Helmholtz free energy, (b) the internal energy, (c) the entropy and (d) the pressure. How does the pressure increase at small V ? Compare this to the expected variation of the pressure assuming the translational energy levels Eq.(11) and comment on your answer.
2. At what temperature does the de Broglie wavelength of He-4 become comparable to the mean-interparticle spacing in the liquid (the density of liquid helium at its boiling point is 0.12 g cm^{-3})? Compare it to that of Ar at its triple-point conditions. What do you conclude regarding a statistical mechanical treatment of the two fluids?
3. Consider a system of 3 *indistinguishable* particles, each of which can occupy any of 3 energy levels ($0, \epsilon, 2\epsilon$). By enumerating the number of possible states $\Omega(E)$ for each value of the total energy E of the system, show that the partition function at temperature T is:

$$\begin{aligned} Q_3 &= \sum_E \Omega(E) \exp(-E/kT) \\ &= 1 + e^{-\beta\epsilon} + 2e^{-2\beta\epsilon} + 2e^{-3\beta\epsilon} + 2e^{-4\beta\epsilon} + e^{-5\beta\epsilon} + e^{-6\beta\epsilon} \end{aligned}$$

where $\beta = 1/kT$.

Calculate the entropy S of the system, and derive its value in the two limits $T \rightarrow \infty$ and $T \rightarrow 0$.

Next calculate the molecular (or single-particle) partition function q . Assuming that Q_3 can be approximated by the ‘Boltzon’ expression $q^N/N!$, obtain the entropy and compare the high and low temperature limits with their exact values.

4. Verify expression Eq.(18) for the vibrational partition function of a harmonic oscillator. Generalize your result to 3D harmonic oscillator. What is the pressure in a system consisting of a harmonic oscillator?

5. In the canonical ensemble, $P_m = e^{-\beta E_m} / Q_N(V, T)$. Prove that

$$\left(\frac{\partial S}{\partial \langle E \rangle} \right)_{V, N} = k\beta$$

and hence deduce that the Lagrange multiplier β is equal to $1/kT$.

6. The microcanonical ensemble. A real isolated system (N, V, E fixed) can be represented by an ensemble in which all systems have the same energy. By maximizing S , show that the ensemble probability distribution P_m is independent of m , and that the microcanonical partition function is $W(N, V, E)$ (the total number of microstates of the system).
7. Fluctuations in the canonical ensemble. Prove that:

$$\langle E^2 \rangle - \langle E \rangle^2 = kT^2 C_V$$

Hence show that the root mean-square energy relative to the mean value is of the order $1/N^{1/2}$. What happens in the limit $N \rightarrow \infty$? How can the narrowness in the spread be accounted for?

Hint: consider $(\partial \langle E \rangle) / (\partial \beta)_{N, V}$.

8. By applying Newton's law to the classical Hamiltonian, prove Eq.(121), i.e. show that an isolated classical system conserves energy.
9. A one-dimensional harmonic oscillator has a classical Hamiltonian given by:

$$H_1 = \frac{p^2}{2m} + \frac{1}{2} kx^2$$

Write down Newton's equation of motion for this system, and solve it explicitly for the given initial conditions $x(t = 0) = x_0$ and $p(t = 0) = p_0$. Show that the corresponding trajectory in two-dimensional phase-space is an ellipse. Calculate the total energy of the oscillator.

10. (For the mathematically inclined). When dealing with the microcanonical ensemble, it is sometime convenient to define a function $\Phi(E)$:

$$\Phi(E) = \frac{1}{N! h^{3N}} \int_{H_N \leq E} d\mathbf{r}^N d\mathbf{p}^N$$

i.e the volume enclosed by the hypersurface $H_N = E$. For the case of an ideal gas, show that:

$$\Phi(E) = C_{3N} \frac{V^N}{N!} (2mE)^{3N/2}$$

where $C_{3N} = \pi^{3N/2} / \Gamma(3N/2 + 1)$, $\Gamma(x + 1) = x\Gamma(x)$ being the Gamma function.

The classical density of states $\Omega(E)$ is given by:

$$\Omega(E) = \frac{\partial \Phi(E)}{\partial E}$$

Obtain Ω , and hence show that the the microcanonical entropy $S(E, V, N)$ is given by:

$$S(E, V, N) = Nk \left(\ln \left[\frac{V}{N\Lambda^3} \right] + \frac{5}{2} \right) + O(\ln(N))$$

[You can assume (although you may like to show it!) that the volume of a $3N$ dimensional sphere of radius $R = \sqrt{2mE}$ is $C_{3N} \times R^{3N}$, where C_{3N} is given above. The Gamma function generalizes the factorial to non-integer arguments. In particular $\Gamma(1/2) = \sqrt{\pi}$.]

11. Verify the internal energy E , entropy S , chemical potential μ , and pressure P given in Eqs.(148,149,150) for an ideal gas.
12. Calculate the entropy of mixing of two ideal gases made up of different molecular species. Initially the two gases, containing respectively N_1 molecules of species 1 (mass m_1) and N_2 molecules of species 2, mass m_2 , are enclosed in two separate compartments of volume V_1 and V_2 respectively, separated by a wall. Calculate the total entropy S_i (initial entropy). The wall is next removed so that the two gases can mix, and occupy the total volume $V_1 + V_2$. After a short equilibration time, the entropy reaches it final value, S_f . Compute $\Delta S = S_f - S_i$. Show that is it always positive and explain why. What happens if the two gases are of the *same* species?

13. Harmonic oscillators The Einstein model considers the N atoms of a crystal to vibrate independently of each other around N lattice sites. The vibrational frequency is assumed to be the same for each of the $3N$ oscillators (3 oscillators being associated with each site, along the x, y, z directions). The total Hamiltonian is

$$H_N = \sum_i^{3N} \left[\frac{p_i^2}{2m} + \frac{m\omega^2 x_i^2}{2} \right]$$

Assuming the oscillators are distinguishable (and classical), calculate Q_N . Hence deduce the Helmholtz free energy, and internal energy. Explain why the free energy is independent of the crystal volume, and show that the specific heat is given by:

$$C_V = 3Nk \quad [\text{Dulong-Petit Law}]$$

Do you expect this to hold at low temperatures? Calculate the kinetic and potential energy contributions to the internal energy. Show that each of the $3N$ degrees of freedom contribute kT to the internal energy, half of which comes from the kinetic and half from the potential energy.

14. Show that for a system containing molecules of a single species, that their chemical potential can be expressed as:

$$\mu = -kT \ln \left(\frac{Z_{N+1}}{Z_N} \frac{1}{(N+1)\Lambda^3} \right)$$

Hence show that the ‘excess’ (i.e. non-ideal) part of the chemical potential, $\mu_{ex} = \mu - \mu_{id}$ can be cast as:

$$\mu_{ex} = -kT \ln \left(\frac{Z_{N+1}}{V \times Z_N} \right)$$

15. Consider a paramagnetic crystal made up of N ions of spin-1/2, carrying a magnetic moment m , fixed at the sites of a lattice, and subjected to an external magnetic field B aligned along the z -axis. Each m_i can have only two projections $\pm m$, along the z -axis. The ions are assumed to be distinguishable, and that the interaction between them can be neglected.

(a) Write down an expression for the total energy E , and total magnetic moment M , in terms of the numbers N_+ and N_- of moments parallel and antiparallel to B . Give the total number of microstates Ω_t , of the N spins.

(b) Calculate the number of microstates $\Omega(E)$ accessible for a given E , and hence obtain an expression for the resulting entropy $S(N, E, B)$, using Stirling's formula. Show that S is extensive and positive. Show that for fixed N, B , $S(E)$ goes through a maximum, and make a schematic plot of $S(E)$.

(c) Calculate the temperature T of the crystal as a function of E . Derive from this the fraction of spins $n_{\pm} = N_{\pm}/N$ parallel and antiparallel to B . Plot $1/T$ as a function of E , and show that the temperature can become negative over a specified interval of energies. What is the explanation of this unphysical behaviour?

(d) Compute next the canonical partition function, Q_N , at a temperature T , and use it to derive the Helmholtz energy A , internal energy E , and specific heat C_V . Show that the relation between E and T is compatible with that found in the microcanonical ensemble (c). Plot schematically the variation of E and C_V with T .

(e) Calculate the mean magnetization of the crystal $\langle M \rangle$, and derive from it the magnetic susceptibility χ

$$\chi = \left(\frac{\partial \langle M \rangle}{\partial B} \right)_{V, T}$$

Plot schematically the variation of $\langle M \rangle$ and χ with B at constant T . Derive the temperature dependence of the zero-field susceptibility $\chi_0 = \lim_{B \rightarrow 0} \chi$ (Curie's law).

16. Fluctuations in the grand-canonical ensemble. Prove that:

$$\langle N^2 \rangle - \langle N \rangle^2 = kT \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{V, T}$$

This result (and that of problem 7) refer to the distribution of N or E over

imaginary systems of the ensemble. What is their meaning for the behaviour of real systems?

17. The maximum-term method All partition functions are sums (or integrals) over states. In most cases the log of the partition function can be, with negligible error, be replaced by the log of the largest term in the sum. Confirm that this is true in the case for which

$$\Xi(M, z) = \sum_{N=0}^M C_N$$

with

$$C_N = M!z^N/N!(M - N)!$$

where z is a fugacity parameter. How does $\log \Xi$ compare with $\log C_{N^*}$ where N^* is the value of N which maximizes C_N . Is there anything surprising about your result?

Hint: it is easier to find the largest $\log C_N$ rather than largest C_N .

18. Equivalence of ensembles (1). Consider an arbitrary open system with grand partition function $\Xi = \sum_N C_N$, and let the largest term in the sum be C_{N^*} . Show (a) that $N^* = \langle N \rangle$, and (b) that the fundamental relation for the grand-canonical ensemble ($PV = kT \log \Xi$) reduces to the canonical result ($A = -kT \log Q_{N^*}$) if the maximum-term approximation is used.
19. Equivalence of ensembles (2). The canonical partition function can be written in the form:

$$Q_N(V, T) = \sum_E W(N, V, E)e^{-\beta E}$$

where W is the microcanonical partition function and the sum is over energy levels (not over states). If E^* is the energy corresponding to the largest term of the sum, show that use of the maximum term method leads to the expression

$$S = k \log W(N, V, E^*)$$

What implication does this result have for the validity of the Boltzmann definition of the entropy?

20. Maxwell distribution of momenta. By integrating over the classical probability density $\rho(\mathbf{r}^N, \mathbf{p}^N)$ over unwanted variables, show that, if there are no external fields, in the canonical ensemble the probability that a given particles lies in \mathbf{r} to $\mathbf{r} + d\mathbf{r}$ with momentum in the range \mathbf{p} and $\mathbf{p} + d\mathbf{p}$ is independent of \mathbf{r} and given by:

$$\rho_1(\mathbf{r}, \mathbf{p})d\mathbf{r}d\mathbf{p} = \frac{\rho}{(2\pi mkT)^{3/2}} e^{-p^2/2mkT} d\mathbf{r}d\mathbf{p}$$

where $\rho = N/V$. What is the normalization of ρ_1 ? Is this result equally applicable to gases and liquids?

21. Ideal gas in the canonical ensemble. Use the thermodynamic relation $\mu = (\partial A/\partial N)_{V,T}$ to show that the chemical potential of an ideal gas is:

$$\mu = kT \log(\rho) + 3kT \log(\Lambda)$$

where $\rho = N/V$. What are the values for z and Ξ (the grand partition function) for the ideal gas? What happens to μ as $V \rightarrow \infty$?

22. Show that the internal energy $\langle E \rangle$ is given in the GCE by the derivative:

$$\langle E \rangle = - \left(\frac{\partial \ln \Xi}{\partial \beta} \right)_{z,V} \quad (433)$$

23. Adsorption of a gas at the surface of a crystal. A simple model of gas adsorption on an ordered surface containing M adsorption sites organized in a square lattice. The molecules of a gas can be adsorbed onto any one of the sites, each site being able to receive up to one molecule (excluded volume effect). An occupied site corresponds to a binding energy ϵ . The adsorbed gas is assumed to be in thermal equilibrium at the temperature T . The molecules are indistinguishable.
- (a) If N molecules are adsorbed, calculate the canonical partition function $Q_N(M, N, T)$.

(b) Compute the grand partition function $\Xi(z, M, T)$, where z is the fugacity. For a two-dimensional system (as an adsorbed layer) the area A and surface pressure Π play the same role as the volume V and pressure P in a 3D system. Express Π as a function z, A, T . Calculate $\langle N \rangle$, and the occupied fraction $f = \langle N \rangle / M$. From these, derive the equation of state $\Pi(f, T)$. Show that the results reduce to the ideal gas limit only in the limit of low coverage.

24. The virial coefficient of the square-well model A somewhat more realistic model of an interatomic potential than the HS model crudely accounts for the attractive part by adding to the hard-core repulsion a square well of depth ϵ and width $\alpha\sigma$ ($\alpha > 1$), i.e.:

$$v(r) = \begin{cases} \infty; & r < \sigma \\ -\epsilon; & \alpha\sigma > r > \sigma \\ 0; & r > \alpha\sigma \end{cases} \quad (434)$$

Calculate $B_2(T)$ for this model. Show that $B_2 < 0$ at low temperatures and goes over to the hard-sphere result as $T \rightarrow \infty$. Determine the Boyle temperature, T_B , at which B_2 changes sign, as a function of the potential parameters ϵ, σ, α . Calculate this temperature for methane by assuming these parameters have values identical to those appropriate to the L-J potential and $\alpha = 1.5$. Show that at high temperatures ($T^* = kT/\epsilon \gg 1$), $B_2(T)$ behaves as:

$$B_2(T) \approx b - \frac{c}{T} \quad (435)$$

and give expressions for b and c . Give a physical argument why $B_2(T)$ is expected to be negative at low temperatures ($T^* < 1$). Which phenomenon will this eventually lead to?

25. Using the parameters a and b appropriate for N_2 , calculate the EOS $P(v)$ (in Pascals) as a function of volume per molecule (in nm^3) along the isotherms $T = 200$ K and 100 K. Where do you expect the critical point T_c to be?
26. Calculate the entropy S of the Ising model near the Curie temperature T_c in zero external field $B = 0$, using the mean-field expression for the free energy.

Calculate S for $T > T_c$ and $T < T_c$, and show that it is continuous across the transition. Derive the zero-field specific heat C_V from the entropy, and show that it exhibits a jump discontinuity at T_c . [Such a discontinuity in C is the signature of a second-order phase transition].

27. By expanding PV/kT and $\langle N \rangle$ to second order in z , show that the equation of state of a weakly degenerate BE or FD gas is:

$$\frac{PV}{NkT} = 1 \mp \frac{n\Lambda^3}{2^{5/2}}$$

where the $-ve$ and $+ve$ signs refer to BE and FD statistics respectively. On purely statistical grounds, the bosons appear to attract each other. What is the physical reason for this behaviour?

28. Verify Eqs.(419,420) for the Helmholtz energy density and pressure of a photon gas (black body radiation) at temperature T .
29. Show that the density of states $g(\epsilon)$ for the translational energy levels in a periodic box of volume V is:

$$g(\epsilon) = \frac{2\pi V}{h^3} \times (2m)^{3/2} \times \epsilon^{1/2}$$

30. Show that for Cu at room temperature, the contribution to C_V of the electrons is far smaller than that due to lattice vibrations, which can be estimated using the Dulong-Petit law (see an earlier problem).