Contents

Chapter 1

What is Statistical Mechanics?

1.1 A few questions

We begin by posing a few questions

- Why are physical properties of steam, water and ice different, though they consist of the same kind of molecule (water)? Why is it that most pure substances show such 'phases', with similar qualitative properties, even though the microscopic constituents are very different?
- What is the origin of ferromegnetism, and why are some aspects of ferromagnetism so similar to those of transitions between phases of water?
- Why is there an 'arrow of time'? Even though microscopic physics is time symmetric (barring exceptions of some quantum mechanical phenomena), why does time seem to 'flow' from past to future?
- What is information, and how is it related to entropy (which you have encountered in thermodynamics)?
- Can all this be understood as consequences of microscopic 'fundamental' laws of Physics, or are these phenomena a consequence of entirely new kinds of laws?

In this course, we will try to address these questions. We will observe that so long as a physical system consists of a small number of constituents, microscopic laws govern dynamics of the system. However, as the number of constituents becomes larger, a completely new set of statistical laws emerge, which are apparently insensitive to the microscopic physics. Further, these new laws indicate an arrow of time, even though such an 'arrow' is absent in the microscopic laws.

1.2 Classical Dynamics

We briefly review classical, Newtonian Physics. The idea is to try and see if an 'arrow of time' emerges from time-symmetric classical laws, as the number of particles becomes 'large' [1]. Later in the course, we will try to see how such a behaviour might emerge from Quantum Mechanics.
Consider a system of N interacting particles whose dynamics is governed by Newton's Laws

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = -\vec{\nabla}_i U \tag{1.1}$$

where $U = \sum_{ij} U_{ij}$ is the total potential energy of the system, equal to the sum over the interaction potential energy of pairs of particles. Given the position and velocity of all the particles at any one

---

[1] By large, we usually mean of the order of Avogadro number, though much smaller systems will typically still show a time-asymmetric behaviour

instant of time, say, $t = 0$, these equations, *in principle*, allow us to uniquely predict these at any other instant of time t, for both $t > 0$ and $t < 0$. That is, given $\vec{r}_i(0)$ and $\vec{v}_i(0) = \dot{\vec{r}}_i(0)$, we can compute $\vec{r}_i(t)$ and $\vec{v}(t) = \dot{\vec{r}}_i(t)$ for $t > 0$ or $t < 0$. Further, these equations are invariant under $t \to -t$. This implies the following: say, we start with a configuration $\vec{r}_i(0), \vec{v}_i(0)$. This evolves to a new configuration $\vec{r}_i(\tau), \vec{v}_i(\tau)$ at instant $t = \tau$ (for definiteness, we assume here that $\tau > 0$). Now, say we start with the configuration $\vec{r'}_i(0) = \vec{r}_i(\tau), \vec{v'}_i(0) = -\vec{v}_i(\tau)$. That is, our new initial configuration is the same as the previous configuration at time τ, except *with velocities of all the particles reversed.* The time-symmetric nature of the equations implies the following: starting with this new initial configuration at $t = 0$, the configuration we obtain at time τ will be the same as the previous configuration at $t = 0$, but again with *all velocities reversed*

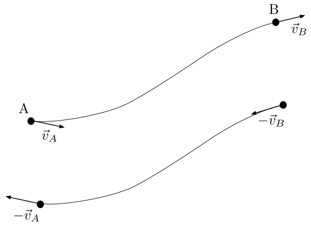


Figure 1.1: Time-Reversal symmetry in Classical Physics

In other words, if we were to record a motion picture of the first evolution and run it backward, we will observe the second evolution. This has the implication that if we were to record the time evolution of a classical system of particles and run it backward, the resulting time-reversed motion picture is realizable in nature. This leads to the following puzzle: if we drop a glass on the floor, it shatters into pieces that eventually come to rest on the floor. If we were to record a motion picture and run it backward, we would see the shattered pieces come together, form the glass, and rise up to our hand. Then, this should be a physically realisable process. However, we never see it happen. We need to be a little careful here. When discussing the time-symmetric nature of Newton's laws, we analysed an isolated system of particles (this is why there exists a conserved energy and a potential energy function for the system). In the 'shattering of glass' example, the glass is not an isolated system. Even if we ignore effects due to air friction, the glass interacts with the floor when it shatters. If we are to construct a time-reversed motion picture, we should also include the dynamics of the floor, which further consists of atoms. When the glass shatters, it loses momentum and energy to the particles constituting the floor, which take away this momentum and energy. One can visialise it as a wave spreading out on the floor, with atoms/molecules of the floor vibrating about their equilibrium positions. If we include this in our motion picture and then run it backward, we should see this wavefront converge, lose momentum and energy to the shattered pieces, till they come together, get 'glued' (atoms form bonds once again to recreate the original solid structure), and rise up to our hand. Why is it that we never see this happen? We shall see that at some level, it can be explained away 'easily' using ideas of probability. However, if we probe deep enough, there will still remain a mystery.

1.3 Phase Space

To resove this paradox, we need a convenient way to visualise the state of a system of N particles. The state is completely described if we specify the position and velocity (equivelently, momentum) of each particle. One way to visualise the system is to construct a 6-dimensional space, with three 'coordinate' axes reserved for position coordinates and three for momentum components. Then, the state of the system at any instant (complete specification of these six numbers for each particle) will be represented as N points in this space. This swarm of points will move about in this space with time, under Newtonian dynamics. However, a much more convenient way to visualise the system is to instead construct a $6N$

dimensional space, with $6N$ ‘coordinate’ axes, $3N$ for position coordinates and another $3N$ for momentum components for all the particles. Then, the state of the system will be a single point in this space. This ‘space’ corresponding to the state of the system is known as *Phase Space*

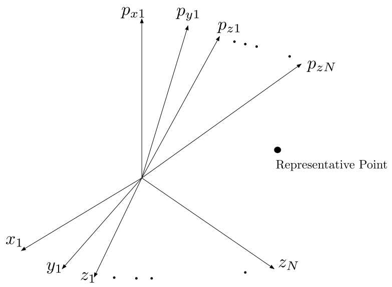


Figure 1.2: Phase Space

As the system evolves, this point will move around in this space (for convenience, we will collectively display the momentum components as ‘p’ and position coordinates as ‘x’)

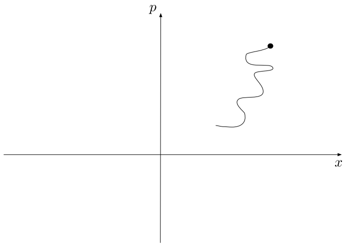


Figure 1.3: Motion of representative point in Phase Space

The region of phase space available for the representative point will depend on the (conserved) energy of the system and the volume available to the system. For instance, for a system of identical particles of mass m constrained within a cubical box of volume V, the position coordinates are constrained to lie within this volume. Further, if we assume that the box is (practically) infinitely massive compared to the system and further assume that the collisions of the partiles with the walls of the box are elastic, then the mechanical energy of the system of particles is conserved. If the system starts out in a configuration such that the total energy is E, then the position coordinates and the momentum components are constrained by the equation

$$\sum_{i=1}^{3N} \frac{p_i^2}{2m} + U\left(q_1, q_2, .., q_{3N}\right) = E \tag{1.2}$$

where we use a convenient notation in which the $3N$ position coordinates of all the particles are labelled $q_1, q_2, .., q_{3N}$ and momentum components lablelled $p_1, p_2, .., p_{3N}$. Then, $q_1 = x_1, q_2 = y_1, q_3 = z_1, q_4 = x_2, q_5 = y_2, ..., q_{3N} = z_N$ and $p_1 = p_{x1}, p_2 = p_{y1}, p_3 = p_{z1}, p_4 = p_{x2}, p_5 = p_{y2}, ..., p_{3N} = p_{zN}$. Equation (1.5) is a single equation in $6N$ variables, whcih, when visualised in the $6N$ dimensional phase space, will result in a $6N - 1$ dimensional ‘hypersurface’ on whcih the representative point moves

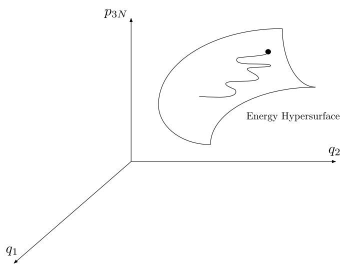


Figure 1.4: Energy Hypersurface

1.4 Microscopic vs Macroscopic

An exact specification of position and momentum of every particle of the system defines what is known as a *Microstate* if the system. The representative point in phase space corresponds to a microstate, since given the location of the point, we know the exact position and momentum of all the particles of the system. However, for a macroscopic system consisting of a large number of particles, it impossible to know the precise position and momentum of every particle. Fortunately, one is not interested in such a detailed description of the system. Instead, what one might be more interested in, for example, could be the spatial density distribution of the particles, and how much momentum such a distribution carries at a given point. In such a description, one effectively treats the system as a fluid-like entity, without worrying about the 'graininess' of the system. It should be noted that in this description, one is not interested in *which* particle occupies a given point and how much momentum it carries. If we could somehow label all the particles, then given any such fluid-like configuration in which at different points in space there is some density of particles, with particles in the neighborhood of that point carrying some momentum, if we just interchanged particle labels (which would change the microstate), we would not change the macroscopic state of this fluid. In other words, given such a macroscopic distribution, there are *many* (we will soon learn to quantify 'how many') microstates corresponding to the same macroscopic description, each differing by a simple interchange of particle between any two points in space (it is assumed when such an interchange occurs, the momentum of the particle occupying a given spatial point remains unchanged, leaving the momentum distribution of the fluid unchanged). Such a macroscopic description of the system is known as a *Macrostate* of the system. What one is interested in is how such a macrostate evolves with time, given the macrostate at any one instant. Then, as soon as we shift our focus from the microstate to the macrostate, we are losing information (we could think of the specification of position and momentum of each particles as encoding information). Given this ignorance about the precise microstate, we are left with a *region* of phase space in which the representative point of the system *could* lie. For definiteness, say we have a gas confined to a container of volume V_0 at some time t_0. Let the energy of the gas atoms be E. The volume occupied by the gas and its total energy specifies a macrostate M_0 of the gas, corresponding to which there can be many microstates. Then, the representative point of the system at t_0 could lie in a region Γ_{M_0} of the phase space, where each point within this region corresponds to a different distribution of positions and velocities of the atoms, such that the volume they occupy is V_0 and their energy being E

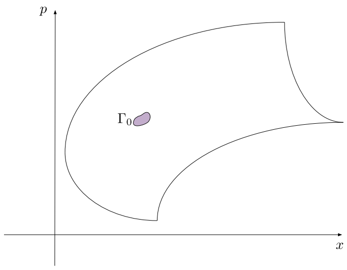


Figure 1.5: Initial Macrostate

Now, say this container is placed within a larger container of volume $V > V_0$, and is opened, so that the gas can escape into the larger container. Experience tells us that the gas *will* expand and eventually homogeneously occupy the larger available volume. However, nothing in the microscopic dynamics of the gas particles prevents the gas contracting to fill a *smaller* volume within the original container (we can visualise another, yet smaller container within the container of volume V_0). For, since we always observe gases expanding spontaneously, if we could record a motion picture of this expansion and run it backward, it would show the gas contract, and will be a possible physical process because of the time-reversal symmetry of the dynamical equations (assumed to be classical). Then, why is it no matter how many times we conduct such an experiment, we will always observe the gas expand to fill the larger volume? Given the initial macrostate M_0, the system's reresentative point could be anywhere within the region Γ_{M_0}. That is, there are many microstates that the system could have been in. Say, the system was in microstate X_0. With time, this microstate will change to $X(t)$. This will also result in a change of the macrostate to $M(t)$ (the gas would occupy a different volume, etc.). Corrsepondingly, the region of phase space within which the representative point of the system could lie will change to $\Gamma_{M(t)}$

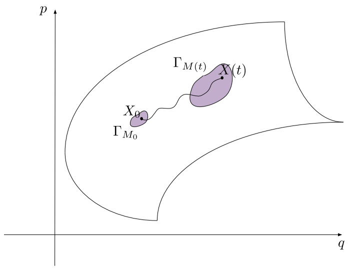


Figure 1.6: Evolution of Macrostate

The question we wish to address is whether this evolution will make the system of gas particles occupy a larger or smaller volume at time t. At this stage, it is useful to relate the 'size' Γ_M of the region of phase space corresponding to a macrostate to the physical volume V occupied by the gas particles. This size will just be the 'volume' of the phase space, given the constraints that all the $3N$ coordinates have to be constrained to lie within region V of physical space and that the total energy of the system is E. Let us assume for simplicity that the interactions between the particles can be ignored. Then, the $3N$ momenta are constrained by

$$\sum_{i=1}^{3N} \frac{p_i^2}{2m} = E \tag{1.3}$$

Then, the volume of he region of phase space will be

$$\Gamma_M = \int_{E,V} dq_1 dq_2 .. dq_{3N} dp_1 dp_2 .. dp_{3N} \tag{1.4}$$

where the subscript E, V in the integral implies the volume and energy constraints. The q and p integrals are independent, the q integral yielding V^N, since for any one particle,

$$\int dxdydz = V$$

Then

$$\Gamma_M = V^N \int_E dp_1 dp_2 .. dp_{3N}$$

At this stage, it is useful to assume that we do not have precise information about the energy of the system (there will always be some uncertainty in this measurement, limited by the measuring apparatus). Say, the energy of the system lies between E and $E + \delta E$. Then, the phase space integral is

$$\Gamma_M = V^N \int_{E,\delta E} dp_1 dp_2 .. dp_{3N}$$

where the subscript $E, \delta E$ implies the constraint

$$E < \sum_{i=1}^{3N} \frac{p_i^2}{2m} < E + \delta E \tag{1.5}$$

which can be written as

$$2mE < p_1^2 + p_2^2 .. + p_{3N}^2 < 2m(E + \delta E) \tag{1.6}$$

The relation (1.6) describes a spherical shell in a $3N$ dimensional space with inner radius $\sqrt{2mE}$ and outer radius $\sqrt{2m(E + \delta E)}$. Therefore, the momentum integral equals the volume of this spherical shell. To compute it, we need to know the expression for the volume of an n-dimensional sphere of radius R. This is given by [2]

$$\mathcal{V}_n(R) = \frac{\pi^{n/2}}{\Gamma(n/2+1)} R^n \tag{1.7}$$

Then, the volume of a spherical shell with radius between R and $R + \delta R$ is

$$\delta \mathcal{V}_n(R) = n \frac{\pi^{n/2}}{\Gamma(n/2+1)} R^{n-1} \delta R \tag{1.8}$$

The momentum integral is then

$$\begin{aligned} \int_{E,\delta E} dp_1 dp_2 .. dp_{3N} &= \delta \mathcal{V}_{3N}(\sqrt{2mE}) \\ &= \frac{\pi^{3N/2}}{\Gamma(3N/2)} (2m)^{3N/2} E^{3N/2-1} \delta E \end{aligned}$$

Then the volume of the 'accessible phase space' corresponding to macrostate with volume V and energy between E and $E + \delta E$ is

$$\Gamma_M = \frac{\pi^{3N/2}}{\Gamma(3N/2)} (2m)^{3N/2} V^N E^{3N/2-1} \delta E \tag{1.9}$$

As the microstate and macrostate of the system change, the energy, being conserved, does not change. However, the volume changes and eqn.(1.9) tells us that as a function of the physical volume V, the 'accessible phase space' volume changes *exponentially* with the number of particles of the system

$$\Gamma_M \propto V^N \tag{1.10}$$

We are now ready to analyse what happens when we open the container of volume V_0, within a larger container of volume V. Will the gas exapnd to fill volume V or contract to occupy a smaller volume?

---

[2] See appendix

When we open the container, the representative point of the system has more of phase space accessible to it, since the available physical volume is increased. This point is constantly moving around, so the question is, will it move such that at time $t > t_0$, it corresponds to a macrostate with volume $V > V_0$, rather than one with volume $V < V_0$?

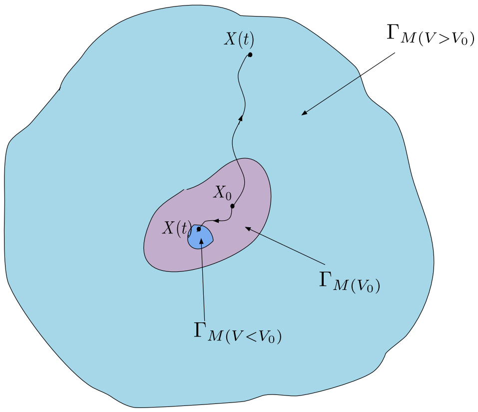


Figure 1.7: Possible Evolutions of Macrostate

The key here is the relative phase space volumes of regions $\Gamma_{M(V_0)}, \Gamma_{M(V<V_0)}$ and $\Gamma_{M(V>V_0)}$. From (1.10), it follows that

$$\frac{\Gamma_{M(V)}}{\Gamma_{M(V_0)}} = \left(\frac{V}{V_0}\right)^N \tag{1.11}$$

where for a macroscopic system, N could be of order N_A or more. If V is even a few percent smaller than V_0, for a macroscopic system, the ratio if vanishingly small. To see this, we take the (natural) logarithm of both sides of eqn(1.11), assuming that $V = V_0 - \delta V$, where $\delta V/V_0 << 1$

$$\begin{aligned}
\ln\left(\frac{\Gamma_{M(V)}}{\Gamma_{M(V_0)}}\right) &= N \ln\left(\frac{V_0 - \delta V}{V_0}\right) \\
&= N \ln\left(1 - \frac{\delta V}{V_0}\right) \\
&\simeq -N\frac{\delta V}{V_0}
\end{aligned}$$

Therefore,

$$\frac{\Gamma_{M(V)}}{\Gamma_{M(V_0)}} \simeq e^{-N(\delta V/V_0)} \tag{1.12}$$

Given $N \sim N_A$, this gives

$$\frac{\Gamma_{M(V)}}{\Gamma_{M(V_0)}} \simeq e^{-10^{23}(\delta V/V_0)} \tag{1.13}$$

No matter how small $\delta V/V$ is on a macrocopic scale, the 10^{23} in the exponent will make the ratio of phase space volumes, for all practical purposes, zero. Then, the tiniest decrease in the physical volume of the gas leads to the phase space volume of the resulting microstate practically vanish. Equivalently, the tiniest increase in physical volume makes the phase space volume of the new macrostate practcally infinitely larger than the original one. Now we see why the gas never contracts, but is always seen to expand. Given the relative sizes of phase space volume corresponding to the expanded and contracted state, any motion of the representative point is overwhelmingly likely to take it to the macrostate corresponding to expansion of the gas. Given the initial macrostate, for which there are many possible microstates, there *will* be microstates which will lead to the gas contract. But, the number of such microstates would be vanishingly

small, compared with the number of microstates which will lead to expansion. This is the idea of *typicality*, at the heart of the foundarions of Statistical Mechanics. *Typical* microstates will lead to expansion of the gas, and *atypical* ones to contraction. Further, an overwhelmingly large fraction of microstates for a given macrostate will be typical, with 'overwhelming' being a gross understatement for macroscopic systems [3]. Then, this is why we see an 'arrow of time': systems evolve preferentially in one direction because their evolution in the other direction would be overwhelmingly improbable. In fact, given the numbers we have just encountered, for all practical purposes, it will *never* happen.
This idea is at the heart of Statistical Mechanics: the idea of *typicality.* Similar arguments exist when one describes microscopic physics using quantum mechanics, which indeed is how microscopic phenomena should, in principle, be described. However, as compelling as these arguments seem, they are unfortunately incomplete. This is because, since there is in principle no preferred direction of time (in the dynamical equations, $t \to -t$ produces no change), we can flip this argument on its head by taking $t \to -t$ in the previous discussion. We will then conclude that given a macrostate of the gas in question, with probability practically unity, it must have evolved from a macrostate in which the gas occupied a *larger* volume. Then, we have the following interesting situation: given a macrostate of the gas, the probability that it evolves into the future such that the gas expands is practically unity, *and* the probability that this macrostate in turn evolved from one in the past in which the gas occupied a larger volume is also practically unity! These subtle issues point to a more non-trivial origin of the arrow of time. For a detailed discussion, see..............

1.5 Evolution towards Equilibrium

Given that the macrostate of a system of particles evolves such that it occupies increasingly larger volumes of phase space, when does this evolution stop? The macrostate at which the evolution stops is called the *Equilibrium Macrostate.* It is clear that this is the macrostate for which, consistent with external constraints (such as volume, etc.), the phase space volume is a maximum. Then, given constraints on a system, the problem of determining the equilibrium properties of the system reduces to an extremisation principle: the principal of maximisation of available phase space volume, subject to constraints. Since the phase space volume is an exponential function of such constraints (the exponent being of the order of Avogadro number), it is mathematically prudent to extremise its logarithm, which tames the sensitivity to any variation (an exponential function with such a large exponent is extremely sensitive to any changes in parameters, even the tiniest change potentially bringing aabout a substantial change in the function). Further, it is useful to visualise the phase space being divided into elementary *cells*, such that the volume can be measured in units of volume of such an elementary cell. It is useful to divide each coordinate into segments of size Δq and each momentum component into segments of size Δp, such that $\Delta q \Delta p = h$, h being an arbitrary, 'small' dimensional quantity (we will eventually identify it as Planck's constant, though here it is just an arbitrary, small quantity with dimensions of Planck's constant). Then, for a system of N particles, the elementary cell in phase space will have volume h^{3N}. The quantity Γ/h^{3N} is the volume of phase space in units of this elementary volume. It is also useful to think of it as the number of microstates accessible to the sustem for a given macrostate, where we assume that the entire region within any one elementary cell is counted as a single microstate. Extremising this ratio will be equivalent to extremising the phase space volume, with the added benefit that this quantity is dimensionless, so the logarithm of this quantity will be a pure number. We now observe that for a system of N identical particles, there is an apparent redundancy in the number of microstates estimated, since this counting includes a simple interchange of particles, all else being same. Since these particles are identical, such microstates are indistinguishable, in the sense that if we could somehow measure the position and momentum of every particle, then such a measurement will yield the same information for any microstates differing simply by

---

[3] The 'atypical' microstates are very sensitive to perturbation. The tiniest change is expected to render them 'typical'. This is easily demonstrated by simulating a system of elastically colliding billiard-balls in a rectangular box. Then, such microstates are isolated in phase space, with neighboring microstates being typical. This can be attributed to chaotic dynamics of the system. Chaos is often cited as the reason for almost all microstates being typical.

interchange of particles [4]. To account for this redundancy, we define the *Thermodynamic Probability* of macrostate M as the number of distinct microstates corresponding to M

$$\mathcal{N}_M = \frac{\Gamma_M}{h^{3N} N!} \tag{1.14}$$

The equilibrium macrostate is obtained by extremising the logarithm of the thermodynamic probability

$$\begin{aligned} I_M &= \ln \mathcal{N}_M \\ &= \ln \left(\frac{\Gamma_M}{h^{3N} N!} \right) \end{aligned} \tag{1.15}$$

Let us compute this for a system of N weakly interacting particles with energy E and occupying volume V. The accessible phase space volume is given by eqn.(1.9). Given that $N >>> 1$, the thermodynamic probability simplifies to

$$\mathcal{N}_M = \frac{1}{N!(3N/2)!} V^N \left(\frac{2m\pi E}{h^2} \right)^{3N/2} \delta E \tag{1.16}$$

Taking the logarithm, we get

$$\ln \mathcal{N}_M \;=\; N \ln \left[V \left(\frac{2m\pi E}{h^2} \right)^{3/2} \right] - \ln N! - \ln(3N/2)! + \ln \delta E$$

We now use *Sterling's Formula*

$$\ln n! \simeq n \ln n - n, \quad n >> 1 \tag{1.17}$$

Also, we can ignore the term $\ln \delta E$ compared with the other terms, which are all proportional to N. Finally, we get

$$I_M = N \ln \left[\frac{V}{N} \left(\frac{4m\pi E}{3h^2 N} \right)^{3/2} \right] + \frac{5}{2} N \tag{1.18}$$

This is an *extensive* quantity, in the sense that under transformations $N \to \lambda N, E \to \lambda E, V \to \lambda V$, $I_M \to \lambda I_M$.

Given such a system of particles, experience suggests that in equilibrium, the system will be distributed homogeneously in both number and energy density. Can we see this as a consequence of the extremisation of the function I_M? Let us place a partition within the volume V, such that it is divided into two parts with volumes V_1 and V_2 with N_1 particles trapped within V_1 and having energy E_1 and N_2 particles trapped in V_2, having energy E_2. If the system was in equilibrium before we put this partition, the quantities N_1, N_2, E_1 and E_2 should be such that the function I_M is a maximum. To compute I_M, we notice that since the two sub-systems are disjointed, the toal number of microstates accessible to the entire system will be a product

$$\mathcal{N}_M = \mathcal{N}_{M_1} \times \mathcal{N}_{M_2} \tag{1.19}$$

where $\mathcal{N}_{M_1}$ and $\mathcal{N}_{M_2}$ are of the form (1.16) with corresponding values of N, V and E. Then

$$\begin{aligned} I_M &= \log \mathcal{N}_M \\ &= \log \mathcal{N}_{M_1} + \log \mathcal{N}_{M_2} \\ &= N_1 \ln \left[\frac{V_1}{N_1} \left(\frac{4m\pi E_1}{3h^2 N_1} \right)^{3/2} \right] + N_2 \ln \left[\frac{V_2}{N_2} \left(\frac{4m\pi E_2}{3h^2 N_2} \right)^{3/2} \right] + \frac{5}{2} N_1 + \frac{5}{2} N_2 \\ &= N_1 \ln \left[\frac{V_1}{N_1} \left(\frac{4m\pi E_1}{3h^2 N_1} \right)^{3/2} \right] + N_2 \ln \left[\frac{V_2}{N_2} \left(\frac{4m\pi E_2}{3h^2 N_2} \right)^{3/2} \right] + \frac{5}{2} N \end{aligned}$$

---

[4] In *principle*, though, such mocrostates *are* distinct in classical physics, because even if particles are identical in intrinsic properties, they are distinguished by their trajectories leading up to a given microstate. In the quantum description, since there are no trajectories, the particles are, even in principle, indistinguishable. Here, we are introducing indistinguishability in an operational sense, so that we don't have to correct the estimation of microstates later when we use quantum mechanics.

I_M is to be extremised with respect to N_1, N_2, E_1, E_2, subject to the constraints $N = N_1 + N_2$ and $E = E_1 + E_2$. Therefore, only N_1 and E_1 are independent variables and we need to extremise with respect to them only. Then, while taking derivatives we need to use

$$\begin{aligned}\frac{\partial N_2}{\partial N_1} &= \frac{\partial(N - N_1)}{\partial N_1} \\ &= -1 \\ \frac{\partial E_2}{\partial E_1} &= \frac{\partial(E - E_1)}{\partial E_1} \\ &= -1\end{aligned}$$

Differentiating I_M with respect to N_1 gives

$$\frac{\partial I_M}{\partial N_1} = \ln\left[\frac{V_1}{N_1}\left(\frac{4m\pi E_1}{3h^2 N_1}\right)^{3/2}\right] - \ln\left[\frac{V_2}{N_2}\left(\frac{4m\pi E_2}{3h^2 N_2}\right)^{3/2}\right]$$

Extremisation with respect to N_1 implies that this is zero, which gives the relation

$$\frac{V_1}{N_1}\left(\frac{E_1}{N_1}\right)^{3/2} = \frac{V_2}{N_2}\left(\frac{E_2}{N_2}\right)^{3/2} \tag{1.20}$$

Differnetiation with respect to E_1 gives

$$\frac{N_1}{E_1} = \frac{N_2}{E_2} \tag{1.21}$$

Together, these equations give

$$\begin{aligned}\frac{N_1}{V_1} &= \frac{N_2}{V_2} \\ \frac{E_1}{V_1} &= \frac{E_2}{V_2}\end{aligned} \tag{1.22}$$

Which tells us that the number density and the energy density in the two halves must be the same. This argument can be recursively used within each such partition, so that eventually we see that in equilibrium, the system of particles must be homogenoeusly distributed.

Chapter 2

Statistical Thermodynamics

2.1 Entropy

We now analyse the physical significance of the quantity I_M. For a system of weakly interacting particles in equilibrium, we constructed I_M and observed that it is an extensive quantity. Further, if we partition such a system of particles into two subsystems and compute I_M, we saw that $I_M = I_{M_1} + I_{M_2}$. Further, if we free the constraints partitioning the system (we allowed N_1, N_2, E_1 and E_2 to vary), the extremisation of I_M with respect to the free parameters results in a configuration that is consistent with our intuitive understanding of equilibrium (the particles are distributed homogeneously). All these properties remind us of *entropy* in thermodynamics, which (a) is extensive (b) is maximised when a constrained is freed such that a new equilibrium configuration is attained. Then, we conjecture the following: given a system in a macrostate M, it has an entropy

$$S = k_B \ln \mathcal{N}_M \tag{2.1}$$

This entropy, which we call *Statistical Enropy*, is defined for all macrostates, not just equilibrium ones. This, then, is a more general definition of entropy compared with thermodynamic entropy, which is defined only for equilibrium configurations. In relating S to I_M, we have introduced a constant k_B such that it has the same dimensions as thermodynamic entropy. We will later identify k_B as Boltzmann constant. Given eqn.(2.1), we can easily demonstrate the additive nature of S for a system formed of n subsystems. Let there be n distinct systems (with possibly different constituents) forming a single systems. Let these systems be in macrostates $M_1, M_2, ..M_n$, such that there are $\mathcal{N}_{M_1}$ microstates corresponding to the macrostate M_1, $\mathcal{N}_{M_2}$ microstates corresponding to the macrostate M_2, and so on. Then, given the macrostate $M_1.M_2..M_n$ of the composite system, the total number of microstates accessible to the entire system will be

$$\mathcal{N}_M = \mathcal{N}_{M_1} \times \mathcal{N}_{M_2} \times .. \times \mathcal{N}_{M_n} \tag{2.2}$$

From this, it follows that the statistical entropy of the system is

$$\begin{aligned} S &= k_B \ln \mathcal{N}_M \\ &= k_B \ln \left(\mathcal{N}_{M_1} \times \mathcal{N}_{M_2} \times .. \times \mathcal{N}_{M_n}\right) \\ &= S_1 + S_2 + .. + S_n \end{aligned} \tag{2.3}$$

where $S_i = k_B \ln \mathcal{N}_{M_i}$ is the statistical entropy of the i^{th} system.

2.2 Temperature

Let us now consider two systems A and B separated by a partition which allows exchange of energy (but not particles). We assume that the total energy of the systems is conserved, equal to $E = E_A + E_B$. The systems, when brought in contact, will exchange energy, till an equilibrium is reached. This equilibrium configuration will be one in which the statistical entropy of the composite system is a maximum, extremised with respect to the energy of either system (since the total energy is conserved, the energy of only one system is an independent variable). This is, of course, the same as maximising the phase space volume.

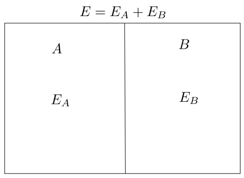


Figure 2.1: Energy exchanged by two systems

Let the energies of the systems in equilibrium be $\bar{E}_A$ and $\bar{E}_B$. For a configuration in which the energy of system A is E_A and that of B is E_B, the total entropy of the system, as a function of, say E_A, is

$$S(E_A) = S_A(E_A) + S_B(E_B) \tag{2.4}$$

where $E_B = E - E_A$ is a function of E_A. At equilibrium,

$$\begin{aligned} \left.\frac{\partial S}{\partial E_A}\right|_{E_A=\bar{E}_A} &= 0 \\ \Longrightarrow \quad \left.\frac{\partial S_A}{\partial E_A}\right|_{E_A=\bar{E}_A} + \left.\frac{\partial S_B}{\partial E_A}\right|_{E_A=\bar{E}_A} &= 0 \\ \Longrightarrow \quad \left.\frac{\partial S_A}{\partial E_A}\right|_{E_A=\bar{E}_A} - \left.\frac{\partial S_B}{\partial E_B}\right|_{E_B=\bar{E}_B} &= 0 \end{aligned}$$

where we have used

$$\frac{\partial S_B}{\partial E_A} = -\frac{\partial S_B}{\partial E_A}$$

since $E_A = E - E_B$. Then, the equilibrium values of energies satisfy the condition

$$\left.\frac{\partial S_A}{\partial E_A}\right|_{E_A=\bar{E}_A} = \left.\frac{\partial S_B}{\partial E_B}\right|_{E_B=\bar{E}_B} \tag{2.5}$$

This tells us that if two systems are allowed to exchange energy, they reach equilibrium when a certain function, the partial derivartive of entropy with respect to energy, takes the same numerical value for both systems. Clearly, this function must be related to the temperature of either system, since we know that systems in thermodynamic equilibrium possess the same temperature. Then, the temperature of a system must be related to the (partial) derivative of the entropy of the system with respect to energy. We employ a definition which (as we will see) agrees with the absolute temperature (Kelvin temperature)

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} \tag{2.6}$$

where the derivative with respect to energy is computed at fixed volume and number of particles (we assumed above that the two systems could exchange energy only). Given any arbitrary system with a certain number of particles, volume and energy, we can, in principle, compute the accessible phase space volume (1.4)

$$\Gamma_M(E, V, N) = \int_{E,V} dq_1 dq_2 .. dq_{3N} dp_1 dp_2 .. dp_{3N}$$

where for a general interacting system, the integrals are subject to constraints of volume and energy (1.5). This allows us to compute the equilibrium statistical entropy as a function of energy, volume and number of particles

$$S(E,N,V) = k_B \ln\left[\frac{\Gamma_M(E,V,N)}{h^{3N}N!}\right] \tag{2.7}$$

Differentiating this with respect to energy gives a relation between energy, volume and temperature of the system.

As an example, let us take a system of N weakly interacting particles, for which the entropy is computed to be (see eqn.(1.18))

$$S = Nk_B \ln\left[\frac{V}{N}\left(\frac{4m\pi E}{3h^2N}\right)^{3/2}\right] + \frac{5}{2}Nk_B \tag{2.8}$$

Differentiating partially with respect to energy E gives

$$\left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{3}{2}\frac{Nk_B}{E}$$

From eqn.(2.6), it follows that

$$E = \frac{3}{2}Nk_BT \tag{2.9}$$

which agrees with the result obtained from Kinetic Theory (with k_B in eqn.(2.1) identified as Boltzmann constant), and justifies eqn.(2.6) for thermodynamic temperature.

Let us go back to the two systems A and B separated by a partition. We have seen that if the systems are allowed to exchange energy, the equilibrium values of energies of the two systems are such that the two systems have the same temperature. Now, let us disturb this equilibrium slightly, such that E_A and E_B are not equal to their equilibrium values $\bar{E}_A$ and $\bar{E}_B$, but infinitesimally away from them. We also assume that these systems are initially insulated from each other (cannot exchange energy). Then, each system is in itself in equilibrium, with temperatures

$$\begin{aligned}\frac{1}{T_A} &= \left(\frac{\partial S_A}{\partial E_A}\right)_{N_A,V_A}\\ \frac{1}{T_B} &= \left(\frac{\partial S_B}{\partial E_B}\right)_{N_B,V_B}\end{aligned} \tag{2.10}$$

Now, we bring them in contact and let them exchange energy. Then, energy of A will change by ΔE_A and that of B will change by ΔE_B, such that $\Delta E_A + \Delta E_B = 0$ (total energy is conserved). Energy will *flow* from one system to the other, this flow being identified as the *heat* flowing from one system to the other. Since this flow is directional, we ask: how is the direction of flow of ‘heat’ related to the temperatures of the two systems? As the systems exchange energy to reach a new equilibrium, the total entropy of the two systems must *increase*, since in the new equilibrium, it is a maximum, and therefore greater than the initial entropy when the systems were insulated. The change in the total entropy of the two systems when heat flows is

$$\begin{aligned}\Delta S &= \Delta S_A + \Delta S_B\\ &= \frac{\partial S_A}{\partial E_A}\Delta E_A + \frac{\partial S_B}{\partial E_B}\Delta E_B\\ &= \left(\frac{1}{T_A} - \frac{1}{T_B}\right)\Delta E_A\end{aligned} \tag{2.11}$$

where we have used $\Delta E_A + \Delta E_B = 0$. Since $\Delta S > 0$, if $T_A > T_B$, then $\Delta E_A < 0$. That is, energy (heat) will flow from system A to system B. This just illustrates that heat flows from a hotter to a cooler object. Here, we have seen it as a consequence of increase of entropy due to a spontaneous process of an isolated system, in which the macrostate changes such that more available phase space volume is covered, which is equivalent to saying that the entropy increases.

2.3 Entropy Extremization

The relation between temperature and entropy deduced above is an example of the principle of extremization of entropy (or, equivalently, volume of available phase space). In general, consider an isolated system in equilibrium, subject to internal constraints which take the form of fixed parameters $\lambda_1, \lambda_2, .., \lambda_n$. In this configurarion, the entropy of the system is a function of parameters λ_i. Say, we set one or more of these constraints free, such that some parameters are now free to change. The new equilibrium configuration will be one in which the free parameters take such values, that the available phase space volume is a maximum. This is the same as saying that the equilibrium values of the parameters are those which extremize the entropy of the system with respect to those parameters. For instance, in the previous section, internal constraints separated the system into subsystems A and B, with parameters being the number of particles, energy and volumes of the subsystems. When the energy of the subsystems was allowed to vary, the equilibrium configuration was the one in which the entropy was extremized with respect to the energy of any one subsystem, and this led to the condition that the tempertures of the subsystems must be equal in the equilibrium configuration.
As another illustration of entropy extremization, let us deduce an expression for the pressure exerted by a system of interacting particles in equilibrium on a container containing the particles. Consider the system confined to a cylindrical chamber with a moving frictionless piston on which an object of mass m is placed. Initially, the piston is locked in place, and the particles have energy E_i and occupy volume V_i. If the lock is removed, the piston will move till a new equilibrium configuration is attained. We ask: what is the new position of the piston? Intuitively, we know that the equilibrium position will be such that the pressure exerted by the system on the piston equals the weight of the object placed on it (assuming the mass of the piston is very small compared with that of the object placed on it). However, let us view this from the point of view of entropy extremization

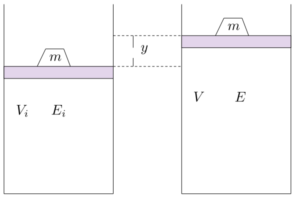


Figure 2.2: Pressure exerted by a system of particles

For a given position of the piston, the energy of the system is $E = E_i - mgy$ where y is the displacement of the piston relative to its initial, locked position. The volume accessible to the system is $V = V_i + Ay$ where A is the area of the piston. Let the entropy of the system of particles when its energy is E and volume V be $S(E, V)$. This expression will depend on the details of interactions betwen the particles. For a given position of the piston, this can be thought of as a function of a single parameter: the position of the piston, y

$$S(E, V) = S(E_i - mgy, V_i + Ay) \tag{2.12}$$

Initially, when the piston is locked, this parameter is fixed ($y = 0$). However, as this parameter is allowed to vary, it will reach an equilibrium value such that the entropy is an extremum with respect to y. The condition for equilibrium is then

$$\begin{aligned} \frac{\partial S}{\partial y} &= 0 \\ \implies \left(\frac{\partial S}{\partial E}\right)_V \frac{\partial E}{\partial y} + \left(\frac{\partial S}{\partial V}\right)_E \frac{\partial V}{\partial y} &= 0 \end{aligned}$$

Since $E = E_i - mgy, V = V_i + Ay$, $\partial E/\partial y = -mg$ and $\partial V/\partial y = A$. Then, the equilibrium condition is

$$\begin{aligned} A\left(\frac{\partial S}{\partial V}\right)_E &= mg\left(\frac{\partial S}{\partial E}\right)_V \\ &= \frac{mg}{T} \end{aligned}$$

where T is the temperature of the system in the equilibrium configuration. Then, we get

$$\frac{mg}{A} = T\left(\frac{\partial S}{\partial V}\right)_E$$

However, in equilibrium, mg/A equals the pressure exerted by the system on the piston. This gives us an expression for the pressure of the system as a function of its energy and volume (equation of state)

$$P = T\left(\frac{\partial S}{\partial V}\right)_{E,N} \tag{2.13}$$

where we have added N (number of particles) to the subscript, previously suppressed. Then, a calculation of the statistical entropy allows us to determine the equation of state of the system of particles. Let us apply (2.13) to a system of weakly interacting particles, for which the entropy is given by eqn.(2.8). Differentiating with respect to volume gives

$$\left(\frac{\partial S}{\partial V}\right)_{E,N} = \frac{Nk_B}{V} \tag{2.14}$$

Then, the pressure of the system is

$$\begin{aligned} P &= T\left(\frac{\partial S}{\partial V}\right)_{E,N} \\ &= T\frac{Nk_B}{V} \end{aligned} \tag{2.15}$$

which gives the equation of state

$$PV = Nk_BT \tag{2.16}$$

2.4 Principle of equal a-priori probabilities

Given that a macroscopic system is in equilibrium does not imply that the system is 'static'. The microstate of the system keeps evolving with time, consistent with dynamical laws. Consequently, its macrostate also keeps evolving. However, we have seen that with 'overwhelming probability' (practically unity), the system will be found in a state in which the accessible phase space volume is a maximum. Still, the macrostate, in principle, will fluctuate about this equilibrium state. What are the odds that it will be found in a macrostate M? From what has been discussed before, it seems 'reasonable' that the probability of a system existing in macrostate M should be proportional to the phase space volume associated with that macrostate. In other words, the ratio of the probability that the system is in macrostates M and M' is

$$\begin{aligned} \frac{P(M')}{P(M)} &= \frac{\Gamma_{M'}}{\Gamma_M} \\ &= \frac{\mathcal{N}_{M'}}{\mathcal{N}_M} \end{aligned} \tag{2.17}$$

where $\mathcal{N}_M$ is the number of microstates corresponding to macrostate M, given by eqn.(3.3). Saying that the probability of a macrostate is proportional to the numner of microstates corresponding to the macrostates is equivalent to assuming that *all microstates are equally probable.* This is the so-called *Principle of equal a* *priori probabilities.* It is not possible to give a rigorous justification of this hypothesis based on dynamical equations. All that can be said is that it is 'reasonable' and experiments support it. How would one verify such a principle experimentally? One way to do it is to predict the amount of fluctuation in a system, about the equilibrium macrostate. For instance, a system of particles in gaseous state enclosed in a volume is expected to be homogeneously distributed throughout the volume. However, there are expected to be density fluctuations, however small. If these fluctuations are measured, they can be tested aginst predictions based on the principle of equal a priori proabability. Assuming this principle, let us caculate the probability of such fluctuations. To keep the discussion general, let us assume there is a parameter λ which, left free, has attained equilibrium value $\bar{\lambda}$. This is the value for which the entropy of the system, as a function of λ, is an extremum. That is, $S'(\bar{\lambda}) = 0$, where $S'(\lambda) = \partial S/\partial\lambda$. Given eqn.(2.17) and (2.1), it follows that the probability of the parameter taking value λ is given by

$$\begin{aligned}\frac{P(\lambda)}{P(\bar{\lambda})} &= \frac{\mathcal{N}_{M(\lambda)}}{\mathcal{N}_{M(\bar{\lambda})}} \\ &= \frac{e^{S(\lambda)/k_B}}{e^{S(\bar{\lambda})/k_B}} \\ &= e^{\left(S(\lambda)-S(\bar{\lambda})\right)/k_B} \qquad (2.18)\end{aligned}$$

Expanding $S(\lambda)$ about $\bar{\lambda}$ and retaining terms upto $\Delta\lambda^2$, we get

$$\begin{aligned}P(\lambda) &= P(\bar{\lambda})\; e^{(1/2k_B)S''(\bar{\lambda})\Delta\lambda^2} \\ &= P(\bar{\lambda})\; e^{-(1/2k_B)\left|S''(\bar{\lambda})\right|\Delta\lambda^2} \qquad (2.19)\end{aligned}$$

which shows a Gaussian form for fluctuations. We will see later that $\Delta\lambda/\bar{\lambda} \sim 1/\sqrt{N}$, where N is the number of particles of the system.

Chapter 3

Statistical Distribution of an Isolated System

3.1 Microcanonical Distribution

We have seen that the microstate of an isolated system of particles keeps fluctuating (driven by dynamical equations), and so does the corresponding macrostate. To determine equilibrium properties of the system, the principle of equal a-priori probabilities in invoked, which essentially states that given constraints on the system, all microstates of the system consistent with the constraints are equally probable. Then, the 'reason' for observing special macrostates and not others (such as a homogeneous distribution of particles of a gas) is that such macrostates are associated with possible microstates whose number is 'overwhelmingly large' compared with the number of microstates for other macrostates. This is equivalent to saying that in equilibrium, macrostates manifest with probabilities proportional to the corresponding phase space volume associated with them. Let us mathematically state this probability distribution more precisely. Say, the (conserved) energy of the system of particles is known to lie in the range $E_0 < E < E_0 + \delta E$. Then, the probability that the system is in a microstate r is given by

$$P_r = \begin{cases} 1/\mathcal{N} & E_r \in [E_0, E_0 + \delta E] \\ 0 & E_r \notin [E_0, E_0 + \delta E] \end{cases} \tag{3.1}$$

where $\mathcal{N}$ is the total number of microstates consistent with the fact that the energy of the system lies in the range $E_0 < E < E_0 + \delta E$, and any other external constraints such as the volume of the system, etc. This probability distribution is called the *Microcanonical Distribution.* At this stage, we have assumed that even though the microstate of the system changes with time due to the dynamics so that it explores the phase space, the representative point of the system is always constrained to lie on a 'surface' of constant energy in the phase space, since the energy of the system is conserved. When we say this, we assume that energy is the *only* conserved quantity. This may *not* be true. In addition to energy, there could be, in principle, other conserved quantities. For instance, consider a system of particles confined to a spherical volume, interacting with the wall through elastic collisions. Spherical symmetry implies that in addition to energy, the *angular momentum* of the system is also conserved. This leads to additional constraints on the position and momentum components of the system (three, one for each component of angular momentum). Then, the representative point is constrained to lie on a surface of a *smaller* dimension than if the volume was not spherical. This will now be a $6N - 4$ dimensional surface, corresponding to fixed value of energy and angular momentum components. For now, we will assume that energy is the only conserved quanti

3.1.1 Maxwell velocity distribution

As an application, let us deduce the Maxwell velocity distribution for a system of weakly interacting particles, confined to a volume V and with energy in range $[E, E + \delta E]$. Let us ask the question: what is the probability that a particle, picked at random, has momentum that lies between $\vec{p}$ and $\vec{p} + d\vec{p}$? Given the distribution (3.1), the probability is

$$P(\vec{p})d^3\vec{p} = \frac{\text{Number of microstates corresponding to this one particle having momentum in this range}}{\text{Total number of mocrostates consistent with energy an volume constraints}} \tag{3.2}$$

Now, the total number of microstates is given by eqn.(3.3)

$$\mathcal{N}(E,V,N) = \frac{\Gamma(E,V,N)}{h^{3N}N!} \tag{3.3}$$

where we have explicitly specified the constaints E, V, N. Now, say we pick particle 1 and calculate the probability that it has momentum in range $\vec{p}$ and $\vec{p}+d\vec{p}$. The volume of the phase space corresponding to this particle having this momentum is

$$\begin{aligned}
\Gamma_{\vec{p}} &= d^3\vec{p}\int d^3\vec{r}_1 d^3\vec{r}_2..d^3\vec{r}_N\ d^3\vec{p}_2 d^3\vec{p}_3..d^3\vec{p}_N \\
&= V d^3\vec{p}\int d^3\vec{r}_2 d^3\vec{r}_3..d^3\vec{r}_N\ d^3\vec{p}_2 d^3\vec{p}_3..d^3\vec{p}_N
\end{aligned} \tag{3.4}$$

The number of microstates, taking into account the division of the phase space into cells and the idential nature of particles $2, 3.., N$, is

$$\mathcal{N}' = \frac{\Gamma_{\vec{p}}}{h^{3N}(N-1)!}$$

However, there *is* no 'particle 1'. Therefore, we need to divide this number by the total number of ways we could have picked this particle (there are N ways). Then, the correct number of microstates is

$$\mathcal{N}_{\vec{p}} = \frac{\Gamma_{\vec{p}}}{h^{3N}N!} \tag{3.5}$$

Then, the probability (3.2) reduces to

$$P(\vec{p})d^3\vec{p} = \frac{\Gamma_{\vec{p}}}{\Gamma(E,V,N)} \tag{3.6}$$

Now, the phase space integral appearing in the expression for $\Gamma_{\vec{p}}$ is subject to the constraint that the $N-1$ particles $2, 3, .., N$ have available volume V and available energy $E - \vec{p}^2/2m$ (since particle 1 has momentum $\vec{p}$ and energy $\vec{p}^2/2m$). Therefore

$$\Gamma_{\vec{p}} = V d^3\vec{p}\,\Gamma\left(E - \vec{p}^2/2m, V, N-1\right) \tag{3.7}$$

Therefore

$$P(\vec{p}) = V \times \frac{\Gamma\left(E - \vec{p}^2/2m, V, N-1\right)}{\Gamma(E,V,N)} \tag{3.8}$$

The expression for $\Gamma(E,V,N)$ is given in eqn.(1.9)

$$\Gamma(E,V,N) = \frac{\pi^{3N/2}}{\Gamma(3N/2)}(2m)^{(3N-1)/2}V^N E^{3(N-1)/2}\delta E \tag{3.9}$$

Using this, the expression for $P(\vec{p})$ reduces to

$$P(\vec{p}) = \frac{1}{(2m\pi E)^{3/2}} \times \frac{(3N/2-1)!}{(3N/2-5/2)!} \times \left(1 - \frac{\vec{p}^2}{2mE}\right)^{3N/2-3} \tag{3.10}$$

Here, we are working with a system with a very large number of particles, with $N >>> 1$. Then, the exponent $3N/2-3$ can be readily approximated to $3N/2$. We will need to be more careful in approximating the factorials, since they are rapidly increasing functions of N. As before, in such a situation, it is best to work with logarithms. Let $\alpha_N = (3N/2-1)!/(3N/2-5/2)!$. Then, using Sterling's approximation

$$\begin{aligned}
\ln \alpha_N &= \ln(3N/2-1)! - \ln(3N/2-5/2)! \\
&\simeq (3N/2-1)\ln(3N/2-1) - (3N/2-1) - (3N/2-5/2)\ln(3N/2-5/2) + (3N/2-5/2) \\
&\simeq (3N/2-1)\ln(3N/2-1) - (3N/2-5/2)\ln(3N/2-5/2) - \frac{3}{2} \\
&\simeq (3N/2-1)\ln(3N/2) - (3N/2-5/2)\ln(3N/2) \\
&\simeq \frac{3}{2}\ln(3N/2)
\end{aligned}$$

where we have ignored certain numerical terms compared with N. Then,

$$\alpha_N \simeq \left(\frac{3N}{2}\right)^{3/2} \tag{3.11}$$

Finally, we get

$$P(\vec{p}) = \left(\frac{3N}{4m\pi E}\right)^{3/2}\left(1-\frac{\vec{p}^2}{2mE}\right)^{3N/2} \tag{3.12}$$

In equilibrium, the total energy of the system is related to its temperature according to (2.9)

$$E = \frac{3}{2}Nk_BT$$

Substituting this, we get

$$P(\vec{p}) = \left(\frac{1}{2m\pi k_BT}\right)^{3/2}\left(1-\frac{\vec{p}^2}{2mk_BT\times(3N/2)}\right)^{3N/2} \tag{3.13}$$

Since $3N/2 >>> 1$,

$$\left(1-\frac{\vec{p}^2}{2mk_BT\times(3N/2)}\right)^{3N/2} \simeq e^{-\vec{p}^2/2mk_BT} \tag{3.14}$$

Then, we finally get the Maxwell distribution (the velocity distribution is obtained by setting $\vec{p} = m\vec{v}$)

$$P(\vec{p})d^3\vec{p} = \left(\frac{1}{2m\pi k_BT}\right)^{3/2} e^{-\vec{p}^2/2mk_BT}d^3\vec{p} \tag{3.15}$$

3.1.2 A first look at the Ising Model

We now consider another system: a system of N 'spins' located on a fixed lattice. Each spin represents the spin angular momentum of an atom located at the lattice site. Since the magnetic moment of an atom is proportional to its spin angular momentum, each atom possesses a magnetic moment, and as a result, the system of spins can possess a net magnetization. The neighboring spins interact with each other through a magnetic interaction (since this interacting strength falls off rapidly with distance, we ignore interaction between spins further away). We assume that each atom has a net (quantum mechanical) spin half, such that any one component of the spin angular momentum can take values $\pm\hbar/2$. In principle, such a system should be treated quantum mechanically. However, we strip it down to a semi-classical model, known as the *Ising Model*. In this model, the spin orientation is restricted to one direction only, along which it can that values $S = \pm\hbar/2$. The interaction between spins located at neighboring lattice sites i and j results in an interaction energy proportional to S_iS_j. This interaction term can be written as

$$E_{int} = -J\sigma_i\sigma_j \tag{3.16}$$

where σ_i is spin of the i^{th} site measured in units of $\hbar/2$, and therefore takes values ± 1. $J > 0$ is a constant of proportionality, determining the strength of interaction. When the neighboring spins are parallel (both 1 or -1), it lowers the energy of the system. This accounts for the negative sign. Now, say we introduce an external magnetic field. Each spin will now interact with this magnetic field, the interaction energy of the i^{th} spin being of the form $E_i = -h\sigma_i$, where h is proportional to the magnetic field (we will identify h as the magnetic field measured in suitable units). The, the total energy of the system is given by

$$E = -J \sum_{<i,j>} \sigma_i \sigma_j - h \sum_i \sigma_i \tag{3.17}$$

A microstate of this system would be a precise specification of the spin of each lattice site $(+1 or - 1)$. A macrostate would be a specification of the total magnetisation of the system

$$M = \frac{1}{N} \sum_i \sigma_i \tag{3.18}$$

Given M, there will be many spin configurations, resulting in the same value of M. The dynamical equations determining how the microstates evolve are not specified. We just assume that there exist such (semiclassical) equations which will lead to the evolution of spin configurations with time, and that in equilibrium, all microstates are equiprobable.
If the external magnetic field is strong enough $(h >> J)$, we can approximate the expression for energy as

$$E \simeq -h \sum_i \sigma_i \tag{3.19}$$

This energy is conserved. The 'motion' of the phase point of the system is subject to constrainsts that the energy of the system is conserved and the external magnetic field has a certain value. Consistent with these constraints, let us calculate the number of microstates. Given the energy and number of spins, the total number of 'up' spins n_+ and number of 'down' spins n_- are fixed by equations

$$\begin{aligned} n_+ + n_- &= N \\ n_+ - n_- &= -E/h \end{aligned} \tag{3.20}$$

Solving these, we get

$$\begin{aligned} n_+ &= \frac{1}{2}\left(N - \frac{E}{h}\right) \\ n_- &= \frac{1}{2}\left(N + \frac{E}{h}\right) \end{aligned} \tag{3.21}$$

The magnetization of the system (in suitable units) is

$$\begin{aligned} M &= \frac{n_+ - n_-}{N} \\ &= -\frac{E}{Nh} \end{aligned} \tag{3.22}$$

such that

$$E = -NhM \tag{3.23}$$

The macrostate of the system is a specification of the number of spins and the magnetization. This is equivalent to a specification of n_+ and n_-. This does not determine which spin is up and which spin is down, just the total number of up spins and the number of down spins. Each spin's spin state keeps fluctuating, such that the microsate of the system keeps changing, for the given fixed macrostate. The total number of microstates is just the number of ways of having n_+ spins up and n_- spins down, with the total number fixed. This is given by

$$\mathcal{N}(E,h,N) = \frac{N!}{n_+!n_-!} \tag{3.24}$$

Note that we do not divide by $N!$ as we did for the case of identical particles in eqn.(3.3). This is because since the spins are localised at distinct points in space, interchanging spins *does* lead to a different configuration. The entropy of the system is given by (2.1)

$$\begin{aligned} S &= k_B \ln \mathcal{N} \\ &= k_B \ln\left(\frac{N!}{n_+!n_-!}\right) \\ &= k_B \left(\ln N! - \ln n_+! - \ln n_-!\right) \\ &\simeq k_B \left(N\ln N - N - n_+\ln n_+ + n_+ - n_-\ln n_- + n_-\right) \\ &= k_B \left(N\ln N - n_+\ln n_+ - n_-\ln n_-\right) \\ &= -k_B\left[n_+\ln\left(\frac{n_+}{N}\right) + n_-\ln\left(\frac{n_-}{N}\right)\right] \end{aligned}$$

Substituting for n_+ and n_-, we get

$$S = -Nk_B\left\{\frac{1}{2}\left(1-\frac{E}{Nh}\right)\ln\left[\frac{1}{2}\left(1-\frac{E}{Nh}\right)\right] + \frac{1}{2}\left(1+\frac{E}{Nh}\right)\ln\left[\frac{1}{2}\left(1+\frac{E}{Nh}\right)\right]\right\} \tag{3.25}$$

The temperature of the system is given by

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial E}\right)_{h,N} \\ &= -\frac{k_B}{2h}\ln\left(\frac{1+E/Nh}{1-E/Nh}\right) \end{aligned} \tag{3.26}$$

Solving for E, we get

$$E = -Nh\tanh\left(\frac{h}{k_BT}\right) \tag{3.27}$$

The magnetization of the system as a function of temperature is obtained from (3.22)

$$M = \tanh\left(\frac{h}{k_BT}\right) \tag{3.28}$$

Following is a plot of magnetization vs temperature

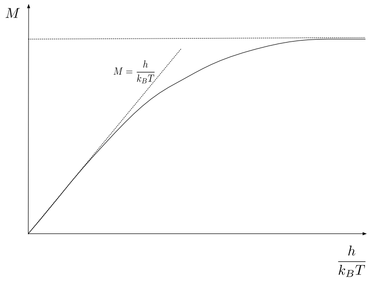


Figure 3.1: Magnetization vs Temperature: Ising Model with strong magnetic field

In presence of an external magnetic field, there is a natural temperature scale $T_0 = h/k_B$. For $T >> T_0$ and $T << T_0$, the behaviour of the magnetization is as follows

$$M \simeq \begin{cases} \frac{h}{k_BT} & T >> T_0 \\ 1 & T << T_0 \end{cases} \tag{3.29}$$

At very low temperatures, almost all spins are aligned along the direction of the magnetic field ($M \simeq 1$) and at very high temperature, the magnetization falls inversely with temperature. The Ising model, in the approximation that the inter-spin interaction can be ignored, is a simple toy model for paramagnetism. However, the behaviour at low temperatures as predicted by (3.29) is non-trivially modified by inter-spin interactions. With interactions included, the model predicts a *spontaneous magnetization* even in absence of an external magnetic field. It therefore is a simple model for ferromagnetism.

3.1.3 Negative Temperatures

An isolated system of N spins allows for an interesting possibility: the potential for the system to have *negative temperatures*. Even though negative temperatures might seem counter-intuitive, given the relation (2.6) between temperature and entropy, if a system is in a state such that an increase in energy of the system results in a decrease in its entropy, the temperature of the system in that state will be negative. Intuitively, one expects the entropy to be an increaing function of energy, as in case of a system of weakly interacting particles (eqn.(2.8). This is because if a system has greater energy, there are many more ways of distributing this energy among the different constituents. However, this argument fails if the system's energy is *bounded from above*. That is, the energy of the system cannot exceed a certain value. In such cases, as one approaches close to the highest attainable energy, the number of microstates (corresponding to a given value of energy) decreases with increasing energy. Therefore, the entropy of the system (proportional to the logarithm of the number of microstates) is a decreasing function of energy. Therefore, it follows from (2.6) that the temperature of the system is negative. In presence of an external magnetic field (ignoring inter-spin interaction), a system of N spins described by the Ising model has this property. Clearly, there is exactly one microstate corresponding to the highest energy: the state in which all the spins are antiparallel to the external magnetic field. The next macrostate (lower in energy) is one in which exactly one spin is parallel to the magnetic field. This macrostates corresponds to N microstates (any one of N spins parallel to the magnetic field). It is easy to see that as we approach the highest energy configuration from below, the number of microstates shrinks. Therefore, we expect this system to exhibit negative temperatures in this region of energy. To see this explicitly, we go back to eqn.(3.26)

$$\frac{1}{T} = -\frac{k_B}{2h} \ln\left(\frac{1 + E/Nh}{1 - E/Nh}\right) \tag{3.30}$$

This equation can be used to plot the temperature of the system as a function of energy

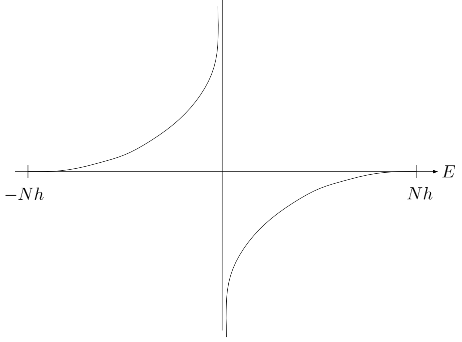


Figure 3.2: Temperture vs Energy for a system of Ising spins (strong magnetic field)

Clearly, the system exhibits negative temperature for the range $0 < E < Nh$. It is instructive to plot the entropy of the system (eqn.(3.25) as a function of energy

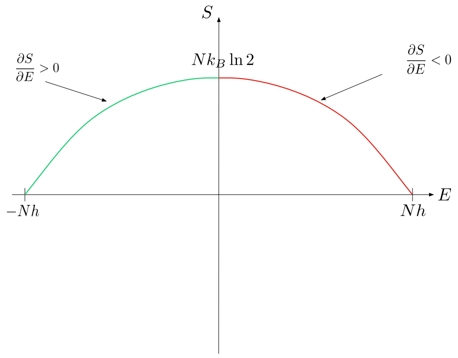


Figure 3.3: Entropy vs Energy for a system of Ising spins (strong magnetic field)

The macrostate with the least energy $E = -Nh$ and magnetization $M = -1$ has exactly one corresponging microstate: all spins parallel to the field h. The entropy of this system is clearly zero. We can increase the energy (and entropy, upto a point) by flipping one spin at a time anti-parallel to the magnetic field. For instance, if we flip one spin antiparallel, there are N possible microstates (any one of N possible spins antiparallel), so the entropy increases from zero the $k_B \ln N$. The maximum entropy configuration is one in which the total energy (nd magnetization) of the system is zero. This corresponds to half of the spins parallel and half anti-parallel to the external field. The entropy of the system for this configuration is $Nk_B \ln 2$. If we flip more spins anti-parallel, we start decreasing the entropy of the system, till it reaches zero when all the spins are anti-parallel. A look at figure 3.2 shows that the temperature of the system diverges as we approach $E = 0$. However, the divergence is in different directions, depending on how $E = 0$ is approached. If we approach from the positive temperature side, the temperature diverges to positive infinity. Approaching from the negative temperature region, it diverges to negative infinity. What does all this mean? If we interpret temperature as being related to the rate of change of entropy with energy, this behaviour is understandable. However, intuitively, we understand temperature in the context of a thermometric measurement, which involves the system coming to equilibrium with a measuring device. What will a thermometer measure, if it is brought 'in contact' with this system? The answer depends on what the properties of the thermometric substance. Any traditional thermometric substance is a system without an upper energy bound. Therefore, it is incapable of attaining negative temperatures. What happens then if such a thermometer is brought in contact with a system at negative temperature? Will heat flow from the thermometer to the system or vice-versa? Let us consider two systems A and B, with A being a conventional system that does not have an upper energy bound and B a system which can attain negative temperatures. For definiteness, let us assume that B is our spin system. Let us assume that initially $T_B < 0$ (when B is not in contact with system A). Now, the two systems are allowed to interact (brought in 'thermal contact'). In which direction will heat flow?

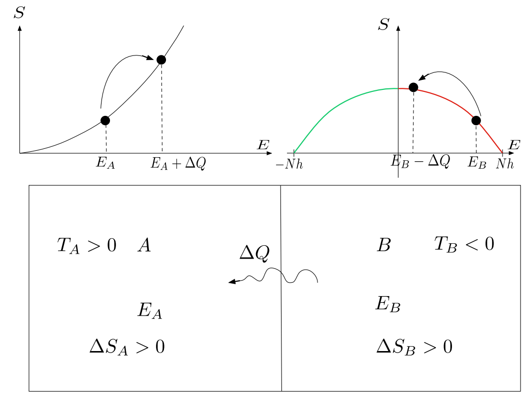


Figure 3.4: Heat flow from negative to positive temperature

Figure 3.4 illustrates the direction of flow of heat and the consequent change in entropy of both systems. The key is that once the systems are allowed to exchange energy, the total entropy of the two systems will increase. The negative temperature system can increase its entropy by losing energy ΔQ and at the same time, the positive temperature system can increase its entropy by gaining this energy. Then, a flow of heat from the negative to the positive temperature system will increase the entropy of both systems, resulting in an overall increase in entropy. On the other hand, a flow of heat from the positive to the negative temperature system will decrease the entropy of both systems, resulting in an overall decrease in entropy, which is not possible. Therefore, heat will flow from the negative to the positive temperature system. Negative temperatures are *hotter* than positive temperatures. When will this heat exchange stop? This exchange will continue till system B is in the negative temperature region. As soon as it enters the positive temperature region ($E_B < 0$), if it loses energy ΔQ to system A, while the entropy of A will incrase, the entropy of B will *decrease*. Since the flow of energy has to be such that the total entropy increases, there will be a tension between the increase of entropy of A and the decrease of entropy of B. At this stage, the systems will behave just like two 'ordinary' positive temperature systems trying to attain thermal equilibrium - the flow of heat from B to A will stop when their tempearatures equalize (both attaining the same positive temperature)

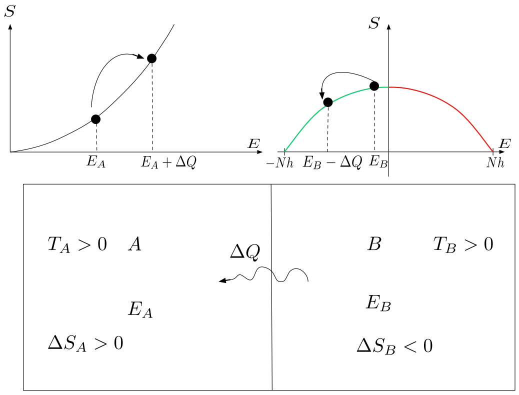


Figure 3.5: Heat flow from positive to positive temperature

Negative temperature systems have very interesting and counter-intuitive properties. Why is it we don't encounter them everyday? For instance, a poramagnetic solid in presence of an external magnetic field should be able to behave like a negative temperature system, if we can excite the atomic spins to high eneough energy. The problem is, that such a system does not just have spin 'degrees of freedom'. Each atom has its own translational degrees of freedom as well, and can oscillate about its equilibrium position. These degrees of freedom correspond to the $6N$ dimensional phase space of position and momentum components of the system of atoms. Further, there is upper bound to the energy this system of atoms can possess by virtue of their translational motion. Then, if we somehow excite the spins to an energy which corresponds to negative temperatures, then the spin degrees of freedom will be at negative temperature, but the translational degrees at positive temperature. Two different dynamical aspects of the same system are at different temperatures! If there is any 'coupling' between these degrees of freedom, the spin degrees will lose energy to the translational degrees, incresing the entropy of both aspects of the system, till eventually both reach the same positive temperature. Therefore, negative temperature states are inherently unstable, though they can be, under controlled conditions, created in the lab.

Chapter 4

Statistical Distribution of Subsystems

4.1 Statistics of subsystems

In nature, no system is truly isolated. Systems in equilibrium are in fact *subsystems* of much larger systems, in equilibrium with the other parts of the larger system. For example, a block of copper in a room can be modelled as a system of copper atoms in the solid phase, in equilibrium with the gases in the surrounding air (ignoring the existence of a support on which the block rests). If we could in principle somehow isolate the room from eveything else, then the system consisting of the air and the copper block would be isolated, with total energy conserved. However, the energy of the block is not conserved, since it exchanges energy continuously with the air molecules due to collisions. Together, though, the air molecules and the copper are in equilibrium, which means they have together attained a macrostate consistent with maximum possible phase space volume (the phase space constructed from coordinates and momentum components of the copper and gas atoms).

Given a subsystem in equilibrium with an environment, we wish to deduce the statistical distribution for this subsystem. For a (hypothetically) isolated system, the probability that the system is in some microstate is given by the microcanonical (3.1), which essentially gives the same probability to every microstate of the system. To deduce this distibution, we will assume the microcanonical distribution for the entire system, and further assume that the number of degrees of freedom of the subsystem, while very large, are much fewer than those of the entire system (and therefore the rest of the system as well), though . In addition, we will also make a very important assumption: the subsystem interacts very 'weakly' with the rest of the system. This is justified if the interaction of the constituents of the subsystem with the constituents of the environment are *short-ranged*. For instance, in the example of the block of copper interacting with gas molecules, any one molecule interacts with a copper atom only when it is within a few angstroms of the atom. The following arguments will *not* be valid if the interactions are long-range, as for example in the case of gravitationally interacting systems (if the system is a star and the subsystem is a part of the star, the particles of the subsystem interact with the particles of the rest of the star through a long-range gravitational interaction). Assuming this to be the case, since the interaction of the subsystem with the environment occurs only at the boundary (being short-range) separating the two, the interaction energy, compared with the energy of the subsystem and the environment, will by smaller by a factor of the ratio of surface area to violume of the subsystem. For a macroscopic system, this will be vanishingly small. As a result, the total energy of the system is, to an excellent approximation, additive

$$E = E^{sub} + E^{env} \tag{4.1}$$

E^{sub} and E^{env} keep fluctuating with time because of energy exchange between the subsystem and the environment. In addition to energy, there could be other conserved quantities. For instance, if the entire system possesses rotational symmetry, the total angular momentum of the system is conserved. Angular momentum of the system is also additive. Therefore, for any one component of angular momentum, we will have an equation of the form (4.1)

$$J = J^{sub} + J^{env} \tag{4.2}$$

Similarly, translational invariance would result in the momentum of the system being conserved, with similar equation for each momentum component.

When we make measurements on macroscopic subsystems in equilibrium with their environment, we observe that properties such as energy, volume, etc. are sharply defined. That is, they appear to be constant, even though the subsystem can exchange energy, share volume, etc. with the environment. The key to this behaviour is that these quantities are *extensive*, that is, scale with the number of particles of the subsystem. The reason behind the constancy of such properties is the idea of *statistical independence.* Two such subsystems interacting with each other (with short range interactions) are statistically independent. That is, the probability that one subsystem is in microstate r and another in microstate r' satisfies

$$P_{rr'} = P_r \times P_r' \tag{4.3}$$

This is true only if the interaction beween the two subsystems is *weak*, which is expected to be true if the interaction is short-range. Then, there is expected to be no correlation between one subsytem being in a certain microstate r and another being in microstate r'. Given this staistical independence, let us show that if the subsystem has a large enough number of particles, fluctuations in extensive (additive) quantities such as energy, volume will be 'very small'. Let us visualise the subsystem in question as consisting of N furthur subsystems $i = 1, 2, ..., N$, where each such subsystem still consists of a very large number of particles, such that it is statistically independent from the others. Let A be an extensive quantity of the subsystem (such as energy, volume, angular momentum, etc.). Then, if A_i is the value of this quantity for the i^{th} subsystem, then it follows that

$$A = \sum_{i=1}^{N} A_i \tag{4.4}$$

The quantity A (and all the A_is) will not be constant, but will fluctuate because of interaction with the environment. Let the average of A be $\bar{A}$ and that of A_i be $\bar{A}_i$. Then,

$$\bar{A} = \sum_{i=1}^{N} \bar{A}_i \tag{4.5}$$

Given the probability distribution of microstates of a subsystem (to be deduced), the standard deviation corresponding to this probability distribution will be a measure of the fluctuation ΔA in A (and similarly, A_is)

$$\Delta A = \sqrt{\Delta A^2} \tag{4.6}$$

where

$$\begin{aligned} \Delta A^2 &= \overline{\left(A - \bar{A}\right)^2} \\ &= \overline{A^2} - \left(\bar{A}\right)^2 \end{aligned} \tag{4.7}$$

Now,

$$\begin{aligned} A^2 &= \left(\sum_{i=1}^{N} A_i\right)^2 \\ &= \sum_{i,j=1}^{N} A_i A_j \end{aligned} \tag{4.8}$$

Therefore

$$\begin{aligned} \overline{A^2} &= \sum_{i,j=1}^{N} \overline{A_i A_j} \\ &= \sum_{i \neq j} \overline{A_i A_j} + \sum_{i=1}^{N} \overline{A_i^2} \end{aligned} \tag{4.9}$$

Since the subsystems are statistically independent, therefore

$$\overline{A_i A_j} = \overline{A_i} \times \overline{A_j} \tag{4.10}$$

Therefore

$$\overline{A^2} = \sum_{i \neq j} \overline{A_i} \times \overline{A_j} + \sum_{i=1}^{N} \overline{A_i^2} \tag{4.11}$$

Further, it follows from (4.5) that

$$\begin{aligned} \left(\overline{A}\right)^2 &= \left(\sum_{i=1}^{N} \bar{A}_i\right)^2 \\ &= \sum_{i,j=1}^{N} \overline{A_i} \times \overline{A_j} \\ &= \sum_{i \neq j} \overline{A_i} \times \overline{A_j} + \sum_{i=1}^{N} \left(\overline{A_i}\right)^2 \end{aligned} \tag{4.12}$$

Then, it follows (from (4.11) and (4.12))

$$\begin{aligned} \Delta A^2 &= \overline{A^2} - \left(\bar{A}\right)^2 \\ &= \sum_{i=1}^{N} \overline{A_i^2} - \sum_{i=1}^{N} \left(\overline{A_i}\right)^2 \\ &= \sum_{i=1}^{N} \left[\overline{A_i^2} - \left(\overline{A_i}\right)^2\right] \\ &= \sum_{i=1}^{N} \Delta A_i^2 \end{aligned} \tag{4.13}$$

where $\Delta A_i = \sqrt{\Delta A_i^2}$ is the fluctuation in A_i. This is the key result. Given a subsystem furthur comprising of N subsystems (each macroscopic), the fluctuation squared of any extensive physical quantity is *additive* because of statistical independence

$$\Delta A^2 = \sum_{i=1}^{N} \Delta A_i^2 \tag{4.14}$$

Given this, it follows that if we scale the entire subsystem by λ (thereby scaling each sub-part by the same factor), the mean of any extensive quantity scales by λ (from (4.5), and the fluctuation *squared* also scales by λ (from (4.14). Therefore, the fluctuation scales as the *square root* of λ

$$\begin{aligned} \overline{A} &\longrightarrow \lambda\, \overline{A} \\ \Delta A &\longrightarrow \sqrt{\lambda}\, \Delta A \end{aligned} \tag{4.15}$$

Therefore, the *relative fluctuation* scales as $1/\sqrt{\lambda}$

$$\frac{\Delta A}{\overline{A}} \longrightarrow \frac{1}{\sqrt{\lambda}} \frac{\Delta A}{\overline{A}} \tag{4.16}$$

From this it immediately follows that for a subsystem consisting of n particles,

$$\frac{\Delta A}{\overline{A}} \propto \frac{1}{\sqrt{n}} \tag{4.17}$$

which, for a macroscopic system, is very small. This is why in equilibrium, we observe that extensive properties of subsystems are very sharply defined.

4.2 The probability distribution

Consider a subsystem A interacting with an environment B. We assume that the number of degrees of freedom of the environment are much larger than those of the subsystem (practically infinite). Let there be n additive, conserved quantities $Q_1, Q_2, ..., Q_n$ for the entire isolated system, of which the subsystem is a small part. Typically, these will be energy, volume and number of particles of the system. However, in case of rotationally and/or translationally invariant systems, these could be angular momentum and/or momentum. The subsystem could 'exchgange' one or more of these quantities with the environment. For instance, if there is a boundary separating the subsystem from the environment which allows only energy exchange, then energy will be the quantitiy that is exchanged. More generally, the subsystem could also exchange particles with the environment, or might share the total volume with the environment. For instance, it could be a system of particles trapped within a membrane (this membrane embedded in an environment of other particles) whose volume could be variable. Then, the number of particles and volume would be such quantities, apart from energy. The key is that (a) the quantities should be conserved for the entire system and (b) they should be additive and (c) should be exchanged between the subsystem and the environment. Let there be m such quantities that satisfy all three assumptions. Let us ask the following question: what is the probability that the subsystem is in a microstate r, in which these quantities have values $Q_1^r, Q_2^r, .., Q_m^r$? Let the total (conserved) values of these quantities, for the entire system, be $Q_1^0, Q_2^0, .., Q_m^0$. Then, given the subsystem is in microstate r, these additive quantities will take values $Q_1^0 - Q_1^r, Q_2^0 - Q_2^r, .., Q_m^0 - Q_m^r$ for the environment. Given these values, there will be many different microstates of the environment, all corresponding to these quantities taking these fixed values. Let the number of such microstates of the environment be $\mathcal{N}^{env}(Q_1^0 - Q_1^r, Q_2^0 - Q_2^r, .., Q_m^0 - Q_m^r)$. Since all microstates of the entire system are equally probable, the probability that the subsystem is in the particular microstate r is given by

$$P_r = \frac{\mathcal{N}^{env}(Q_1^0 - Q_1^r, Q_2^0 - Q_2^r, .., Q_m^0 - Q_m^r)}{\mathcal{N}(Q_1^0, Q_2^0, .., Q_m^0)} \tag{4.18}$$

where $\mathcal{N}(Q_1^0, Q_2^0, .., Q_m^0)$ is the total number of microstates of the entire system corresponding to conserved values $Q_1^0, Q_2^0, .., Q_m^0$. Let the entropy function for the environment be $S^{env}(Q_1, Q_2, .., Q_m)$, which will be a function of the values $Q_1, Q_2, ...Q_m$ of the additive quantities for the environment. This is related to $\mathcal{N}^{env}$ by (2.1)

$$S^{env}(Q_1, Q_2, .., Q_m) = k_B \ln \mathcal{N}^{env}(Q_1, Q_2, .., Q_m) \tag{4.19}$$

which implies

$$\mathcal{N}^{env}(Q_1, Q_2, .., Q_m) = e^{[S^{env}(Q_1, ..Q_m)/k_B]} \tag{4.20}$$

Therefore

$$P_r = \frac{1}{\mathcal{N}(Q_1^0, .., Q_m^0)} e^{\left[S^{env}(Q_1^0 - Q_1^r, ..Q_m^0 - Q_m^r)/k_B\right]} \tag{4.21}$$

The term S^{env} in the exponent can be expanded in a Taylor series about the point $Q_1^r = Q_2^r = .. = Q_m^r = 0$. Retaining terms upto order Q_i^r, we get

$$\begin{aligned} S^{env}(Q_1^0 - Q_1^r, ..Q_m^0 - Q_m^r) &\simeq S^{env}(Q_1^0, Q_2^0, ..Q_m^0) - \left(\frac{\partial S^{env}}{\partial Q_1^0}\right) Q_1^r - \left(\frac{\partial S^{env}}{\partial Q_2^0}\right) Q_2^r ... - \left(\frac{\partial S^{env}}{\partial Q_m^0}\right) Q_m^r \\ &= S^{env}(Q_1^0, Q_2^0, ..Q_m^0) - \alpha_1 Q_1^r - \alpha_2 Q_2^r .. - \alpha_m Q_m^r \end{aligned} \tag{4.22}$$

where $\alpha_1, \alpha_2, .., \alpha_m$ are constants depending on the properties of the environment. Finally, we get the general probability distribution for the subsystem

$$P_r = c\ e^{-\left(\alpha_1 Q_1^r + \alpha_2 Q_2^r .. + \alpha_m Q_m^r\right)/k_B} \tag{4.23}$$

where

$$c = \frac{1}{\mathcal{N}(Q_1^0, .., Q_m^0)} e^{S^{env}(Q_1^0, Q_2^0, ..Q_m^0)/k_B} \tag{4.24}$$

is a constant, in the sense that it is independent of the properties of the subsystem. In the Taylor expansion (4.22), we ignored the second and higher order terms. We will see that these terms are suppressed by inverse powers of the total number of particles of the subsystem. Therefore, for a macroscopic subsystem, these terms can usually be ignored.

4.3 The canonical distribution

The most common situation is one in which a subsystem, in equilibrium with an environment, exchanges energy with it. In this situation, the probability distribution (4.23) reduces to

$$P_r = ce^{-\beta E_r} \tag{4.25}$$

where E_r is the energy of the subsystem in microstate r. The constant β is easily seen to be related to the temperature of the environment (and also of the subsystem in equilibrium with it)

$$\begin{aligned} \beta &= \frac{1}{k_B}\left(\frac{\partial S^{env}}{\partial E_0}\right) \\ &= \frac{1}{k_B T} \end{aligned} \tag{4.26}$$

The constant of proportionality in (4.25) is set by the constraint $\sum_r P_r = 1$, which gives

$$c = \frac{1}{\sum_r e^{-\beta E_r}} \tag{4.27}$$

With this, the probability distribution reduces to

$$P_r = \frac{1}{Z} e^{-\beta E_r} \tag{4.28}$$

where

$$Z = \sum_r e^{-\beta E_r} \tag{4.29}$$

is a sum over all microstates of the subystem, weighed by $e^{-\beta E_r}$. This sum is central to the statistical properties of the subsystem, and is known as the *Partition Function* for the subsystem. If the partition function can be calculated for a system, all equilibrium properties of the system can be computed. Since the sum is over all possible microstates of the system, the partition function is a function only of the constraints on the system, the temperature of the system (through β), and constants associated with the particles of the system and their interactions. It cannot depend on position or momentum (or spin) of the particles, since these dynamical quantities are summed over in the sum over microstates). Typically, the constraints are the number of particles of the system and the volume of the system (for systems interacting with external fields, the constraints also include the field strength). The probability distribution (4.28) is known as the *Canonical Distribution.* It describes the statistical behaviour of a system in thermal equilibrium with an environment at temperature T, with $\beta = 1/k_B T$.

Let us now justify dropping the second and higher order terms in the Taylor expansion (4.22). If the subsystem exchanges energy with the environment, the expansion is

$$S^{env}(E_0 - E_r) = S^{env}(E_0) - \beta E_r + \frac{1}{2!}\left(\frac{\partial^2 S^{env}}{\partial E_0^2}\right) E_r^2 + ... \tag{4.30}$$

The term quadratic in E_r can be written as

$$\frac{1}{2!}\left(\frac{\partial^2 S^{env}}{\partial E_0^2}\right) E_r^2 \quad = \quad \frac{1}{2!}\left(E_0^2 \frac{\partial^2 S^{env}}{\partial E_0^2}\right)\left(\frac{E_r}{E_0}\right)^2 \tag{4.31}$$

If the number of particles of the environment is N_{env} and the number of particles of the susbsystem is N, then quantity $E_0^2(\partial^2 S^{env}/\partial E_0^2)$ in the parenthesis scales as N_{env} (The entropy and energy of the environment both scale as N_{env}). The ratio E_r/E_0 scales as N/N_{env}. Therefore, the second order term scales as N^2/N_{env}. Since the number of particles of the environment are practically infinitely larger than that of the subsystem, this ratio is vanishingly small. Higher order terms will be suppressed by even higher (inverse) powers of N_{env}. This justifies ignoring second and higher order terms in the expansion (4.22). Given the distribution (4.25), the mean value and fluctuation about the mean of any physically interesting quantity can be computed. Let f be such a physical 'observable'. In microstate r, it will take value f_r. Then, the average of f will be

$$\begin{aligned} \overline{f} &= \sum_r P_r f_r \\ &= \frac{1}{Z}\sum_r f_r e^{-\beta E_r} \end{aligned} \tag{4.32}$$

The fluctuation (squared) about this mean value will be

$$\begin{aligned} \Delta f^2 &= \overline{f^2} - (\overline{f})^2 \\ &= \frac{1}{Z}\sum_r \left(f_r^2 - \overline{f}^2\right) e^{-\beta E_r} \end{aligned} \tag{4.33}$$

As an example, let us compute the mean energy of the system. This energy will fluctuate, since the subsystem constantly exchanges energy with the environment. The mean energy of the system is

$$\begin{aligned} \overline{E} &= \frac{1}{Z}\sum_r E_r e^{-\beta E_r} \\ &= -\frac{1}{Z}\sum_r \frac{\partial}{\partial\beta} e^{-\beta E_r} \\ &= -\frac{1}{Z}\frac{\partial}{\partial\beta}\sum_r e^{-\beta E_r} \\ &= -\frac{1}{Z}\frac{\partial Z}{\partial\beta} \end{aligned} \tag{4.34}$$

This can be written as

$$\overline{E} = -\frac{\partial}{\partial\beta}\ln Z \tag{4.35}$$

Here, the partial derivative with respect to β is computed keeping other quantities on which the partition function depends constant (such as volume, number of particles, external field strengths, etc.). Equation (4.35) gives the mean energy of the system as a function of other parameters (such as volume, number of particles, etc.) and the temperature. The fluctuation (squared) in energy is

$$\Delta E^2 = \overline{E^2} - \overline{E}^2 \tag{4.36}$$

where

$$
\begin{aligned}
\overline{E^2} &= \frac{1}{Z}\sum_r E_r^2 e^{-\beta E_r} \\
&= \frac{1}{Z}\sum_r \frac{\partial^2}{\partial\beta^2} e^{-\beta E_r} \\
&= \frac{1}{Z}\frac{\partial^2}{\partial\beta^2}\sum_r e^{-\beta E_r} \\
&= \frac{1}{Z}\frac{\partial^2 Z}{\partial\beta^2} \\
&= \frac{\partial}{\partial\beta}\left[\frac{1}{Z}\frac{\partial Z}{\partial\beta}\right] - \frac{\partial}{\partial\beta}\left(\frac{1}{Z}\right)\frac{\partial Z}{\partial\beta} \\
&= \frac{\partial}{\partial\beta}\left[\frac{\partial}{\partial\beta}\ln Z\right] + \frac{1}{Z^2}\left(\frac{\partial Z}{\partial\beta}\right)^2 \\
&= \frac{\partial^2}{\partial\beta^2}\ln Z + \left(\frac{\partial}{\partial\beta}\ln Z\right)^2 \qquad (4.37)
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\Delta E^2 &= \overline{E^2} - \overline{E}^2 \\
&= \frac{\partial^2}{\partial\beta^2}\ln Z \qquad (4.38)
\end{aligned}
$$

It is observed that the mean and fluctuation of energy can be computed if we know the form of the partition function. This is generally true. The partition function contains all the information about averages and fluctuations of physically interesting quantities. Therefore, it form the backbone of computational statistical mechanics.

Given that the mean energy $\overline{E}$ is an extensive quantity and β an intensive quantity, if we scale the system size by λ, $\overline{E} \longrightarrow \lambda\overline{E}$, $\beta \longrightarrow \beta$. Therefore, $\ln Z \longrightarrow \lambda \ln Z$. Then, the logarithm of the partition function scales with the size of the system. Then, it follows from (4.38) that $\Delta E \longrightarrow \sqrt{\lambda}\Delta E$. Therefore

$$\frac{\Delta E}{\overline{E}} \longrightarrow \frac{1}{\sqrt{\lambda}}\frac{\Delta E}{\overline{E}}$$

This implies that

$$\frac{\Delta E}{\overline{E}} \propto \frac{1}{\sqrt{N}} \qquad (4.39)$$

where N is the number of particles of the subsystem. This agrees with what we had anticitpated in section 4.1 (see eqn.(4.17). Therefore, the energy of a macroscopic system in equilibrium with an environment is sharply defined. It is (apart from microscopic fluctuations) as if the system is isolated from the environment. If we could carve out this subsystem out of the entire system, keep it separated and isolated, and then put it back, nothing would change, apart from microscopic fluctuations. In this sense, properties of a macrocopic system in thermal equilibrium with an environment with mean energy $\overline{E}$, and the same system isolated, with conserved energy equal to $\overline{E}$, are the same.

To compute other equilibrium properties such as entropy, pressure (in case of a gas) or magnetization (for a magnetic system), we reorganize the sum occuring in the partition function. The sum over microstates can be written as

$$
\begin{aligned}
Z &= \sum_r e^{-\beta E_r} \\
&= \sum_E \mathcal{N}(E) e^{-\beta E} \qquad (4.40)
\end{aligned}
$$

where $\mathcal{N}(E)$ is the number of microstates corresponding to a given value of energy of the system. This way, the sum over microstates has been converted into a sum over all possible energies of the system, with $\mathcal{N}(E)$ being a 'degeneracy factor' accounting for the fact that many microstates can have the same energy. Next, we observe that $\mathcal{N}(E)$ is usually an exponentially increasing function of energy. For example, for a system of N weakly interacting particles, $\mathcal{N}(E) \sim E^{3N/2}$ (eqn.(1.16)). In general, we expect $\mathcal{N}(E) \sim E^{\alpha N}$ where α is some numerical factor. Then, the sum in (4.40) over a function of energy of the form $f(E) = E^{\alpha N} e^{-\beta E}$. Let us show that this function is sharply peaked at a particular value of energy, the same as the mean energy $\overline{E}$ of the system given by (4.35). We rewrite the function $f(E)$ as

$$\begin{aligned} f(E) &= e^{\alpha N \ln E - \beta E} \\ &= e^{g(E)} \end{aligned} \tag{4.41}$$

where $g(E) = \alpha N \ln E - \beta E$. This function has a maximum at a certain value of energy $\tilde{E} = \alpha N/\beta$, with $g''(\tilde{E}) = -\beta^2/\alpha N < 0$. Expanding $g(E)$ about $\tilde{E}$ and retaining terms upto second order

$$g(E) \simeq g(\tilde{E}) - \frac{\beta^2}{2\alpha N}(E - \tilde{E})^2 \tag{4.42}$$

The

$$\begin{aligned} f(E) &\simeq e^{g(\tilde{E})}\, e^{-(\beta^2/2\alpha N)(E-\tilde{E})^2} \\ &= e^{g(\tilde{E})} e^{-\frac{1}{\sigma^2}(\frac{E-\tilde{E}}{\tilde{E}})^2} \end{aligned} \tag{4.43}$$

This is a Gaussian distribution, with width

$$\sigma \simeq \sqrt{\frac{\alpha N}{\beta^2 \tilde{E}^2}} \tag{4.44}$$

Since $\tilde{E}$ scales with the number of particles of the system, $\sigma \sim 1/\sqrt{N}$. For a macroscopic system, this is vanishingly small. Therefore, the function $f(E)$ sharply peaks at $E = \tilde{E}$ and is practcally zero elsewhere. Therefore, in the sum (4.40), there is practically just one term, so that

$$Z \simeq \mathcal{N}(\tilde{E}) e^{-\beta \tilde{E}} \tag{4.45}$$

From (4.35), it immediately follows that this sharply defined energy is just the mean energy $\overline{E}$ of the system. Since $\mathcal{N}(\overline{E})$ is the number of microstates corresponding to this energy, the entropy of the system is

$$S = k_B \ln \mathcal{N}(\overline{E}) \tag{4.46}$$

Taking the logarithm of the partition function gives

$$\ln Z = \ln \mathcal{N}(\overline{E}) - \beta \overline{E} \tag{4.47}$$

Using (4.46), we get

$$S = k_B \ln Z + \frac{\overline{E}}{T} \tag{4.48}$$

Since equation (4.35) gives $\overline{E}$ as a function of temperature and constraints such as volume, etc., equation (4.48) gives the entropy of the system as a function of temperature and constraints. It is useful to recast (4.48) in the following form

$$F = E - TS \tag{4.49}$$

where the quantity

$$F = -k_B T \ln Z \tag{4.50}$$

is known as the *Helmholtz function* or the *Helmholtz Free Energy* (we will now stop using a 'bar' on top of equilibrium values of energy, etc., since they are practically constant). We shall see later why it is called 'free energy'. For now, we show that this function can be used to compute all equilibrium properties of the system. For definiteness, let us take a system of N particles in the gas phase, occupying volume V and at temperature T. The partition function for this system will be a function of T, V and N (apart from microscopic parameters such as masses of particles, interaction strength parameters, etc.). Then $F = F(T, V)$ (suppressing the dependence on the number of particles). Under a tiny change in temperature and volume, F, E and S will change. The change in F is given by

$$dF = dE - TdS - SdT \tag{4.51}$$

Visualising S as a function of E and V,

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E}\right)_V dE + \left(\frac{\partial S}{\partial V}\right)_E dV \\ &= \frac{1}{T}\, dE + \frac{P}{T}\, dV \end{aligned} \tag{4.52}$$

where we have used (2.13). This gives

$$TdS = dE + PdV \tag{4.53}$$

This is just the *First Law of Thermodynamics.* Substituting this in (4.51), we get

$$dF = -PdV - SdT \tag{4.54}$$

From this, it follows that

$$\begin{aligned} P &= -\left(\frac{\partial F}{\partial V}\right)_T \\ S &= -\left(\frac{\partial F}{\partial T}\right)_V \end{aligned} \tag{4.55}$$

This shows that the pressure and entropy of the system can be determined from F directly. Once P and S are know, the energy of the system can be computed using

$$E = F + TS \tag{4.56}$$

As an example, let us once again consider a system of N weakly interacting particles in equilibrium with an environment at temperature T. We wish to determine the mean energy of the system, it entropy and its equation of state (relation between pressure, volume and temperature). First, we need to calculate the partition function (4.29), for which we need to identify the microstates, the energy of a given microstate. As before, the microstate is a specification of position and momentum of every particle of the system (taking into account redundancy due to interchange of identical particles). The energy of a microstate is given by

$$E = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m} \tag{4.57}$$

The sum over microstates can be replaced by an integral

$$\sum_r \longrightarrow \frac{1}{h^3 N!} \int d^3\vec{r}_1 ... d^3\vec{r}_N d^3\vec{p}_1 .. d^3\vec{p}_N \tag{4.58}$$

where $d^3\vec{r}_i = dx_i dy_i dz_i$ and $d^3\vec{p}_i = dp_{xi} dp_{yi} dp_{zi}$ and the indistinguishability of the particles and discretization of phase space have been taken into account. The partition function is then given by

$$Z = \frac{1}{h^{3N} N!} \int d^3\vec{r}_1 ... d^3\vec{r}_N d^3\vec{p}_1 .. d^3\vec{p}_N e^{-\beta \sum_{i=1}^{N} \vec{p}_i^2 / 2m} \tag{4.59}$$

We note that the position and momentum integrals for each particle are independent, since the exponent factorizes. Further, they are all the same

$$\begin{aligned}
Z &= \frac{1}{h^{3N}N!}\int d^3\vec{r}_1...d^3\vec{r}_N d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\vec{p}_1^2/2m}e^{-\beta\vec{p}_1^2/2m}..e^{-\beta\vec{p}_N^2/2m} \\
&= \frac{1}{h^{3N}N!}\left(\int d^3\vec{r}_1\int d^3\vec{p}_1 e^{-\beta\vec{p}_1^2/2m}\right)..\left(\int d^3\vec{r}_N\int d^3\vec{p}_N e^{-\beta\vec{p}_N^2/2m}\right) \\
&= \frac{1}{h^{3N}N!}\left(\int d^3\vec{r}\int d^3\vec{p}e^{-\beta\vec{p}^2/2m}\right)^N \\
&= \frac{1}{h^{3N}N!}Z_1^N \qquad (4.60)
\end{aligned}$$

where

$$Z_1 = \int d^3\vec{r}\int d^3\vec{p}e^{-\beta\vec{p}^2/2m} \qquad (4.61)$$

can be thought of as the partition function for a single particle. The position integral is just a volume integral over the entire available volume, so that it is equal to V. Further, the momentum integral factorizes into integrals for each component, each integral being the same, equal to a Gaussian integral

$$\begin{aligned}
\int d^3\vec{p}e^{-\beta\vec{p}^2/2m} &= \int dp_x dp_y dp_z e^{-\beta p_x^2/2m}e^{-\beta p_y^2/2m}e^{-\beta p_z^2/2m} \\
&= \left(\int dp e^{-\beta p^2/2m}\right)^3 \\
&= \left(\sqrt{\frac{2m\pi}{\beta}}\right)^3 \\
&= \left(\frac{2m\pi}{\beta}\right)^{3/2} \qquad (4.62)
\end{aligned}$$

Therefore

$$Z_1 = V\left(\frac{2m\pi}{\beta}\right)^{3/2} \qquad (4.63)$$

Then, the N-particle partition function (4.59) is

$$Z = \frac{1}{N!}V^N\left(\frac{2m\pi}{h^2\beta}\right)^{3N/2} \qquad (4.64)$$

To calculate the Helmholtz function, we calculate the logarithm of the partition function

$$\begin{aligned}
\ln Z &= N\ln V + N\ln\left(\frac{2m\pi}{h^2\beta}\right)^{3/2} - \ln N! \\
&= N\ln V + N\ln\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} - N\ln N + N \\
&= N\ln\left(\frac{V}{N}\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right) + N \qquad (4.65)
\end{aligned}$$

Then

$$F = -Nk_B T\ln\left(\frac{V}{N}\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right) - Nk_B T \qquad (4.66)$$

The pressure of the system is given by (4.55)

$$\begin{aligned}
P &= -\left(\frac{\partial F}{\partial V}\right)_T \\
&= \frac{Nk_B T}{V} \qquad (4.67)
\end{aligned}$$

which results in the familiar equation of state (2.16)

$$PV = Nk_BT$$

The entropy of the system is

$$\begin{aligned} S &= -\left(\frac{\partial F}{\partial T}\right)_V \\ &= Nk_B \ln\left(\frac{V}{N}\left(\frac{2m\pi k_BT}{h^2}\right)^{3/2}\right) + Nk_B + \frac{3}{2}Nk_B \\ &= Nk_B \ln\left(\frac{V}{N}\left(\frac{2m\pi k_BT}{h^2}\right)^{3/2}\right) + \frac{5}{2}Nk_B \end{aligned} \tag{4.68}$$

Finally, the energy of the system is calculated using (4.56)

$$\begin{aligned} E &= F + TS \\ &= -Nk_BT + \frac{5}{2}Nk_BT \\ &= \frac{3}{2}Nk_BT \end{aligned} \tag{4.69}$$

which agrees with (2.9). The expression for entropy is essentially the same as obtained in section 2.2 (eqn.(2.8)).
Let us deduce the position and momentum distribution for a single particle of this system using the Canonical probability distribution. The probability of a microstate, which corresponds to the positions and momenta of the N particles precisely defined, is given by eqn.(4.28). For this system,

$$E_r = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m} \tag{4.70}$$

and the partition function is given by (4.59) Then, the probability of a microstate in which particle 1 has position $\vec{r}_1$ and momentum $\vec{p}_1$, particle 2 has position $\vec{r}_2$ and momentum $\vec{p}_2$ ...... is given by

$$P(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2..) = \frac{1}{Z} e^{-\beta \sum_i \vec{p}_i^2/2m} \tag{4.71}$$

Then, the probability that a particle (say particle 1) has momentum $\vec{p}$ and position $\vec{r}$, irrespective of the positions and momenta of the other particles is given by

$$\begin{aligned} P(\vec{r},\vec{p})d^3\vec{r}d^3\vec{p} &= \frac{1}{Z}\frac{1}{h^{3N}N!} d^3\vec{r}d^3\vec{p} \int d^3\vec{r}_2..d^3\vec{r}_N d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\sum_i \vec{p}_i^2/2m} \\ &= d^3\vec{r}d^3\vec{p}\; e^{-\beta\vec{p}^2/2m}\; \frac{1}{Z}\frac{1}{h^{3N}N!}\int d^3\vec{r}_2..d^3\vec{r}_N d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\sum_{i=2}^{N}\vec{p}_i^2/2m} \end{aligned} \tag{4.72}$$

It follows (from calculation similar to one in (4.60)) that

$$\int d^3\vec{r}_2..d^3\vec{r}_N d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\sum_{i=2}^{N}\vec{p}_i^2/2m} = Z_1^{N-1} \tag{4.73}$$

where Z_1 is given by (4.61) and (4.63). It follows that

$$\begin{aligned} P(\vec{r},\vec{p})d^3\vec{r}d^3\vec{p} &= \frac{1}{Z_1} d^3\vec{r}d^3\vec{p}\; e^{-\beta\vec{p}^2/2m} \\ &= \frac{1}{V}\left(\frac{1}{2m\pi k_BT}\right)^{3/2} d^3\vec{r}d^3\vec{p}\; e^{-\beta\vec{p}^2/2m} \end{aligned} \tag{4.74}$$

The position probability distribution is obtained by integrating over momentum

$$\begin{aligned} P(\vec{r})d^3\vec{r} &= d^3\vec{r}\int d^3\vec{p}P(\vec{r},\vec{p}) \\ &= \frac{1}{V}d^3\vec{r} \end{aligned} \tag{4.75}$$

which shows that the position probability is a constant. Therefore, the probability of detecting a particle anywhere is the same, implying that in equilibrium the system will exhibit uniform density.
The momentum distribution is obtained by integrating over the position variable

$$\begin{aligned} P(\vec{p})d^3\vec{p} &= d^3\vec{p}\int d^3\vec{r}P(\vec{r},\vec{p}) \\ &= \left(\frac{1}{2m\pi k_BT}\right)^{3/2} d^3\vec{p}\; e^{-\beta\vec{p}^2/2m} \end{aligned} \tag{4.76}$$

which is just the Maxwell velocity distribution.
Next, let us analyze a system of N spins (the Ising model in the strong field limit) in equilibrium with an environment at temperature T. A microstate of the system is a specification of spin labels $\{\sigma_1, \sigma_2, .., \sigma_N\}$. The energy of the system for a given microstate is (see eqn.(3.19))

$$E(\sigma_1,\sigma_2,..,\sigma_N) = -\left(\sum_{i=1}^{N}\sigma_i\right)h \tag{4.77}$$

The partition function for the system is given by (4.29)

$$\begin{aligned} Z &= \sum_{\sigma_1\sigma_2..\sigma_N} e^{-\beta E(\sigma_1,\sigma_2,..,\sigma_N)} \\ &= \sum_{\sigma_1\sigma_2..\sigma_N} e^{-\beta E(\sigma_1)}e^{-\beta E(\sigma_2)}..e^{-\beta E(\sigma_N)} \\ &= \left(\sum_{\sigma_1} e^{-\beta E(\sigma_1)}\right)\left(\sum_{\sigma_2} e^{-\beta E(\sigma_2)}\right)..\left(\sum_{\sigma_N} e^{-\beta E(\sigma_N)}\right) \\ &= \left(\sum_{\sigma=-1}^{+1} e^{-\beta E(\sigma)}\right)^N \\ &= Z_1^N \end{aligned} \tag{4.78}$$

where $E(\sigma) = -\sigma h$ and Z_1 can be visualised as the partition function for a single spin, and is given by

$$\begin{aligned} Z_1 &= \sum_{\sigma=-1}^{+1} e^{-\beta E(\sigma)} \\ &= \sum_{\sigma=-1}^{+1} e^{\beta h\sigma} \\ &= e^{-\beta h} + e^{\beta h} \\ &= 2\cosh(\beta h) \end{aligned} \tag{4.79}$$

The partition function for the entire system is

$$\begin{aligned} Z &= Z_1^N \\ &= 2^N\cosh^N(\beta h) \end{aligned} \tag{4.80}$$

The mean energy of the system can be computed using (4.35)

$$\begin{aligned}\overline{E} &= -\frac{\partial}{\partial\beta}\ln Z \\ &= -N\frac{\partial}{\partial\beta}\ln\cosh(\beta h) \\ &= -Nh\tanh(\beta h) \qquad (4.81)\end{aligned}$$

which is, expectedly, the same as (3.27).

Let us now consider a system of N interacting particles with interaction potential energy $U(\vec{r}_1, \vec{r}_2, .., \vec{r}_N)$. The partition function for this system is

$$\begin{aligned}Z &= \frac{1}{h^{3N}N!}\int d^3\vec{r}_1...d^3\vec{r}_N d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\left(\sum_{i=1}^N \vec{p}_i^2/2m + U(\vec{r}_1,..,\vec{r}_N)\right)} \\ &= \frac{1}{h^{3N}N!}\int d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\sum_{i=1}^N \vec{p}_i^2/2m}\int d^3\vec{r}_1...d^3\vec{r}_N e^{-\beta U(\vec{r}_1,\vec{r}_2,..,\vec{r}_N)} \qquad (4.82)\end{aligned}$$

where we observe that the position and momentum integrals factorize. The momentum integrals have been evaluated already

$$\begin{aligned}\int d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\sum_{i=1}^N \vec{p}_i^2/2m} &= \left(\int d^3\vec{p}e^{-\beta\vec{p}^2/2m}\right)^N \\ &= \left(\frac{2m\pi}{\beta}\right)^{3N/2} \qquad (4.83)\end{aligned}$$

Then

$$Z = \frac{1}{h^{3N}N!}\left(\frac{2m\pi}{\beta}\right)^{3N/2} Z_V \qquad (4.84)$$

where

$$Z_V = \int d^3\vec{r}_1...d^3\vec{r}_N e^{-\beta U(\vec{r}_1,\vec{r}_2,..,\vec{r}_N)} \qquad (4.85)$$

can be thought of as the position part of the partition function. This part contains all the information about interactions between the particles, and encodes within it the possibility of the interesting phenomena of phase transitions.

It is interesting to compute the momentum probability distribution for a single particle. As before, the probability that a particle has momentum $\vec{p}$ and position $\vec{r}$, irrespective of the positions and momenta of the other particles is given by

$$\begin{aligned}P(\vec{r},\vec{p})d^3\vec{r}d^3\vec{p} &= \frac{1}{Z}\frac{1}{h^{3N}N!}d^3\vec{r}d^3\vec{p}\int d^3\vec{r}_2..d^3\vec{r}_N d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\left(\sum_{i=2}^N \vec{p}_i^2/2m + U(\vec{r}_1,\vec{r}_2,..,\vec{r}_N)\right)} \\ &= \frac{d^3\vec{r}d^3\vec{p}\ e^{-\beta\vec{p}^2/2m}}{Zh^{3N}N!}\int d^3\vec{r}_2..d^3\vec{r}_N e^{-\beta U(\vec{r},\vec{r}_2,..,\vec{r}_N)}\int d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\sum_{i=2}^N \vec{p}_i^2/2m} \qquad (4.86)\end{aligned}$$

The momentum probability is calculated by integrating over $\vec{r}$

$$\begin{aligned}P(\vec{p})d^3\vec{p} &= d^3\vec{p}\int d^3 r P(\vec{r},\vec{p}) \\ &= \frac{d^3\vec{p}\ e^{-\beta\vec{p}^2/2m}}{Zh^{3N}N!}\int d^3\vec{r}d^3\vec{r}_2..d^3\vec{r}_N e^{-\beta U(\vec{r},\vec{r}_2,..,\vec{r}_N)}\int d^3\vec{p}_2..d^3\vec{p}_N e^{-\beta\sum_{i=2}^N \vec{p}_i^2/2m} \qquad (4.87)\end{aligned}$$

Given the form of the partition function (4.84), it is easy to see that the position integral term will cancel and we are left with the Maxwell distribution as before

$$P(\vec{p})d^3\vec{p} = \left(\frac{1}{2m\pi k_B T}\right)^{3/2} d^3\vec{p}\ e^{-\beta\vec{p}^2/2m} \qquad (4.88)$$

which is very interesting. It tells us that even for a system of interacting particles in equilibrium, the velocity distribution is Maxwellian.

Chapter 5

The Partition Function

5.1 Factorizing the Partition Function

The partition function of a system in equilibrium with an environment at temperature T has the following important property: if the system can be viewed as a collection of weakly interacting subsystems, the partition function for the system is a product of the partition fucntions of these subsystems. Let a system consist of weakly interacting subsystems $A, B, C....$ A microstate of the system is then a complete specification of the microstates of these subsystems. If r is a microstate of the system, it is just the specification $\{r_A, r_B, r_C, ...\}$ where $r_A, r_B, ..$ are the microstates of $A, B, ...$ Since these subsystems are weakly interacting, the total energy of the system in microstate r is

$$E_r = E_{r_A} + E_{r_B} + ... \tag{5.1}$$

Then the partition function of the system is

$$\begin{aligned} Z &= \sum_r e^{-\beta E_r} \\ &= \sum_{r_A, r_B, ..} e^{-\beta(E_{r_A} + E_{r_B} + ...)} \\ &= \sum_{r_A} e^{-\beta E_{r_A}} \sum_{r_B} \sum_{r_B} e^{-\beta E_{r_B}} .. \\ &= Z_A Z_B ... \end{aligned} \tag{5.2}$$

where $Z_A, Z_B, ..$ are partition functions of subsystems $A, B, ...$ This property also applies to different *degrees of freedom* of the same system, so long as they interact weakly with each other. For instance, for a system of spins on a lattice, so long as the spin degree of each entity interacts only weakly with its lattice degree of freedom, the partition function will be a product of the partition functions associated with these degrees of freedom. We observe something similar in case of degrees of freedom of the system associated with different directions in space. As an example, consider a system of N weakly interacting particles. We observed in eqn.(4.60) that the partition function of the system is given by

$$Z = \frac{1}{h^{3N} N!} Z_1^N \tag{5.3}$$

where Z_1 could be visualised as the partition function of a single particle. This happened because the particles were assumed to interact weakly, so that the energy of the system could be written (approximately) as the sum of energies of the particles. If the interaction between particles cannot be ignored, the energy is *not* additive and it makes no sense to visualise a single particle as a subsystem with its own partition function. For the weakly interacting system, however, we also observed that the 'single particle partition fucntion' further factorised into three equal pieces (in the momentum integrals), corresponding to three independent directions of motion. Similarly, for a system of weakly interacting spins (Ising Model in the

strong field limit), the partion function once again factorized (equation (4.78)).
As a consequence of factorization of the partition function, the Helmholtz function of a system consisting of weakly interacting subsystems is the sum of the Helmholtz functions for the subsystems

$$\begin{aligned} F &= -k_BT\ln Z \\ &= -k_BT\ln\left(Z_AZ_B...\right) \\ &= -k_BT\left(\ln Z_A+\ln Z_B+..\right) \\ &= F_A+F_B+...... \end{aligned} \tag{5.4}$$

As an application, let us consider a system of two weakly interacting species of particles (A and B), each species further consisting of weakly interacting particles of one kind. For instance, this could be a system of two different inert gases. Let the system be confined to volume V and at temperature T. The partition function of the system is given by

$$Z = Z_AZ_B \tag{5.5}$$

where it follows from eqn.(4.64) that

$$\begin{aligned} Z_A &= \frac{1}{N_A!}V^{N_A}\left(\frac{2m_A\pi k_BT}{h^2}\right)^{3N_A/2} \\ Z_B &= \frac{1}{N_B!}V^{N_B}\left(\frac{2m_B\pi k_BT}{h^2}\right)^{3N_B/2} \end{aligned} \tag{5.6}$$

where m_A and M_B are masses of the particles of the two species and N_A, N_B the number of particles. The Helmhotz function of the system is

$$\begin{aligned} F &= F_A+F_B \\ &= -N_Ak_BT\ln\left(\frac{V}{N_A}\left(\frac{2m_A\pi k_BT}{h^2}\right)^{3/2}\right)-N_Bk_BT\ln\left(\frac{V}{N_B}\left(\frac{2m_B\pi k_BT}{h^2}\right)^{3/2}\right) \\ &\quad -(N_A+N_B)k_BT \end{aligned} \tag{5.7}$$

The pressure and entropy of the system are given by (4.55)

$$\begin{aligned} P &= -\left(\frac{\partial F}{\partial V}\right)_T \\ &= -\left(\frac{\partial F_A}{\partial V}\right)_T-\left(\frac{\partial F_B}{\partial V}\right)_T \\ &= \frac{N_Ak_BT}{V}+\frac{N_Bk_BT}{V} \\ &= P_A+P_B \end{aligned} \tag{5.8}$$

where $P_A = N_Ak_BT/V$ and $P_B = N_Bk_BT/V$ are the *partial pressures* of the two systems (pressure exerted if it were the only species present).This is just the Law of partial pressures. Note that $P = P_A + P_B$ even if each species consists of interacting particles, so long as the two species interact only weakly with each other. This is just a consequence of the additive nature of the Helmholtz functions.

5.2 Partition function and specific heat

Let us now consider a striking property of the partition function: its behaviour under change of degrees of freedom. In this context, let us ask: what allows us to treat atoms as rigid ‘billiard balls’ when computing equilibrium properties of gases? For instance, if we wish to compute equilibrium properties of an inert gas under conditions of ‘high temperature’ and ‘low pressure’, we model the system as a collection of spherical 'particles' interacting through hard sphere interactions, ignoring the internal structure of the atoms. That is, we compute the partition function assuming that a microstate of the system is a specification of the position and momentum of each atom, implicitly ignoring their internal structure. Why don't we consider the system as a collection of protons, neutrons and electrons in equilibrium, with a microstate being specified by the position and velocity of the centre of mass of each atom, and also the quantum 'motion' of the nucleons and the electrons about the centre of mass? The answer lies in the inherent discreteness imposed by quantum mechanics, in addition to the Boltzmann probability distribution for microstates of a system. Say, we model an atom as a rigid nuclear core (ignoring internal nuclear dynamics) and electrons which occupy quantized energy states. Then, a microstate of the system is a specification of the position and momentum of the rigid core, along with the electronic energy states. Let us now ask: what is the probability of an electron being in an excited state relative to the probability of it occupying the lowest energy (ground) state at temperature T? The probability is given by the Boltzmann distribution

$$\frac{P_{exc}}{P_{grnd}} = e^{-\Delta\epsilon/k_B T} \tag{5.9}$$

where $\Delta\epsilon$ is the energy difference between the ground and excited states, and is of the order of an electron volt. Then, this relative probability is negligible, unless the temperature T is of the order of $T_{elec} = \Delta\epsilon/k_B$, which is of the order of $T_{elec} \sim 10000^0 K$. The discreteness in energy introduced by quantum mechanics has introduced a natural temperature scale T_{elec}, and for $T < T_{elec}$, just about all the electrons in such a gas system are expected to be in their ground states. That is, their dynamics are 'frozen' because there is not eneough thermal energy available to excite them. However, it is an excellent approximation to ignore the discreteness of the energy associated with the centre of mass motion of the rigid core, since these energy states are just the 'particle in a box' states, the box here being the container in which the gas is enclosed. Given the macroscopic size of the volume of this container, the discreteness in these energy states can be ignored (the energy eigenvalues depend on volume as $V^{-2/3}$). There being no discreteness in this energy spectrum, there is no inherent temperature scale associated with it, and therefore these translational dynamics are excited at even the lowest of temperatures. Then, at 'ordinary' temperatures, an atom is well approximated as a rigid, spherical 'lump' with no internal dynamics. But, what about molecules? Is it a good approximation to consider molecules as rigid objects which, in addition to translational motion, undergo rigid rotations? Or, do we need to consider a molecule as a system of atoms, with the dynamics of individual atoms relevant? Again, we expect there to be intrinsic temperature scales with each possible internal dynamics of a molecule. Unless the temperature is high enough top excite these (quantum) dynamics, they will be 'frozen'.

What would be a possible empirical signature of such internal dynamics being 'unfrozen' due to an increase in temperature? Say, we have a system with two different kinds of degrees of freedom, with an energy scale (and therefore temperature scale) associated with each degree. The microstate of the system is a specification of each of these degrees. Let us denote this microstate with two indices rr' with r being a specification of the state of the degree of freedom with an associated temperature scale T_0 which is much lower than that associated with the other degree, a state of which is specified by index r'. Let this higher temperature scale be T_1. Then, for $T_0 << T << T_1$, just about all the states associated with the first degree will be excited, and none of the states associated with the second degree will be excited. The partition function of the system in this temperature range will be

$$Z \simeq \sum_r e^{-\beta E_{rr'_0}} \tag{5.10}$$

where r'_0 is the lowest energy state associated with the second degree of freedom. As long as $T << T_1$, the equilibrium properties of the system, such as its energy, entropy, etc. will recieve a contribution only from the first degree of freedom, the second one effectively 'frozen'. However, as the temperature of the system is raised and becomes comparable to T_1, suddenly the second degree of freedom will 'wake up' and the microstates associated with it will be excited. One will now have a transitional region of temperatures in which the nature of the partition function will start changing as it starts receiving contributions from these microstates. When $T >> T_1$, effectively all the mocrostates associuated with this aspect of the system will be excited, so that the partition function will change to

$$Z \to \sum_{r,r'} e^{-\beta E_{rr}} \tag{5.11}$$

What could be a possible empirical signature of this transition? The observable macroscopic quantity most directly associated with the partition function is the (mean) energy of the system, given by (4.35)

$$\overline{E} = -\frac{\partial}{\partial \beta} \ln Z$$

However, the energy of a system is not easily measured. However, the *response* to the energy to a change in temperature *is* measarable in terms of the *heat capacity* of the system (at constant volume)

$$C_v = \left(\frac{\partial \overline{E}}{\partial T} \right)_V \tag{5.12}$$

However, it is useful to define a related quantity which is independent of the size of the system. Therefore, we define the *Molar Specific Heat* as the heat capacity of one mol of the system

$$c_v = \left(\frac{\partial \overline{E}}{\partial T} \right)_V, \quad N = N_A \tag{5.13}$$

where N_A is Avogadro number. As the nature of the partition function changes (with the frozen degrees coming alive), the dependence of the energy of the system on its temperature will change fundamentally. This can be detected by measuring the response of the energy of the system to its temperature, captured by the specific heat of the system. Therefore, the specific heat of a system can be used as a tool to discover microscopic degrees of freedom of a system in equilibrium. A sudden change in the behaviour of the specific heat of a system is a signature that the system possesses more degrees of freedom than one excpected it to have, and therefore, can be used to discover existence of possibly unknown degees of freedom.

5.3 The diatomic molecule

As a concrete example of degress of freedom 'waking up' with an increase in temperature, consider a gas of diatomic molecules, such as H_2. We assume that conditions are such that interaction between molecules can be ignored. Then, the partition function of a system of N molecules will be of the form

$$Z = \frac{1}{N! h^{3N}} Z_1^N \tag{5.14}$$

where Z_1 is the partition function for a single molecule. We assume that the two atoms in the molecule are identical, and the temperature is well below that at which internal electronic excitations of the individual atoms become significant. Then, each atom in the molecule can be visualised as a point-like entity, its internal structure not probed. To compute the partition function for a molecule, we ned the expression for energy of a molecule in terms of its microstates. If classical physics were a good approximation, a microstate of the molecule is a precise description of the position and momentum of each atom in the molecule. Let us choose a Cartesian coordinate system, such that (x_1, y_1, z_1) and (x_2, y_2, z_2) are coordinates of the two atoms. Then, the exnergy of the molecule has the form

$$E = \frac{1}{2} m \left(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2 \right) + \frac{1}{2} m \left(\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2 \right) + V(r) \tag{5.15}$$

where m is the mass of each atom and V is the potential energy of the pair of atoms, which depends on their relative separation $r = \sqrt{x^2 + y^2 + z^2}$, where $x = x_1 - x_2, y = y_1 - y_2$ and $z = z_1 - z_2$ are the relative coordinates. Let X, Y, Z be the coordinates of the centre of mass of the two atoms. Then

$$\begin{aligned} X &= \frac{x_1 + x_2}{2} \\ Y &= \frac{y_1 + y_2}{2} \\ Z &= \frac{z_1 + z_2}{2} \end{aligned} \tag{5.16}$$

We can eliminate x_1, y_1, z_1 and x_2, y_2, z_2 in favour of the relative coordinates x, y, z and the centre of mass coordinates X, Y, Z. Then, the expression for energy becomes

$$E = \frac{1}{2}M\left(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2\right) + \frac{1}{2}\mu\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + V(r) \tag{5.17}$$

where $M = 2m$ is the total mass of the molecule and $\mu = m/2$ its ‘reduced mass’.

The bond length r_0 of the molecule is just the equilibrium separation between the two atoms, corresponding to the minimum of $V(r)$. Expanding $V(r)$ about this equilibrium separation and retaining terms upto second order, we get

$$\begin{aligned} V(r) &\simeq V(r_0) + \frac{1}{2}V''(r_0)(r - r_0)^2 \\ &= V(r_0) + \frac{1}{2}k(r - r_0)^2 \end{aligned} \tag{5.18}$$

where $k = V''(r_0)$. We now introduce spherical polar coordinates (r, θ, ϕ) to describe the relative separation of the two atoms

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned} \tag{5.19}$$

With this, the expression for energy of the system becomes (after dropping the constant term $V(r_0)$

$$E = \frac{1}{2}M\left(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2\right) + \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}k(r - r_0)^2 + \frac{1}{2}\mu r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.20}$$

Finally, we introduce a radial coordinate $\eta = r - r_0$, following which the energy expression becomes

$$E = \frac{1}{2}M\left(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2\right) + \frac{1}{2}\mu\dot{\eta}^2 + \frac{1}{2}\mu\omega^2\eta^2 + \frac{1}{2}\mu r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.21}$$

where $\omega = \sqrt{k/\mu}$. In this form, there is a clear separation of the energy expression into three terms. The first piece corresponds to energy associated with the translational motion of the molecule

$$\begin{aligned} E_{trans} &= \frac{1}{2}M\left(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2\right) \\ &= \frac{1}{2M}\vec{P}^2 \end{aligned} \tag{5.22}$$

The second term corresponds to the vibrational motion of the two atoms about their equilibrium separation

$$\begin{aligned} E_{vib} &= \frac{1}{2}\mu\dot{\eta}^2 + \frac{1}{2}\mu\omega^2\eta^2 \\ &= \frac{1}{2\mu}p_\eta^2 + \frac{1}{2}\mu\omega^2\eta^2 \end{aligned} \tag{5.23}$$

where $p_\eta = \mu\dot{\eta}$ is the momentum associated with the radial motion. The third piece is associated with the rotational motion of the molecule about its centre of mass. To see this, we first approximate r by the equlibrium separation r_0 in the third piece. Then, this piece becomes

$$E_{rot} = \frac{1}{2}\mu r_0^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.24}$$

In the centre of mass frame, assuming for a moment that the separation between the atoms is fixed, equal to the bond length r_0, the only possible motion is rotational, with the centre of mass as rest. The kinetic energy associated with this motion is just

$$E_{rot} = \frac{1}{2}m\vec{v}_1^2 + \frac{1}{2}m\vec{v}_2^2 \tag{5.25}$$

where $\vec{v}_1$ and $\vec{v}_2$ are the velocities of the two atoms in the centre of mass frame. However, since $m_1\vec{v}_1 + m_2\vec{v}_2 = 0$, we can express these velocities in terms of the relative velocity $\vec{v} = \vec{v}_1 - \vec{v}_2$ of the two atoms

$$\begin{aligned} \vec{v}_1 &= \frac{1}{2}\vec{v} \\ \vec{v}_2 &= -\frac{1}{2}\vec{v} \end{aligned} \tag{5.26}$$

Then, the kinetic energy of the system is

$$\begin{aligned} E_{rot} &= \frac{1}{4}m\vec{v}^2 \\ &= \frac{1}{\mu}(v_x^2 + v_y^2 + v_z^2) \\ &= \frac{1}{2}\mu\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) \end{aligned} \tag{5.27}$$

which is the same as (5.24), assuming the separation between the atoms is fixed. Then, the expression for energy is additive in the three kinds of motions, which in this approximation ($r \simeq r_0$ in the third term) are completely independent

$$E = E_{trans} + E_{vib} + E_{rot} \tag{5.28}$$

Then, the molecular partition function factorizes

$$Z_1 = Z_{trans}\ Z_{vib}\ Z_{rot} \tag{5.29}$$

Let us compute these three pieces separately. The translational partition function is essentially the same as that we had calculated before (4.63)

$$Z_{trans} = V\left(\frac{2M\pi}{\beta}\right)^{3/2} \tag{5.30}$$

since we have argued that the translational motion of the molecule is expected to be well-approximated by classical physics.

For a moment, let us assume that the vibrational and rotational motions are also well-described by classical physics. The vibrational partition function is computed to be

$$\begin{aligned} Z_{vib} &= \frac{1}{h}\int_{-\infty}^{\infty} d\eta\ dp_\eta e^{-\beta\left[p_\eta^2/2\mu + (1/2)\mu\omega^2\eta^2\right]} \\ &= \frac{1}{h}\int_{-\infty}^{\infty} d\eta\ e^{-(\beta\mu\omega^2/2)\eta^2}\int_{-\infty}^{\infty} dp_\eta\ e^{-(\beta/2\mu)p_\eta^2} \\ &= \frac{1}{h}\left(\frac{2\pi}{\mu\beta\omega^2}\right)^{1/2}\left(\frac{2\pi\mu}{\beta}\right)^{1/2} \\ &= \frac{2\pi}{\beta\omega h} \end{aligned} \tag{5.31}$$

To compute the partition function for the rotational motion, we need to express the rotational energy in terms of canonical momenta and coordinates. Since the energy is purely kinetic, the Lagrangian for the system is

$$L = \frac{1}{2}\mu r_0^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.32}$$

The momentum conjugate to coordinate θ is given by

$$\begin{aligned} p_\theta &= \frac{\partial L}{\partial \dot{\theta}} \\ &= \mu r_0^2 \dot{\theta} \end{aligned} \tag{5.33}$$

and the moimentum conjugate to ϕ is

$$\begin{aligned} p_\phi &= \frac{\partial L}{\partial \dot{\phi}} \\ &= \mu r_0^2 \sin^2\theta \dot{\phi} \end{aligned} \tag{5.34}$$

These equations allow us to eliminate $\dot{\theta}$ and $\dot{\phi}$ in favour of p_θ and p_ϕ

$$\begin{aligned} \dot{\theta} &= \frac{p_\theta}{\mu r_0^2} \\ \dot{\phi} &= \frac{p_\phi}{\mu r_0^2 \sin^2\theta} \end{aligned} \tag{5.35}$$

Then, the energy (Hamiltonian) of the system expressed in terms of coordinates (θ, ϕ) and conjugate momenta $p_\theta, p_\phi)$ is

$$\begin{aligned} E &= p_\theta \dot{\theta} + p_\phi \dot{\phi} - L \\ &= \frac{p_\theta^2}{2\mu r_0^2} + \frac{p_\phi^2}{2\mu r_0^2 \sin^2\theta} \end{aligned} \tag{5.36}$$

The partition function is calculated to be

$$Z_{rot} = \frac{1}{h^2} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_{-\infty}^{\infty} dp_\theta \int_{-\infty}^{\infty} dp_\phi e^{-(\beta/2\mu r_0^2)\ p_\theta^2}\ e^{-(\beta/2\mu r_0^2 \sin^2\theta)\ p_\phi^2} \tag{5.37}$$

The ϕ integral gives 2π and the momentum integrals are both Gaussian. This gives

$$\begin{aligned} Z_{rot} &= \frac{1}{h^2} 2\pi \int_0^{\pi} d\theta \left(\frac{2\pi\mu r_0^2}{\beta}\right)^{1/2} \left(\frac{2\pi\mu r_0^2}{\beta}\sin^2\theta\right)^{1/2} \\ &= \frac{1}{h^2}\frac{4\pi^2\mu r_0^2}{\beta}\int_0^{\pi} d\theta \sin\theta \\ &= \frac{1}{h^2}\frac{8\pi^2\mu r_0^2}{\beta} \end{aligned} \tag{5.38}$$

The total partition function for N molecules is

$$Z = \frac{1}{N! h^{3N}}\ Z_{trans}^N \times Z_{vib}^N \times Z_{rot}^N \tag{5.39}$$

Therefore

$$\ln Z = N\ln(Z_{trans}) + N\ln(Z_{vib}) + N\ln(Z_{rot}) - \ln(N! h^{3N}) \tag{5.40}$$

The mean energy of the diatomic gas is given by

$$\begin{aligned} \overline{E} &= -\frac{\partial}{\partial\beta}\ln Z \\ &= -N\frac{\partial}{\partial\beta}\ln(Z_{trans}) - N\frac{\partial}{\partial\beta}\ln(Z_{vib}) - N\frac{\partial}{\partial\beta}\ln(Z_{rot}) \\ &= N k_B T^2 \left[\frac{\partial}{\partial T}\ln(Z_{trans}) + \frac{\partial}{\partial T}\ln(Z_{vib}) + \frac{\partial}{\partial T}\ln(Z_{rot})\right] \end{aligned} \tag{5.41}$$

where partial derivatives with respect to $\beta = 1/(k_B T)$ have been traded for partial derivatives with respect to T. Clearly,

$$\overline{E} = \overline{E}_{trans} + \overline{E}_{vib} + \overline{E}_{rot} \tag{5.42}$$

where $\overline{E}_{trans} = N k_B T^2 (\partial Z_{trans}/\partial T)$, etc. The mean energy of the system is the sum of mean energies associated with translational, vibrational and rotational degrees of freedom. As observed in section(5.2), the specific heat of a system gives a glimpse into microscopic degrees of freedom of a system in thermal equilibrium. The molar heat capacity of the diatomic gas is

$$\begin{aligned} c_v &= \left(\frac{\partial \overline{E}}{\partial T}\right)_V, \quad N = N_A \\ &= c_{trans} + c_{vib} + c_{rot} \end{aligned} \tag{5.43}$$

where

$$\begin{aligned} c_{trans} &= \left(\frac{\partial \overline{E}_{trans}}{\partial T}\right)_V, \quad N = N_A \\ &= N_A k_B \frac{1}{\partial T}\left[T^2 \frac{\partial}{\partial T} \ln(Z_{trans})\right] \\ &= R\,\frac{1}{\partial T}\left[T^2 \frac{\partial}{\partial T} \ln(Z_{trans})\right] \end{aligned} \tag{5.44}$$

where $R = N_A k_B$ is the so-called ‘gas constant’. Similarly,

$$\begin{aligned} c_{vib} &= R\,\frac{1}{\partial T}\left[T^2 \frac{\partial}{\partial T} \ln(Z_{vib})\right] \\ c_{rot} &= R\,\frac{1}{\partial T}\left[T^2 \frac{\partial}{\partial T} \ln(Z_{rot})\right] \end{aligned} \tag{5.45}$$

Given the epressions for Z_{trans}, Z_{vib} and Z_{rot}, we get

$$\begin{aligned} c_{trans} &= \frac{3}{2} R \\ c_{vib} &= R \\ c_{rot} &= R \end{aligned} \tag{5.46}$$

which gives the total molar heat capacity to be $c_v = 7R/2$, a constant. Following is an experimentally determined molar specific heat for hydrogen gas

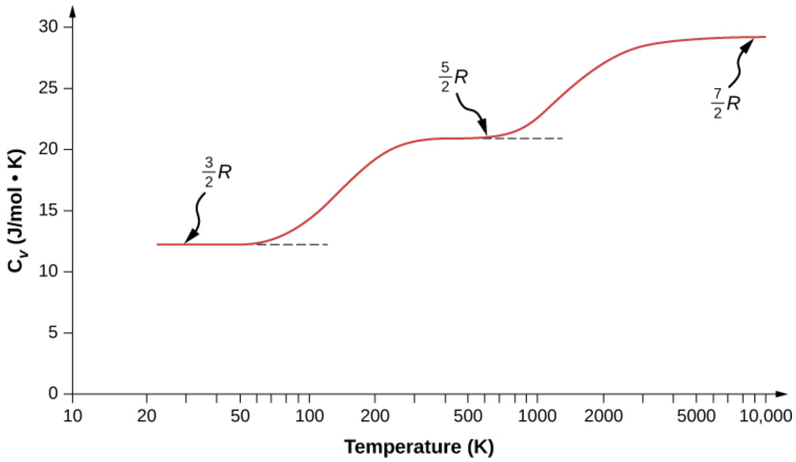


Figure 5.1: Variation of molar specific heat of hydrogen gas with temperature (shown on a logarithmic scale)

The figure shows an interesting behaviour. Given the values of c_{trans}, c_{vib} and c_{rot} (eqns.(5.46)), it appears that below $T \sim 50^0 K$, only translational degrees of motion are active, contributing $3R/2$ to molar specific heat. It is as if the vibrational and rotational motion of the molecules is 'frozen'. Above $T \sim 100^0 K$, there is a change in the behaviour of the heat capacity. It starts to increase up to about $T \sim 300^0 K$, after which it saturates to a constant value $c_v = 3R/2 + R = 5R/2$, till about $T \sim 1000^0 K$, after which there is again a rise in the heat capacity, till it once again saturates at about $T \sim 3000^0 K$ to $c_v = 3R/2 + R + R = 7R/2$. This tells us that there are two intrinsic temperature scales, one about $T_1 \sim 100^0 K$ and the other about $T_2 \sim 1000^0 K$. These are clearly temperature scales associated with vibrational and rotational motion of the diatomic molecule. Further, it is clear that quantum mechanics is at work here, the discrete energy spacing between energy states of rotational and vibrational degrees inducing the two temperature scales. Which of the two corresponds to T_1 and which to T_2 will be explored next.

5.4 A first look at quantum statistical mechanics

The statistical mechanics of quantum systems is a fairly intricate topic, and an open area of research. Instead of going through the foundations of quantum statistical mechanics, in this introductory course, we will take the following point of view: given a system, described by quantum mechanics, in thermal equilibrium with an environment at temperature T, all equlibrium (classical) properties of the system are still computed exactly as for classical systems, except that the microstates are now quantum microstates. Therefore, the partition function is still the quantity of fundamental interest, except that the sum over microstates is now a sum over quantum states. Here, we immediately encounter a problem. Assuming (incorrectly) that the state of the system in equilibrium is represented by a vector in a Hilbert space, what are these microstates? Since all observables in quantum mechanics are represented by Hermitian operators, with eigenvectors of these operators furnishing alternative bases in terms of which the state of the system can be 'resolved', one could take the (complex) expansion coefficients in a given basis set as a specification of the quantum microstate of the system. However, given that there are in principle several such observables that one could come up with (which have physical relevance, apart from mathematically being Hermitian), there seems to be an ambiguity as to which basis set is to be preferred. Since the partition function involves a sum over states with the summand being exponential functions of energy, the eigenvectors of the Hamiltonian present themselves as a preferred basis. Then, the quantum partition function is to be computed as the sum (4.29)

$$Z = \sum_r e^{-\beta E_r} \tag{5.47}$$

where E_r are the energy eigenvalues of the system, which depend on the Hamiltonian of the system. For instance, for a system of weakly interacting particles confined to a volume V, the single particle partition function will involve sum over energy eigenstates corresponding to the particle being confined to a 'box' of volume V. These eigenvalues are well known from any elementary course in quantum mechanics, and we will use them at some point. Once the partition function has been so computed, all else, which corresponds to extracting equlibrium information about macroscopic, *classical* observables (mean energy, pressure, entropy, etc.) can be obtained as before from the Helmholtz function (4.50)

$$F = -k_B T \ln Z \tag{5.48}$$

As mentioned, this is at best an *operational* point of view. What exactly *are* the foundations of quantum statistical mechanics? What replaces the ideas of phase space, maximization of phase space volume for isolated systems and principle of equal a-priori probabilities, from which we deduced the Boltzmann distribution? The current understanding seems to be the following: the state of a system in equilibrium with an environment is *not* described by a vector, but rather a *density operator*. The reason is that in quantum mechanics, a subsystem of a larger system is *entangled* with its environment. If observations are made on this subsystem, the results of these observations are consistent with the subsystem, instead of being in a state which can be represented as a vector (a 'pure' state), are consistent with the subsystem being in an *ensemble* of states each of which can be reprsented as a vector. That is, the subsystem behaves as if it is 'fluctuating' between quantum states which can be visualised as vectors, with a set of probabilities describing the relative weights of these 'pure' states. The key problem at the heart of the foundations of quantum statistical mechanics seems to be to justify from first principles why a subsystem, in equilibrium at temperature T, behaves as if it is described by an ensemble of energy eigenstates, which occur with the Boltzmann probability distribution. We shall gloss over these questions and assume that there is some justification which allows us to visualise a subsystem in equilibrium as an ensemble of energy eigenstates occuring with the Boltzmann probability distribution. That is, the probability that the subsystem is in energy eigenstate $|E_r\rangle$ is

$$P_r = \frac{1}{Z} e^{-\beta E_r} \tag{5.49}$$

Using this approach, let us analyse the behaviour of the molar specific heat of a diatomic molecule. We have seen that it is a poor approximation to assume that the vibrational and rotational degrees of freedom are well described by classical physics. Let us consider the single particle vibrational partition function Z_{vib}. The expression for energy of the system associated with vibrational motion is given by (5.23)

$$E_{vib} = \frac{1}{2\mu} p_\eta^2 + \frac{1}{2}\mu\omega^2\eta^2 \tag{5.50}$$

which is just the expression for energy of a simple harmonic oscillator. If we promote this to an operator in quantum mechanics, the energy eigenvalues are well known

$$\epsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{5.51}$$

where $n = 0, 1, 2, ...$ is a natural number. The partition function is then

$$\begin{aligned} Z_{vib} &= \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(n+1/2)} \\ &= e^{-\beta\hbar\omega/2} \sum_{n=0}^{\infty} x^n \end{aligned} \tag{5.52}$$

where $x = e^{-\beta\hbar\omega}$. The series is simple to sum since $x < 1$. The result is

$$Z_{vib} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \tag{5.53}$$

Next, let us compute the rotational partition function. the expression for rotational energy of the molecule is (5.24)

$$E_{rot} = \frac{1}{2}\mu r_0^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.54}$$

Let us compute the angular momentum of the diatomic molecule about its centre of mass. Let the positions of the two masses relative to the centre of mass be $\vec{r}_1$ and $\vec{r}_2$ and their velocities in the centre of mass frame be $\vec{v}_1$ and $\vec{v}_2$. The angular momentum vector for the system about the centre of mass is

$$\begin{aligned} \vec{J} &= m\vec{r}_1 \times \vec{v}_1 + m\vec{r}_2 \times \vec{v}_2 \\ &= \frac{m}{2}\vec{r} \times \vec{v} \end{aligned} \tag{5.55}$$

where $\vec{r} = \vec{r}_1 - \vec{r}_2 = 2\vec{r}_1$ and $\vec{v} = \vec{v}_1 - \vec{v}_2$. Since the motion is purely rotational, $\vec{v}$ is orthogonal to $\vec{r}$. Then, the magnitude of angular momentum is

$$J = \mu r_0 v \tag{5.56}$$

Using spherical polar coordinates for relative position of the two atoms, the relative position is given by

$$\vec{r} = r_0 \hat{r} \tag{5.57}$$

where $\hat{r}$ is a unit vector along the (relative) radial direction. Therefore, the relative velocity is

$$\begin{aligned} \vec{v} &= \frac{d\vec{r}}{dt} \\ &= r_0 \left(\dot{\theta}\ \hat{\theta} + \sin\theta \dot{\phi}\ \hat{\phi} \right) \end{aligned} \tag{5.58}$$

where $\hat{\theta}$ and $\hat{\phi}$ are unit vectors associated with the two angular directions. Then

$$v^2 = r_0^2 \left(\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2 \right) \tag{5.59}$$

The magnitude squared of the angular momentum is then

$$\begin{aligned} J^2 &= \mu^2 r_0^2 v_2 \\ &= \mu^2 r_0^2 \left(\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2 \right) \\ &= 2\mu r_0^2 E_{rot} \end{aligned} \tag{5.60}$$

which gives

$$E_{rot} = \frac{J^2}{2\mu r_0^2} \tag{5.61}$$

Again, this expression is promoted to a corresponding operator expression in which the classical magnitude (squared) of the angular momentum is replaced by a Hermitian operator $\hat{J}^2$, which has eigenvalues of the form

$$J_l^2 = l(l+1)\hbar^2, \quad l = 0, 1, 2, ... \tag{5.62}$$

Correspondingly, the energy eigenvalues are

$$\epsilon_l = \frac{\hbar 2}{2\mu r_0^2}\ l(l+1) \tag{5.63}$$

Further, there is a *degeneracy* associated with each value of l: there are $2l+1$ quantum states all with the same value of l. Physically, this is because for a fixed magnitude of angular momentum, there are $2l+1$ values that its component in any one direction can possess. Since the energy depends only on the magnitude of the angular momentum, all these states will be *degenerate*. Then, the partition function corresponding to rotations is given by

$$Z_{rot} = \sum_{l=0}^{\infty} (2l+1) e^{-(\beta\hbar^2/2\mu r_0^2)\ l(l+1)} \tag{5.64}$$

Clearly, there are natural temperature scales associated with the vibrational and rotational motion. A look at equation (5.52) shows that for vibrations, this scale is $T_v \sim \hbar\omega/k_B$ and for rotations, it is $T_r \sim \hbar^2/2\mu r_0^2 k_B$. For the H_2 molecule, $\omega \sim 1.3 \times 10^{14} s^{-1}$. This gives $T_v \sim 1000^0 K$, which is of the order of the temperature scale T_2 observed in the behaviour of the heat capacity of H_2 in section (5.3). The bond length of H_2 molecule is $r_0 \sim 0.7\mathring{A}$, which gives $T_r \sim 100^0 K$, which we identify with the temperature scale T_1 associated with the behaviour of the heat capacity of hydrogen.

Given the exact expressions for Z_{vib} and Z_{rot}, let us estimate their form when the temperature is much larger than the corresponding characteris tic temperature. In the limit $T >> T_v$, we can approximate Z_{vib} as

$$\begin{aligned} Z_{vib} &= \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \\ &\simeq \frac{1}{\beta\hbar\omega} \\ &= \frac{2\pi}{\beta\omega h} \end{aligned} \tag{5.65}$$

which agrees with (5.31). Next, we compute Z_{rot} in the limit $T >> T_r$. Z_{rot} can be written as

$$\sum_{l=0}^{\infty}(2l+1)e^{-T_r/T\ l(l+1)}\Delta l \tag{5.66}$$

where we have inserted $\Delta l = 1$ in the sum. Define a variable $x = l(l+1)T_r/T$ such that as l changes by one unit, x changes infinitesimally (in the limit $T >> T_r$). The change in x for a unit change in l is $dx = (2l+1)\Delta l T_r/T$, so that $(2l+1)\Delta l = (T/T_r)dx$. Substituting in (5.66), we get

$$\begin{aligned} Z_r &\simeq \frac{T}{T_r}\int_0^{\infty} e^{-x} \\ &= \frac{T}{T_r} \\ &= \frac{8\pi^2\mu r_0^2}{h^2\beta} \end{aligned} \tag{5.67}$$

which agrees with (5.38).

We can compute the exact behaviour of the molar specific heats associated with the vibrational and rotational degrees of freedom, using partition function expressions (5.53) and (5.64). The agrrement with figure (5.1) is excellent.

5.5 The Equipartition Principle

Consider a system in equilibrium at temperature T whose energy can be expressed in the form

$$E(Q,\alpha) = \lambda Q^2 + \tilde{E}(\alpha) \tag{5.68}$$

where Q is a coordinate or momentum component and α is the set of all coordinates and momentum components, *except* Q. Further, it is assumed that the parameter $\lambda > 0$. The Equipartition Principle states that in equilibrium, the mean value of λQ^2 is $k_B T/2$. Any such coordinate or momentum component is referred to as a *degree of freedom* of the system. Then, the degrees of freedom of a system is the set of all coordinates and momentum components which precisely specify its state. For example, a system of N point-like particles has $6N$ degrees of freedom, since a complete specification of the state of the system requires a specification of $3N$ coordinates and $3N$ momentum components.

In equilibrium, the probability of the system being in microstate (Q,α) is given by the canonical distribution

$$\begin{aligned} P(Q,\alpha) &= \frac{1}{Z}e^{-\beta\tilde{E}(Q,\alpha)} \\ &= \frac{1}{Z}e^{-\beta\lambda Q^2}e^{-\beta\tilde{E}(\alpha)} \end{aligned} \tag{5.69}$$

where Z is the partition function, given by (supressing factors arising out of phase space discretization and identical nature of constituents)

$$\begin{aligned} Z &= \int dQ\int d\alpha\ e^{-\beta\lambda Q^2}e^{-\beta\tilde{E}(\alpha)} \\ &= \int dQ\ e^{-\beta\lambda Q^2}\int d\alpha\ e^{-\beta\tilde{E}(\alpha)} \end{aligned} \tag{5.70}$$

The probability that the degree of freedom takes value Q, irrespective of values taken by the set α, is given by

$$\begin{aligned} P(Q) &= \int d\alpha\ P(Q,\alpha) \\ &= \frac{1}{Z}e^{-\beta\lambda Q^2}d\alpha e^{-\beta\tilde{E}(Q,\alpha)} \\ &= \frac{1}{Z_Q}e^{-\beta\lambda Q^2} \end{aligned} \tag{5.71}$$

where

$$Z_Q = \int dQ\ e^{-\beta\lambda Q^2} \tag{5.72}$$

can be though of as the partition function associated with the degree of freedom Q. Clearly, the probability distribution for Q given by eqn.(5.71) is 'Boltzmannian'. The mean value of the energy term associated with Q can now be computed

$$\begin{aligned}\overline{\lambda Q^2} &= \frac{1}{Z_Q}\int dQ\ (\lambda Q^2)\ e^{-\beta\lambda Q^2}\\ &= -\frac{1}{Z_Q}\frac{\partial}{\partial\beta}\left(\int dQ\ e^{-\beta\lambda Q^2}\right)\\ &= -\frac{1}{Z_Q}\frac{\partial Z_Q}{\partial\beta}\\ &= -\frac{\partial \ln Z_Q}{\partial\beta}\end{aligned} \tag{5.73}$$

which is similar to eqn.(4.35). Z_Q is easily computed (Gaussian integral) to give

$$Z_Q = \sqrt{\frac{\pi}{\beta\lambda}} \tag{5.74}$$

Using this, the expression for mean value of the energy term is

$$\overline{\lambda Q^2} = \frac{1}{2}k_B T \tag{5.75}$$

which proves the principle.
As a simple application of the Equipartition Principle, consider a system of N interacting point-like particles, such that the energy of the system is

$$E = \frac{\vec{p}_1^{\,2}}{2m} + \frac{\vec{p}_2^{\,2}}{2m} + .. + \frac{\vec{p}_N^{\,2}}{2m} + U(\vec{r}_1, .., \vec{r}_N) \tag{5.76}$$

Let us compute the mean kinetic energy of particle 1. It is given by

$$\begin{aligned}\overline{K}_1 &= \frac{1}{2m}\,\overline{\vec{p}_1^{\,2}}\\ &= \frac{1}{2m}\,\overline{p_{1x}^2} + \frac{1}{2m}\,\overline{p_{1y}^2} + \frac{1}{2m}\,\overline{p_{1z}^2}\end{aligned} \tag{5.77}$$

Clearly, each term in (5.77) is a candidate for applying the Equipartition Principle, since each term is quadratic in the momentum component (with positive coefficient) and the energy of the system can be written as such a term plus everything else. Then, $\overline{p_{1x}^2}/2m = \overline{p_{1y}^2}/2m = \overline{p_{1z}^2}/2m = k_B T/2$. Then, the mean kinetic energy per particle of the system is

$$\overline{K}_1 = \frac{3}{2}k_B T \tag{5.78}$$

It immediately follows that the mean kinetic energy of the system of particles is

$$\overline{K} = \frac{3}{2}Nk_B T \tag{5.79}$$

Then, the temperature of a system of interacting particles in equilibrium is a direct measure of the mean kinetic energy of the system. For a weakly interacting system (such that the potential energy contribution can be ignored), the mean energy of the system is given by (5.79), which is what we computed in (4.69).

As another example, let us consider a diatomic gas (such as H_2). The expression for energy for a molecule is given by

$$E = \frac{\vec{P}^2}{2M} + E_{vib} + E_{rot} \tag{5.80}$$

We have observed that it is an excellent approximation to consider the centre of mass degrees $\vec{R}$ (position of centre of mass) and $\vec{P}$ (total momentum of molecule) as classical degrees. It immediately follows that

$$\overline{\frac{\vec{P}^2}{2M}} = \frac{3}{2}k_BT \tag{5.81}$$

Then, the temperature of a diatomic gas is a measure of the mean translational kinetic energy per molecule. It is now easy to understand why the specific heat of such a system rises when a new degree of freedom 'awakens'. The heat capacity of a system is given by (5.13), which essentially translates to the ratio of energy exchanged by the system and a subsequent change in its temperature

$$c_v \simeq \frac{\Delta E}{\Delta T} \tag{5.82}$$

Say, we introduce some energy ΔE into the system before the rotational threshold has been reached. At this point, only the translational degrees are active. Then, since the mean translational energy is proportional to the temperature of the system (eqn.(5.81)), the temperature of the system rises in proportion, such that c_v does not change. However, as the rotational threshold is reached, part of this energy goes into exciting the now accessible rotational excited states, and the remaining fraction of this energy goes into increasing the translational kinetic energy. since the temperature of the system is proportional to this piece of energy only, there is a smaller change in temperature than would have been if the entire energy had gone into increasing the translational kinetic energy. As a result, the specific heat is *more.* The same thing happens when the vibrational degrees are excited.

It should be noted that the Equipartition Principle is a *classical* principle, applicable only of the energy term corresponding to a degree of freedom is *quadratic* in the degree, and is not applicable to quantum degrees of freedom.

Chapter 6

Free Energy

6.1 Helmholtz Free Energy

We have seen that an isolated system undergoes spontaneous processes that result in maximisation of entropy. If the system is initially in equilibrium with a certain parameter λ fixed, if this parameter is allowed to vary, it attains an equilibrium value $\bar{\lambda}$ sauch that $S(\lambda)$ is a maximum for $\lambda = \bar{\lambda}$. Technically, there is a probability distribution for the parameter λ, given by (2.19)

$$P(\lambda) = P(\bar{\lambda})\ e^{-(1/k_B)\left|S''(\bar{\lambda})\right|\Delta\lambda^2} \tag{6.1}$$

which is a Gaussian, peaked at $\lambda = \bar{\lambda}$, with a width which scales with the number of particles of the system as

$$\frac{\Delta\lambda}{\bar{\lambda}} \sim \frac{1}{\sqrt{N}} \tag{6.2}$$

What if the system is not isolated (as is always the case), but in thermal equilibrium with an environment at temperature T? Let the system undergo a spontaneous process such that an initially fixed parameter λ is made 'free'. As the parameter attains anew equilibrium value $\bar{\lambda}$, the system will exchange energy ΔE with the environment. As a result, the entropy of the environment will change by $\Delta S_{env} = -\Delta E/T$. Then, if the change in entropy of the system during the process is ΔS, since the system along with the environment can be considered 'isolated', it follows that

$$\begin{aligned} \Delta S + \Delta S_{env} &> 0 \\ \implies \Delta S - \frac{\Delta E}{T} &> 0 \\ \implies \Delta(E - TS) &< 0 \end{aligned} \tag{6.3}$$

Therefore, the direction of the spontaneous process corresponds to a decrease in the quantity $F = E - TS$, and therefore the new equilibrium will correspond to its minimum. But, the quantity F is just the Helmholtz function encountered in section 4.3 (eqn.(4.49)). We had observed that all equilibrium properties of a system in equilibrium can be computed from this single function (which is directly related to the partition function throught (4.50)). Now, we observe that this function is fundamentally important in another way: a decrease in F indicates the direction of the 'arrow of time' for a system in equilibrium with an environment. Then, F is to a system interacting with an environment what S is to an isolated system. Clearly, since in principle there are no isolated systems, the function F (or its generalisations) lies at the heart of Statistical Mechanics.

Let us make this principle of minimisation of F more precise. For a system in equilibrium at temperature T, the probability of a microstate r is giben by the canonical distribution (4.28)

$$P_r = \frac{e^{-\beta E_r}}{Z} \tag{6.4}$$

Then, the probability of a 'free' parameter taking a certain value λ is

$$P(\lambda) = \frac{1}{Z}\,\mathcal{N}(\lambda)e^{-\beta E(\lambda)} \tag{6.5}$$

where $\mathcal{N}(\lambda)$ is the number of microstates with the free parameter equal to λ and $E(\lambda)$ their energy. This can be expressed as

$$\begin{aligned} P(\lambda) &= \frac{1}{Z}\,e^{S(\lambda)/k_B-\beta E(\lambda)} \\ &= \frac{1}{Z}\,e^{-(E(\lambda)-TS(\lambda))/k_BT} \\ &= \frac{1}{Z}e^{-F(\lambda)/k_BT} \end{aligned} \tag{6.6}$$

where

$$F(\lambda) = E(\lambda) - TS(\lambda) \tag{6.7}$$

is the value of the Helmholtz function for the parameter equal to λ. The exponent can be expanded in a Taylor series. As usual, retaining terms up to second order, we get

$$P(\lambda) = \frac{1}{Z}e^{-F(\bar{\lambda})/k_BT}e^{-(1/2k_BT)F''(\bar{\lambda})(\lambda-\bar{\lambda})^2} \tag{6.8}$$

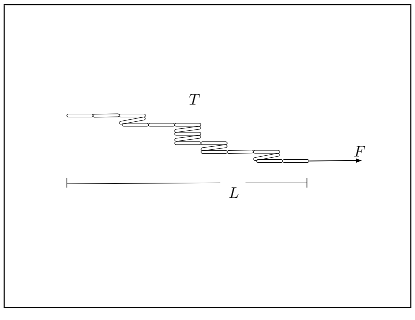
which is a Gaussian with a characteristic width which scales as $\sim 1/\sqrt{N}$. For a macroscopic system, the probability of λ being equal to $\bar{\lambda}$ is essentially unity. The function $F(\lambda)$ is computed as before through the partition function (which depends on the constraint λ aprt from temperature and other fixed constraints)

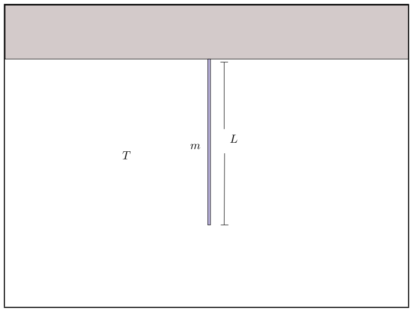
$$F(\lambda) = -k_BT\ln Z(\lambda) \tag{6.9}$$

We now look at a few of the many examples of free energy minimisatio.

6.2 Elasticity of a rubber band

As a simple (but interesting) application of free energy minimisatio, let us analyze the origin of the elasticity of a rubber band from the point of view of minimization of free energy. It is observed that when a rubber band is heated, it contracts, instead of expanding. If a mass is suspended from the rubber band, heating it will result in the mass being raised and work being done by the rubber band. To model this phenomena, we visualise a rubber band as a one-dimensional chain of N links, each of length l_0 which with equal energy point parallel or antiparallel to the previous link (all in one dimension). The total length L of the rubber band is the distance bewtween the beginning of the first link and the end of the last link. We assume that the rubber band has mass m and suspended from a support in an environment with themperature T. We wish to determine the length of the rubber band as a function of the ambient temperature





The chain has no internal energy, buy has entropy. Referring to the above illustration, if the number of links to the 'right' is n_R and the number to the 'left' is n_L, then

$$\begin{aligned} L &= l_0(n_R - n_L) \\ N &= n_R + n_L \end{aligned} \tag{6.10}$$

where N is the total number of links and L is the length of the rubber band. Given the length L, the entropy of the system is

$$\begin{aligned} S(L) &= k_B \ln\left(\frac{N!}{n_R! n_L!}\right) \\ &= -Nk_B\left[\frac{n_L}{N}\ln\left(\frac{n_L}{N}\right) + \frac{n_R}{N}\ln\left(\frac{n_R}{N}\right)\right] \\ &= -\frac{Nk_B}{2}\left[\left(1+\frac{L}{Nl_0}\right)\ln\left\{\frac{1}{2}\left(1+\frac{L}{Nl_0}\right)\right\} + \left(1-\frac{L}{Nl_0}\right)\ln\left\{\frac{1}{2}\left(1-\frac{L}{Nl_0}\right)\right\}\right] \end{aligned} \tag{6.11}$$

The energy of the system is just the potential energy of the rubber band due to gravity. Choosing the origin of potential energy at the point from which the band is suspended, the potential energy of the band is $E(L) = -mgL/2$. Then, the free energy of the system is

$$\begin{aligned} F(L) &= E(L) - TS(L) \\ &= -\frac{mgL}{2} - TS(L) \end{aligned} \tag{6.12}$$

The equilibrium length of the band is given by the condition $F'(\bar{L}) = 0$. Given the expression for F, this results in

$$\begin{aligned} -\frac{mg}{2} &= TS'(\bar{L}) \\ &= -\frac{k_B T}{2l_0}\ln\left(\frac{1+L/Nl_0}{1-L/Nl_0}\right) \end{aligned} \tag{6.13}$$

which is solved to give

$$L = Nl_0 \tanh\left(\frac{mgl_0}{2k_BT}\right) \tag{6.14}$$

At 'hight' temperatures ($T >> T_0 = mgl_0/2k_B$), this approximates to

$$L \simeq L_{max}\frac{T_0}{T} \tag{6.15}$$

where $L_{max} = Nl_0$ is the maximum possible length.

6.3 Mean Field Theory of Ferromagnetism

As another illustration of free energy minimization, let us revisit the Ising model, first in in the strong field limit, and then finally in the zero field limit. For a spin system in thermal equilibrium at temperature T, we had computed the mean magnetization using both the microcanonical and the canonical distribitions (in the latter, we had computed the mean energy, which is equivalent to a computation of the mean magentization in the strong field limit). Let us, however, view this in the light of of minimization of free energy. The magnetization of the system is not fixed, but will fluctuate, since the system exchanges energy with the environment. Let us compute the free energy fo a fixed value of magnetization M (given by (3.18). Fixing M fixes the number of 'up' and 'down' spins to

$$\begin{aligned} n_+ &= \frac{N(1+M)}{2} \\ n_- &= \frac{N(1-M)}{2} \end{aligned} \tag{6.16}$$

where N is the total number of spins. for fixed M, the entropy of the system is

$$\begin{aligned} S &= k_B \ln\left(\frac{N!}{n_+! n_-!}\right) \\ &= -Nk_B\left[\frac{n_+}{N}\ln\left(\frac{n_+}{N}\right) + \frac{n_-}{N}\ln\left(\frac{n_-}{N}\right)\right] \\ &= -Nk_B\left[\frac{(1+M)}{2}\ln\left(\frac{1+M}{2}\right) + \frac{(1-M)}{2}\ln\left(\frac{1-M}{2}\right)\right] \end{aligned} \tag{6.17}$$

The energy of the system, given M, is $E(M) = -MNh$ where h is the external field. The free energy is, therefore, given by

$$\begin{aligned} F(M) &= E(M) - TS(M) \\ &= -MNh + Nk_BT\left[\frac{(1+M)}{2}\ln\left(\frac{1+M}{2}\right) + \frac{(1-M)}{2}\ln\left(\frac{1-M}{2}\right)\right] \end{aligned} \tag{6.18}$$

The derivative of F w.r.t M gives

$$F'(M) = -Nh + \frac{Nk_BT}{2}\ln\left(\frac{1+M}{1-M}\right) \tag{6.19}$$

The equilibrium magnetization $\overline{M}$ is given by $F'(\overline{M}) = 0$, which gives

$$\overline{M} = \tanh\left(\frac{h}{k_BT}\right) \tag{6.20}$$

which agrees with our previous calculation(s). It is easy to check that $F''(\overline{M}) > 0$.
Let us now consider the Ising model with spin-spin interaction included. Is this a viable model of ferromagnetism? The expression for the total energy of the system is (eqn.(3.17))

$$E = -J\sum_{<i,j>}\sigma_i\sigma_j - h\sum_i \sigma_i \tag{6.21}$$

which can be written as

$$E = -J\sum_{<i,j>}\sigma_i\sigma_j - NhM \tag{6.22}$$

Unfortunately, with spin-spin interaction taken into account, the energy is not just a function of magentization, since a given magnetization only fixed to number of 'up' and 'down' spins, whereas the interaction between spins depends on how these 'up' and 'down' spins are distributed. To get around this problem, we use the so-called *Mean Field* approximation. In this approximation, we assume that each spin interacts with an *average* magnetization produced due to all the spins. This approximation, therefore, ignores the local variations in the magnetization (fluctuations). To mathematically incorporate this approximation, we write

$$\sigma_i = M + \Delta\sigma_i \tag{6.23}$$

where $\Delta\sigma_i$ is the fluctuation (assumed small) of the i^{th} spin about the mean magnetization M. Then,

$$\begin{aligned}
E &= -J\sum_{<ij>}(M+\Delta\sigma_i)(M+\Delta\sigma_j) - NhM \\
&= -J\sum_{<ij>}\left(M^2 + M(\Delta\sigma_i+\Delta\sigma_j) + \Delta\sigma_i\Delta\sigma_j\right) - NhM \\
&\simeq -J\sum_{<ij>}\left(M^2 + M(\Delta\sigma_i+\Delta\sigma_j)\right) - NhM \\
&= -JM^2\sum_{<ij>}1. - JM\sum_{<ij>}(\sigma_i+\sigma_j-2M) - NhM \\
&= +JM^2\sum_{<ij>}1. - 2JM\sum_{<ij>}\sigma_i - NhM
\end{aligned} \tag{6.24}$$

where we have ignored terms quadratic in $\Delta\sigma$. The sum over nearest neighbours can be written as

$$\begin{aligned}
\sum_{<ij>}1. &= \frac{1}{2}\sum_i\sum_{i'} \\
&= \frac{q}{2}\sum_i \\
&= \frac{Nq}{2}
\end{aligned} \tag{6.25}$$

where $\sum_{i'}$ is the sum over nearest neighbours of the i^{th} spin and q is the number of nearest neighbours of any one spin. Similarly

$$\begin{aligned}
\sum_{<ij>}\sigma_i &= \frac{1}{2}\sum_i\sigma_i\sum_{i'} \\
&= \frac{q}{2}\sum_i\sigma_i \\
&= \frac{NqM}{2}
\end{aligned} \tag{6.26}$$

Substituting in the expression for energy gives (in the mean field approximation)

$$E(M) = -\frac{1}{2}JNqM^2 - NhM \tag{6.27}$$

For a given M, the entropy of the system, as before, is

$$S = -Nk_B\left[\frac{(1+M)}{2}\ln\left(\frac{1+M}{2}\right) + \frac{(1-M)}{2}\ln\left(\frac{1-M}{2}\right)\right] \tag{6.28}$$

Then, the free energy of the system, as a function of M, is

$$\begin{aligned}
F(M) &= E(M) - TS(M) \\
&= -\frac{1}{2}JNqM^2 - NhM + Nk_BT\left[\frac{(1+M)}{2}\ln\left(\frac{1+M}{2}\right) + \frac{(1-M)}{2}\ln\left(\frac{1-M}{2}\right)\right]
\end{aligned} \tag{6.29}$$

The equilibrium value of M is given by minimising this expression. Differentiating F w.r.t. M and equating to zero gives the following relation

$$\overline{M} = \tanh\left(\frac{Jq\overline{M}+h}{k_BT}\right) \tag{6.30}$$

This is a transcendental equation with solutions obtained as points of intersection of curves $f(x) = x$ and $g(x) = \tanh\left(\frac{Jqx+h}{k_BT}\right)$. Let us focus on the zero-field situation. In absence of a magnetic field, the equartion reduces to

$$\overline{M} = \tanh\left(Jq\beta\overline{M}\right) \tag{6.31}$$

Does this equation possess a solution $\overline{M} \neq 0$? If it does, it would correspond to a *spontaneous* magnetization in absence of an external magnetic field, and would 'explain' the origin of ferromagnetism. For such a solution to exist, the curves corresponding to equation $f(M) = M$ and $g(M) = \tanh(Jq\beta M)$ should intersect at non-zero value(s) of M. Clearly, $M = 0$ is always a solution. The following shows two possibilities, corresponding to $Jq\beta < 1$ and $Jq\beta > 1$

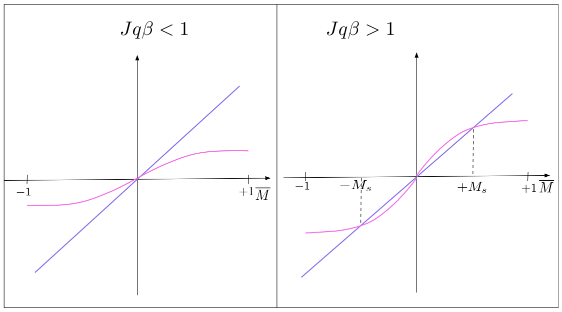


Figure 6.1: Spontaneous Magnetization for $T > T_C = Jq/k_B$

The function $g(M) = \tanh(Jq\beta M)$ saturates to ±1 for large absolute values of the argument. It will intersect the curve $f(M) = M$ for $M \neq 0$ if the slope of $g(M)$ is greater than 1 at $M = 0$. Differentiating $g(M)$ w.r.t. M gives

$$g'(M) = Jq\beta\,\mathrm{sech}^2(Jq\beta M) \tag{6.32}$$

At $M = 0$, this gives $g'(0) = Jq\beta$. Then, the condition for intersection of the two curves for $M \neq 0$ is $Jq\beta > 1$ which gives $T < T_C$, where $T_C = Jq/k_B$. This tells us that there exists a special temperature $T = T_C$ below which there are three possible solutions to equation (6.31), $M = 0, M = +M_s$ and $M = -M_s$. However, for $T > T_C$, there is only one solution, $M = 0$, corresponding to no magnetization. Which of the solutions is the correct one for $T < T_C$? Equation (6.31) corresponds to an *extremum* of the free energy $F(M)$. The correct solution is one which corresponds to a *minimum*. To see which of these corresponds to a minimum, we observe from continuity that for $T < T_C$ but 'close' to it, M will be small. Let us determine the form of $F(M)$ for $|M| << 1$ For $|M| << 1$, we expand the logarithms in eqn.(6.29) upto order M^2

$$\begin{aligned}\ln(1+M) &\simeq M - \frac{M^2}{2} + \frac{M^3}{3} - \frac{M^4}{4} \\ \ln(1-M) &\simeq -M - \frac{M^2}{2} - \frac{M^3}{3} - \frac{M^4}{4}\end{aligned} \tag{6.33}$$

Retaining terms up to order M^4, the expression for free energy reduces to

$$F(M) \simeq -Nk_BT\ln 2 + \frac{Nk_B(T-T_C)}{2}M^2 + \frac{Nk_BT}{12}M^4 \tag{6.34}$$

We now analyze the behaviour of $f(M) = F(M)/N$ for $T > T_C$ and $T < T_C$.

$$f(M) = f(0) + \frac{k_B(T-T_C)}{2}M^2 + \frac{k_BT}{12}M^4 \tag{6.35}$$

where $f(0) = -k_B T \ln 2$. The first and second derivatives of $f(M)$ are given by

$$\begin{aligned} f'(M) &= k_B(T - T_C)M + \frac{k_B T}{3} M^3 \\ &= (k_B T)M\left[\left(1 - \frac{T_C}{T}\right) + \frac{M^2}{3}\right] \end{aligned} \tag{6.36}$$

$$f''(M) = k_B T\left[\left(1 - \frac{T_C}{T}\right) + M^2\right] \tag{6.37}$$

For $T > T_C, f'(M) = 0$ has only one solution, $M = 0$, for which $f''(0) > 0$. Therefore, a zero magnetization corresponds to a minimum of the free energy. On the other hand, for $T < T_C$, there are three possible solutions to $f'(M) = 0$: $M = 0, M = +M_s, M = -M_s$ where

$$M_s = \sqrt{3\left(\frac{T_C}{T} - 1\right)}, \quad T < T_C \tag{6.38}$$

The second derivative of f at these values of M is given by

$$\begin{aligned} f''(0) &= k_B T\left(1 - \frac{T_C}{T}\right) < 0 \\ f''(\pm M_s) &= 2k_B T\left(\frac{T_C}{T} - 1\right) > 0 \end{aligned} \tag{6.39}$$

Clearly, for $T < T_C$, $M = 0$ corresponds to a *maximum* of f, whereas both $M = \pm M_s$ correspond to *degenerate* minima (since $f(+M_s) = f(-M_s)$. The following figure illustrates the behaviour of $f(M)$ for $T > T_C$ and $T < T_C$ (ignoring the constant term $f(0)$)

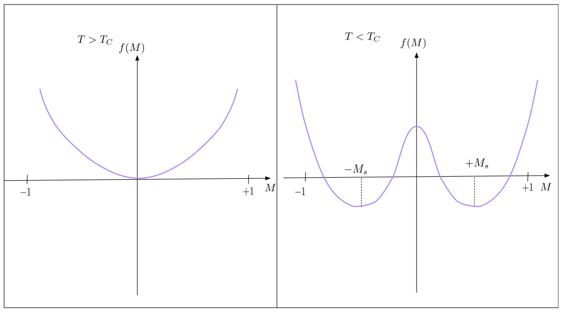


Figure 6.2: Behaviour of free energy for $T > T_C$ and $T < T_C$

As the temperature of the spin system is lowered, for $T > T_C$, there is no magnetization. However, for $T < T_C$, there is sontaneous magnetization with magnitude M_s

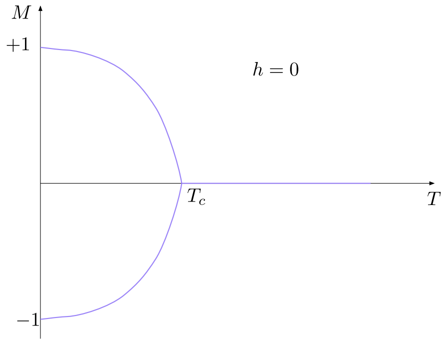


Figure 6.3: Zero filed Magnetization vs T

What decides whether the spontaneous magnetization is positive or negative? In absence of any external magnetic field, both directions are realized with the formation of *domains* of either direction. However, the tiniest external field will force the magnetization along the direction of the external field. To see this, let us take a closer look at the 'equation of state' (eqn.(6.30) in presence of a non-zero, but small magnetic field h. We wish to compute the response of the system to a small change in external magnetic field, for very small magnetic fields (approaching zero). The measure of this response is the magnetic susceptibility of the system $\chi_m = \partial M/\partial h$. Differentiating eqn.(6.30) w.r.t. h

$$\frac{\partial M}{\partial h} = \frac{1}{k_BT}\left(Jq\frac{\partial M}{\partial h} + 1\right)\operatorname{sech}^2\left(\frac{JqM+h}{k_BT}\right) \tag{6.40}$$

Let us now compute the 'zero filed susceptibility', which is obtained by taking the limit $h \to 0$

$$\begin{aligned}\frac{\partial M}{\partial h} &= \frac{1}{k_BT}\left(Jq\frac{\partial M}{\partial h} + 1\right)\operatorname{sech}^2\left(\frac{JqM}{k_BT}\right) \\ &= \left(\frac{Jq}{k_BT}\right)\frac{\partial M}{\partial h}\operatorname{sech}^2\left(\frac{JqM}{k_BT}\right) + \left(\frac{1}{k_BT}\right)\operatorname{sech}^2\left(\frac{JqM}{k_BT}\right)\end{aligned} \tag{6.41}$$

This can be solved for $\chi_m = \partial M/\partial h$

$$\begin{aligned}\chi_m &= \frac{\operatorname{sech}^2\left(\frac{JqM}{k_BT}\right)}{k_BT - qJ\operatorname{sech}^2\left(\frac{JqM}{k_BT}\right)} \\ &= \frac{1-\tanh^2\left(\frac{JqM}{k_BT}\right)}{k_BT - qJ\left(1-\tanh^2\left(\frac{JqM}{k_BT}\right)\right)} \\ &= \frac{1-M^2}{k_BT - Jq(1-M^2)}\end{aligned} \tag{6.42}$$

For $T > T_c, M = 0$. Therefore

$$\chi_m = \frac{1}{k_B(T-T_c)}, \quad T > T_c, h = 0 \tag{6.43}$$

For $T < T_c$ (but close to T_c), using (6.38)

$$\chi_m = \frac{1}{2k_B(T_c - T)}, \quad T < T_c, h = 0 \tag{6.44}$$

Clearly, the zero filed susceptibility *diverges* at T_c. This means that the system is extremely sensitive to changes in external field, about zero filed. This implies that the *tiniest* magnetic field close to T_c will magnetize the system along the direction of that field. The behaviour of the zero field susceptibility is plotted

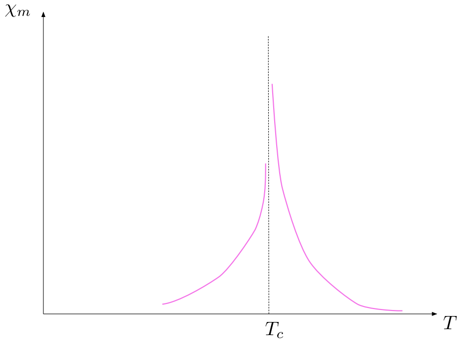


Figure 6.4: Zero field magnetic susceptibility vs T

The Ising model, therefore, contains within it the possibility of spontaneous magnetization, in the 'mean field' approximation. Does this agree with exact computational results, and more importantly, experiments? The mean field approximation correctly anticipates the onset of ferromagnetism in the Ising model. However, it predicts the existence of ferromagnetism in any dimension. Exact solutions show that there is no onset of spontaneous magnetization in one-dimensional systems, and that the precise dependence of the magnetization with temperature is dimension dependent. Further, the temperature T_c is not exactly that predicted in the mean-field approximation. Clearly, mean-field theory does not capture some important aspect of this transition from paramagnetism to ferromagnetism. We shall revisit this issue later.

6.4 ChemicalEquilibrium

Next, we turn to chemical reactions and chemical equilibrium. Consider a chemical reaction of the form

$$n_A\ A + n_B\ B \rightleftharpoons n_C\ C \tag{6.45}$$

where n_A molecules of A combine with n_B molecules of B to give n_C molecules of C. The reaction is reversible and the species are in equilibrium, with number concentrations $[A], [B]$ and $[C]$ (number of molecules per unit volume) of the three species . The equilibrium can be shifted by changing, say, the temperature of the system. In prctice, chemical reactions usually occur at fixed temperature and pressure, and not fixed volume and temperature. However, since the canonical probability distribution corresponds to fixed volume, let us consider the above reaction as occuring in an enclosure of fixed volume and temperature. We assume here that the reactants and products are in gas phase, so that they occupy the entire available volume. We further assume that molecules of any one species interact only weakly with each other and with the other molecules. the energy of any one molecule has the form

$$E = E_T + E_{int} \tag{6.46}$$

where E_T is the translational kinetic energy and E_{int} is the internal energy (due to electronic, rotational, vibrational degrees of freedom). Then, the partition function for any one species has the form

$$Z = \frac{1}{N!} Z_T^N Z_{int}^N \tag{6.47}$$

Correspondingly, the Helmholtz free energy has the form

$$F(T, V, N) = -Nk_BT \ln\left(\frac{V}{N}\left(\frac{2m\pi k_BT}{h^2}\right)^{3/2}\right) - Nk_BT + Nf_0 \tag{6.48}$$

where the first two terms arise due to the translational part of the partition function (eqn.(4.66) and the second term due to the internal degrees of freedom. Here,

$$f_0 = -k_B T \ln Z_{int} \tag{6.49}$$

Since the three species are weakly interacting, therefore the Helmholtz free energy of the system is

$$F = F_A(T, V, N_A) + F_B(T, V, N_B) + F_C(T, V, N_C) \tag{6.50}$$

Say, the system is in equilibrium at the given temperature, with concentrations $[A], [B]$ and $[C]$. If the number of molecules of the species change by $\Delta N_A, \Delta N_B$ and ΔN_C respectively, the change in the free energy will be

$$\begin{aligned} \Delta F &= \Delta F_A + \Delta F_B + \Delta F_C \\ &= \frac{\partial F_A}{\partial N_A}\Delta N_A + \frac{\partial F_B}{\partial N_B}\Delta N_B + \frac{\partial F_C}{\partial N_C}\Delta N_C \end{aligned} \tag{6.51}$$

The smallest change in the numbers is clearly $\Delta N_A = -n_A, \Delta N_B = -n_B$ and $\Delta N_C = +n_C$ respectively. For this change,

$$\Delta F = -\mu_A n_A - -\mu_B n_B + \mu_C n_C \tag{6.52}$$

where we encounter a new quantity, the (partial) derivative of F w.r.t. the number of particles

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{V,T} \tag{6.53}$$

This quantity, which is a measure of the response of the Helmholtz free energy to a change in number of particles of a system, is known as the *chemical potential* of the system. We will take a closer look at the chemical potential and its physical interpretation later. For now, we observe that it is an *intensive* quantity, just like pressure and temperature, since it is a ratio of two extensive quantities.

If the system is in chemical equilibrium (with the concentrations of reactants and products corresponding to a mininum of free energy), then $\Delta F = 0$. This gives the following condition for chemical equilibrium for the reaction (6.45)

$$n_C \mu_C - n_A \mu_A - n_B \mu_C = 0 \tag{6.54}$$

To see what this condition implies about the relative concentrations of reactants and products, we compute the chemical potential for any one species from the free energy expression (6.48)

$$\begin{aligned} \mu &= \left(\frac{\partial F}{\partial N}\right)_{V,T} \\ &= k_B T \ln\left(\frac{N}{V}\left(\frac{h^2}{2m\pi k_B T}\right)^{3/2}\right) + f_0 \end{aligned} \tag{6.55}$$

Then, condition (6.54) reduces to (after a little algebra)

$$\frac{[C]^{n_C}}{[A]^{n_A}[B]^{n_B}} = \left(\frac{Q_C^{n_C}}{Q_A^{n_A} Q_B^{n_B}}\right) e^{-\Delta f^0/k_B T} \tag{6.56}$$

where

$$Q = \left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} \tag{6.57}$$

is the so-called *quantum concentration* and

$$\Delta f^0 = n_C f_C^0 - n_A f_A^0 - n_B f_B^0 \tag{6.58}$$

The *quantum concentration* has a simple interpretation, if we express in terms of the *thermal deBroglie* wavelength

$$\lambda_T = \frac{h}{\sqrt{2m\pi k_B T}} \tag{6.59}$$

The thermal deBroglie wavelength of a particle is the deBroglie wavelength associated with it due to thermal energy $\epsilon_T \sim k_B T$ and corresponding momentum $p_T = \sqrt{2m\epsilon_T} \sim \sqrt{2mk_B T}$

$$\lambda_T = \frac{h}{p_T} \tag{6.60}$$

In terms of this, the quantum concentration is

$$Q = \frac{1}{\lambda_T^3} \tag{6.61}$$

which has dimension of inverse length cube (hence 'concentration').

Equation (6.56) gives the relative concentrations of reactants and products in thermal equilibrium. At a given temperature T, the ratio on the left hand side of the equation is a constant. This is known as the *Law of Mass Action.* In addition to giving this law, minimisation of free energy also gives the precise temperature dependence of this constant. Let us isolate the temperature dependence

$$\frac{[C]^{n_C}}{[A]^{n_A}[B]^{n_B}} \sim T^{3/2(n_C - n_A - n_B)} e^{-\Delta f^0/k_B T} \tag{6.62}$$

The internal free energy change (6.58) will depend explicitly on the internal degrees of freedom of the molecules of reactants and products. Assuming that the temperature is low enough such that the internal excitations are 'frozen', the molecules will be in their respective ground states. Then,

$$Z_{int} \simeq e^{-\beta E_0} \tag{6.63}$$

where E_0 is the ground state energy of the molecule. Then

$$\Delta f^0 = n_C E_0^C - n_A E_0^A - n_B E_0^A \tag{6.64}$$

As an example, consider the following reaction for ammonia synthesis

$$3H_2 + N_2 \rightleftharpoons 2NH_3 \tag{6.65}$$

This is an exothermic reaction, with $\Delta f^0 = -92.4 kJ/mol$. Here, $n_A = 3, n_B = 1, n_C = 2$. Then,

$$\frac{[NH_3]^2}{[H_2]^3[N_2]} \sim T^{-3} e^{+92.4\times 10^3/RT} \tag{6.66}$$

where R is the gas constant. This equation shows that lowering the temperature will shift the equilibrium towards incresed concentration of ammonia and vice-versa.

As another example, consider the equilibrium between atomic and molecular hydrogen

$$H_2 \rightleftharpoons 2H \tag{6.67}$$

For this reaction, eqn.(6.56) reduces to

$$\frac{[H]^2}{[H_2]} = \frac{Q_H^2}{Q_{H_2}} e^{-\Delta E/k_B T} \tag{6.68}$$

where ΔE is the dissociation energy of H_2 molecule. Given that $m_{H_2} = 2m_H$, it follows that $Q_{H_2} = 2^{2/3} Q_H$. Therefore

$$\begin{aligned} \frac{[H]^2}{[H_2]} &= \frac{1}{8} Q_{H_2} e^{-\Delta E/k_B T} \\ \Longrightarrow \frac{[H]}{[H_2]} &= 2^{-3/2} \frac{1}{\sqrt{[H_2]/Q_{H_2}}} e^{-\Delta E/2k_B T} \end{aligned} \tag{6.69}$$

There are two factors on the right hand side of eqn.(6.69). The exponential factor supresses the formation of atomic hydrogen at temperatures smaller than the characteristic temperature scale $T_0 \sim \Delta E/2k_B \sim 25,000^0 K$. This is just a consequence of the Boltzmann factor, since the energy of the dissociated atoms is larger, and so dissociation is supressed exponentially. The other factor, $1/\sqrt{[H_2]/Q_{H_2}}$ can be interpreted as follows: $Q_{H_2}/[H_2]$ can be thought of as the cube of the number of deBroglie wavelengths that can fit between two hydrogen molecules. At typical temperature and densities, this is a very small number. As a result, the pre-factor of the exponential is very large. The pre-factor can be though to arise due to entropy. Dissociation is entropically favorable, since all things being equal, two hydrogen atoms will have more entropy than a hydrogen molecule (expecially with internal degrees 'frozen').

Chapter 7

Open Systems

7.1 The Gibbs distribution

We have seen that a system in equilibrium with an environment with which it can exchange energy satisfies the canonical distribution (4.28)

$$P_r = \frac{e^{-\beta E_r}}{Z} \tag{7.1}$$

In general, an ‘open’ system can exchange not just energy, but share volume, exchange particles, or interact with ‘fields’ in the external environment. For instance, phase transitions can involve a change in the volume of the system (with external pressure fixed), such as water vapour condensing to form liquid water at atmospheric pressure. Given a change in the external temperature and pressure, the volume of the system ‘responds’ by attaining an equilibrium value. A ferromagnetic material, interacting with an external magnetic field, responds by attaining an equilibrium magnetization. A catalytic surface adsorbs particles depending on external conditions of temperature and pressure. The particles are constantly exchanged with the environment, and in equilibrium, a certain mean number of particles will adhere to the surface of the catalyst. A polymer under tension will attain an equilibrium length depending on the external temperature and tension. In all these examples, there is an external ‘field’ (such as pressure, tension, external magnetic field) which couples to an extensive property of the system (volume, length, magnetization), this property responding to changes in the external temperature and the external field. The probability distribution is no longer canonical. To get the new probability distribution, we can take two alternative approaches: (a) revert back to visualising the system and its environment as a single isolated system in equilibrium, and derive the probability distribution for the system assuming the microcanonical distribution for the system plus environment (b) include the interaction with the external field in the expression for energy of the system, and directly use the canonical Boltzmann distribution. Let us start with approach (b), which is the natural approach for certain situations, and later we will take approach (a) for a specific situation which is not so intuitively visualised as situation (b).
We assume that interaction with the external field modifies the energy of the system to

$$E' = E - fx \tag{7.2}$$

where E is the energy of the system if the external interaction is absent (apart from thermal energy exchange with the environment), f is the external field and x is an extensive parameter belonging to the system which couples to field f. In general, the field can be a vector, in which case the interaction term will have the form $(-\vec{f}\cdot\vec{x})$ where $\vec{f}$ is the vector field and $\vec{x}$ is a vector parameter. As an example, $\vec{f}$ could be a magnetic field and $\vec{x}$ would be the magnetization $\vec{M}$. Note that we are assuming that the systems couples to the environment *linearly* through parametr x.
Given (7.2), the probability distribution for a microstate r of the system becomes

$$P_r = \frac{1}{Z_G} e^{-\beta(E_r - f x_r)} \tag{7.3}$$

where x_r is the value of the parameter x in mocrostate r. The partition function is now

$$Z_G = \sum_r e^{-\beta(E_r - f x_r)} \tag{7.4}$$

Such a probability distribution is called a *Gibbs distribution* and the corresponding partition function a *Gibbs Partition Function.*

Given this distribution, the mean value of x can be computed to be

$$\begin{aligned} \overline{x} &= \sum_r x_r P_r \\ &= \frac{1}{Z_G} \sum_r x_r \; e^{-\beta(E_r - f x_r)} \\ &= \frac{1}{\beta} \frac{1}{Z_G} \left(\frac{\partial Z_G}{\partial f} \right)_T \\ &= \frac{1}{\beta} \left(\frac{\partial \ln Z_G}{\partial f} \right)_T \end{aligned} \tag{7.5}$$

where the derivative of the partition function is computed at fixed temperature. The fluctuation (squared) in x is

$$\begin{aligned} \Delta x^2 &= \overline{x^2} - (\overline{x})^2 \\ &= \frac{1}{\beta^2} \left(\frac{\partial^2 \ln Z_G}{\partial f^2} \right)_T \end{aligned} \tag{7.6}$$

Given the extensivity of $\ln Z_G$ (since $\overline{x}$ is extensive), it follows that

$$\frac{\Delta x}{\overline{x}} \sim \frac{1}{\sqrt{N}} \tag{7.7}$$

where N is the number of particles of the system. This is something that we anticipate by now, and expect relative fluctuations to be 'small' for macroscopic systems. However, let us rewrite eqn.(7.6) as

$$\frac{\Delta x}{\overline{x}} = \sqrt{\frac{(k_B T)\ \kappa}{\overline{x}}} \tag{7.8}$$

where

$$\kappa = \frac{1}{\overline{x}} \left(\frac{\partial \overline{x}}{\partial f} \right)_T \tag{7.9}$$

In this form, it is still clear that the relaitve fluctuation scales as $1/\sqrt{N}$, since $\overline{x}$ scales as N and the parameter κ is intensive. However, in this form, we establish a connection between the microscopic statistical fluctuations in parameter $\overline{x}$ to a 'response' in $\overline{x}$ due to a change in the external field f, through the directly experimentally accessible 'response function' κ. Naively, one expects the relative fluctuations to be small. However, we will see that under special values of temperature and external field f, the response function κ can diverge, leading to fluctuations themselves diverging. Such a point in parameter space (T, f) is known as a *critical point.* Then, adding external fields which interact with a system (in addition to thermal interaction with the environment) is not only natural to analyse the behaviour of certain systems (such as particles at constant external pressure, ferromagnet in external magnetic field, etc.) but is also a useful tool to analyse the behaviour of statistical fluctuations in the system even in absence of such fields. In the latter case, a field is added as an auxiliary mathematical tool to 'probe' the behaviour of these fluctuations even in absence of such a field (in which case one computes the 'zero-field' value of κ).

7.2 Rubber Bands, Revisited

As an example of the Gibbs distribution, let us go back to the simple model of a rubber band considered in section 6.2) as a system of N links. Let us assume that the rubber band kept under tension by pulling one end with a constant external force f. If, under the influence of the force f, the band stretches by ΔL, the work done on the band is $\Delta W = f\Delta L$. If this band was insulated from the external environment, this would result in a change in energy of the band by $\Delta E = f\Delta L$. Then, the quantity $E' = E - fL$ would be conserved. Now, if the system is in a heat bath at temperature T, The probability distribution of microstates is

$$P_r = \frac{1}{Z_G} e^{-\beta(E_r - fL_r)} \tag{7.10}$$

where the sum is over microstates of the links, with each link having two possible 'states', pointing to the 'left', or to the 'right'. In our simple model which ignores any dependence of internal energy of the band on the orientation of the links, E_r is just a constant, which can be ignored. Then, the probability distribution is

$$P_r = \frac{1}{Z_G} e^{-\beta f L_r} \tag{7.11}$$

with

$$Z_G = \sum_r e^{-\beta f L_r} \tag{7.12}$$

As before, the length $L = l_0(n_R - n_L)$ with l_0 being the length of any one link, n_R the number of links pointing to the right and n_L to the left (it is assumed that the band is pulled to the right). Formally, this system is similar to the Ising model, where instead of spins, each link can be in to states $\sigma = \pm 1$, such that

$$L = l_0 \sum_i \sigma_i \tag{7.13}$$

A microstate of the system is just a specification of the set $\{\sigma_i\}$. Then, the partition function of the system is

$$\begin{aligned} Z_G &= \sum_\sigma e^{-\beta f l_0 \sum_i \sigma_i} \\ &= Z_1^N \end{aligned} \tag{7.14}$$

where

$$\begin{aligned} Z_1 &= \sum_{\sigma=-1}^{+1} e^{-\beta f l_0 \sigma} \\ &= 2\cosh(\beta f l_0) \end{aligned} \tag{7.15}$$

Then,

$$Z_G = 2^N \cosh^N(\beta f l_0) \tag{7.16}$$

The mean length of the band is given by (7.5)

$$\begin{aligned} \overline{L} &= \frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial f}\right)_T \\ &= \frac{1}{\beta} N \frac{\partial}{\partial f} \cosh(\beta f l_0) \\ &= N l_0 \tanh(\beta f l_0) \end{aligned} \tag{7.17}$$

which can be written as

$$L = L_{max} \tanh\left(\frac{f l_0}{k_B T}\right) \tag{7.18}$$

which, for 'large' temperatures, reduces to

$$L \simeq L_{max} \frac{f l_0}{k_B T} \tag{7.19}$$

7.3 The NPT Distribution

As another example, let us consider a system of interacting particles at constant external pressure P and temperature T. If this system was not exchanging thermal energy with the environment, then if the volume of the system was changed by ΔV, its energy would change by

$$\Delta E = -P\Delta V \tag{7.20}$$

This is the same as $\Delta(E + PV) = 0$. Then, the effective conserved energy would be

$$E' = E + PV \tag{7.21}$$

If now the system exchanges thermal energy with the environment at temperature T, the probability of microstate r will be

$$P_r = \frac{1}{Z_G} e^{-\beta(E_r + PV_r)} \tag{7.22}$$

where V_r is the volume of the system in microstate r and

$$Z_G = \sum_r e^{-\beta(E_r + PV_r)} \tag{7.23}$$

The distribution (7.22) is often called the *NPT distribution*, since it corresponds to fixed N, P and T (as opposed to the canonical distribution for a system of particles, often called the *NVT distribution* for obvious reasons).

Given the external pressure P and temperature T, the mean volume of the system will be (eqn.(7.5))

$$\overline{V} = -\frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial P}\right)_{N,T} \tag{7.24}$$

The *specific volume* (volume per particle) of the system is

$$\begin{aligned} v &= \frac{\overline{V}}{N} \\ &= -\frac{1}{N\beta}\left(\frac{\partial \ln Z_G}{\partial P}\right)_{N,T} \end{aligned} \tag{7.25}$$

Depending on the temperature and pressure, the same system of particles can have a different specific volume (or equivalently, density), depending on its phase. The volume fluctuation is given by (7.6)

$$\begin{aligned} \Delta V^2 &= \frac{1}{\beta^2}\left(\frac{\partial^2 \ln Z_G}{\partial P^2}\right)_{N,T} \\ &= -\frac{1}{\beta^2}\beta\left(\frac{\partial \overline{V}}{\partial P}\right)_{N,T} \\ &= -k_B T\left(\frac{\partial \overline{V}}{\partial P}\right)_{N,T} \\ &= (k_B T)\ \kappa_T\ \overline{V} \end{aligned} \tag{7.26}$$

where

$$\kappa_T = -\frac{1}{\overline{V}}\left(\frac{\partial \overline{V}}{\partial P}\right)_{N,T} \tag{7.27}$$

is the *isothermal compressibility* of the system, a macroscopic, intensive property which is experimentally easily measurable. The relative volume fluctuation is

$$\frac{\Delta V}{\overline{V}} = \sqrt{\frac{(k_B T)\ \kappa_T}{\overline{V}}} \tag{7.28}$$

which can be expressed in terms of specific volume

$$\frac{\Delta v}{\overline{v}} = \sqrt{\frac{(k_B T)\ \kappa_T}{N\ v}} \tag{7.29}$$

As discussed, this shows two things: (a) the microscopic statistical fluctuations in volume are directly related to the macroscopic, directly measurable κ_T (b) the relative fluctuation scales as $1/\sqrt{N}$. Again, even though we expect this relative fluctuation to be negligible for macroscopic systems, this expectation is in fact rooted in the assumption that κ_T is a smooth function of external temperature and pressure. We will see that it need not be so.
To get a feel for the probability distribution (7.22), let us apply it to a system of weakly interacting particle with energy given by (4.57). The sum appearing in the partition function (7.23) can be rearranged as

$$Z_G = \int_0^\infty dV e^{-\beta PV} Z(N,V,T) \tag{7.30}$$

where $Z(N,V,T)$ is the canonical partition function when the volume of the system is V. For a system of weakly interacting particles, the canonical partition function is given by (4.64)

$$Z(N,V,T) = \frac{1}{N!} V^N \left(\frac{2m\pi k_B T}{h^2}\right)^{3N/2} \tag{7.31}$$

Then, the Gibbs partition function reduces to

$$\begin{aligned} Z_G &= \frac{1}{N!}\left(\frac{2m\pi k_B T}{h^2}\right)^{3N/2} \int_0^\infty dV V^N e^{-\beta PV} \\ &= \left(\frac{2m\pi k_B T}{h^2}\right)^{3N/2} \left(\frac{k_B T}{P}\right)^{N+1} \end{aligned} \tag{7.32}$$

The specific volume is given by (7.25)

$$\begin{aligned} v &= -\frac{1}{N\beta}\left(\frac{\partial \ln Z_G}{\partial P}\right)_{N,T} \\ &= -\frac{k_B T}{N}(-N-1)\frac{\partial \ln P}{\partial P} \\ &\simeq \frac{k_B T}{P} \quad N >> 1 \end{aligned} \tag{7.33}$$

which gives the familiar ‘ideal gas’ equation of state

$$Pv = k_B T \tag{7.34}$$

The NPT distribution is the most natural distribution to analyze phase transitions in particle systems. It is amenable to Monte Carlo techniques and can be used to computatioanlly explore gas to liquid to solid transitions.

7.4 Ferromagnetism revisited

As another example, let us take the Ising model in presence of an external magnetic field. The expression for energy (eqn.(3.17)) can be written as

$$E' = E - h\mathcal{M} \tag{7.35}$$

where

$$E = -J \sum_{<ij>} \sigma_i \sigma_j \tag{7.36}$$

is the spin-spin interaction energy and

$$\mathcal{M} = \sum_i \sigma \tag{7.37}$$

is the total magnetic moment of the system (the magnetization $M = \mathcal{M}/N$). The probability of microstate r is

$$P_r = \frac{1}{Z_G} e^{-\beta(E_r - h\mathcal{M}_r)} \tag{7.38}$$

where

$$Z_G = \sum_r e^{-\beta(E_r - h\mathcal{M}_r)} \tag{7.39}$$

The mean magnetic moment of the system at given external temperature T and magnetic field h is obtained from (7.5)

$$\overline{\mathcal{M}} = \frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial h}\right)_{N,T} \tag{7.40}$$

The specific magnetization is just $M = \mathcal{M}/N$

$$M = \frac{1}{N\beta}\left(\frac{\partial \ln Z_G}{\partial h}\right)_{N,T} \tag{7.41}$$

The fluctuation in $\mathcal{M}$ is

$$\begin{aligned}\Delta\mathcal{M}^2 &= \frac{1}{\beta^2}\left(\frac{\partial^2 \ln Z_G}{\partial h^2}\right)_{N,T} \\ &= \frac{1}{\beta^2}\beta\left(\frac{\partial\overline{\mathcal{M}}}{\partial h}\right)_{N,T} \\ &= k_B T\left(\frac{\partial\overline{\mathcal{M}}}{\partial h}\right)_{N,T} \qquad (7.42)\end{aligned}$$

Therefore, the fluctuation in the magnetization M is

$$\begin{aligned}\Delta M^2 &= \frac{1}{N^2}\Delta\mathcal{M}^2 \\ &= \frac{1}{N^2}k_B T\left(\frac{\partial\overline{\mathcal{M}}}{\partial h}\right)_{N,T} \\ &= \frac{1}{N}k_B T\left(\frac{\partial M}{\partial h}\right)_{N,T} \\ &= \frac{1}{N}(k_B T)\ \chi_m \qquad (7.43)\end{aligned}$$

where χ_m is the magnetic susceptibility. The relative fluctuation of magnetization is

$$\frac{\Delta M}{M} = \sqrt{\frac{(k_B T)\ \chi_m}{N\ M}} \tag{7.44}$$

which, once again, is directly related to the experimentally accessible χ_m. Again, one might naively conclude that fluctuations in magnetization should be imperceptible for macroscopic systems. However, as we saw in section 6.3, we expect χ_m to *diverge* at $h = 0, T = T_C$. Therefore, fluctuations at such *critical points* cannot be assumed to be small, and the mean field approximation is not valid close to this point. We will see that something similar happens for a system of interacting particles at a critical point corresponding to special values of temperature and pressure.

7.5 Energy Fluctuations

We have observed that if a statistical system couples to an external environment through a *linear* interaction of an extensive parameter (x) with an external field (f), the (relative) fluctuation in the parameter is related to the *response* of the parameter to a small change in the external field.
What is the response function corresponding to energy fluctuations? In the canonical distribution, energy is already coupled to an external 'field': temperature. Given the canonical distribution, we have seen that the mean energy and fluctuations in energy are given by (eqns (4.35) and (4.38)

$$\begin{aligned} \overline{E} &= -\frac{\partial}{\partial \beta} \ln Z \\ \Delta E^2 &= \frac{\partial^2}{\partial \beta^2} \ln Z \end{aligned} \tag{7.45}$$

Then

$$\begin{aligned} \Delta E^2 &= -\frac{\partial \overline{E}}{\partial \beta} \\ &= k_B T^2 \frac{\partial \overline{E}}{\partial T} \\ &= N k_B T^2 c_v \end{aligned} \tag{7.46}$$

where c_v is the specific heat of the system (at constant volume, since derivatives are taken at fixed volume)

$$c_v = \frac{1}{N} \left(\frac{\partial \overline{E}}{\partial T} \right)_V \tag{7.47}$$

Then, the relative fluctuation in energy of the system is

$$\frac{\Delta E}{\overline{E}} = \sqrt{\frac{(N k_B T^2)\ c_v}{\overline{E}^2}} \tag{7.48}$$

Given that $\overline{E}$ scales as N, it is clear that the relative fluctuation scales as $1/\sqrt{N}$. However, more importantly, the specific heat of the system (response function) is a direct measure of these statistical fluctuations. Again, a possible divergence in specific heat would be a signature of large, diverging fluctuations.

7.6 The Grand Canonical Distribution

An open system can couple to the environment linearly in an interesting way: by exchanging particles with the environemnt. The extensive parameter of the system coupled to the environment in this situation is N, the number of particles. As before, the probability distribution changes. The question is: what is the external field with which this parameter couples? The external field in this case happens to be the *chemical potential*, encountered in section 6.4 (eqn.(6.53))

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{V,T} \tag{7.49}$$

To see the connection, we need to first have a better insight into the significance of the chemical potential. Consider a system consisting of two sub-parts of fixed volumes (V_1 and V_2) at fixed temperature T. Let the sub-parts exchange particles (of the same species). The energy of a particle could be different in these parts because of a different environment. For instance, there could be different kinds of fields in these two parts with which the particle could interact, altering its energy as it moves from one to the other. Given a fixed total number of particles N, in equilibrium, what fraction of particles will be in the two regions?

The answer can be viewed from the point of view of free energy minimisation. Let there be N_1 particles in the first region and N_2 in the second. The total free energy of the system is then

$$F = F_1(N_1, V_1, T) + F_2(N_2, V_2, T) \tag{7.50}$$

We can view F as a function of variable N_1 ($N_2 = N - N_1$). Extremising F w.r.t. N_1 will given the equilibrium values of N_1 and N_2. Differentiating F w.r.t. N_1 gives

$$\begin{aligned} \frac{\partial F}{\partial N_1} &= \frac{\partial F_1}{\partial N_1} + \frac{\partial F_2}{\partial N_1} \\ &= \frac{\partial F_1}{\partial N_1} - \frac{\partial F_2}{\partial N_2} \quad (\text{since } \partial/\partial N_1 = -\partial/\partial N_2) \\ &= \mu_1 - \mu_2 \end{aligned} \tag{7.51}$$

where μ_1 and μ_2 are the chemical potentials of the two sub-parts. Equilibrium therefore requires that the chemical potentials of the two parts should be the same. Then, chemical potential is a kind of 'pressure', a difference in which results in exchange of particles.

We have seen that the chemical potential of a system is a response of the free energy of a system to a change in the number of particles, at fixed volume and temperature. We can also visualise it as a response of the entropy of the system to a change in particle number, but at fixed volume and *energy*, as opposed to temperature. To see this, we consider a variation in the free energy of a system, in response not just to a variation in volume and temperature (as was done in eqn.(4.54), but also, in addition, to a change in the number of particles. Given that $F = E - TS$, a change in N, V, T produces a chhange

$$dF = dE - TdS - SdT \tag{7.52}$$

where dE and dS are corresponding changes in energy and entropy of the system. Visualising S as a function of E, V and N, we get

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E}\right)_{N,V} dE + \left(\frac{\partial S}{\partial V}\right)_{N,E} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN \\ &= \frac{1}{T}\, dE + \frac{P}{T}\, dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN \end{aligned} \tag{7.53}$$

Substituting back in (7.52) gives

$$dF = -PdV - SdT - T\left(\frac{\partial S}{\partial N}\right)_{E,V} dN \tag{7.54}$$

This immediately gives

$$\left(\frac{\partial F}{\partial N}\right)_{V,T} = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \tag{7.55}$$

which gives

$$\mu = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \tag{7.56}$$

We are now equipped to deduce the probability distribution for a system which can exchange, in addition to energy, particles with the environment. We start by visualising the system plus the environment as a single isolated system in equilibrium with a microcanonical probability distribution for microstates. Let the system be in a microstate r corresponding to which it has N_r particles and energy E_r (a microstate is a precise specification of the number of particles and their positions and momenta). Then the number of microstates accessible to the environment is $\mathcal{N}^{env}(N - N_r, E - E_r)$ where N is the total number of particles of the system plus environment and E is the total energy. Then, since all microstates of the system plus environment are equally probable, it follows that the probability of microstate r is

$$P_r = \frac{\mathcal{N}^{env}(N - N_r, E - E_r)}{\mathcal{N}(N, E)} \tag{7.57}$$

where $\mathcal{N}(N, E)$ is the total number of microstates of the system plus environment. Given the entropy of the environment is $S^{env} = k_B \ln \mathcal{N}^{env}$, it follows that (Taylor expanding the entropy and retaining terms upto first order)

$$\begin{aligned} \mathcal{N}^{env}(N - N_r, E - E_r) &= e^{S^{env}(N-N_r,E-E_r)/k_B} \\ &= e^{S^{env}(N,E)/k_B} e^{(1/k_B)(-E_r\ \partial S^{env}/\partial E -\ N_r\ \partial S^{env}/\partial N)} \\ &= e^{S^{env}(N,E)/k_B} e^{-\beta(E_r - \mu N_r)} \end{aligned} \tag{7.58}$$

where $\mu = -T(\partial S^{env}/\partial N)$ is the chemical potential of the environment. Then, the probability distribution is

$$P_r = \frac{1}{Z_G}\ e^{-\beta(E_r - \mu N_r)} \tag{7.59}$$

with

$$Z_G = \sum_r e^{-\beta(E_r - \mu N_r)} \tag{7.60}$$

Clearly, the distribution (7.59) is of the general form (7.3), with the extensive parameter x identified as the number of particles and the external field f as the chemical potential of the environment. The probability distribution (7.59) is known as the *Grand Canonical Distribution.*
The mean number of particles of the system is (eqn.(7.5))

$$\overline{N} = \frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial \mu}\right)_{V,T} \tag{7.61}$$

and the fluctuation in particle number is (eqn.(7.6))

$$\begin{aligned} \Delta N^2 &= \frac{1}{\beta^2}\left(\frac{\partial^2 \ln Z_G}{\partial \mu^2}\right)_{V,T} \\ &= \frac{1}{\beta}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{V,T} \end{aligned} \tag{7.62}$$

As a simple application, we consider a system of weakly interacting particles in equilibrium at temperature T and pressure P. We consider an imaginary region of volume V within this system of particles. Particles will constantly enter and exit this region. We ask: on an average, how many particles are likely to be found in this region? The probability distribution in this case is clearly Grand canonical, with the mean number of particles given by eqn.(7.61). To compute the mean number of particles, we need to compoute the Grand canonical partition function (7.63). We can organise the sum as follows

$$\begin{aligned} Z_G &= \sum_r e^{-\beta(E_r - \mu N_r)} \\ &= \sum_{N=0}^{\infty} e^{\beta \mu n} Z(N, V, T) \end{aligned} \tag{7.63}$$

where $Z(N, V, T)$ is the canonical partition function for N particles. For a weakly interacting system, this is given by (4.64)

$$\begin{aligned} Z(N, V, T) &= \frac{1}{N!} V^N \left(\frac{2m\pi k_B T}{h^2}\right)^{3N/2} \\ &= \frac{1}{N!}\left[V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]^N \end{aligned} \tag{7.64}$$

Then, the Grand partition function is

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} e^{\beta\mu n} Z(N,V,T) \\ &= \sum_{N=0}^{\infty} \frac{1}{N!}\left[V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]^N e^{\beta\mu n} \\ &= \sum_{N=0}^{\infty} \frac{1}{N!}\left[e^{\beta\mu} V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]^N \end{aligned} \tag{7.65}$$

This is just an exponential series, which sums to

$$Z_G = e^{e^{\beta\mu} V\left(2m\pi k_B T/h^2\right)^{3/2}} \tag{7.66}$$

The mean particle number can now be computed

$$\begin{aligned} \overline{N} &= \frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial \mu}\right)_{V,T} \\ &= \frac{1}{\beta}\frac{\partial}{\partial\mu}\left[e^{\beta\mu} V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]_{V,T} \\ &= e^{\beta\mu} V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} \end{aligned} \tag{7.67}$$

At this point, we need an expression for the chemical potential. Note that in the Grand canonical distribution, μ is the chemical potential of the environment with which the system exchanges particles. Here, the environment, as the system, consists of the same species of weakly interacting particles. Then, we need the expression for the chemical potential of a system of weakly interacting particles at temperature T and pressure P. In section 6.4), we had computed the expression for the chemical potential of such a system (in that section, each 'particle' was a molecule with possible internal degrees of freedom). There, the system was at constant temperature T and occupied a fixed volume V (eqn.(6.55)

$$\mu = k_B T \ln\left(\frac{N}{V}\left(\frac{h^2}{2m\pi k_B T}\right)^{3/2}\right) + f_0 \tag{7.68}$$

Here, the system is at pressure P and has no internal degrees of freedom (or they are 'frozen'). We can eliminate the number density N/V in favour of the pressure of the system using the ideal gas relation (2.16). further, we can ignore the 'internal' contribution f_0. Then, in terms of pressure and temperature, the chemical potential is given by

$$\mu = k_B T \ln\left(\frac{P}{k_B T}\left(\frac{h^2}{2m\pi k_B T}\right)^{3/2}\right) \tag{7.69}$$

Then,

$$e^{\beta\mu} = \frac{P}{k_B T}\left(\frac{h^2}{2m\pi k_B T}\right)^{3/2} \tag{7.70}$$

Substituting in the expression for $\overline{N}$ gives

$$\begin{aligned} \overline{N} &= e^{\beta\mu} V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} \\ &= \frac{PV}{k_B T} \end{aligned} \tag{7.71}$$

which is the same as

$$PV = \overline{N}\ k_B T \tag{7.72}$$

This is what one expects, except that here the number of particles is not fixed and there will be fluctuations about the mean value $\overline{N}$. The fluctuation is given by (7.62)

$$\begin{aligned}
\Delta N^2 &= \frac{1}{\beta}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{V,T} \\
&= \frac{1}{\beta}\frac{\partial}{\partial \mu}\left[e^{\beta\mu} V \left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]_{V,T} \\
&= e^{\beta\mu} V \left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} \\
&= \overline{N}
\end{aligned} \tag{7.73}$$

Therefore

$$\frac{\Delta N}{\overline{N}} = \frac{1}{\sqrt{\overline{N}}} \tag{7.74}$$

The Grand canonical distribution is important to the study of surface chemistry. As an example, consider a surface exposed to a gas. Under certain conditions (depending on the property of the surface and the gas), the gas atoms (or molecules) can get *adsorbed* to the surface, creating a surface layer. It is observed that an increase in the pressure of the gas (at fixed temperature) results in increased coating of the surface with gas atoms, which saturates at high pressures. A simple model which accounts for this visualises the surface as consisting of $\mathcal{N}$ adsorption sites, each of which can accomodate at most one atom. In equilibrium, at temperature T and pressure P (of the gas), a certain (mean) number of atoms $\overline{N}$ will occupy a fraction of the available sites $\mathcal{N}$. To compute this fraction, we compute the Grand canonical partition function for the system of atoms trapped to the surface. It is assumed that any atom trapped to the surface has energy $E = -\epsilon$, where $\epsilon > 0$. That is, the surface sites act as potential 'traps' for atoms. In more complicated models, there can be more than one of such trapped states with different energies. The Grand partition function for the system can be computed using equation (7.63), with the modification that there is an upper limit to the number of atoms that can stick to the surface (the limit being $\mathcal{N}$)

$$Z_G = \sum_{N=0}^{\mathcal{N}} e^{\beta\mu N}\ Z_C(N, \mathcal{N}, T) \tag{7.75}$$

where $Z_C(N, \mathcal{N}, T)$ is the canonical partition function for the system of atoms trapped on the surface, occupying N sites out of the available $\mathcal{N}$ sites (note that $\mathcal{N}$ plays the role of volume). μ is the chemical potential of the gas to which the surface is exposed. For simplicity, we assume that the gas is ideal (weakly interacting atoms). A microstate of the system of N atoms trapped on the surface is just a specification of the sites that are occupied (it is assumed in this simple model that the atoms cannot move on the surface, therefore momentum variables are not relevant to the microstates). The energy of each microstate is simply $E = -N\epsilon$ and the number of such (degenerate) microstates is $\binom{\mathcal{N}}{N}$. Then, the Canonical partition function is

$$\begin{aligned}
Z_C &= \sum_r e^{-\beta E_r} \\
&= \binom{\mathcal{N}}{N} e^{N\beta\epsilon}
\end{aligned} \tag{7.76}$$

The Grand partition function reduces to

$$\begin{aligned} Z_G &= \sum_{N=0}^{\mathcal{N}} \binom{\mathcal{N}}{N} e^{\beta\mu N} e^{N\beta\epsilon} \\ &= \sum_{N=0}^{\mathcal{N}} \binom{\mathcal{N}}{N} \left[e^{\beta(\mu+\epsilon)}\right]^N \\ &= \left(1+e^{\beta(\mu+\epsilon)}\right)^{\mathcal{N}} \end{aligned} \tag{7.77}$$

The mean number of particles adsorbed is given by (7.61)

$$\begin{aligned} \overline{N} &= \frac{1}{\beta}\left(\frac{\partial \ln Z_G}{\partial \mu}\right)_{\mathcal{N},T} \\ &= \mathcal{N}\frac{e^{\beta(\mu+\epsilon)}}{1+e^{\beta(\mu+\epsilon)}} \end{aligned} \tag{7.78}$$

which gives

$$\frac{\overline{N}}{\mathcal{N}} = \frac{1}{1+e^{-\beta\mu}e^{-\beta\epsilon}} \tag{7.79}$$

For a system of weakly interacting particles, $e^{\beta\mu}$ is given by eqn.(7.70). Then,

$$\begin{aligned} e^{-\beta\mu} &= \frac{k_BT}{P}\left(\frac{2m\pi k_BT}{h^2}\right)^{3/2} \\ &= \left(\frac{k_BT}{P}\right)\lambda_T^{-3} \end{aligned} \tag{7.80}$$

where λ_T is the thermal deBroglie wavelength (eqn.(6.59), section 6.4). Substituting, we get

$$\frac{\overline{N}}{\mathcal{N}} = \frac{P}{P+K(T)} \tag{7.81}$$

where

$$K(T) = \frac{k_BT}{\lambda_T^3}e^{-\epsilon/k_BT} \tag{7.82}$$

is a temperature dependent factor. The plot of $\overline{N}/\mathcal{N}$ vs P looks as follows

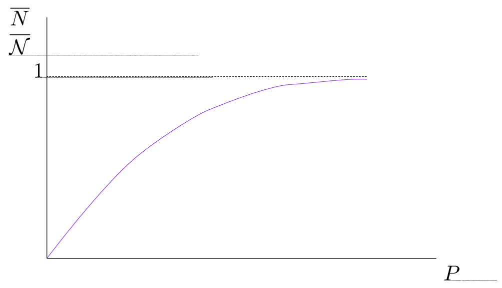


Figure 7.1: Langmuir Isotherm

Such isotherms are called *Langmuir Isotherms*. Qualitatively, the plotted isotherm seems to have the characteristics of emirically observed coating of surfaces with gas atoms. For some gas/surface interfaces, there is good agreement with experiments.

7.7 The Gibbs and Landau Potentials

In the canonical probability distribution (section 4.3) we found that the partition function readily gives the mean energy of the system, but to compute other equilibrium properties (such as pressure and entropy), it is natural to define a 'thermodynamic potential' or 'free energy' (the Helmholtz funtion F) which is directly related to the partition function, and in terms of which all equilibrium variables of interest can be directly computed

$$\begin{aligned} F(N,V,T) &= -k_B T \ln Z \\ P &= -\left(\frac{\partial F}{\partial V}\right)_T \\ S &= -\left(\frac{\partial F}{\partial T}\right)_V \\ E &= F + TS \end{aligned} \tag{7.83}$$

Further, this thermodynamic potential has the following interpretation: If the system is destabilized, the new equilibrium corresponds to a minimum of this free energy. It turns out that associated with every statistical probability distribution is a free energy function which: (a) is directly related to the partition function for that distribution (b) all equilibrium properties can be computed directly through this function by taking suitable partial derivatives and (c) equilibrium of the system corresponds to a minimum of this function.

Let us start with the NPT distribution (section 7.3) for a system of particles under constant temperature and pressure conditions

$$\begin{aligned} P_r &= \frac{1}{Z_G} e^{-\beta(E_r + PV_r)} \\ Z_G &= \sum_r e^{-\beta(E_r + PV_r)} \end{aligned} \tag{7.84}$$

In parallel with the arguments leading to the introduction of the Helmholtz function in section 4.3, we reorganize the sum in the partition function

$$\begin{aligned} Z_G &= \sum_r e^{-\beta(E_r + PV_r)} \\ &= \sum_{E,V} \mathcal{N}(E,V) e^{-\beta(E+PV)} \end{aligned} \tag{7.85}$$

where $\mathcal{N}(E,V)$ is the number of microstates of the system corresponding to energy E and volume V. This is directly related to the entropy of the system

$$S = k_B \ln \mathcal{N}(E,V) \tag{7.86}$$

Then

$$\begin{aligned} Z_G &= \sum_{E,V} e^{S(E,V)/k_B} e^{-\beta(E+PV)} \\ &= \sum_{E,V} e^{-\beta G(E,V)} \end{aligned} \tag{7.87}$$

where

$$G(E,V) = E + PV - TS(E,V) \tag{7.88}$$

Similar to the discussion in section 4.3, for a macroscopic system, the sum in 7.87 receives a contribution from sharply defined values of E and V (which, in principle, fluctuate), such that

$$Z_G(N,P,T) \simeq e^{-\beta G(E,V)} \tag{7.89}$$

where we denote these sharp values by E and V. Taking the logarithm gives

$$G = -k_B T \ln Z_G \tag{7.90}$$

Note that the partition function here is a function of three variables:N, P, T. Therefore, G is also a function of these three variables. However, G is directly related to the (mean) energy, volume and entropy as

$$G(N, P, T) = E + PV - TS \tag{7.91}$$

The function G is known as the *Gibbs Free Energy* or the *Gibbs Potential*, and is the analog of the Helmholtz function for this probability distribution. Given G (which can be directly computed through the partition function), we can compute all other functions. To see this, consider an infinitesimal variation in the external temperature and pressure (the control variables). This will lead to an infinitesimal change in the Gibbs function, mean energy, volume and entropy, such that

$$dG = dE + PdV + VdP - TdS - SdT \tag{7.92}$$

As in section 4.3, we can visualise S to be a function of N, E and V, such that under a change in P and V (resulting in a change in E by dE),

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E}\right)_V dE + \left(\frac{\partial S}{\partial V}\right)_E dV \\ &= \frac{1}{T}\, dE + \frac{P}{T}\, dV \end{aligned} \tag{7.93}$$

Substituting in equation (**??**), we get

$$\begin{aligned} dG &= dE + PdV + VdP - T\left(\frac{1}{T}\, dE + \frac{P}{T} dV\right) - SdT \\ &= VdP - SdT \end{aligned} \tag{7.94}$$

from which it follows that

$$\begin{aligned} V &= \left(\frac{\partial G}{\partial P}\right)_T \\ S &= -\left(\frac{\partial G}{\partial T}\right)_P \end{aligned} \tag{7.95}$$

which gives the mean volume and entropy directly as partial derivatives of the Gibbs function. The mean energy can be computed from (7.103)

$$E = G - PV + TS \tag{7.96}$$

Arguments similar to those in section 6.1 show that equilibrium corresponds to a minimum of the function G.

Next, we take a look at the Grand canonical distribution, for which the probability distribution and the partition function are

$$\begin{aligned} P_r &= \frac{1}{Z_G}\, e^{-\beta(E_r - \mu N_r)} \\ Z_G &= \sum_r e^{-\beta(E_r - \mu N_r)} \end{aligned} \tag{7.97}$$

Once again, we can reorganize the sum in the partition function

$$\begin{aligned} Z_G &= \sum_r e^{-\beta(E_r - \mu N_r)} \\ &= \sum_{N,E} \mathcal{N}(E, N) e^{-\beta(E - \mu N)} \end{aligned} \tag{7.98}$$

where $\mathcal{N}(E,N)$ is the number of microstates corresponding to N particles with energy E. As before, this can be expressed in terms of entropy, such that

$$\begin{aligned} Z_G &= \sum_{E,V} e^{S(E,V)/k_B} e^{-\beta(E-\mu N)} \\ &= \sum_{E,V} e^{-\beta\mathcal{G}(E,N)} \end{aligned} \tag{7.99}$$

where

$$\mathcal{G} = E - TS - \mu N \tag{7.100}$$

Again, the contribution in the sum (7.99) comes from sharply defined values of E and N, such that

$$Z_G(V,T,\mu) \simeq e^{-\beta\mathcal{G}(E,N)} \tag{7.101}$$

where we denote these sharp values by E and N. Taking the logarithm gives

$$\mathcal{G} = -k_B T \ln Z_G(V,T,\mu) \tag{7.102}$$

The function $\mathcal{G}$ is (throught the partition function) a function of three variables: V, T and μ. Further, it is directly related to the (mean) energy, entropy and particle number as

$$\mathcal{G}(V,T,\mu) = E - TS - \mu N \tag{7.103}$$

the function $\mathcal{G}$ is known as the *Grand Free Energy* or the *Landau Potential.* Given $\mathcal{G}$ (computed through the partition function), we can compute all other functions. To see this, consider an infinitesimal variation in the external temperature, volume and chemical potential (the control variables). This will lead to an infinitesimal change in the Landau potential, mean energy, entropy and particle number such that

$$d\mathcal{G} = dE - TdS - SdT - d\mu N - Nd\mu \tag{7.104}$$

Visualising the entropy as a function of E, V and N, the change in entropy due to a change in V, T and μ is (see section 7.6, eqn(7.105))

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E}\right)_{N,V} dE + \left(\frac{\partial S}{\partial V}\right)_{N,E} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN \\ &= \frac{1}{T}\,dE + \frac{P}{T}\,dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN \\ &= \frac{1}{T}\,dE + \frac{P}{T}\,dV - \frac{\mu}{T}dN \end{aligned} \tag{7.105}$$

where we have used the expression for the chemical potential as derivative of entropy with respect to particle number, eqn.(7.56). Substituting back in eqn.(7.104) gives

$$d\mathcal{G} = -PdV - SdT - Nd\mu \tag{7.106}$$

from which it follows that

$$\begin{aligned} P &= -\left(\frac{\partial \mathcal{G}}{\partial V}\right)_{T,\mu} \\ S &= -\left(\frac{\partial \mathcal{G}}{\partial T}\right)_{V,\mu} \\ N &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{V,T} \end{aligned} \tag{7.107}$$

Finally, the mean enrgy of the system is given by

$$E = \mathcal{G} + TS + \mu N \tag{7.108}$$

To get a feel for the Landau potential, let us compute the pressure, mean number of particles and the chemical potential for a system of weakly interacting particles, for which the Grand partition function was computed in eqn.(7.66)

$$Z_G = e^{e^{\beta\mu} V \left(2m\pi k_B T/h^2\right)^{3/2}} \tag{7.109}$$

The Landau potential is then given by (7.102)

$$\begin{aligned} \mathcal{G} &= -k_B T \ln Z_G(V, T, \mu) \\ &= -k_B T e^{\beta\mu} \left(\frac{V}{\lambda_T^3}\right) \end{aligned} \tag{7.110}$$

where λ_T is the thermal deBroglie wavelength (eqn.(6.59). The pressure of the system is computed to be

$$\begin{aligned} P &= -\left(\frac{\partial \mathcal{G}}{\partial V}\right)_{T,\mu} \\ &= k_B T e^{\beta\mu} \left(\frac{1}{\lambda_T^3}\right) \end{aligned} \tag{7.111}$$

and the mean number of particles is

$$\begin{aligned} N &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{V,T} \\ &= e^{\beta\mu} \left(\frac{V}{\lambda_T^3}\right) \end{aligned} \tag{7.112}$$

Dividing the equations for pressure and mean number of particles gives us the equation of state

$$PV = N k_B T \tag{7.113}$$

Further, equation (7.111) directly gives an expression for the chemical potential for a system of weakly interacting particles

$$\mu = k_B T \ln \left(\frac{P}{k_B T} \left(\frac{h^2}{2m\pi k_B T} \right)^{3/2} \right) \tag{7.114}$$

which we had obtained in equation (7.69) through more tedious means.
For a system with a variable number of particles, approach to equilibrium involves a decrease in the Landau potential, with equilibrium corresponding to its minimum value.

7.8 The Lattice Gas

The Grand canonical distribution is a very powerful way of visualising the statistical mechanics of macroscopic systems. As a simple example, let us demonstrate that a system of interacting particles can be 'mapped' to the Ising model in the Grand canonical picture. Consider a system of interacting particles in equilibrium at temperature T. We consider an imaginary, fixed volume V in space and ask the following: as the external pressure and temperature of the system are varied, how does the mean number of particles in this volume change? This is equivalent to analyzing the variation in density of the system as a function of temperature and pressure. Clearly, since the volume and temperature are fixed and the number of particles can vary, the Grand canonical distribution is the natural distribution to answer this question.

The Grand partition function for the system is

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} e^{\beta\mu N}\ Z_C(N,V,T) \\ &= \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{h^{3N}N!}\int d^{3N}\vec{r}\ d^{3N}\vec{p}\ e^{-\beta(\sum_i \vec{p}_i^2/2m+U(\vec{r}_1,..,\vec{r}_N))} \end{aligned} \tag{7.115}$$

where $U(\vec{r}_1,..,\vec{r}_N)$ is the interaction potential energy of the system. The momentum integrals are easily evaluated, giving

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{h^{3N}N!}(2m\pi k_B T)^{3N/2}\int d^{3N}\vec{r}e^{-\beta U(\vec{r}_1,..,\vec{r}_N)} \\ &= \sum_{N=0}^{\infty} \frac{z^N}{N!}Q_N \end{aligned} \tag{7.116}$$

where

$$z = e^{\beta\mu}\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2} \tag{7.117}$$

and

$$Q_N = \int d^{3N}\vec{r}e^{-\beta U(\vec{r}_1,..,\vec{r}_N)} \tag{7.118}$$

The interaction between a pair of particles is assumed to be short ranged, such that there is a natural length scale σ. If the separation between the particles is less than σ, the particles strongly repel each other. If the distance is larger than σ, they attract upto separation of the the order of a few σ. To model this behaviour, we discretize volume V into $\mathcal{N}$ elementary 'cells'. By assuming that a cell can be occupied by at most one particle, we can replicate the effect of strong repulsion at short distance. Further, we can replicate the effect of attraction between a pair at short distances beyond σ by assuming that if two particles occupy neighbouring cells, they have an interaction potential energy $-\epsilon$, the negative sign corresponding to attraction. Beyond neighbouring cells, the interaction is asumed to be zero. Then, for a fixed number of particles N, a microstate corresponds to a specification of which particle occupies which cell. Note that in Q_N, the particles are *distinguishable*, the inherent indistinguishability accounted for the $1/N!$ term in eqn.(7.116). The occupation of the i^{th} cell can be described in terms of an *occupation number* n_i which is zero if the cell is unoccupied and one if it is occupied. Let us define a configuration $\mathcal{C}$ by a specification of numbers n_i, $i = 1, 2, .., \mathcal{N}$. Given a configuration $\mathcal{C}$ of N particles, the potential energy of the system can be written as

$$U(\mathcal{C}) = -\epsilon \sum_{<i,j>} n_i n_j \tag{7.119}$$

where $\sum_i n_i = N$ and $< i, j >$ denotes nearest neighbours. When we sum over microsates of the N particles, there is a $N!$ degeneracy in energy due to interchange of particles, which leaves energy unchnaged. Then,

$$Q_N = N!\sum_{\mathcal{C}} e^{-\beta U(\mathcal{C})} \tag{7.120}$$

In eqn.(7.116), the term z^N can be written as

$$\begin{aligned} z^N &= e^{N\ln z} \\ &= e^{\ln z\sum_i n_i} \end{aligned} \tag{7.121}$$

The Grand partition function now reduces to

$$Z_G = \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} \tag{7.122}$$

where

$$E(\mathcal{C}) = -\epsilon \sum_{<i,j>} n_i n_j - \frac{\ln z}{\beta} \sum_i n_i \tag{7.123}$$

where $n_i = \pm 1$. Let us introduce a 'spin' variable σ_i associated with the i^{th} lattice site, which takes values $\sigma_i = \pm 1$. In terms of this variable, the occupation number can be written as

$$n_i = \frac{1}{2}(1 + \sigma_i) \tag{7.124}$$

such that $n_i = 1$ corresponds to $\sigma_i = 1$ and $n_i = 0$ corresponds to $\sigma_i = -1$. Substituting, we get

$$\begin{aligned} E(\mathcal{C}) &= -\frac{\epsilon}{4} \sum_{<i,j>} (1+\sigma_i)(1+\sigma_j) - \frac{\ln z}{2\beta} \sum_i (1+\sigma_i) \\ &= -\frac{\epsilon}{4} \sum_{<i,j>} \sigma_i \sigma_j - \frac{\epsilon}{2} \sum_{<i,j>} \sigma_i - \frac{\ln z}{2\beta} \sum_i \sigma_i + \text{constant terms} \end{aligned} \tag{7.125}$$

Let q be the number of nearest neighbours of any one cell. Then

$$\sum_{<i,j>} \sigma_i = \frac{q}{2} \sigma_i \tag{7.126}$$

Therefore (dropping constant terms)

$$E(\mathcal{C}) = -\frac{\epsilon}{4} \sum_{<i,j>} \sigma_i \sigma_j - \left(\frac{q\epsilon}{4} + \frac{\ln z}{2\beta} \right) \sum_i \sigma_i \tag{7.127}$$

which is identical to the expression for energy for the Ising model (eqn.(3.17)), if we make the identification

$$\begin{aligned} J &\longleftrightarrow \frac{\epsilon}{4} \\ h &\longleftrightarrow \frac{q\epsilon}{4} + \frac{\ln z}{2\beta} \end{aligned} \tag{7.128}$$

What is the analog of magnetization in this model? In the Ising model, the magnetization for a microstate $\mathcal{C}$ is given by

$$M = \frac{1}{\mathcal{N}} \sum_i \sigma_i \tag{7.129}$$

where $\mathcal{N}$ is the total number of sites in our lattice model. Corresponding to configuration $\mathcal{C}$, the total number of particles occupying $\mathcal{N}$ sites is

$$\begin{aligned} N &= \sum_i n_i \\ &= \frac{1}{2} \sum_i (1+\sigma_i) \\ &= \frac{\mathcal{N}}{2} + \frac{1}{2} \sum_i \sigma_i \end{aligned} \tag{7.130}$$

Then,

$$\frac{N}{\mathcal{N}} = \frac{1}{2}(1+M) \tag{7.131}$$

Let the volume of an elementary cell be v_0 and the mass of a particle be m_0. Then, the density of the gas is

$$\begin{aligned} \rho &= \frac{m_0 N}{\mathcal{N} v_0} \\ &= \left(\frac{m_0}{v_0} \right) \frac{N}{\mathcal{N}} \end{aligned} \tag{7.132}$$

Therefore, $N/\mathcal{N}$ is just the density of the system measured in units of m_0/v_0. Therefore, the analog of magneization is density of the system of particles

$$\rho = \frac{1}{2}(1 + M) \tag{7.133}$$

In the Ising model, we have observed that the system spontaneously magnetizes below a critical temperature T_C. What does this imply for the lattice gas? To see this, let us revisit the Ising model and plot isotherms corresponding to $T > T_C$ and $T < T_C$. An isotherm in the Ising model is just a plot of the magnetization M versus the magnetic field h at constant temperature. For $T > T_C$, the system is in a paramagnetic phase, with no spontaneous magnetization (in absence of h). Then, an isotherm will typically look as follows

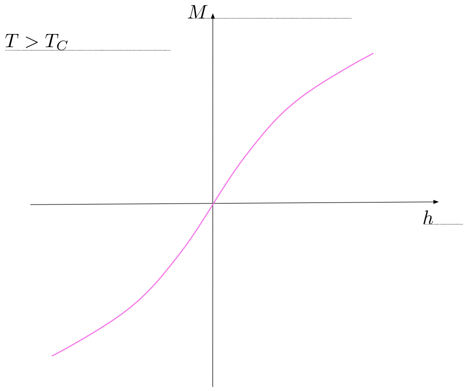


Figure 7.2: Isotherm for Paramagnetic phase

The variation of M with h is smooth, with $M = 0$ when $h = 0$. However, for $T < T_C$, there can be a non-zero magnetization even in absence of a magnetic field. There is an ambiguity in the direction of magnetization, with both directions possible (see figure 6.3). In presence of the magnetic field, though, the magnetization is in the same direction as the magnetic field. A plot of magnetization vs magnetic field for $T < T_C$ looks as follows

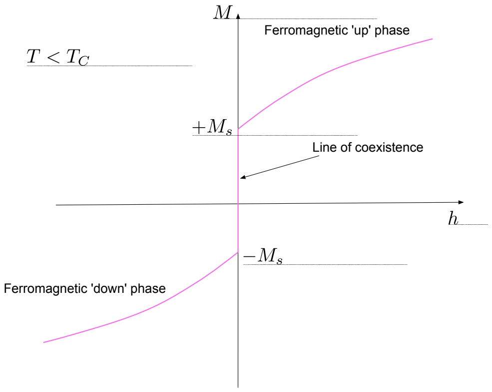


Figure 7.3: Isotherm for Paramagnetic phase

In the figure, one cas see that there are two 'phases': ferromagnetic 'up' (more spins 'up') and ferromagnetic 'down' (more spins 'down') phases. However, at zero external field, both phases coexist in the ferromagnetic substance, with the formation of domains of each kind. What is the analog of this in the lattice gas? We have seen that the analog of magnetization is density of the system of particles. Given the correspondence in eqns.(7.128), the quantity corresponding to the magnetic field h is (apart from the constant $q\epsilon/4$) is

$$\begin{aligned} h &\longrightarrow \frac{1}{2\beta}\ln z \\ &= \frac{1}{2}\mu + \frac{3}{2}\ln\left(\frac{2m\pi k_B T}{h^2}\right) \end{aligned} \tag{7.134}$$

Along an isotherm, a changing h corresponds to a changing μ. Then, the isotherm for a system of particles for $T < T_C$ will look as follows

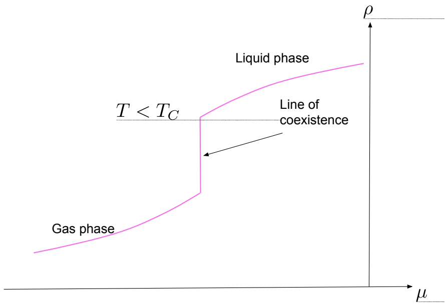


Figure 7.4: Lattice Gas Isotherm for $T < T_C$

A change in chemical potential at fixed temperature and volume (same as fixed $\mathcal{N}$) is equivalent to a change in pressure of the system. Then, the isotherm can be viewed as than demonstrating a change in density with pressure. Clearly, there is a discontinuity in density at a specific pressure, which corresponds to a change from a gas to a liquid state. The line of coexistence corresponds to both phases coexisting for $T < T_C$. The equation of this line is just $h = 0, T < T_C$, which for a system of particles translates to (through correspondence 7.128)

$$\mu + \frac{q\epsilon}{2} + \frac{3}{2}k_B T\ln\left(\frac{2m\pi k_B T}{h^2}\right) = 0 \tag{7.135}$$

What is the interpretation of the critical temperature T_C and what is the analog of the paramagnetic phase (for $T > T_C$)? Just as fluctuations in magnetization for a magnetic system are large for $h = 0, T = T_C$ (which is a 'critical point'), similar behaviour exists in a system of particles where density fluctuations are large and the system loses homogeneity. The critical point is obtained from eqn.(7.135) by setting $T = T_C$. for $T > T_C$, the analog of tha paramagnetic phase is the so-called 'fluid phase' in which there is no distinction between vapour and liquid.

7.9 Density variation in an external field

We have observed that if a system consists of two subsystems with different chemical potentials, there is a flow of particles from the subsystem with a higher chemical potential to one with a lower chemical potential (section 7.6). This 'diffusion' stops at equilibrium when the chemical potentials equalize. In general, if there is a variation in the chemical potential of a system of particles from point to point, diffusion occurs till the chemical potential at all points is the same. However, in general, this does not imply that the density of the system is uniform at all point. In presence of external fields (electromagnetic or gravitational, for example), the chemical potential, apart from density, is also a function of the 'potential' associated with these fields. To keep the chemical potential uniform, the density must change from point to point because of the variation in the potential associated with such fields. As a simple example, consider a syste of weakly interacting particles in presence of a uniform external gravitartional field, such as that near the surface of the Earth. The density of the system in such a situation is not uniform, but changes with height above the surface of the Earth. This variation in density, in equilibrium, can be visualised as a consequence of the chemical potential depending on the potential associated with the gravitational field (apart from the density of the system). To see this. consider an imaginary slice of volume ΔV at height z above the surface of the Earth. The thickness of this slice is assumed to be much smaller than z, such that z can be assumed to be a constant in this slice. Each particle in this slice has a gravitational potential energy $U = mgz$ where g is the acceleration due to gravity. Particles are free to enter and exit this imaginary slice. Therefore, the system satisfies the Grand canonical probability distribution. In equilibrium, there will be a (sharply defined) mean number of particles ΔN in this slice, which can be obtained from the Landau potential

$$\Delta N = -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{\Delta V,T} \tag{7.136}$$

The Grand partition function for the system is given by eqn.(7.63)

$$Z_G = \sum_{N=0}^{\infty} e^{\beta\mu n} Z(N, \Delta V, T) \tag{7.137}$$

where $Z(N, \Delta V, T)$ is the Canonical partition function for this system, given by

$$\begin{aligned} Z &= \frac{1}{h^{3N}N!}\int d^3\vec{r}_1...d^3\vec{r}_N d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\left[\sum_{i=1}^N \vec{p}_i^2/2m + \sum_{i=1}^N mgz\right]} \\ &= e^{N\beta mgz} \times \frac{1}{h^{3N}N!}\int d^3\vec{r}_1...d^3\vec{r}_N d^3\vec{p}_1..d^3\vec{p}_N e^{-\beta\sum_{i=1}^N \vec{p}_i^2/2m} \\ &= e^{N\beta mgz}\frac{1}{N!}\left[\Delta V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]^N \end{aligned} \tag{7.138}$$

The Grand partition function is then

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty}\frac{1}{N!}\left[e^{\beta(\mu-mgz)}\Delta V\left(\frac{2m\pi k_B T}{h^2}\right)^{3/2}\right]^N \\ &= e^{\left[e^{\beta(\mu-mgz)}\Delta V\left(2m\pi k_B T/h^2\right)^{3/2}\right]} \end{aligned} \tag{7.139}$$

The Landau potential is

$$\begin{aligned} \mathcal{G} &= -k_B T \ln Z_G \\ &= -k_B T e^{\beta(\mu-mgz)}\left(\frac{\Delta V}{\lambda_T^3}\right) \end{aligned} \tag{7.140}$$

where λ_T is the thermal deBroglie wavelength (eqn.(6.59). The mean number of particles is

$$\begin{aligned} \Delta N &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{\Delta V,T} \\ &= \Delta V e^{\beta\mu} e^{-\beta mgz} \times \frac{1}{\lambda_T^3} \end{aligned} \tag{7.141}$$

Them, the number density $\rho(z)$ od the system is

$$\rho(z) = e^{\beta\mu} e^{-\beta mgz} \times \frac{1}{\lambda_T^3} \tag{7.142}$$

This gives the following expression for the chemical potential

$$\mu = mgz + k_B T \ln\left(\frac{\rho(z)}{\lambda_T^3}\right) \tag{7.143}$$

In equilibrium, μ should be independent of z. Therefore, the right hand side of equation (7.143) is idependent of z. In particular, its value at any z should be equal to that at $z = 0$

$$mgz + k_B T \ln\left(\frac{\rho(z)}{\lambda_T^3}\right) = k_B T \ln\left(\frac{\rho(0)}{\lambda_T^3}\right) \tag{7.144}$$

This gives the variation of density with z

$$\rho(z) = \rho(0)\ e^{-mgz/k_B T} \tag{7.145}$$

This is just the so-called *barometric equation*, since it also gives the variation with pressure with height above the surface of the Earth (assuming that the temperature does not vary with height, which is quite incorrect!).

7.10 Critical point for stellar equilibrium

Stars start out as predominantly 'clouds' of hydrogen gas (GMC or Giant Molecular Clouds), held in equlibrium (upto a critical density) by pressure due to collisions between gas molecules. This pressure counters the gravitational attraction, which by itself would lead to gravitational collapse. Beyond a certain critical density, the pressure due to simple collisions is not enough to counter gravitational collapse, and the collapse proceeds, till a 'protostar' forms, leading to nuclear fusion acting as a source of energy, increasing the pressure, creating a new equilibrium. Here, we roughly estimate an expression for this critical density. We model the hydrogen gas as a system of weakly interacting particles, interacting through collisions and mutual gravitational attraction. We use the Grand canonical ensemble, and pose the following question. Imagine the cloud as a spherical region of radius R and volume $V = (4/3)\pi R^3$ at temperature T, free to exchange particles with a hypothetical environment. What is the density of this system in equilibrium? The density of the system will in fact change with distance from the centre. What we are interested in is the *mean density* $\overline{\rho} = M/V$ where M is the mass of the cloud. The Grand canonical partition function for the system is

$$Z_G = \sum_{N=0}^{\infty} e^{\beta\mu N}\ Z_C(N, V, T) \tag{7.146}$$

where $Z_C(N, V, T)$ is the canonical partition function

$$Z_C(N, V, T) = \frac{1}{h^{3N} N!} \int d^{3N}\vec{r}\ d^{3N}\vec{p}\ e^{-\beta(\sum_i \vec{p}_i^2/2m + U(\vec{r}_1,..,\vec{r}_N))} \tag{7.147}$$

where U is the gravitational potential energy of the system. We visualise the system as a continuous system with mass density function $\rho(\vec{r})$. Then, the gravitational potential energy of the system is

$$U = -\frac{1}{2} G \int d^3\vec{r} d^3\vec{r}'\ \frac{\rho(\vec{r})\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \tag{7.148}$$

The first simplification we make is to assume that we can replace $\rho(\vec{r})$ in eqn.(7.148) by the mean density. Effectively, this is equivalent to computing the gravitational potential energy of a uniform sphere of radius R and mass M (a simple exercise in mechanics!) which gives

$$U = -\frac{3}{5}\frac{M^2 G}{R} \tag{7.149}$$

Here, $M = N m_{H_2}$ where m_{H_2} is the mass of a hydrogen molecule. The next simplifying assumption is to substitute M^2 by $M \times \overline{M}$ where as before $M = N m_{H_2}$, and where $\overline{M}$ is the *mean mass* of the cloud, given by $\overline{M} = \overline{N} m_{H_2}/V$ where $\overline{N}$ is the mean number of hydrogen molecules (computed using the Landau potential). The momentum integrals in Z_C are easily computed, so that

$$Z_C = \frac{1}{N!} V^N \left(2m\pi k_B T/h^2\right)^{3N/2} e^{(3/5)\beta N m_H \overline{M} G/R} \tag{7.150}$$

Substituting in the expression for Z_G gives

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{V}{\lambda_T^3}\right)^N e^{N\beta(\mu + 3m_H \overline{M} G/5R)} \\ &= e^{\left[e^{\beta(\mu+3m_H\overline{M}G/5R)} V/\lambda_T^3\right]} \end{aligned} \tag{7.151}$$

The Landau potential is

$$\begin{aligned} \mathcal{G} &= -k_B T \ln Z_G \\ &= -k_B T \left(\frac{V}{\lambda_T^3}\right) e^{\beta(\mu + 3m_H \overline{M} G/5R)} \\ &= -k_B T \left(\frac{V}{\lambda_T^3}\right) e^{\beta\left[\mu + (3m_H/5)\overline{M} G (4\pi/3V)^{1/3}\right]} \end{aligned} \tag{7.152}$$

where we have used $R = (3V/4\pi)^{1/3}$. The mean number of particles is

$$\begin{aligned} \overline{N} &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{V,T} \\ &= \left(\frac{V}{\lambda_T^3}\right) e^{\beta\left[\mu + (3m_H/5)\overline{M} G (4\pi/3V)^{1/3}\right]} \end{aligned} \tag{7.153}$$

The pressure of the system is

$$\begin{aligned} P &= -\left(\frac{\partial \mathcal{G}}{\partial V}\right)_{T,\mu} \\ &= \left(\frac{k_B T}{\lambda_T^3}\right) e^{\beta\left[\mu + (3m_H/5)\overline{M} G (4\pi/3V)^{1/3}\right]} \left[1 - \frac{1}{k_B T} \frac{m_{H_2} \overline{M} G}{5} \left(\frac{4\pi}{3}\right)^{1/3} \frac{1}{V^{1/3}}\right] \end{aligned} \tag{7.154}$$

Dividing the pressure equation by the (mean) particle number equation, we get the equation of state

$$P = \frac{\overline{N} k_B T}{V} - \frac{\overline{N} m_{H_2} G \overline{M}}{5} \left(\frac{4\pi}{3}\right)^{1/3} V^{-4/3} \tag{7.155}$$

We now need to check for possible criticality. Is it possible for the compressibility of the system to diverge? Equivalently, under what conditions

$$\frac{\partial P}{\partial V} = 0 \tag{7.156}$$

Taking the derivative of P with respect to V and setting it to zero gives the critical radius beyond which pressure due to collisions cannot prevent gravitational collapse

$$R_C = \left(\frac{45}{16\pi}\right)^{1/2} \sqrt{\frac{k_B T}{m_{H_2}^2 G \rho}} \tag{7.157}$$

where ρ is the number density.

Chapter 8

Statistical Mechanics of Radiation

8.1 Microstates of Electromagnetic Fields

We now analyse the statistical mechanics of a completely different kind of system: electromagnetic fields. Given a system of atoms or molecules in thermal equilibrium at some temperature, there will always be a distribution of electric and magnetic fields associated with such a system of particles, since atoms, consisting of charged particles, can emit and absorb radiation through quantum mechanical transitions. Consider a system of atoms occupying a region of space with volume V and at temperature T. It is convenient to visualise this region of space as one without boundaries, since boundaries come with their own set of conditions that need to be imposed on the physical system enclosed within them. Then, we take this region of space to be a torus, constructed as follows: we start with a cube of edge L ($V = L^3$) and identify the opposite faces to form a torus. Mathematically, if we choose a coordinate system x, y, z such that one edge of the cube has coordinates $x, y, z = 0$. Then, constructing a torus out of this cube is equivalent to forcing 'periodic boundary conditions' on any function $f(x, y, z)$ defined on this region

$$\begin{aligned} f(x+L, y, z) &= f(x, y, z) \\ f(x, y+L, z) &= f(x, y, z) \\ f(x, y, z+L) &= f(x, y, z) \end{aligned} \tag{8.1}$$

This region of space will be filled with electromagnetic fields in equilibrium with matter at temperature T. Then, we can consider the system as two interacting subsystems, particles and fields, interacting with each other through exchange of energy. The probability of the fields being in amicrostate r will be given by a suitable statistical distribution. The question is: what are the mocrostates of this system and what is the appropriate statistical distribution? A system of electromagnetic fields is just a distribution of electric and magentic fields satisfying Maxwell's equations. These fileds are confined to this toroidal region, with a microstate being a specification of distributions $\vec{E}(x, y, z)$ and $\vec{B}(x, y, z)$. As always, we consider this system in isolation (in absence of any other system) and determine the dependence of the energy of this system on its microstates. Then, if say the canonical distribution is to be used, the probability of a microstate will be given by eqn.(4.28). Since this system of fields is being considered in isolation (no charged particles around), it will satisfy source-free Maxwell equations

$$\begin{aligned} \vec{\nabla} \cdot \vec{E} &= 0 \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} &= 0 \\ \vec{\nabla} \times \vec{B} &= \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \end{aligned} \tag{8.2}$$

These equations describe a *constained* system. Therefore, to identify the physical degrees of freedom, we introduce scalar and vector potential ϕ and $\vec{A}$ such that

$$\begin{aligned} \vec{E} &= -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \vec{\nabla} \times \vec{A} \end{aligned} \tag{8.3}$$

with these potentials now identified as the dynamic degrees of freedom. However, the 'gauge freedom' in the system implies that there is a redundancy in this description, since given a pair of potentials $(\phi, \vec{A})$, another pair $(\phi', \vec{A}')$ related to the first through

$$\begin{aligned} \vec{A}' &= \vec{A} + \vec{\nabla}\lambda \\ \phi' &= \phi - \frac{\partial \lambda}{\partial t} \end{aligned} \tag{8.4}$$

will give the same electric and magnetic field distributions for any scalar function λ. The fix is to choose a gauge, which is just a restriction of the potentials. Given a pair $(\phi, \vec{A})$, we can always choose a λ such that the new pair $(\phi', \vec{A}')$ satisfies the conditions

$$\begin{aligned} \phi' &= 0 \\ \vec{\nabla} \cdot \vec{A}' &= 0 \end{aligned} \tag{8.5}$$

This is just one of infinite possible choices, picked here for its simplicity. This choice is known as choosing the *Coulomb gauge*. Then, the microstates of the system are all possible distributions $\vec{A}$ on the torus, restricted by

$$\vec{\nabla} \cdot \vec{A} = 0 \tag{8.6}$$

Any such distribution must satisfy the periodic boundary conditions

$$\begin{aligned} \vec{A}(x+L, y, z, t) &= \vec{A}(x, y, z) \\ \vec{A}(x, y+L, z, t) &= \vec{A}(x, y, z) \\ \vec{A}(x, y, z+L, t) &= \vec{A}(x, y, z) \end{aligned} \tag{8.7}$$

To understand the implications of this periodicity, let us consider a function $f(x,t)$ such that $f(x+L,t) = f(x,t)$. Since the function is periodic in x with period L, it can be expanded in a Fourier series

$$f(x,t) = \sum_{n=0}^{\infty} \left[a_n(t) \cos\left(\frac{2n\pi x}{L}\right) + b_n(t) \sin\left(\frac{2n\pi x}{L}\right) \right] \tag{8.8}$$

where a_n, b_n are Fourier coefficients. The series can be equivalently written as a complex series of the form

$$\begin{aligned} f(x,t) &= \sum_{n=-\infty}^{\infty} c_n(t) e^{i2\pi n x/L} \\ &= \sum_{n=-\infty}^{\infty} c_n(t) e^{ik_n x} \end{aligned} \tag{8.9}$$

where $k_n = 2\pi n/L$. The complex coefficients c_n are given by

$$c_n(t) = \frac{1}{L} \int_0^L dx\ f(x,t) e^{-ik_n x} \tag{8.10}$$

This can be easily extended to periodic functions of three variables x, y, z with period L

$$\begin{aligned} f(x,y,z,t) &= \sum_{l,m,n} c_{l,m,n}(t)\ e^{ik_l x} e^{ik_m y} e^{ik_n z} \\ &= \sum_{l,m,n} c_{l,m,n}(t)\ e^{i(k_l x + k_m y + k_n z)} \end{aligned} \tag{8.11}$$

where $k_l = 2\pi l/L, k_m = 2\pi m/L$ and $k_n = 2\pi n/L$ and

$$c_{l,m,n}(t) \quad = \quad \frac{1}{L^3}\int_0^L dx \int_0^L dy \int_0^L dz \; f(x,y,z,t) \; e^{-i(k_l x + k_m y + k_n z)} \tag{8.12}$$

It is convenient to define a vector

$$\vec{k} = \frac{2\pi l}{L}\;\hat{i} + \frac{2\pi m}{L}\;\hat{j} + \frac{2\pi n}{L}\;\hat{k} \tag{8.13}$$

where $\hat{i},\hat{j},\hat{k}$ are unit vectors along x,y,z directions respectively. Then, the Fourier expansion can be written as

$$f(\vec{r},t) = \sum_{\vec{k}} c_{\vec{k}}(t)\; e^{i\vec{k}\cdot\vec{r}} \tag{8.14}$$

with

$$c_{\vec{k}}(t) = \frac{1}{V}\int d^3\vec{r}\; f(\vec{r},t)\; e^{-i\vec{k}\cdot\vec{r}} \tag{8.15}$$

Given that the function f is real, $f^*(\vec{r},t) = f(\vec{r},t)$. This implies

$$\sum_{\vec{k}} c^*_{\vec{k}}\; e^{-i\vec{k}\cdot\vec{r}} \quad = \quad \sum_{\vec{k}} c_{\vec{k}}\; e^{i\vec{k}\cdot\vec{r}} \tag{8.16}$$

Since the sum over $\vec{k}$ is over all positive and negative values of $\vec{k}$, on the left hand side of equation(8.16), we can replace $\vec{k}$ by $-\vec{k}$ in the summand. Then, eqn.(8.16) reduces to

$$\sum_{\vec{k}} c^*_{-\vec{k}}\; e^{i\vec{k}\cdot\vec{r}} \quad = \quad \sum_{\vec{k}} c_{\vec{k}}\; e^{i\vec{k}\cdot\vec{r}} \tag{8.17}$$

which leads to the condition $c^*_{\vec{k}} = c_{-\vec{k}}$.

Going back to electromagnetic fields, the vector potential can be similarly expanded in a Fourier series

$$\vec{A}(\vec{r},t) = \sum_{\vec{k}} \vec{c}_{\vec{k}}(t)\; e^{i\vec{k}\cdot\vec{r}} \tag{8.18}$$

where $\vec{c}_{\vec{k}}$ are vector Fourier coefficients. We now implement the gauge condition $\vec{\nabla}\cdot\vec{A} = 0$. Given eqn.(8.18), the divergence of $\vec{A}$ is

$$\vec{\nabla}\cdot\vec{A} = i\sum_{\vec{k}} \vec{k}\cdot\vec{c}_{\vec{k}}(t)\; e^{i\vec{k}\cdot\vec{r}} \tag{8.19}$$

Then, the gauge condition $\vec{\nabla}\cdot\vec{A} = 0$ is equivalent to the condition

$$\vec{k}\cdot\vec{c}_{\vec{k}} = 0 \quad \forall\vec{k} \tag{8.20}$$

This just means that every Fourier mode is orthogonal to the associated direction of propagation of the corresponding plane wave $e^{i\vec{k}\cdot\vec{r}}$, given by the direction of vector $\vec{k}$.

A microstate of the system is the specification of these Fourier coefficients. Now, we need an expression for energy of the system in terms of these coefficients. The energy of a system of fields confined to volume V is given by

$$E = E_E + E_M \tag{8.21}$$

where

$$E_E = \frac{\epsilon_0}{2}\int d^3\vec{r}\;\vec{E}^2 \tag{8.22}$$

and

$$E_M = \frac{1}{2\mu_0}\int d^3\vec{r}\;\vec{B}^2 \tag{8.23}$$

In the Coulomb gauge

$$\begin{aligned} \vec{E} &= -\frac{\partial \vec{A}}{\partial t} \\ \vec{B} &= \vec{\nabla} \times \vec{A} \end{aligned} \tag{8.24}$$

Using the Fourier expansion of the vector potential, we get

$$\frac{\partial \vec{A}}{\partial t} = \sum_{\vec{k}} \dot{\vec{c}}_{\vec{k}}(t)\ e^{i\vec{k}\cdot\vec{r}} \tag{8.25}$$

and

$$\vec{\nabla} \times \vec{A} = i \sum_{\vec{k}} \vec{k} \times \vec{c}_{\vec{k}}(t)\ e^{i\vec{k}\cdot\vec{r}} \tag{8.26}$$

We are now ready to compute the expression for the energy of the system. The electric contribution is given by

$$\begin{aligned} E_E &= \frac{\epsilon_0}{2} \int d^3\vec{r}\ \vec{E}^2 \\ &= \frac{\epsilon_0}{2} \int d^3\vec{r}\ \frac{\partial \vec{A}}{\partial t} \cdot \frac{\partial \vec{A}}{\partial t} \\ &= \frac{\epsilon_0}{2} \int d^3\vec{r}\ \sum_{\vec{k},\vec{k}'} \dot{\vec{c}}_{\vec{k}} \cdot \dot{\vec{c}}_{\vec{k}'}\ e^{i\vec{k}\cdot\vec{r}}\ e^{i\vec{k}\cdot\vec{r}'} \\ &= \frac{\epsilon_0}{2} \sum_{\vec{k},\vec{k}'} \dot{\vec{c}}_{\vec{k}} \cdot \dot{\vec{c}}_{\vec{k}'} \int d^3\vec{r}\ e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} \end{aligned} \tag{8.27}$$

Given the form of $\vec{k}$, it is easy to show that

$$\int d^3\vec{r}\ e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} = V\delta_{\vec{k},-\vec{k}'} \tag{8.28}$$

where $V = L^3$. To see this, we observe that given $\alpha = 2n\pi/L$ and $\beta = 2m\pi/L$

$$\begin{aligned} \int_0^L dx\ e^{i(\alpha+\beta)x} &= \int_0^L dx\ e^{(2\pi i/L)(n+m)x} \\ &= \begin{cases} 0 & n+m \neq 0 \\ L & n+m = 0 \end{cases} \\ &= L\ \delta_{n,-m} \end{aligned} \tag{8.29}$$

and that integral in eqn.(8.28) is a product of three integrals of this type (integrals over x, y and z). Equation (8.28) is just an expression of orthogonality of functions $e^{i\vec{k}\cdot\vec{r}}$. Given this, the expression for the electric part of energy reduces to

$$\begin{aligned} E_E &= \frac{\epsilon_0}{2} V \sum_{\vec{k}} \dot{\vec{c}}_{\vec{k}} \cdot \dot{\vec{c}}_{-\vec{k}} \\ &= \frac{\epsilon_0}{2} V \sum_{\vec{k}} \dot{\vec{c}}_{\vec{k}} \cdot \dot{\vec{c}}_{\vec{k}}^{\,*} \\ &= \frac{\epsilon_0}{2} V \sum_{\vec{k}} \left| \dot{\vec{c}}_{\vec{k}} \right|^2 \end{aligned} \tag{8.30}$$

where we have used $\vec{c}_{-\vec{k}} = \vec{c}_{\vec{k}}^{\,*}$. Similar calculation for magnetic part of the energy gives

$$\begin{aligned}
E_M &= \frac{1}{2\mu_0}\int d^3\vec{r}\,\left(\vec{\nabla}\times\vec{A}\right)^2 \\
&= -\frac{1}{2\mu_0}\sum_{\vec{k},\vec{k}'}(\vec{k}\times\vec{c}_{\vec{k}})\cdot(\vec{k}'\times\vec{c}_{\vec{k}'})\int d^3\vec{r}\;e^{i(\vec{k}+\vec{k}')\cdot\vec{r}} \\
&= -\frac{V}{2\mu_0}\sum_{\vec{k},\vec{k}'}(\vec{k}\times\vec{c}_{\vec{k}})\cdot(\vec{k}'\times\vec{c}_{\vec{k}'})\;\delta_{\vec{k},-\vec{k}'} \\
&= \frac{V}{2\mu_0}\sum_{\vec{k}}(\vec{k}\times\vec{c}_{\vec{k}})\cdot(\vec{k}\times\vec{c}_{-\vec{k}}) \\
&= \frac{V}{2\mu_0}\sum_{\vec{k}}(\vec{k}\times\vec{c}_{\vec{k}})\cdot(\vec{k}\times\vec{c}_{\vec{k}}^{\,*}) \\
&= \frac{V}{2\mu_0}\sum_{\vec{k}}\left|\vec{k}\times\vec{c}_{\vec{k}}\right|^2 \qquad (8.31)
\end{aligned}$$

Using eqn.(8.20), this reduces to

$$\begin{aligned}
E_M &= \frac{V}{2\mu_0}\sum_{\vec{k}}k^2\left|\vec{c}_{\vec{k}}\right|^2 \\
&= \frac{\epsilon_0}{2}V\sum_{\vec{k}}\omega_{\vec{k}}^2\left|\vec{c}_{\vec{k}}\right|^2 \qquad (8.32)
\end{aligned}$$

where $\omega_{\vec{k}} = c\left|\vec{k}\right|$ and we have used $\mu_0\epsilon_0 = 1/c^2$.

The expression for the total energy becomes

$$E = \frac{\epsilon_0}{2}V\sum_{\vec{k}}\left\{\left|\dot{\vec{c}}_{\vec{k}}\right|^2 + \omega_{\vec{k}}^2\left|\vec{c}_{\vec{k}}\right|^2\right\} \qquad (8.33)$$

Let us write the complex fourier coefficients $\vec{c}_{\vec{k}}$ as

$$\vec{c}_{\vec{k}} = \vec{a}_{\vec{k}} + i\;\vec{b}_{\vec{k}} \qquad (8.34)$$

where $\vec{a}_{\vec{k}}$ and $\vec{b}_{\vec{k}}$ are real. Given that $\vec{c}_{-\vec{k}} = \vec{c}_{\vec{k}}^{\,*}$, it follows that

$$\begin{aligned}
\vec{a}_{-\vec{k}} &= \vec{a}_{\vec{k}} \\
\vec{b}_{-\vec{k}} &= -\vec{b}_{\vec{k}} \qquad (8.35)
\end{aligned}$$

Furthur, $\left|\vec{c}_{\vec{k}}\right|^2 = \vec{a}_{\vec{k}}^2 + \vec{b}_{\vec{k}}^2$ and $\left|\dot{\vec{c}}_{\vec{k}}\right|^2 = \dot{\vec{a}}_{\vec{k}}^2 + \dot{\vec{b}}_{\vec{k}}^2$. Then,

$$E = \frac{\epsilon_0}{2}V\sum_{\vec{k}}\left\{\left(\dot{\vec{a}}_{\vec{k}}^2 + \omega_{\vec{k}}^2\;\vec{a}_{\vec{k}}^2\right) + \left(\dot{\vec{b}}_{\vec{k}}^2 + \omega_{\vec{k}}^2\;\vec{b}_{\vec{k}}^2\right)\right\} \qquad (8.36)$$

Equation(8.36) has the following interpretation. Each term in the sum is of the form (apart from the factor of $\epsilon_0 V$

$$\frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega^2x^2 \qquad (8.37)$$

which can be thought of as the energy of a harmonic oscillator with frequency ω and unit mass. Then, eqn.(8.36) seems to suggest that electromagnetic radiation enclosed in some volume is mathematically equivalent to a collection of *two* types of harmnonic oscillaltors (corresponding to the a and b terms), each type itself being a collection of oscillaotrs of different angular frequencies $\omega_{\vec{k}} = c\left|\vec{k}\right|$ where $\vec{k}$ is the wave number vector associated with a plane wave of the form $e^{i\vec{k}\cdot\vec{r}}$. Note, however, that since $\vec{a}_{-\vec{k}} = \vec{a}_{\vec{k}}$ and $\vec{b}_{-\vec{k}} = \vec{b}_{\vec{k}}$, and further since $\omega_{\vec{k}} = \omega_{-\vec{k}}$, we can visualise the set of $\vec{k}$ vectors to be divided into two subsets, one set being the negative of the other. Then, we associate the oscillators of type a with one set and oscillators of type b with the other. It is in fact useful to write the expression for energy as follows

$$E = \epsilon_0 V \sum_{\vec{k}} \left(\dot{\vec{a}}^2_{\vec{k}} + \omega^2_{\vec{k}}\, \vec{a}^2_{\vec{k}} \right) \tag{8.38}$$

where we identify $\vec{a}_{-\vec{k}}$ with $\vec{b}_{\vec{k}}$ (and don't assume that $\vec{a}_{-\vec{k}} = \vec{a}_{\vec{k}}$). To make the identification with a system of harmonic oscillators complete, we need to resolve vectors $\vec{a}_{\vec{k}}$ into scalar components. Here, we use the gauge condition (8.20) which is equivalent to the condition

$$\vec{k} \cdot \vec{a}_{\vec{k}} = 0 \tag{8.39}$$

Given a plane wave mode with wave vector $\vec{k}$, we take two orthogonal unit vectors $\hat{\epsilon}_{\vec{k},1}$ and $\hat{\epsilon}_{\vec{k},2}$, such that $\hat{\epsilon}_{\vec{k},1} \cdot \hat{\epsilon}_{\vec{k},2} = \vec{k} \cdot \hat{\epsilon}_{\vec{k},1} = \vec{k} \cdot \hat{\epsilon}_{\vec{k},2} = 0$. Then, the set $\hat{\epsilon}_{\vec{k},1}, \hat{\epsilon}_{\vec{k},2}, \hat{k}$ forms an orthonormal set. Then, given (8.39), we can resolve the vector $\vec{a}_{\vec{k}}$ as follows

$$\vec{a}_{\vec{k}} = a_{\vec{k},1}\, \hat{\epsilon}_{\vec{k},1} + a_{\vec{k},2}\, \hat{\epsilon}_{\vec{k},2} \tag{8.40}$$

where $a_{\vec{k},\mu}; \mu = 1, 2$ are scalar functions of time. Then, the expression for energy reduces to

$$E = \epsilon_0 V \sum_{\mu=1}^{2} \sum_{\vec{k}} \left(\dot{a}^2_{\vec{k},\mu} + \omega^2_{\vec{k}}\, a^2_{\vec{k},\mu} \right) \tag{8.41}$$

We can absorb the factor $\epsilon_0 V$ into the definition of the functions $a_{\vec{k},\mu}$ (and also introduce a factor of 1/2, such that finally, the energy has the form

$$E = \sum_{\mu=1}^{2} \sum_{\vec{k}} \left(\frac{1}{2} \dot{a}^2_{\vec{k},\mu} + \frac{1}{2} \omega^2_{\vec{k}}\, a^2_{\vec{k},\mu} \right) \tag{8.42}$$

Now, the identification is complete. Electromagnetic fileds enclosed in a closed region of volume V can be thought of as an infinite collection of harmonic oscillators, two oscillators associated with each propagation vector $\vec{k}$ in the plane wave expansion of the field configuration: one oscillator for each independent direction of polarization. Further, both oscilaators corresponding to a given $\vec{k}$ have frequency $\omega_{\vec{k}} = c\left|\vec{k}\right|$.

Given this interpretation, let us compute the mean energy of a system of elctromagnetic fields enclosed in volume V. It follows from eqn.(8.42) that the mean energy of the system is the sum of mean energies of the individual harmonic oscillators. The energy of each harmonic oscillator is of the form

$$E_{\vec{k},\mu} = \frac{1}{2}\, p^2_{\vec{k},\mu} + \frac{1}{2}\, \omega^2_{\vec{k}} a^2_{\vec{k},\mu} \tag{8.43}$$

where $p_{\vec{k},\mu} = \dot{a}_{\vec{k},\mu}$ is the momentum of the oscillator. Then, the mean energy of the oscillator equals the sum of the mean kinetic energy and the mean potential energy. However, each is quadratic in the corresponding momentum/coordinate. Therefore, it follows from the principle of equipartition (section 5.5) that the mean kinetic and potential energy are each equal to $k_B T/2$. Therefore, the mean energy of each oscillator is $k_B T$. Then, the mean energy of electromagnetic fields at temperature T equals

$$\overline{E} = k_B T \times \text{number of oscillators} \tag{8.44}$$

Let us compute the total number of oscillators. There are two oscillators for every wave mode with wave vector $\vec{k}$, given by eqn.(8.13)

$$\vec{k} = \frac{2\pi l}{L}\ \hat{i} + \frac{2\pi m}{L}\ \hat{j} + \frac{2\pi n}{L}\ \hat{k} \tag{8.45}$$

where l, m, n take integer values. Since there is an infinite number of such modes, there is an infinity of oscillators, implying that the mean electromagnetic energy is infinite! This is clearly unphysical. How does one interpret this? Further, what is the physical interpretation of these oscillators? It turns out that quantum mechanics provides a very simple interpretation.

8.2 Quantum Theory of Radiation

Since quantum mechanics is believed to be the fundamental way to describe the physical world, in principle, all systems are quantum mechanical. Often, though, classical physics is an excellent approximation. We saw in section 5.4 that though it sufficed to describe translational degrees of freedom of atoms and molecules classically, to describe electronic, vibrational and rotational dynamics, it was essential to use quantum mechanics. So, let us try to quantise the system of harmonic oscillators that a system of electromagnetic fields seems to consist of. That is, we consider electromagnetic fields as a quantum system. This is equivalent to treating each term in the energy expression (8.42) as an operator, the Hamiltonian operator corresponding to a given mode $\vec{k}$ and its associated polarisation

$$\hat{H} = \sum_{\mu=1}^{2} \sum_{\vec{k}} \hat{H}(\vec{k}, \mu) \tag{8.46}$$

where $\hat{H}$ is the Hamiltonian operator for the entire system and $\hat{H}(\vec{k}, \mu)$ is the Hamiltonian operator for the harmonic oscillator corresponding to mode $\vec{k}$ and polarisation μ which has angular frequency $\omega_{\vec{k}} = c\left|\vec{k}\right|$. We know that the eigenvectors of $\hat{H}(\vec{k}, \mu)$ (quantum states of well-defined energy) are labelled by natural numbers $n_{\vec{k},\mu}$, corresponding to eigenvalues $(n_{\vec{k},\mu} + 1/2)\hbar\omega_{\vec{k}}$

$$\hat{H}(\vec{k}, \mu)\left|n_{\vec{k},\mu}\right\rangle = \left(n_{\vec{k},\mu} + \frac{1}{2}\right)\hbar\omega_{\vec{k}}\ \left|n_{\vec{k},\mu}\right\rangle \tag{8.47}$$

Since the total Hamiltonian $\hat{H}$ is the sum of Hamiltonians $\hat{H}(\vec{k}, \mu)$, its eigenvectors are just direct products of eigenvectors of Hamiltonians $\hat{H}(\vec{k}, \mu)$. In other words, these eigenvectors correspond to all these oscllators being in some specific energy states. Then, if the oscillators are in quantum states $\left|n_{\vec{k}_1,\mu_1}\right\rangle, \left|n_{\vec{k}_2,\mu_2}\right\rangle, \left|n_{\vec{k}_3,\mu_3}\right\rangle, ...$ then the system is a quantum state which is the sumultaneous eigenstate of the different oscillator Hamiltonians. This eigenstate is represented as $\left|n_{\vec{k}_1,\mu_1}, n_{\vec{k}_2,\mu_2}, n_{\vec{k}_3,\mu_3}, ...\right\rangle$ with eigenvalue

$$E_{n_{\vec{k}_1,\mu_1}, n_{\vec{k}_2,\mu_2}, n_{\vec{k}_3,\mu_3},...} = \left(n_{\vec{k}_1,\mu_1} + \frac{1}{2}\right)\hbar\omega_{\vec{k}_1} + \left(n_{\vec{k}_2,\mu_2} + \frac{1}{2}\right)\hbar\omega_{\vec{k}_2} + \left(n_{\vec{k}_3,\mu_3} + \frac{1}{2}\right)\hbar\omega_{\vec{k}_3} + ... \tag{8.48}$$

The lowest energy state of the electromagnetic field corresponds to all the oscillators being in their respective ground states, with $n_{\vec{k},\mu} = 0\ \forall\ \vec{k}, \mu$. This energy state, unfortunately, has infinite energy, as is clear from eqn.(8.48). The way around this problem is to redefine the 'zero' of energy such that the ground state energy is zero. This of course involves subtracting an infinite amount of energy from (8.48), and is clearly problematic. We shall not pursue this further here, but it needs to be mentioned that the quantum mechanical description of electromagnetic fields is an example of a *Quantum Field Theory*, and such theories are plagued with problems of infinities. However, there is a consistent way of resolving these problems, but this is not the place for it.

Then, we find that the quantum mechanical description of electromagnetic fields gives rise to quantum states which describe ground, first, second...excited states of a set of harmonic oscillators, with the energy of the system characterised by quantum numbers $n_{\vec{k},\mu}$ such that the energy of the system is

$$E\left\{n_{\vec{k},\mu}\right\} = \sum_{\mu=1}^{2}\sum_{\vec{k}} n_{\vec{k},\mu}\ \hbar\omega_{\vec{k}} \tag{8.49}$$

The gound state of the system corresponds to all the oscillators in their ground state, and has zero energy (because of the infinite subtraction carried out). The first excited state has energy $\hbar\omega_{\vec{k}}$ where $\left|\vec{k}\right| = 2\pi/L$, and corresponds to one oscillator, with frequency $\omega_{\vec{k}}$ in an excited state. In fact, there are six degenerate ground states, corresponding to $(l = 1, m = 0, n = 0)$, $(l = 0, m = 1, n = 0)$, $(l = 0, m = 0, n = 1)$ (there are two polarisations for each such state, which makes it six states). However, what exactly is the interpretation of these oscillator quantum states? To discover it, we compute the momentum of the electromagnetic fields. The momentum of a system of electromagnetic fields is proportional to the integral of the Poynting vector over all space. We shall not pursue the calculation here, but is is not difficult to show that if we compute this integral, it can be written, once again, in terms of our harmonic oscillator variables. Once we quantise the resulting expression, this momentum vector becomes an operator, whose eigenvectors are just the energy eigenvectors $\left|n_{\vec{k}_1,\mu_1}, n_{\vec{k}_2,\mu_2}, n_{\vec{k}_3,\mu_3}, ...\right\rangle$ with eigenvalues

$$\vec{P}\left\{n_{\vec{k},\mu}\right\} = \sum_{\mu=1}^{2}\sum_{\vec{k}} n_{\vec{k},\mu}\ \hbar\vec{k} \tag{8.50}$$

The ground state has zero momentum. but, consider a quantum state corresponding to $n_{\vec{k},\mu} = 1$ for some $\vec{k}, \mu$ (and zero for all others. This state has energy and momentum

$$\begin{aligned} E_{\vec{k},\mu} &= \hbar\ \omega_{\vec{k}} \\ \vec{P}_{\vec{k},\mu} &= \hbar\ \vec{k} \end{aligned} \tag{8.51}$$

Given that $\omega_{\vec{k}} = c\vec{k}$, it follows that

$$E_{\vec{k},\mu} = \left|\vec{P}_{\vec{k},\mu}\right|\ c \tag{8.52}$$

Which is the relation between energy and momentum of a massless particle (a particle moving with the speed of light). Then, the most natural interpretation of such a quantum state is that it describes a *photon* with momentum $p = \hbar\ \vec{k}$ and energy $E = \hbar\ \omega_{\vec{k}} = h\ \nu_{\vec{k}}$ where $\nu = \omega/2\pi$ is the frequency of the 'oscillator', now identified as the frequancy of the photon. Next, consider a state with $n_{\vec{k}_1,\mu_1} = 1, n_{\vec{k}_2,\mu_2} = 1$ for some $\vec{k}_1, \vec{k}_2, \mu_1, \mu_2$ and all other quantum numbers zero. This state has energy $E = h\ \nu_{\vec{k}_1} + h\ \nu_{\vec{k}_2}$ and momentum $\vec{P} = \hbar\ \vec{k}_1 + \hbar\ \vec{k}_2$. This state is interpreted as a state with *two* photons, one with energy/momentum $h\nu_{\vec{k}_1}/\hbar\vec{k}_1$ and the other with energy/momentum $h\nu_{\vec{k}_2}/\hbar\vec{k}_2$. Then, the state $\left|n_{\vec{k}_1,\mu_1}, n_{\vec{k}_2,\mu_2}, n_{\vec{k}_3,\mu_3}, ...\right\rangle$ is one corresponding to $n_{\vec{k}_1,\mu_1}$ photons with energy $h\nu_{\vec{k}_1}$, momentum $\hbar\vec{k}_1$ and polarisation μ_1, $n_{\vec{k}_2,\mu_2}$ photons with energy $h\nu_{\vec{k}_2}$ and momentum $\hbar\vec{k}_2$, and so on. We now need to adress the 'polarisation' of a photon and what physical quantity it corresponds to. It turns out that the polarisation of a photon is a measure of its quantum spin. Just as an electron has two intrinsic spin states, so does a photon. These spin states have spin $\pm\hbar$. In case of an electron, the spin can have components $\pm\hbar/2$ along any direction in space. However, in case of a photon, the component of spin is well-defined *only along the direction of momentum.* The spin of a photon can be parallel or antiparallel to the direction of momentum, with values $\pm\hbar$ (plus for direction along momentum and minus for direction opposite to momentum). Then, the polarisation of a photon is related to its spin state. However, the polarisation vectors we constructed are *not* the ones corresponding to these spin values. Instead, the spin states are linear superpositions of these two polarisation states. From now on, we will take the label μ to stand for the two spin states.

8.3 The Planck Distribution

Now, we turn back to the statistical distribution of electromagnetic fields at temperature T. The microstates of the system are quantum states, each specified by the number of photons corresponding to wave vectors $\vec{k}$ (a measure of momentum) given by eqn.(8.13). The probability distribution for these microstates is just the canonical distribution (4.28)

$$P\left\{n_{\vec{k},\mu}\right\} = \frac{1}{Z} e^{-\beta E\left\{n_{\vec{k},\mu}\right\}} \tag{8.53}$$

with

$$Z = \sum_{\left\{n_{\vec{k},\mu}\right\}} e^{-\beta E\left\{n_{\vec{k},\mu}\right\}} \tag{8.54}$$

where $E\left\{n_{\vec{k},\mu}\right\}$ is given by (8.49). The partition function is computed to be

$$\begin{aligned} Z &= \sum_{\left\{n_{\vec{k},\mu}\right\}} e^{-\beta E\left\{n_{\vec{k},\mu}\right\}} \\ &= \sum_{n_{\vec{k}_1,\mu_1}, n_{\vec{k}_2,\mu_2},..} e^{-\beta\left(n_{\vec{k}_1,\mu_1} h\nu_{\vec{k}_1} + n_{\vec{k}_2,\mu_2} h\nu_{\vec{k}_2} + ...\right)} \\ &= \sum_{n_{\vec{k}_1,\mu_1}} \sum_{n_{\vec{k}_1,\mu_1}} ...e^{-\beta\left(n_{\vec{k}_1,\mu_1} h\nu_{\vec{k}_1}\right)} e^{-\beta\left(n_{\vec{k}_2,\mu_2} h\nu_{\vec{k}_2}\right)} ... \\ &= \Pi_{\vec{k},\mu} Z_{\vec{k},\mu} \end{aligned} \tag{8.55}$$

where

$$\begin{aligned} Z_{\vec{k},\mu} &= \sum_{n_{\vec{k},\mu}=0}^{\infty} e^{-\beta n_{\vec{k},\mu} h\nu_{\vec{k}}} \\ &= \sum_{n_{\vec{k},\mu}=0}^{\infty} \left(e^{-\beta h\nu_{\vec{k}}}\right)^{n_{\vec{k},\mu}} \\ &= \frac{1}{1 - e^{-\beta h\nu_{\vec{k}}}} \end{aligned} \tag{8.56}$$

For a given $\vec{k}$, for each polarisation $\mu = 1, 2$, the partition function gives the same value. Since the total partition function involves products od partition functions corresponding to different $\vec{k}$ and μ, it follows that

$$\begin{aligned} Z &= \Pi_{\vec{k},\mu} \frac{1}{1 - e^{-\beta h\nu_{\vec{k}}}} \\ &= \Pi_{\vec{k}} \left(\frac{1}{1 - e^{-\beta h\nu_{\vec{k}}}}\right)^2 \end{aligned} \tag{8.57}$$

where for each $\vec{k}$, the product of the partition functions for the two polarisation states is simply equal to the square of the partition function for any one polarisation.

At this point, we need to compute the Helmholtz function for the system, from which all equilibrium properties follow. This involves computing the logarithm of the partition function

$$\ln Z = -2 \sum_{\vec{k}} \ln\left(1 - e^{-\beta h\nu_{\vec{k}}}\right) \tag{8.58}$$

It is instructive, however, to directly compute the mean energy of the system through the partition function (eqn.(4.35))

$$\begin{aligned} \overline{E} &= -\frac{\partial}{\partial\beta}\ln Z \\ &= 2\sum_{\vec{k}} \frac{h\nu_{\vec{k}} e^{-\beta h\nu_{\vec{k}}}}{1-e^{-\beta h\nu_{\vec{k}}}} \\ &= 2\sum_{\vec{k}} \frac{h\nu_{\vec{k}}}{e^{\beta h\nu_{\vec{k}}}-1} \end{aligned} \tag{8.59}$$

The above equation has a simple interpretation: the mean energy of a system of electromagnetic fileds in equilibrium is the sum over mean energies corresponding to different modes $\vec{k}$. The factor of two arises because of photon spin which has two states, and the energy is independent of spin. At this point, we need to compute the sum over modes $\vec{k}$. The mean energy associated with any mode depends on the frequency of the mode. Then, it is natural to convert the sum over modes to sum over frequencies. Let us define a function $g(\nu)$, termed 'density of states', as the number of modes lying in frequency range ν and $\nu + d\nu$ per unit frequency range. Further, let $N(\nu)$ be the total number of modes upto frequency ν. Then

$$N(\nu) = \int_0^\nu d\nu g(\nu) \tag{8.60}$$

Then, the number of modes in frequency range ν and $\nu + d\nu$ is

$$dN(\nu) = g(\nu)d\nu \tag{8.61}$$

The total number of modes upto frequency ν is given by the number of sets of integers l, m, n which satisfy the inequality

$$|k| \le \frac{2\pi}{c}\nu \tag{8.62}$$

where $\vec{k}$ is given by eqn.(8.13) Then, the number of such modes is

$$N(\nu) = 2\sum_{l,m,n} \Delta l \Delta m \Delta n \tag{8.63}$$

where $\Delta l = \Delta m = \Delta n = 1$ and the values of l, m, n are subject to constraint (8.62). Let $\lambda = c/\nu$. This is the wavelength upto which the total number of modes is to be computed. The number of modes can be written as

$$\begin{aligned} N(\nu) &= 2\left(\frac{L}{\lambda}\right)^3 \sum_{l,m,n} \left(\frac{\lambda\Delta l}{L}\right)\left(\frac{\lambda\Delta m}{L}\right)\left(\frac{\lambda\Delta n}{L}\right) \\ &= 2V\frac{\nu^3}{c^3}\sum_{\alpha,\beta,\gamma} \Delta\alpha\Delta\beta\Delta\gamma \end{aligned} \tag{8.64}$$

where $\alpha = \lambda l/L, \beta = \lambda m/L$ and $\gamma = \lambda n/L$. So long as $\lambda << L$, we can approximate discrete changes in α, β, γ by continuous changes. Assuming this to be true, we get

$$N(\nu) = 2V\frac{\nu^3}{c^3}\int d\alpha d\beta d\gamma \tag{8.65}$$

The constraint (8.62) translates $\alpha^2 + \beta^2 + \gamma^2 \le 1$. Then, the integral over α, β, γ is just the volume of a unit sphere, $4\pi/3$. This gives

$$N(\nu) = \frac{8\pi V}{3c^3}\nu^3 \tag{8.66}$$

Then, the density of states is

$$\begin{aligned} g(\nu) &= \frac{dN(\nu)}{d\nu} \\ &= \frac{8\pi V}{c^3}\nu^2 \end{aligned} \tag{8.67}$$

The sum over modes can now be replaced by an integral over ν

$$2\sum_{\vec{k}} f(\nu) \longrightarrow \int_0^\infty d\nu g(\nu) f(\nu) \tag{8.68}$$

for any function $f(\nu)$ of frequency, assuming the integral converges.

We can now compute the mean energy of electromagnetic system at temperature T

$$\begin{aligned} \overline{E} &= 2\sum_{\vec{k}} \frac{h\nu_{\vec{k}}}{e^{\beta h\nu_{\vec{k}}}-1} \\ &= \frac{8\pi Vh}{c^3}\int_0^\infty d\nu \left(\frac{\nu^3}{e^{\beta h\nu}-1}\right) \end{aligned} \tag{8.69}$$

Since this energy depends on the volume confining the electromagnetic radiation, it is more useful to study the energy *density*

$$\begin{aligned} \mathcal{U} &= \frac{\overline{E}}{V} \\ &= \frac{8\pi h}{c^3}\int_0^\infty d\nu \left(\frac{\nu^3}{e^{\beta h\nu}-1}\right) \end{aligned} \tag{8.70}$$

Let us analyse the integrand. Equation(8.70) tells us that the energy of electromagnetic radiation at temperature T is spread over all frequencies, with the energy (density) in the range ν and $\nu + d\nu$ given by

$$\mathcal{U}(\nu)d\nu = \frac{8\pi h}{c^3}\left(\frac{\nu^3}{e^{\beta h\nu}-1}\right)d\nu \tag{8.71}$$

The energy distribution function $\mathcal{U}(\nu)$ describes the famous Planck distribution.

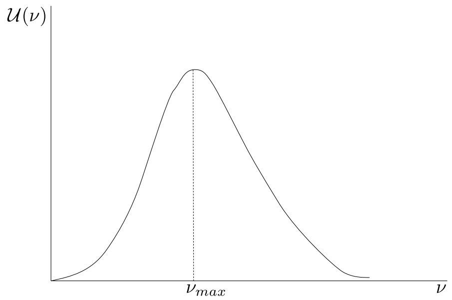


Figure 8.1: The Planck distribution

The energy peaks at a specific frequency, which corresponds to the maximum of the function $\mathcal{U}(\nu)$. This frequency ν_{max} satisfies

$$h\nu_{max} \simeq 2.82 k_B T \tag{8.72}$$

Equivalently, the wavelength corresponding to maximum energy is given by

$$\lambda_{max} = \frac{hc}{2.82\ T} \tag{8.73}$$

As the temperature changes, the position of the peak of the curve shifts. With increase in temperature, the peak shifts to higher frequency/lower wavelength.

It is instructive to take the classical limit of equation (8.71). This is obtained in the limit $h \to 0$. In this limit, we can approximate the denominator of the integrand as

$$e^{\beta h\nu} - 1 \simeq \beta h\nu \tag{8.74}$$

Substituting in eqn.(8.71), we get the classical limit

$$\mathcal{U}(\nu)d\nu \to \frac{8\pi}{c^3}\nu^2 d\nu\ k_B T \tag{8.75}$$

We see from eqn.(8.67) that we can write the above eequation in the form

$$E(\nu)d\nu = g(\nu)d\nu\ k_B T \tag{8.76}$$

This equation has a simple interpretation: Each mode has the same energy, $k_B T$, and the total energy in frequency range ν and $\nu + d\nu$ is given by the above expression. This is just what we had observed in eqn.(8.44) when we had attempted to use a classical description of electromagnetic phenomena.

How do we verify eqn.(8.71)? To verify, we need to do a spectral analysis of the electromagnetic energy. A physical quantity direcly related to energy density and more accessible experimentally is the *intensity* of radiation. In electrodynamics, the physical quantity representing intensity of electromagnetic radiation is the *Poynting vector*

$$\vec{S} = \epsilon_0 c^2 \vec{E} \times \vec{B} \tag{8.77}$$

A fourier decomposition similar to (8.18) gives

$$\vec{S} = \sum_{\vec{k}} \vec{S}_{\vec{k}} \tag{8.78}$$

where

$$\begin{aligned}
\vec{S}_{\vec{k}} &= \epsilon_0 c^2 \vec{E}_{\vec{k}} \times \vec{B}_{\vec{k}} \\
&= \epsilon_0 c^2 \left|\vec{E}_{\vec{k}}\right| \left|\vec{B}_{\vec{k}}\right| \hat{k} \\
&= \epsilon_0 c \vec{E}^2_{\vec{k}}\ \hat{k}
\end{aligned} \tag{8.79}$$

where $\hat{k}$ is a unit vector along $\vec{k}$. The energy density associated with mode $\vec{k}$ is

$$\begin{aligned}
\mathcal{U}_{\vec{k}} &= \frac{\epsilon_0}{2}\vec{E}^2_{\vec{k}} + \frac{1}{2\mu_0}\vec{B}^2_{\vec{k}} \\
&= \epsilon_0 \vec{E}^2_{\vec{k}}
\end{aligned} \tag{8.80}$$

Clearly, the intensity of radiation corresponding to a given $\vec{k}$ is proportional to the energy density corresponding to that $\vec{k}$

$$\left|\vec{S}_{\vec{k}}\right| = c\ \mathcal{U}_{\vec{k}} \tag{8.81}$$

Say, we have an object at temperature T which is emitting electromagnetic radiation from its surface. Let us ask: how much electromagnetic energy per unit area per unit time does it emit from its surface in frequency range ν and $\nu + d\nu$? This is the same as the intensity of radiation emitted by the object. The flux of radiation through the surface corresponding to mode $\vec{k}$ is given by

$$\begin{aligned}\Phi_{\vec{k}} &= \vec{S}_{\vec{k}} \cdot \hat{n} \\ &= \left|\vec{S}_{\vec{k}}\right| \cos\theta \end{aligned} \tag{8.82}$$

where $\hat{n}$ is a unit normal to the surface at the point from which the radiation is emitted

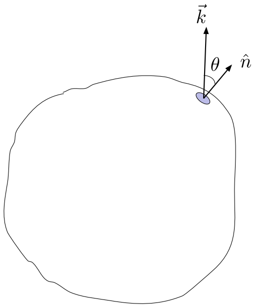


Figure 8.2: Radiation flux for mode $\vec{k}$

To compute the flux corresponding to frequency ν, we need to integrate the above expression over half the solid angle of 4π, for $\left|\vec{k}\right| = 2\pi\nu/c$. This is given by

$$\begin{aligned}\Phi_\nu &= \frac{\left|\vec{S}_{\vec{k}}\right|}{4\pi} \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta \sin\theta\cos\theta \\ &= \frac{1}{4}\left|\vec{S}_{\vec{k}}\right| \\ &= \frac{c}{4}\,\mathcal{U}_{\vec{k}} \end{aligned} \tag{8.83}$$

Then the intensity of radiation in frequency range ν and $\nu + d\nu$ is

$$I(\nu)d\nu = \frac{c}{4}\,\mathcal{U}(\nu)d\nu \tag{8.84}$$

Clearly, the intensity of radiation is a direct measure of the energy density. Let us now compute the total intensity emitted from the surface of the object over the entire spectrum. This is obtained by integrating the above expression over all frequencies

$$\begin{aligned} I &= \int_0^\infty d\nu I(\nu) \\ &= \frac{c}{4}\left(\frac{8\pi h}{c^3}\right)\int_0^\infty d\nu \left(\frac{\nu^3}{e^{\beta h\nu}-1}\right) \end{aligned} \tag{8.85}$$

Let $\beta h\nu = x$. Then

$$I = \frac{c}{4}\left(\frac{8\pi h}{c^3}\right)\frac{1}{\beta^4 h^4}\int_0^\infty dx \frac{x^3}{e^x - 1} \tag{8.86}$$

The integral over x can be computed as an infinite series

$$\begin{aligned} J &= \int_0^\infty dx\ \frac{x^3}{e^x - 1} \\ &= \int_0^\infty dx\ \frac{x^3 e^{-x}}{1 - e^{-x}} \\ &= \int_0^\infty dx\ x^3 e^{-x} \sum_{n=0}^\infty e^{-nx} \\ &= \sum_{n=0}^\infty \int_0^\infty dx\ x^3 e^{-(n+1)x} \\ &= \sum_{n=1}^\infty \int_0^\infty dx\ x^3 e^{-nx} \\ &= \sum_{n=1}^\infty \frac{1}{n^4} \int_0^\infty dt\ t^3 e^{-t} \\ &= \sum_{n=1}^\infty \frac{1}{n^4} \Gamma(4) \\ &= 6\ \zeta(4) \\ &= \frac{\pi^4}{15} \end{aligned} \tag{8.87}$$

where ζ is the Riemann Zeta function. Substituting the in intensity expression, we finally get

$$\begin{aligned} I &= \frac{c}{4}\left(\frac{8\pi h}{c^3}\right)\frac{1}{\beta^4 h^4}\frac{\pi^4}{15} \\ &= \sigma T^4 \end{aligned} \tag{8.88}$$

where

$$\sigma = \frac{2\pi^5 k_B^4}{15 h^3 c^2} \tag{8.89}$$

Equation(8.88) is known as *Stefan's Law*. The constant σ is *Stefan's constant* and is numerically equal to (approximately) $5.67 \times 10^{-8} W/m^2/K^4$. Given eqns.(8.88) and (8.84), we can compute the total energy density oif electromagnetic radiation

$$\mathcal{U} = \left(\frac{4}{c}\right)\sigma T^4 \tag{8.90}$$

Since photons carry momentum, radiation exerts pressure on surfaces which absorb/reflect it. Let us compute the radiation pressure exerted on a perfectly absorbing/perfectly reflecting surface. We start with the relation between energy and momentum of a photon. Given a photon with energy E, the magnitude of its momentum is given by

$$|\vec{p}| = \frac{E}{c} \tag{8.91}$$

Then, it follows that a mode corresponding to propagation vector $\vec{k}$ is associated with a *momentum density* given by

$$\begin{aligned} \vec{\mathcal{P}}_{\vec{k}} &= \frac{\mathcal{U}_{\vec{k}}}{c}\ \hat{k} \\ &= \frac{1}{c^2}\vec{S}_{\vec{k}} \end{aligned} \tag{8.92}$$

where $\mathcal{U}_{\vec{k}}$ is the energy density associated with this mode and $\vec{S}_{\vec{k}}$ is the corresponding Poynting vector. Now, let us consider a surface on which this mode is incident. Let us assume that this surface perfectly absorbs all the radiation falling on it. The momentum flowing per unit time per unit area across a surface perpendicular to $\vec{k}$ equals $c\vec{\mathcal{P}}_{\vec{k}} = \vec{S}_{\vec{k}}/c$ (since the momentum is flowing with speed c). Then, if θ is the angle between the normal to the surface and the propagation vector $\vec{k}$, the magnitude of the momentum absorbed by the surface per unit time per unit area equals $(1/c)\vec{S}_{\vec{k}} \cdot \hat{k} = (1/c)\left|\vec{S}_{\vec{k}}\right| \cos\theta = \mathcal{U}_{\vec{k}} \cos\theta$. The pressure exerted on the surface due to this mode is the normal component of this momentum flow per unit time per unit area, and is given by

$$\begin{aligned} P_{\vec{k}} &= \mathcal{U}_{\vec{k}} \cos\theta \times \cos\theta \\ &= \mathcal{U}_{\vec{k}} \cos^2\theta \end{aligned} \tag{8.93}$$

Had the surface been perfectly reflecting, the pressure would be twice this much. The total pressure exerted on the surface due to modes with the same magnitude $\left|\vec{k}\right|$ is obtained by integrating the above expression over half a solid angle

$$\begin{aligned} P_{|\vec{k}|} &= \frac{1}{4\pi}\int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta \sin\theta P_{\vec{k}} \\ &= \frac{1}{6}\mathcal{U}_{\vec{k}} \end{aligned} \tag{8.94}$$

Then, the pressure experienced by the surface in the frequency range ν and $\nu + d\nu$ is

$$P_{abs}(\nu)d\nu = \frac{1}{6}\mathcal{U}(\nu)d\nu \tag{8.95}$$

The total pressure due to all the frequencies is then

$$P_{abs} = \frac{1}{6}\mathcal{U} \tag{8.96}$$

where $\mathcal{U}$ is the total energy density. For a perfectly reflecting surface, the pressure is twice this much

$$P_{ref} = \frac{1}{3}\mathcal{U} \tag{8.97}$$

For radiation in thermal equilibrium, it follows from (8.90) that

$$P_{abs} = \frac{2}{3c}\sigma T^4 \tag{8.98}$$

and

$$P_{ref} = \frac{4}{3c}\sigma T^4 \tag{8.99}$$

8.4 Radiation Thermodynamics

Given radiation in equilibrium, what is the *intrinsic* pressure due to radiation? Is it the pressure measured by a perfectly absorbing or a perfectly reflecting surface? To investigate this, we need to compute the equation of state of the system. As usual, this is obtained from the Helmholtz function, (4.50)

$$\begin{aligned} F &= -k_B T \ln Z \\ &= k_B T \times 2\sum_{\vec{k}} \ln\left(1 - e^{-\beta h\nu_{\vec{k}}}\right) \\ &= k_B T \int_0^\infty d\nu g(\nu) \ln\left(1 - e^{-\beta h\nu}\right) \\ &= k_B T \frac{8\pi V}{c^3} \int_0^\infty d\nu \nu^2 \ln\left(1 - e^{-\beta h\nu}\right) \\ &= k_B T \frac{8\pi V}{c^3} \frac{(k_B T)^3}{h^3} \int_0^\infty dx x^2 \ln\left(1 - e^{-x}\right) \end{aligned} \tag{8.100}$$

On integration by parts, this reduces to

$$F = -k_B T \frac{8\pi V}{3c^3} \frac{(k_B T)^3}{h^3} \int_0^\infty dx \frac{x^3 e^{-x}}{1 - e^{-x}} \tag{8.101}$$

The integral has been computed before, and equals $\pi^4/15$. Then

$$\begin{aligned} F &= -k_B T \frac{8\pi V}{3c^3} \frac{(k_B T)^3}{h^3} \times \frac{\pi^4}{15} \\ &= -\frac{4\sigma}{3c} V T^4 \end{aligned} \tag{8.102}$$

The pressure of the system is

$$\begin{aligned} P &= -\left(\frac{\partial F}{\partial V}\right)_T \\ &= \frac{4\sigma}{3c} T^4 \\ &= \frac{1}{3}\mathcal{U} \end{aligned} \tag{8.103}$$

This is the pressure experienced by a *perfectly reflecting* surface. We can also obtain an expression for the entropy of the system

$$\begin{aligned} S &= -\left(\frac{\partial F}{\partial T}\right)_V \\ &= \frac{16\sigma}{3c} V T^3 \end{aligned} \tag{8.104}$$

Given the form of the Helmholtz function in (8.102), it is clear that it does not depend on the number of photons. This is not surprising, since the number of photons is *not* conserved, unlike the number of atoms/molecules in equilibrium with an environment. This is because photons can be absorbed and emitted in arbitrary numbers by atoms/molecules/charged particles with which they are in equilibrium. Then, it follows that the chemical potential of a system of photons in equilibrium is *zero*, since it is a measure of variation in F due to a variation in particle number. Then, for a system of photons, the canonical and grand canonical distributions are in fact equivalent.

Chapter 9

Quantum Perfect Gases

9.1 Quantum Indistinguishability

In section (5.4), we applied quantum mechanics to rotational and vibrational degrees of freedom of molecules, but treated translational degrees of freedom as classical. The justification is as follows: even though in principle all degrees of freedom are quantum mechanical and lead to discreteness of the energy spectrum, however, the ability of experiments to resolve this discreteness depends on the temperature. If the typical energy spacing of the spectrum is of the order ΔE, unless temperature is such that $k_B T \lesssim \Delta E$, this discreteness is not thermodynamically resolvable (say, through the specific heat of the system). For translational degrees of freedom, the energy spacing is of the order $\Delta E \sim \hbar^2/(2mV^{2/3})$ where V is the volume to which a particle/atom/molecule is constrained (this comes from treating the system as a 'particle in a box'). For a hydrogen atom confined to a volume $V \sim 1m^3$, this corresponds to a characteristic temperature of order $T = \Delta E/k_B \sim 10^{-17}\ {}^\circ K$ which is too low to be attained in realistic physical situations. Therefore, discreteness induced due to translational degrees is not resolvable. However, quantum effects can creep in due to a fundamentally different reson. Given a high enough density (or low enough temperature), *quantum indistinguishability* becomes important. If we view particles as fuzzy wavefunctions moving around in space, if there is appreciable overlap in these wavefunctions, it becomes impossible, even in principle, to distinguish between structurally identical particles (by structurally identical, we mean particles with the same charge, mass, spin, or any other internal degrees of freedom). We can estimate the density/temperature at which this indistinguishability is important by comparing the thermal de-Broglie wavelength of particles to their mean spacing. A reasonable estimate is given by the condition $\lambda(T) \gtrsim (V/N)^{1/3}$ where $\lambda(T) = \hbar/\sqrt{2m\pi k_B T}$. At a given temperature (which fixed $\lambda(T)$), a high enough density can result in appreciable wavefunction overlap. Conversely, at a given density, a low enough temperature can result in the same. To understand quantum indistinguishability, let us consider an experiment involving scaterring of two particles, say, electrons. Two emitters emit these particles, which we label emitters A and B. The particles are allowed to interact, after which they are detected by detectors labelled C and D

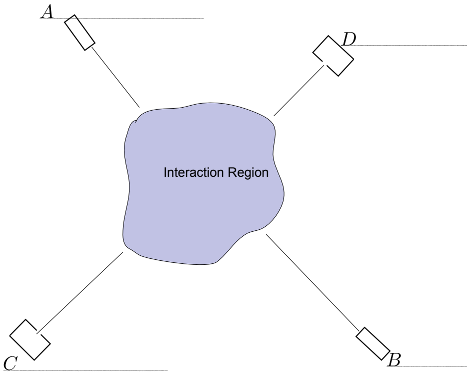


Figure 9.1: Particle Scattering

Even though the particles are intrinsically indistinguishable, we can still label them as 'particle A' and 'particle B', based on their initial positions (confined to the corresponding emitter) at the time of emission. Now, we ask the question: which particle has been detected by detector C and which has been detected by detector D? Since the particles are intrinsically identical, there seems to be no way of telling which is which. However, if the particles are described by classicla Physics, we *can*, in principle, tell. This is because they have well-defined trajectories, and they could have landed in the corresponding detectors in two alternative ways, only *one* of the alternatives having been taken, determined uniquely by their initial positions and velocities

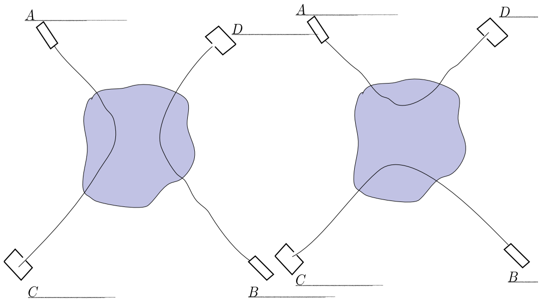


Figure 9.2: Alternative trajectories

Then, clasically, all particles are distinct, even though they may be structurally identical. However, since there are no trakectories in a quantum mechanical description, this distinguishability is lost. The particles, at the time of emission, can be assigned wavefunctions localised around detectors A and B. As the wavefunctions evolve and overlap, the particles interact. When a detection is made at detectors C and D, a particle each is detected at C and D. This measurment corresponds to a wavefunction *collapse*, which is *not* deterministic in quantum mechanics. As a result, given the final locations of the particles, it is impossible, even in principle, to extract information about which particle originated at A and which originated at B. This absence of trajectories and collapse of the wavefunction makes structurally identical particles *truly* identical in a quantum mechanical description. It in fact turns out that the vector space of states of two identical particles is completely different from that of two distinct particles. The former vector space has a smaller dimension, and can be 'carved out' of the vector space of two distinct particles. To see how this works out, say we detect a particle at point x_1 and another at point x_2. If these are distinct particles, we can label quantum states of thes particles as an ordered pair, the first slot reserved for particle of type a and the second for particle of type b. Then, given this measurement, the two possible states corresponding to this measurment are $|x_1, x_2\rangle$ and $|x_2, x_1\rangle$. To decide which of these is the correct state, we will need to measure some intrinsic property that distinguishes these particles, a measurment of which will decide between these two states. However, if the particles are structurlly identical, there is no way to distinguishing between these alternatives. Then, such a measurement results in a state which is desribed by a pair of numbers x_1 ans x_2 that is *not* ordered. Let us label this state as $|\psi(x_1, x_2)\rangle$. This state can be written as a superposition of states $|x_1, x_2\rangle$ and $|x_2, x_1\rangle$

$$|\psi(x_1, x_2)\rangle = \alpha\,|x_1, x_2\rangle + \beta\,|x_2, x_1\rangle \tag{9.1}$$

Since the pair (x_1, x_2) is not ordered, the states $|\psi(x_1, x_2)\rangle$ and $|\psi(x_2, x_1)\rangle$ are physically the same. Since physically same states can still differ by a phase, it follows that

$$|\psi(x_2, x_1)\rangle = \gamma\,|\psi(x_1, x_2)\rangle \tag{9.2}$$

where γ is a phase. Using eqn.(9.1), it follows that

$$\alpha\,|x_2, x_1\rangle + \beta\,|x_1, x_2\rangle = \gamma\alpha\,|x_1, x_2\rangle + \gamma\beta\,|x_2, x_1\rangle \tag{9.3}$$

form which we get

$$\begin{aligned} \gamma\alpha &= \beta \\ \gamma\beta &= \alpha \end{aligned} \tag{9.4}$$

which is satisfied if $\gamma^2 = 1$, that is, $\gamma = \pm 1$. If we choose $\gamma = +1$, we get the following (normalised) state, which is *symmetric* under $x_1 \leftrightarrow x_2$

$$|\psi(x_1, x_2)\rangle_S = \frac{1}{\sqrt{2}}\left(|x_1, x_2\rangle + |x_2, x_1\rangle\right) \tag{9.5}$$

On the other hand, the choice $\gamma = -1$ gives us a state *antisymmetric* under $x_1 \leftrightarrow x_2$

$$|\psi(x_1, x_2)\rangle_A = \frac{1}{\sqrt{2}}\left(|x_1, x_2\rangle - |x_2, x_1\rangle\right) \tag{9.6}$$

Clearly, $|\psi(x_2, x_1)\rangle_S = +\,|\psi(x_1, x_2)\rangle_S$ and $|\psi(x_2, x_1)\rangle_A = -\,|\psi(x_1, x_2)\rangle_S$. In general, if a measurement of an observable yields a pair of numbers (eigenvalues) a and b for two identical particles, the resulting state is either a symmetric or antisymmetric superposition

$$\begin{aligned} |\psi(\alpha, \beta)\rangle_S &= \frac{1}{\sqrt{2}}\left(|\alpha, \beta\rangle + |\beta, \alpha\rangle\right) \quad \text{Symmetric State} \\ |\psi(\alpha, \beta)\rangle_A &= \frac{1}{\sqrt{2}}\left(|\alpha, \beta\rangle - |\beta, \alpha\rangle\right) \quad \text{Antisymmetric State} \end{aligned} \tag{9.7}$$

What decides if a species of particles ‘chooses’ symmetric or antisymmetric states? It turns out that all particles with integer intrinsic spin have symmetric states and those with half-integer spins have antisymmetric states. This is a consequence of the celebrated *Spin statistics theorem* which is a natural consequence of describing nature using *both* quantum mechanics and relativity. A discussion of this theorem is *way* beyond the scope of these lectures....

Particles with symmetric states are called *Bosons* and those with antisymmetric states are called *Fermions*. Composites of elementary particles (atoms, molecules, etc.) act as bosons or Fermions depending on their composite spin. electrons, protons and netrons are Fermions (they are spin-1/2 particles). A hydrogen atom, though, is a Boson, since it consists of two Fermionic particles (giving composite integer spin).

The symmetriszation/antisymmetrization can be extended to a system of N particles which is such that one particle is in state $|\alpha_1\rangle$, another in state $|\alpha_2\rangle$,... If the system is Bosonic/Fermionic, the resulting state is a completely symmetric/antisymmetric superposition of distinct particle states of the form $|\alpha_1, \alpha_2, .., \alpha_N\rangle$

$$|\psi(\alpha_1, \alpha_2, .., \alpha_N)\rangle_S = \frac{1}{\sqrt{N!}}\left[|\alpha_1, \alpha_2, .., \alpha_N\rangle + \text{permutations}\right] \quad \text{Bosons} \tag{9.8}$$

$$|\psi(\alpha_1, \alpha_2, .., \alpha_N)\rangle_A = \frac{1}{\sqrt{N!}}\left[|\alpha_1, \alpha_2, .., \alpha_N\rangle + (-1)^p \text{permutations}\right] \quad \text{Fermions} \tag{9.9}$$

where p is the number of permutations taking a term in the superposition to the first term. For example, given three particles in states $|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle$,

$$|\psi(\alpha_1, \alpha_2, \alpha_3)\rangle_S = \frac{1}{\sqrt{3!}}\left[|\alpha_1, \alpha_2\alpha_3\rangle + |\alpha_2, \alpha_1\alpha_3\rangle + |\alpha_1, \alpha_3, \alpha_2\rangle + |\alpha_3, \alpha_2, \alpha_1\rangle\right. \tag{9.10}$$
$$+ \left.|\alpha_3, \alpha_1, \alpha_2\rangle + |\alpha_2, \alpha_3, \alpha_1\rangle\right]$$

$$|\psi(\alpha_1, \alpha_2, \alpha_3)\rangle_A = \frac{1}{\sqrt{3!}}\left[|\alpha_1, \alpha_2, \alpha_3\rangle - |\alpha_2, \alpha_1, \alpha_3\rangle - |\alpha_1, \alpha_3, \alpha_2\rangle - |\alpha_3, \alpha_2, \alpha_1\rangle\right. \tag{9.11}$$
$$+ \left.|\alpha_3, \alpha_1, \alpha_2\rangle + |\alpha_2, \alpha_3, \alpha_1\rangle\right]$$

The symmetry/antisymmetry of quantum states has remarkable observable implications. Say, we have a system of N particles in a state in which we know that one particle is in state $|\alpha_1\rangle$, another in $|\alpha_2\rangle$,..., one in state $|\alpha_N\rangle$. This state can be prepared, for instance, by measuring some observable A for the N particle system, which has eigenvalues $\alpha_1, \alpha_2, ...$ We can visualise the measurement process measuring the value of A for each particle, and the detector detecting the value of A for the particles, recording numbers $\alpha_1, \alpha_2, ..\alpha_N$. Say, we do not know if the particles are identical or distince, or if they are identical, whether they are Bosons or Fermions. To determine this, we measure another observable B which has eigenvalues $\beta_1, \beta_2, ...$ To do this, we prepare an ensemble (collection) of identical states prepared through the measurement of A, and measure B for each state. Each measurment will give a set of N number $\beta_1, \beta_2, ...\beta_N$. We collect the statistical data arising out of many such measurements on B, all carried out on the same quantum state. After collecting this data, we count the number of times all the values of β were the same, and compute the probability of this happening statistically. Let us analyze this probability assuming the particles are Bosons/Fermions. If the particles are Fermions, this probability is zero, since it is impossible to measure the same value of B for all the particles, since such a state would be manifestly symmetric. On the other hand, if the particles are bosons, the symmetric state corresponding to all particles possessing value β for B is simply

$$|\phi\rangle_S = |\beta, \beta, .., \beta\rangle \tag{9.12}$$

The probability amplitude that this happens is

$$\begin{aligned}
\langle\phi|\, \psi(\alpha_1, .., \alpha_N)\rangle_S &= \frac{1}{\sqrt{N!}}\left[\langle\beta, \beta, ..\beta|\, \alpha_1, \alpha_2, ..\alpha_N\rangle + \text{permutations}\right] \\
&= \frac{1}{\sqrt{N!}}\left[\langle\beta|\, \alpha_1\rangle \langle\beta|\, \alpha_2\rangle .. \langle\beta|\, \alpha_N\rangle + \text{permutations}\right] \\
&= \frac{1}{\sqrt{N!}} \times N! \langle\beta|\, \alpha_1\rangle \langle\beta|\, \alpha_2\rangle .. \langle\beta|\, \alpha_N\rangle \\
&= \sqrt{N!}\langle\beta|\, \alpha_1\rangle .. \langle\beta|\, \alpha_N\rangle
\end{aligned} \tag{9.13}$$

The probability of this happening is

$$\begin{aligned}
P_S(\beta) &= |\langle\phi|\, \psi(\alpha_1, .., \alpha_N)\rangle_S|^2 \\
&= N! \,|\langle\beta|\, \alpha_1\rangle|^2 \,|\langle\beta|\, \alpha_2\rangle|^2 ... |\langle\beta|\, \alpha_N\rangle|^2
\end{aligned} \tag{9.14}$$

Had these particles been all distinct, the measurement of A would have produced only *one* of the states in the superposition in eqn.(9.8). Say, that state was $|\alpha_1, \alpha_2, .., \alpha_N\rangle$. Then, a measurement of B would give the state $|\beta, \beta, .., \beta\rangle$ with probability

$$\begin{aligned}
P(\beta) &= |\langle\beta, \beta, .., \beta|\, \alpha_1, \alpha_2, ..\alpha_N\rangle|^2 \\
&= |\langle\beta|\, \alpha_1\rangle|^2 \,|\langle\beta|\, \alpha_2\rangle|^2 ... |\langle\beta|\, \alpha_N\rangle|^2
\end{aligned} \tag{9.15}$$

Then, it follows that

$$P_S(\beta) = N!\ P(\beta) \tag{9.16}$$

That is, for a system of Bosons, the probability of measuring the same value of B for all the particles is $N!$ times the corresponding probability for a system of distinct particles. Therefore, Bosons have a tendency to evolve to the same state, relative to distince particles. Of course, Fermions have a tendency to evolve *away* from similar states. This is experimentally observable.

Quantum indistinguishability can have interesting consequences. Consider, for instance, two weakly interacting identical spin-half particles in one-dimensional 'box' of length L (this is just a 'toy' example, there can be no spin in one dimension). A single particle in the box has energy eigenstates $|n\rangle$ with eigenvalues $E_n = n^2\pi^2\hbar^2/2mL^2$. Therefore, if the particles were distinguishable, the eigenstates would be of the form $|n_1, n_2\rangle$ with eigenvalues $E(n_1, n_2) = (n_1^2 + n_2^2)\pi^2\hbar^2/2mL^2$. Given that the particles are spin-half, the ground state is the following anti-symmetric state

$$|E_0\rangle = \frac{1}{\sqrt{2}} \left(\ |1\uparrow, 1\downarrow\rangle - |1\downarrow, 1\uparrow\rangle \ \right) \tag{9.17}$$

with energy $E_0 = 2\pi^2\hbar^2/2mL^2$. Here, $\uparrow$ and $\downarrow$ denote 'up' and 'down' spin states. The ground state of the system has zero net spin. Now, let us say we tried to flip one of the spins such that they became parallel. For definiteness, say we try to flip both spins up. The energy of the system does not depend on the spin of the particles. However, if the spins are parallel, the system cannot be in the ground state, since it will not be possible to have an antisymmetric state. The lowest energy of the system with both spins parallel is that corresponding to the antisymmetric state

$$|E_1\rangle = \frac{1}{\sqrt{2}} \left(\ |1\uparrow, 2\uparrow\rangle - |2\uparrow, 1\uparrow\rangle \ \right) \tag{9.18}$$

with energy $E_1 = 5\pi^2\hbar^2/2mL^2$. Then, making the spins parallel takes energy, even though the spins are not interacting. A similar effect is responsible for ferromagnetism, where the interaction between magnetic moments of atoms is too weak to account for the high Curie temperature T_C. The 'effective interaction' between neighbouring spins, arising because of the antisymmetry electron states, is responsible for the high value of T_C.

What are the statistical consequences of identical nature of particles for a system of particles in thermal equilibrium? Since all the statistical properties can be deduced from the partition function, we must focus on the difference that identical nature of particles makes to the partition function. As a simple example, consider a weakly interacting two-particle system, each particle having three energy states $|\epsilon_1\rangle, |\epsilon_2\rangle$ and $|\epsilon_3\rangle$. If the particles were distinct, the possible quantum microstates would be the energy eigenstates $|\epsilon_1, \epsilon_1\rangle, |\epsilon_2, \epsilon_2\rangle, |\epsilon_3, \epsilon_3\rangle, |\epsilon_1, \epsilon_2\rangle, |\epsilon_2, \epsilon_1\rangle, |\epsilon_1, \epsilon_3\rangle, |\epsilon_3, \epsilon_1\rangle, |\epsilon_2, \epsilon_3\rangle$ and $|\epsilon_3, \epsilon_2\rangle$. in thermal equilibrium at temperature T, these would result in the partition function

$$\begin{aligned} Z_D &= e^{-2\beta\epsilon_1} + e^{-2\beta\epsilon_2} + e^{-2\beta\epsilon_3} + 2e^{-\beta(\epsilon_1+\epsilon_2)} + 2e^{-\beta(\epsilon_2+\epsilon_3)} + 2e^{-\beta(\epsilon_1+\epsilon_3)} \\ &= \left(e^{-\beta\epsilon_1} + e^{-\beta\epsilon_2} + e^{-\beta\epsilon_3}\right)^2 \end{aligned} \tag{9.19}$$

where the subscript 'D' is for 'distinct'. Clearly, the partition function is the square of the single-particle partition function, since the energy of the system is additive and a microstate of the system is a specification of the microstate of each individual particle. If the particles are Bosons, there will be six possible (symmetric) microstates

$$\begin{aligned}
|\psi(\epsilon_1,\epsilon_1)\rangle_S &= |\epsilon_1,\epsilon_1\rangle \\
|\psi(\epsilon_2,\epsilon_2)\rangle_S &= |\epsilon_2,\epsilon_2\rangle \\
|\psi(\epsilon_3,\epsilon_3)\rangle_S &= |\epsilon_3,\epsilon_3\rangle \\
|\psi(\epsilon_1,\epsilon_2)\rangle_S &= \frac{1}{\sqrt{2}}\left(|\epsilon_1,\epsilon_2\rangle + |\epsilon_2,\epsilon_1\rangle\right) \\
|\psi(\epsilon_2,\epsilon_3)\rangle_S &= \frac{1}{\sqrt{2}}\left(|\epsilon_2,\epsilon_3\rangle + |\epsilon_3,\epsilon_2\rangle\right) \\
|\psi(\epsilon_1,\epsilon_3)\rangle_S &= \frac{1}{\sqrt{2}}\left(|\epsilon_1,\epsilon_3\rangle + |\epsilon_3,\epsilon_1\rangle\right) \qquad (9.20)
\end{aligned}$$

giving the partition function

$$Z_S = e^{-2\beta\epsilon_1} + e^{-2\beta\epsilon_2} + e^{-2\beta\epsilon_3} + e^{-\beta(\epsilon_1+\epsilon_2)} + e^{-\beta(\epsilon_2+\epsilon_3)} + e^{-\beta(\epsilon_1+\epsilon_3)} \qquad (9.21)$$

If the particles are Fermions, there will be only three (antisymmetric) microstates

$$\begin{aligned}
|\psi(\epsilon_1,\epsilon_2)\rangle_A &= \frac{1}{\sqrt{2}}\left(|\epsilon_1,\epsilon_2\rangle - |\epsilon_2,\epsilon_1\rangle\right) \\
|\psi(\epsilon_2,\epsilon_3)\rangle_A &= \frac{1}{\sqrt{2}}\left(|\epsilon_2,\epsilon_3\rangle - |\epsilon_3,\epsilon_2\rangle\right) \\
|\psi(\epsilon_1,\epsilon_3)\rangle_A &= \frac{1}{\sqrt{2}}\left(|\epsilon_1,\epsilon_3\rangle - |\epsilon_3,\epsilon_1\rangle\right) \qquad (9.22)
\end{aligned}$$

giving rise to the partition function

$$Z_A = e^{-\beta(\epsilon_1+\epsilon_2)} + e^{-\beta(\epsilon_2+\epsilon_3)} + e^{-\beta(\epsilon_1+\epsilon_3)} \qquad (9.23)$$

Clearly, the Bosonic and Fermionic partition functions are not factorizable. This is because even though the energy is additive, the particles are not distinct.

9.2 Quantum Perfect Gases

A 'perfect' gas is a system of weakly interacting identical particles in equilibrium. In the classical limit, this system is often called an 'Ideal gas'. We are interested in a quantum mechanical description of this system. Whether classical physics is a good approximation depends on the relative values of the thermal deBroglie wavelength $\lambda(T)$ and the mean spacing between particles, given by $(V/N)^{1/3}$. The classical limit of this system is given by the condition

$$\lambda(T) << \frac{1}{\rho^{1/3}} \qquad (9.24)$$

where $\rho = N/V$ is the number density. As discussed, in this limit, there is expected to be little overlap between wavefunctions of different particles and therefore they can in principle be distinguished by tracking their trajectories. At a given temperarure, if the density of the system is low enough, this condition is satisfied. Conversely, at a given density, a low enough temperature can also lead to this condition. We are interested in primarily two situations: (a) one, in which the condition (9.24) is *just* violated, such that first order quantum corrections are importanrt and (b) the complement of this condition, when the thermal deBroglie wavelength is much larger than the mean particle spacing. Either way, we need to start with the microstates of the system and compute the partition function. A convenient way to label microstates is to use the so-called 'occupation number representation'. First, we construct the 'single particle states' which are energy eigenstates of a single particle confined to the given volume (occupied by the system of particles). These are just energy states corresponding to a 'particle in a box', where we assume that the volume is in the shape of a cubical box of side L and volume $V = L^3$. The energy eigenstates $|\epsilon_\alpha\rangle$ are labelled by three integers n_x, n_y, n_z such that the energy of the particle in this state is

$$\epsilon_\alpha = \frac{\pi^2\hbar^2}{2mV^{2/3}}\left(n_x^2 + n_y^2 + n_z^2\right) \tag{9.25}$$

where n_x, n_y, n_z are whole numbers $1, 2, 3, ...$ Even though the spectrum is discrete, this discreteness is not important, since typical temperatures cannot resolve it (the natural temperature scale associated with this discreteness, as discussed, before, is far lower than experimentally accessible temperatures). However, as we shall see, there is an important exception where this discreteness *is* fundamentally important. since the particles are weakly interacting, their energies are additive and therefore the energy eigenstates of the system are symmetric/antisymmetric superpositions of 'direct product' states such as $|\epsilon_1, \epsilon_2, .., \epsilon_N\rangle$ which is a state in which the first particle is in energy state $|\epsilon_1\rangle$, the second particle in state $|\epsilon_2\rangle$.... and the N^{th} particle in state $|\epsilon_N\rangle$. However, in this superposition, since a given energy term ϵ_i appears the same number of times in each term of the superposition (only the particle that has that energy term is different in the superposition), therefore, one can lable these states by jyst specifying how many particles occupy a given single particle state. In this representation (the occupation number representation), such a state is replaced by the ordered set $(n_{\alpha_1}, n_{\alpha_2}, n_{\alpha_3}, ..., n_{\alpha_k}, ...)$ where each slot is reserved for a single particle state and the entry in the slot gives the number of particles in that state. since there is an infinite number of single particle states, this is an infinitely long list. However, the occupation numbers n_α corresponding to singla particle states $|\epsilon_\alpha\rangle$ are subject to the constaraint

$$\sum_\alpha n_\alpha = N \tag{9.26}$$

The microstates of the system are all such possible states, specified by a collection of occupation numbers $\{n_\alpha\}$ of single particle states. The energy of such a microstate is given by

$$E_{\{n_\alpha\}} = \sum_\alpha n_\alpha \epsilon_\alpha \tag{9.27}$$

The partition function of the system is

$$\begin{aligned} Z &= \sum_{\{n_\alpha\}} e^{-\beta E_{\{n_\alpha\}}} \\ &= \sum_{\{n_\alpha\}} e^{-\beta \sum_\alpha n_\alpha \epsilon_\alpha} \end{aligned} \tag{9.28}$$

Had the constraint (9.26) not been there, this partition function would have factorised into a product of partition functions for single particle states. The number constraint (9.26) makes the computation of the partition function difficult. To get around the problem, we use the Grand Canonical distribution corresponding to particles occupying volume V at temperature T and chemical potential μ, such that the mean number of particles in this distribution equals N. The Grand partition function is

$$\begin{aligned} Z_G &= \sum_{\{n_\alpha\}} e^{-\beta E_{\{n_\alpha\}} + \beta\mu N_{\{n_\alpha\}}} \\ &= \sum_{\{n_\alpha\}} e^{-\beta \sum_\alpha n_\alpha \epsilon_\alpha + \beta\mu \sum_\alpha n_\alpha} \\ &= \sum_{\{n_\alpha\}} e^{-\beta \sum_\alpha n_\alpha (\epsilon_\alpha - \mu)} \end{aligned} \tag{9.29}$$

where $N_{\{n_\alpha\}} = \sum_\alpha n_\alpha$. Since there is no restriction on $N_{\{n_\alpha\}}$ in the Grand canonical distribution, therefore, the occupation numbers $\{n_\alpha\}$ are unconstrained. This allows us to factorise the partition function

$$Z_G = \Pi_\alpha Z_\alpha \tag{9.30}$$

where

$$Z_\alpha = \sum_{n_\alpha} e^{-\beta n_\alpha(\epsilon_\alpha - \mu)} \tag{9.31}$$

For Bosons, the occupation number for a given single particle state can vary from zero to infinity. For Fermions, it can only be one or zero (apart from a multiplicative spin factor which we will take into account separately). The Boson,

$$\begin{aligned} Z_\alpha^B &= \sum_{n_\alpha=0}^{\infty} \left[e^{-\beta(\epsilon_\alpha-\mu)}\right]^{n_\alpha} \\ &= \frac{1}{1-e^{-\beta(\epsilon_\alpha-\mu)}} \end{aligned} \tag{9.32}$$

For Fermions,

$$\begin{aligned} Z_\alpha^F &= \sum_{n_\alpha=0}^{1} \left[e^{-\beta(\epsilon_\alpha-\mu)}\right]^{n_\alpha} \\ &= 1+e^{-\beta(\epsilon_\alpha-\mu)} \end{aligned} \tag{9.33}$$

The Landau Potential for the system is given by

$$\begin{aligned} \mathcal{G} &= -k_B T \ln Z_G \\ &= -k_B T \sum_\alpha \ln Z_\alpha \end{aligned} \tag{9.34}$$

It is useful the write the expression for $\mathcal{G}$ for both Bosons and Fermions together

$$\mathcal{G} = \pm\frac{1}{\beta}\sum_\alpha \ln\left(1 \mp e^{-\beta(\epsilon_\alpha-\mu)}\right) \tag{9.35}$$

where the upper sign is for Bosons and the lower sign for Fermions. The mean number of particles of the system (which is equl to N, given the equivalence of the canonical and the grand canonical distributions in the thermodynamoc limit) is give by

$$\begin{aligned} N &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{T,V} \\ &= \sum_\alpha \frac{1}{e^{\beta(\epsilon_\alpha-\mu)} \mp 1} \end{aligned} \tag{9.36}$$

Equation (9.36) is to be used to determine the chemical potential of the system as a function of temperature and density. We can compute the equation of state by (partially) differentiating $\mathcal{G}$ with respect to volume, but it is useful to take a brief detour and deduce a couple of useful results, one relating the chemical potential of a system to change in energy of the system as a result of a change in the number of particles, and the other relating the Landau potential directly to the pressure of the system.

We can always visualise the entropy of a system as a function of its energy, volume and number of particles, $S = S(E,V,N)$. Let us consider a change in the entropy of the system as a result of an infinitesimal change in the energy, number of particles and volume of the system This change is

$$\begin{aligned} dS &= \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} \\ &= \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN \end{aligned} \tag{9.37}$$

where we have used eqn.(7.56) for the chemical potential. Clearly, we can always change the energy of the system and the number of particles together, such that the entropy and the volume of the system do not change. In such a situation, given $dS = dV = 0$, the change in energy and the number of particles are related by

$$dE = \mu dN \tag{9.38}$$

which implies

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} \tag{9.39}$$

since the change involves keeping volume and entropy fixed. This is the expression we wanted for the chemical potential. Next, we visualise the energy of the system as a function of S, V and N, all extensive quantities. Since energy is also extensive, this implies

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N) \tag{9.40}$$

where λ is a scaling factor. Differentiating the above equation with respect to λ and taking $\lambda = 1$ after differentiation, we get

$$\left(\frac{\partial E}{\partial S}\right)_{N,V} S + \left(\frac{\partial E}{\partial V}\right)_{S,N} V + \left(\frac{\partial E}{\partial N}\right)_{S,V} N = E \tag{9.41}$$

Given equation(2.6), it follows that

$$\left(\frac{\partial E}{\partial S}\right)_{N,V} = T \tag{9.42}$$

Furthur, the derivative of E with respect to N is μ. We now need to interpret the significance of the derivative of energy with respect to volume (at fixed entropy and number of particles). To do this we go back to (9.37), this time keeping entropy and the number of particles fixed. This gives us

$$\left(\frac{\partial E}{\partial V}\right)_{S,N} = -P \tag{9.43}$$

Then, eqn.(9.41) reduces to

$$TS - PV + \mu N = E \tag{9.44}$$

Since $\mathcal{G} = E - TS - \mu N$, therefore we get

$$\mathcal{G} = -PV \tag{9.45}$$

Then, once the Landau potential is computed, it immediately gives the eqiation of state (after expressing μ as a function of temperature and density of the system through (9.36)). Equation (9.36) can be written as

$$N = \sum_{\alpha} n_{\alpha} \tag{9.46}$$

where

$$n_{\alpha} = \frac{1}{e^{\beta(\epsilon_{\alpha} - \mu)} \mp 1} \tag{9.47}$$

can be interpreted as the mean number of particles occupying the single particle state $|\epsilon_{\alpha}\rangle$.

If we wish to compute first (or higher order) quantum effects, we need a 'small parameter' in terms of which we can expand about the classical results. To identify the samll parameter, we need to cast the condition (9.24) as a condition on the chemical potential, through which the density appears implicitly. At a given temperature, the classical limit corresponds to low density. If there are very few particles occupying a given volume, given the infinite number of single particle states, it seems reasonable that most of the microstates will correspond to particles occupying distinct single particle states. Assuming this to be true, the mean number of particles per single particle state will be very small. Then, it seems reasonable that at sufficiently low densities, the mean occupation numbers $n_{\alpha} << 1$. To test this conjecture, we reduce this to a condition on the chemical potential. Since this condition must be true for all $|\alpha\rangle$, it follows that in the classical limit, $e^{-\beta\mu} >> 1$. Given eqn.(7.56), it is clear that the chemical potential for a classical system of particles is negative, since entropy can only increase due to an increase in the number of particles (at fixed energy and volume). Then, the exponent in $e^{-\beta\mu}$ is positive. This condition can also be written as

$$e^{\beta\mu} << 1 \tag{9.48}$$

which also identifies $e^{\beta\mu}$ as the 'small parameter' in terms of quantum corrections to physical quantities can be computed on expanding about the classical result. To check it this condition is physically resonable, we observe that in this approximation,

$$n_\alpha \simeq e^{-\beta(\epsilon_\alpha-\mu)} \tag{9.49}$$

Substituting this in the expression for the Landau potential (eqn.9.35), we get

$$\begin{aligned} \mathcal{G} &\simeq \pm\frac{1}{\beta}\sum_\alpha \ln\left(1\mp n_\alpha\right) \\ &\simeq -\frac{1}{\beta}\sum_\alpha n_\alpha \\ &= -Nk_BT \end{aligned} \tag{9.50}$$

where we have used $\ln(1\mp n_\alpha)\simeq \mp n_\alpha$. Using eqn.(9.45), we get the classical equation of state, which justifies eqn.(9.48) as the condition for the classical limit.

We now get back to computing the Landau potential (eqn.(9.35)) and the chemical potential (9.36). As discussed, since the quantum discreteness in single particle states is not important, we can relace the sum over single particle states by integrals. The sum is over quantum numbers n_x, n_y, n_z appearing in the energy eigenvalue expression (9.25), subject to constraint

$$n_x^2+n_y^2+n_z^2=\left(\frac{2mV^{2/3}}{\pi^2\hbar^2}\right)\epsilon \tag{9.51}$$

which resembles the equation of a sphere of radius

$$r=\sqrt{\left(\frac{2mV^{2/3}}{\pi^2\hbar^2}\right)\epsilon} \tag{9.52}$$

We replace the sum by integrals, and recognizing that the summands in eqn.(9.35) and (9.36) depend only on energy (and therefore sum of squares of n_x, n_y, n_z), we get, by converting the integral into a spherical integral

$$\begin{aligned} \sum_{n_x,n_y,n_z} f(\epsilon) &\longleftrightarrow \int dn_x dn_y dn_z f(\epsilon) \\ &= \frac{1}{8}4\pi\int_0^\infty dr r^2 f(\epsilon) \\ &= \frac{4\pi\sqrt{2}Vm^{3/2}}{(2\pi\hbar)^3}\int_0^\infty d\epsilon\ \epsilon^{1/2}f(\epsilon) \end{aligned} \tag{9.53}$$

We now introduce a factor recognising the intrinsic spin of the particles. If the spin is s, there are $2s+1$ spin states, increasing the single particle states by this factor. Finally, we can write

$$\sum_{n_x,n_y,n_z} f(\epsilon) \longleftrightarrow \int_0^\infty d\epsilon\ g(\epsilon)f(\epsilon) \tag{9.54}$$

where $g(\epsilon)$ is the 'density of states' (number of single particle states in the energy interval ϵ and $\epsilon+d\epsilon$ per unit interval), given by

$$g(\epsilon)=\left(\frac{4\pi\sqrt{2}(2s+1)Vm^{3/2}}{(2\pi\hbar)^3}\right)\epsilon^{1/2} \tag{9.55}$$

9.2.1 Weakly Degenerate gas

We now compute the first order quantum corrections to the Landau potential and the chemical potential, and consequently obtain the first order quantum correction to the equation of state. We take $e^{\beta\mu}$ as a small parameter in terms of which we can expand, retaining terms up to next to leading order. Up to this order, the logarithm appearing in eqn.(9.35) can be approximated by

$$\ln\left(1 \mp e^{-\beta(\epsilon_\alpha-\mu)}\right) \simeq \mp e^{-\beta(\epsilon_\alpha-\mu)} - \frac{1}{2}e^{-2\beta(\epsilon_\alpha-\mu)} \tag{9.56}$$

Then, the Landau potential reduces to

$$\begin{aligned} \mathcal{G} &= -k_BT\sum_\alpha e^{-\beta(\epsilon_\alpha-\mu)} \mp \frac{k_BT}{2}e^{2\beta\mu}\sum_\alpha e^{-2\beta\epsilon_\alpha} \\ &= \mathcal{G}_{class} \mp \frac{k_BT}{2}e^{2\beta\mu}\sum_\alpha e^{-2\beta\epsilon_\alpha} \end{aligned} \tag{9.57}$$

where $\mathcal{G}_{class}$ is the classical Landau potential, given by

$$\begin{aligned} \mathcal{G}_{class} &= -k_BTe^{\beta\mu}\sum_\alpha e^{-\beta\epsilon_\alpha} \\ &= -k_BTe^{\beta\mu}\int_0^\infty d\epsilon\; g(\epsilon)e^{-\beta\epsilon} \end{aligned} \tag{9.58}$$

Also,

$$\sum_\alpha e^{-2\beta\epsilon_\alpha} = \int_0^\infty d\epsilon\; g(\epsilon)e^{-2\beta\epsilon} \tag{9.59}$$

In the integral, we change variable from ϵ to $\epsilon' = 2\epsilon$. Since $g(\epsilon) \propto \epsilon^{1/2}$, we get

$$\begin{aligned} \sum_\alpha e^{-2\beta\epsilon_\alpha} &= \int_0^\infty \frac{d\epsilon'}{2}\,\frac{g(\epsilon')}{2^{1/2}}e^{-\beta\epsilon'} \\ &= \left(\frac{1}{2}\right)^{3/2}\int_0^\infty d\epsilon\; g(\epsilon)e^{-\beta\epsilon} \end{aligned} \tag{9.60}$$

where we have just replaced the (dummy) variable ϵ' by ϵ. The epression for the Landau potential reduces to

$$\mathcal{G} = \mathcal{G}_{class} \mp k_BTe^{2\beta\mu}\left(\frac{1}{2}\right)^{5/2}\int_0^\infty d\epsilon\; g(\epsilon)e^{-\beta\epsilon} \tag{9.61}$$

Given eqn.(9.58), we can write this as

$$\mathcal{G} = \mathcal{G}_{class}\left[1 \pm \frac{e^{\beta\mu}}{2^{5/2}}\right] \tag{9.62}$$

where

$$\begin{aligned} \mathcal{G}_{class} &= -k_BTe^{\beta\mu}\int_0^\infty d\epsilon\; g(\epsilon)e^{-\beta\epsilon} \\ &= -k_BTe^{\beta\mu}\left(\frac{4\pi\sqrt{2}(2s+1)Vm^{3/2}}{(2\pi\hbar)^3}\right)\int_0^\infty d\epsilon\; \epsilon^{1/2}e^{-\beta\epsilon} \\ &= -\frac{k_BT(2s+1)V}{\lambda^3(T)}e^{\beta\mu} \end{aligned} \tag{9.63}$$

where the integral has been computed using Gamma functions. Finally,

$$\begin{aligned} \mathcal{G} &= -\frac{k_BT(2s+1)V}{\lambda(T)^3}e^{\beta\mu}\left[1 \pm \frac{e^{\beta\mu}}{2^{5/2}}\right] \\ &= -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[e^{\beta\mu} \pm \frac{e^{2\beta\mu}}{2^{5/2}}\right] \end{aligned} \tag{9.64}$$

The number of particles is constrained through

$$\begin{aligned} N &= -\left(\frac{\partial \mathcal{G}}{\partial \mu}\right)_{T,V} \\ &= \frac{(2s+1)V}{\lambda(T)^3}\left[e^{\beta\mu} \pm \frac{e^{2\beta\mu}}{2^{3/2}}\right] \end{aligned} \tag{9.65}$$

This equation is to be used to compute $\mu(T,\rho)$. Let $e^{\beta\mu} = x$. In terms of x, the Landau potential and particle number equations reduce to

$$\mathcal{G} = -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[x \pm \frac{x^2}{2^{5/2}}\right] \tag{9.66}$$

$$N = \frac{(2s+1)V}{\lambda^3(T)}\left[x \pm \frac{x^2}{2^{3/2}}\right] \tag{9.67}$$

It is assumed that $x << 1$. The equation for the number of particles can be written as

$$x \pm \frac{x^2}{2^{3/2}} = \frac{N\lambda^3}{(2s+1)V} \tag{9.68}$$

Let $\alpha = N\lambda^3/(2s+1)V$. Clearly, $\alpha << 1$ since $x << 1$. Then, we can expand x in this small parameter. To leading order, $x = \alpha$ (ignoring x^2). Then, to next to leading order, x must be of the form

$$x = \alpha + \lambda\alpha^2 \tag{9.69}$$

where λ is to be determined. Substituting this in the equation for x (and retaining terms upto α^2 only), we get

$$\alpha + \lambda\alpha^2 \pm \frac{\alpha^2}{2^{3/2}} = \alpha \tag{9.70}$$

which gives

$$\lambda = \mp\frac{1}{2^{3/2}} \tag{9.71}$$

such that

$$x = \alpha\left[1 \mp \frac{\alpha}{2^{3/2}}\right] \tag{9.72}$$

The chemical potential can now be obtained

$$\begin{aligned} \frac{\mu}{k_BT} &= \ln x \\ &\simeq \ln\alpha + \ln\left(1 \mp \frac{\alpha}{2^{3/2}}\right) \end{aligned} \tag{9.73}$$

Then

$$\mu \approx k_BT\left[\log\alpha \mp \frac{\alpha}{2^{3/2}}\right] \tag{9.74}$$

This gives the chemical potential to first order. Substituting for x in the Landau potential, and retaining terms upto order α^2

$$\begin{aligned}
\mathcal{G} &= -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[x \pm \frac{x^2}{2^{5/2}}\right] && (9.75)\\
&= -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[\alpha \mp \frac{\alpha^2}{2^{3/2}} \pm \frac{\alpha^2}{2^{5/2}}\right] && (9.76)\\
&= -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[\alpha \mp \alpha^2\left(\frac{1}{2^{3/2}} - \frac{1}{2^{5/2}}\right)\right] && (9.77)\\
&= -\frac{k_BT(2s+1)V}{\lambda^3(T)}\alpha\left[1 \mp \frac{\alpha}{2^{5/2}}\right] && (9.78)
\end{aligned}$$

Substituting for α, we get

$$\mathcal{G} = -Nk_BT\left[1 \mp \frac{\lambda^3(T)}{2^{5/2}(2s+1)(V/N)}\right]$$

Given that $\mathcal{G} = -PV$, we get the equation of state upto first order

$$PV = Nk_BT\left[1 \mp \frac{\lambda^3(T)}{2^{5/2}(2s+1)(V/N)}\right] \tag{9.79}$$

In terms of temperature and (number) density, we get

$$P = \rho k_BT\left[1 \mp \frac{\lambda^3(T)}{2^{5/2}(2s+1)}\,\rho\right] \tag{9.80}$$

which shows that at a given temperature and density, a system of Bosons/Fermions will have a lower/higher pressure compared with a classical system.

For completeness, let us also compute the first order quantum corrections to energy and entropy. To calculate entropy, we strt with the original expression for $\mathcal{G}$ in terms of the chemical potential

$$\mathcal{G} = -\frac{k_BT(2s+1)V}{\lambda^3(T)}\left[x \pm \frac{x^2}{2^{5/2}}\right]$$

where $x = e^{\mu/k_BT}$. The entropy is given by

$$S = -\left(\frac{\partial \mathcal{G}}{\partial T}\right)_{V,\mu}$$

The derivative will involve differentiation of $\lambda(T)$ and $x = e^{\mu/k_BT}$ w.r.t. T. Given that $\lambda(T) = 2\pi\hbar/\sqrt{2\pi m k_BT}$, we get

$$\begin{aligned}
S &= -\left(\frac{\partial \mathcal{G}}{\partial T}\right)_{V,\mu}\\
&= \frac{5}{2}\frac{k_B(2s+1)V}{\lambda^3}\left[x \pm \frac{x^2}{2^{5/2}}\right] - \frac{k_B(2s+1)V}{\lambda^3}\frac{\mu}{k_BT}\left[x \pm \frac{x^2}{2^{3/2}}\right] && (9.81)
\end{aligned}$$

Substituting for x and μ (and retaining terms upto order α^2), one finally gets

$$S = S_{cl} \mp \frac{Nk_B}{4}\frac{\lambda^3(T)}{2^{3/2}(2s+1)}\,\rho \tag{9.82}$$

where S_{cl} is the classical ideal gas entropy. to calculate the mean energy of the system, it is easire to calculate the change in mean energy relative to the classical value. We start with the expression

$$E = TS - PV + \mu N \tag{9.83}$$

The change in energy is (T and V are fixed)

$$\begin{aligned}\Delta E &= T\Delta S - V\Delta P + N\Delta\mu \\ &= T\Delta S - \Delta(PV) + N\Delta\mu \qquad (9.84)\\ & \qquad (9.85)\end{aligned}$$

Since we have calculated $\mathcal{G}, S$ and μ, we have

$$\begin{aligned}\Delta S &= \mp\frac{Nk_B}{4}\frac{\lambda^3}{2^{3/2}(2s+1)}\ \rho \\ \Delta(PV) &= \mp Nk_BT\frac{\lambda^3}{2^{5/2}(2s+1)}\ \rho \\ \Delta\mu &= \mp k_BT\frac{\lambda^3}{2^{3/2}(2s+1)}\ \rho \qquad (9.86)\end{aligned}$$

Substituting, we get

$$\begin{aligned}\Delta E &= T\Delta S - \Delta(PV) + N\Delta\mu \qquad (9.87)\\ &= \mp\frac{Nk_BT}{4}\frac{\lambda^3(T)}{2^{3/2}(2s+1)}\ \rho \pm Nk_BT\frac{\lambda^3(T)}{2^{5/2}(2s+1)}\ \rho \mp Nk_BT\frac{\lambda^3(T)}{2^{3/2}(2s+1)}\ \rho \\ &= \mp\frac{3}{4}Nk_BT\frac{\lambda^3(T)}{2^{3/2}(2s+1)}\ \rho \qquad (9.88)\end{aligned}$$

Therefore

$$\begin{aligned}E &= E_{cl} + \Delta E \\ &= \frac{3}{2}Nk_BT \mp \frac{3}{4}Nk_BT\frac{\lambda^3(T)}{2^{3/2}(2s+1)}\ \rho\end{aligned}$$

9.2.2 Strongle Degenerate Fermi gas

We now analyze a system of identical particles at the other extreme, when classical physics is a very poor approximation. That is,

$$\lambda(T) >> \frac{1}{\rho^{1/3}} \qquad (9.89)$$

Let us first consider a system of Fermions. The Landau potential for the system is

$$\mathcal{G} = -\frac{1}{\beta}\sum_\alpha \ln\left(1 + e^{-\beta(\epsilon_\alpha-\mu)}\right) \qquad (9.90)$$

The chemical potential is given by the constraint

$$\begin{aligned}N &= \sum_\alpha n_\alpha \\ &= \sum_\alpha \frac{1}{e^{\beta(\epsilon_\alpha-\mu)}+1} \qquad (9.91)\end{aligned}$$

As an extreme case of condition (9.89), let us see what happens if for a fixed density ρ, we take the limit $T \longleftrightarrow 0$. In this limit, let the chemical potential approach the limit μ_0. It is easy to see from eqn.(9.91) that as $T \longrightarrow 0$, the mean single particle state occupation number has the following behaviour

$$n_\alpha = \begin{cases} 1 & \epsilon_\alpha < \mu_0 \\ 0 & \epsilon_\alpha > \mu_0 \end{cases} \qquad (9.92)$$

This can be represented graphically as follows

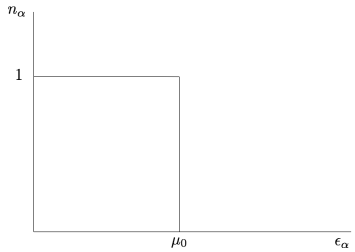


Figure 9.3: Occupation number for Fermions at absolute zero

The interpretation of eqn.(9.92) is simple: At absolute zero, all single particle states upto energy $\epsilon_F = \mu_0$ are completely filled. In essence, the entire system is in a *single* quantum state: an antisymmetric superposition of the ground state, first excited state,..so on. Clearly, the entropy of the system at absolute zero is zero, since there is a single microstate at this temperature. The highest single particle state energy ϵ_F is called *Fermi energy.* It can be computed using eqn.(9.91)

$$\begin{aligned} N &= \sum_\alpha n_\alpha \\ &= \int d\epsilon g(\epsilon) n(\epsilon) \\ &= \int_0^{\epsilon_F} d\epsilon g(\epsilon) \\ &= \left(\frac{4\pi\sqrt{2}(2s+1)Vm^{3/2}}{(2\pi\hbar)^3} \right) \int_0^{\epsilon_F} d\epsilon \ \epsilon^{1/2} \\ &= \left(\frac{4\pi\sqrt{2}(2s+1)Vm^{3/2}}{(2\pi\hbar)^3} \right) \times \frac{2}{3} \ \epsilon_F^{3/2} \end{aligned} \tag{9.93}$$

Inverting this gives the Fermi energy

$$\epsilon_F = \frac{(2\pi\hbar)^2}{2m} \left(\frac{3}{4\pi(2s+1)} \right)^{2/3} \rho^{2/3} \tag{9.94}$$

The fermi energy sets a natural temperature scale, the Fermi temperature

$$T_F = \frac{\epsilon_F}{k_B} \tag{9.95}$$

Since ϵ_F is also the chemical potential at absolute zero, we observe that $\mu_0 > 0$, which is interesting, since at least for classical systems, we expect the chemical potential to be negative, as argued in section 9.2. Clearly, quantum mechanics does not respect that. To see that chemical potential of quantum systems can be positive, it is most convenient to use the following expression for the chemical potential (eqn.(9.39))

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{S,V} \tag{9.96}$$

We can interpret the above equation as follows: the chemical potential is the change in energy of a system if one particle is added while keeping the entropy (and volume) unchanged. Let us consider a very small quantum system of two particles with single particle states with (non-degenerate) energy $0, \epsilon, 2\epsilon, 3\epsilon, 4\epsilon, ...$ Let the total energy of the system be 3ϵ. For Bosons, there are two microstates corresponding to this energy, which, in the occupation number representation, are $(1,0,1,0,0,0,..)_B$ and $(0,1,1,0,0,0,....)_B$.

The entropy of the system is then $S_B = k_B \ln 2$. If the system is Fermionic, there are once again two microstates, $(1,0,1,0,0,0,..)_F$ and $(0,1,1,0,0,0,....)_F$, again corresponding to entropy $S_F = k_B \ln 2$. Let us add one particle, without changing the energy. For Bosons, the number of microstates increases to three: $(2,0,1,0,0,0,....)_B, (1,1,1,0,0,0,.....)_B$ and $(0,3,0,0,0,0,....)_B$. This increases the entropy to $S' = k_B \ln 3$. For Fermions, there is only one three-particle microstate with total energy $3\epsilon : (1,1,1,0,0,0,.....)_F$ with entropy $S' = 0$. Now we try to exchange energy with the system to bring the entropy back to its original value, $S = k_B \ln 2$. For Bosons, this is accomplished by reducing the energy of the system by ϵ, resulting in two microstates with total energy $2\epsilon : (2,0,1,0,0,0,...)_B$ and $(1,2,0,0,0,0,...)_B$ and entropy $S = k_B \ln 2$. For Fermions, we cannot lower the energy (why?). The minimum energy we need to add is 2ϵ, resulting in two microstates with total energy $5\epsilon : (1,0,1,1,0,0,0,...)_F$ and $(1,1,0,0,1,0,0,...)_F$ with entropy $S = k_B \ln 2$. Then, the chemical potential for the Bosonic system is $\mu_B = -\epsilon$ and for the Fermionic system $\mu_F = +2\epsilon$, which is positive.

A system of Fermions at absolute zero exerts a non-zero pressure. To see this, we compute the Landau potential at absolute zero. In eqn.(9.90), as $T \longrightarrow 0$, the term $e^{-\beta(\epsilon_\alpha - \mu)}$ is zero unless $\epsilon_\alpha < \mu_0$. For $\epsilon_\alpha > \mu_0$, the term is zero and the logarithm is zero as well. Therefore, we need to sum over only those single particle states with energy less than the Fermi energy. For such states, as we take the limit $T \longrightarrow 0$, the exponential factor $e^{-\beta(\epsilon_\alpha - \mu)}$ grows without bound, such that the logarithm, in this limit, gives

$$\begin{aligned} \ln\left(1 + e^{-\beta(\epsilon_\alpha - \mu)}\right) &\simeq \ln\left(e^{-\beta(\epsilon_\alpha - \mu_0)}\right) \\ &= -\beta\left(\epsilon_\alpha - \mu_0\right) \end{aligned} \tag{9.97}$$

Then, as $T \longrightarrow 0$, the expression for the Landau potential reduces to

$$\begin{aligned} \mathcal{G} &= \sum_\alpha (\epsilon_\alpha - \mu_0) \\ &= \int_0^{\epsilon_F} d\epsilon\; g(\epsilon)(\epsilon - \epsilon_F) \\ &= \left(\frac{4\pi\sqrt{2}(2s+1)Vm^{3/2}}{(2\pi\hbar)^3}\right)\int_0^{\epsilon_F} d\epsilon\; \epsilon^{1/2}(\epsilon - \epsilon_F) \\ &= -\frac{2}{3}\left[\frac{3}{10}N\frac{(2\pi\hbar)^2}{m}\left(\frac{3}{4\pi(2s+1)}\right)^{2/3}\rho^{2/3}\right] \end{aligned} \tag{9.98}$$

Since $\mathcal{G} = -PV$, we get the equation of state

$$PV = \frac{2}{3}\left[\frac{3}{10}N\frac{(2\pi\hbar)^2}{m}\left(\frac{3}{4\pi(2s+1)}\right)^{2/3}\rho^{2/3}\right] \tag{9.99}$$

The mean energy of the system at absolute zero is easy to compute, given (9.92)

$$\begin{aligned} E &= \sum_\alpha n_\alpha \epsilon_\alpha \\ &= \int_0^{\epsilon_F} d\epsilon\; g(\epsilon)\epsilon \\ &= \frac{3}{10}N\frac{(2\pi\hbar)^2}{m}\left(\frac{3}{4\pi(2s+1)}\right)^{2/3}\rho^{2/3} \end{aligned} \tag{9.100}$$

From eqns.(9.99) and (9.100), it follows that

$$PV = \frac{2}{3}E \tag{9.101}$$

9.2.3 Electrons in a metal: a strongly degenerate Fermionic system

We now look at a physical example of a system of weakly interacting Fermions (effectively) at absolute zero. Consider electrons in a metal. Electrons in ametal are loosely bound and can propagate over the crystal lattice. Let us compute the number density of electrons in a typical metal, say, copper. For an electrons, the fspin factor $(2s+1) = 2$. The number density of electrons in copper is about $\rho \sim 8.5 \times 10^{22}$. Given the electron mass, the Fermi temperature (9.100) is about $T_F \sim 85,000^\circ K$! It is then an excellent approximation to assume that effectively, electrons in metals are close to absolute zero.

Electrons in a metal behave as if they are weakly interacting particles. This is a purely quantum effect, arising because of a combination of three reasons: (a) Even though electrons possess charge, a metal is on the whole electrically neutral. Therefore, charge on an electron is effectively 'shielded' by the positive ionic cores (b) Due to the periodicity of the metal lattice, electrons propagate effectively as 'waves' (the wavefunction is highly delocalized. These wavefunctions are called 'Bloch waves'). Because of this delocalization, any one electron cannot 'resolve' the localized charge of the ionic cores. As a result, it only sees an effective charge $+e$ distributed over the entire lattice (c) The electrons states do interact with each other, but cannot alter their quantum states by scattering off other electrons, as close to absolute zero, those states are already occupied by other electrons (being Fermions, they cannot occupy states already occupied). How can we test this (seemingly paradoxical) hypothesis, that even though electrons interact strongly through charge, they are effectively 'free' in a metal? Once again, a measurement of the heat capacity of the system can shed light on the microscopic Physics, which is not directly accessibel. Therefore, modelling electrons in ametal as a system of weakly interacting particles, we compute the heat capacity of the system at close to absolute zero. At this point, we will need to acknowledge that the system is not really at absolute zero, but at temperature T such that $T << T_F$. Then, we need to compute thermodynamic quantities as expansions in the small parameter T/T_F. Before we do a rigorous analysis, we can qualitatively estimate the behaviour of the heat capacity as a function of temperature. At absolute zero, the highest single particle state occupied by electrons has energy ϵ_F. At $T > 0$, we expect electrons to get excited to higher energies. Given $T << T_F$, how high an energy do we expect the electrons to get excited to? The available energy is of the order $k_B T$. Electrons occupying low energy states cannot absorb this energy and get excited, since the higher states are already occupied by electrons. However, electrons with energy between $\epsilon_F - k_B T$ and ϵ_F can get excited by $k_B T$. Then, qualitatively, the occupation number vs energy plot will change to the following

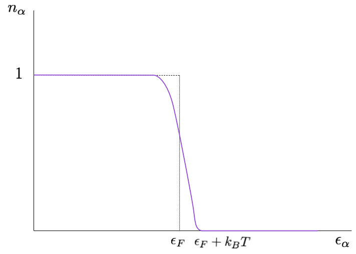


Figure 9.4: Occupation number for Fermions at $T << T_F$

Again, qualitatively, the number of electrons that will be excited will be of the order of

$$N_e \simeq N \frac{T}{T_F} \tag{9.102}$$

These are the only electrons which are 'dynamical'. Classically, the heat capacity per particle associated with translational degrees of freedom is $3/2 k_B$. Therefore, since we expect N_e electrons to participate in the dynamics, the total heat capacity of the system will be

$$\begin{aligned} C_v &\simeq \frac{3}{2}N_e k_B T \\ &= \frac{3}{2}k_B N \frac{T}{T_F} \end{aligned} \tag{9.103}$$

The molar heat capacity is then

$$c_v \simeq \left(\frac{3}{2}R\right)\frac{T}{T_F} \tag{9.104}$$

which is expected to vary linearly with temperature. Since $T << T_F$, the major contribution to c_v comes from lattice vibrations. However, at very low temperatures (T much less than the Debye temperature), the lattice contribution to c_v falls as $\sim T^3$ whereas that due to electrons falls linearly with T. Therefore, to confirm that validity of the 'free electron model', the specific heat needs to be measured at very low temperatures. At very low temperatures, linear behaviour *is* empirically observed.

Let us now do a more rigorous analysis, involving an expansion in powers of the small parameter T/T_F. We will work with a system of electrons, for which the spin factor $2s+1=2$. The Landau potential for the system is

$$\begin{aligned} \mathcal{G} &= -\frac{1}{\beta}\sum_{\alpha}\ln\left(1+e^{-\beta(\epsilon_\alpha-\mu)}\right) \\ &= \frac{-k_B T\sqrt{2}\,Vm^{3/2}}{\pi^2\hbar^3}\int_0^\infty d\epsilon\ \epsilon^{1/2}\ln\left(1+e^{-\beta(\epsilon-\mu)}\right) \end{aligned} \tag{9.105}$$

Integrating by parts, we get

$$\mathcal{G} = -\frac{2}{3}\frac{\sqrt{2}\,Vm^{3/2}}{\pi^2\hbar^3}\int_0^\infty d\epsilon\frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)}+1} \tag{9.106}$$

The constraint equation for the chemical potential is

$$\begin{aligned} N &= \sum_{\alpha}\frac{1}{e^{\beta(\epsilon_\alpha-\mu)}+1} \\ &= \frac{\sqrt{2}\,Vm^{3/2}}{\pi^2\hbar^3}\int_0^\infty d\epsilon\frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)}+1} \end{aligned} \tag{9.107}$$

In either equation, we need to compute integrals of the form

$$I = \int_0^\infty d\epsilon\frac{f(\epsilon)}{e^{\beta(\epsilon-\mu)}+1} \tag{9.108}$$

Introducing variable of integration $x=\beta(\epsilon-\mu)$

$$I = k_B T\int_{-\beta\mu}^\infty dx\ \frac{f(\mu+k_B T x)}{e^x+1} \tag{9.109}$$

The integral can be split into two integrals

$$I = k_B T\int_{-\beta\mu}^0 dx\ \frac{f(\mu+k_B T x)}{e^x+1} + k_B T\int_0^\infty dx\ \frac{f(\mu+k_B T x)}{e^x+1} \tag{9.110}$$

Using transformation $x \longrightarrow -x$ in the first integral gives

$$I = k_B T\int_0^{\beta\mu} dx\ \frac{f(\mu-k_B T x)}{e^{-x}+1} + k_B T\int_0^\infty dx\ \frac{f(\mu+k_B T x)}{e^x+1} \tag{9.111}$$

Now

$$\frac{1}{e^{-x}+1} = 1-\frac{1}{e^x+1} \tag{9.112}$$

Then

$$I = k_BT\int_0^{\beta\mu} dx f(\mu - k_BTx) - k_BT\int_0^{\beta\mu} dx\ \frac{f(\mu - k_BTx)}{1+e^x} + k_BT\int_0^{\infty} dx\ \frac{f(\mu + k_BTx)}{e^x+1} \tag{9.113}$$

In the first integral, we introduce a variable $\epsilon = \mu - k_BTx$. In the second integral, since $\beta\mu >> 1$ ($T << T_F$), we can push the upper limit to infinity. Then

$$I = \int_0^{\mu} d\epsilon\ f(\epsilon) + k_BT\int_0^{\infty} dx\left[\frac{f(\mu + k_BTx) - f(\mu - k_BTx)}{1+e^x}\right] \tag{9.114}$$

Expanding $f(\mu + k_BTx) - f(\mu - k_BTx)$ in a Taylor series gives

$$f(\mu + k_BTx) - f(\mu - k_BTx) = 2\sum_{n=0}^{\infty}\frac{f^{(2n+1)}(\mu)}{(2n+1)!}(k_BT)^{2n+1}x^{2n+1} \tag{9.115}$$

Then

$$I = \int_0^{\mu} d\epsilon\ f(\epsilon) + 2k_BT\sum_{n=0}^{\infty}\frac{f^{(2n+1)}(\mu)}{(2n+1)!}(k_BT)^{2n+1} \times \int_0^{\infty} dx\frac{x^{2n+1}}{1+e^x} \tag{9.116}$$

First few terms in the expansion are

$$I = \int_0^{\mu} d\epsilon\ f(\epsilon) + 2(k_BT)^2 f'(\mu)\int_0^{\infty} dx\frac{x}{1+e^x} + \frac{1}{3}(k_BT)^4 f'''(\mu)\int_0^{\infty} dx\frac{x^3}{1+e^x} + ... \tag{9.117}$$

The integrals can all be evaluated in terms of the Riemann zeta function

$$\begin{aligned}\int_0^{\infty} dx\frac{x}{1+e^x} &= \frac{\pi^2}{12}\\ \int_0^{\infty} dx\frac{x^3}{1+e^x} &= \frac{7\pi^4}{120}\end{aligned} \tag{9.118}$$

This gives

$$I = \int_0^{\mu} d\epsilon\ f(\epsilon) + \frac{\pi^2}{6}(k_BT)^2 f'(\mu) + \frac{7\pi^4}{360}(k_BT)^4 f'''(\mu) + ... \tag{9.119}$$

Let us apply this expansion to determine equailibrium properties of the system. The Landau potential is given by eqn.(9.106). In that expression, we recognize $f(\epsilon) = \epsilon^{1/2}, f'(\epsilon) = 3\epsilon^{1/2}/3, ...$ This gives, upto first order

$$\begin{aligned} I &= \int_0^{\mu} d\epsilon\ \epsilon^{1/2} + \frac{\pi^2}{4}(k_BT)^2\mu^{1/2}\\ &= \frac{2}{5}\mu^{5/2} + \frac{\pi^2}{4}(k_BT)^2\mu^{1/2}\end{aligned} \tag{9.120}$$

Then

$$\mathcal{G} = -\frac{2}{3}\frac{\sqrt{2}\ Vm^{3/2}}{\pi^2\hbar^3}\left[\frac{2}{5}\mu^{5/2} + \frac{\pi^2}{4}(k_BT)^2\mu^{1/2}\right] \tag{9.121}$$

The particle number constraint is

$$\begin{aligned} N &= -\left(\frac{\partial\mathcal{G}}{\partial\mu}\right)_{T,V}\\ &= \frac{V(2m)^{3/2}}{3\pi^2\hbar^3}\left[\mu^{3/2} + \frac{\pi^2}{8}(k_BT)^2\mu^{-1/2}\right]\end{aligned} \tag{9.122}$$

Therefore

$$\frac{3\pi^2\hbar^3}{(2m)^{3/2}}\,\rho = \mu^{3/2}\left[1+\frac{\pi^2}{8}\left(\frac{k_BT}{\mu}\right)^2\right] \tag{9.123}$$

The left hand side of the above equation is just $\epsilon_F^{3/2}$. Then,

$$\epsilon_F^{3/2} = \mu^{3/2}\left[1+\frac{\pi^2}{8}\left(\frac{k_BT}{\mu}\right)^2\right] \tag{9.124}$$

This gives (binomial approximation)

$$\begin{aligned} \mu &= \epsilon_F\left[1-\frac{\pi^2}{12}\left(\frac{k_BT}{\mu}\right)^2\right] \\ &\simeq \epsilon_F\left[1-\frac{\pi^2}{12}\left(\frac{k_BT}{\epsilon_F}\right)^2\right] \end{aligned} \tag{9.125}$$

where in the second equation, to the given approximation, μ on the right hand side has been replaced by ϵ_F. the above equaiton determines the chemical potential of the system.

To compute the equation of state, we compute the Landau potential to this order. for this, we need to compute $\mu 65/2$ and $\mu^{1/2}$. These are computed to be

$$\begin{aligned} \mu^{5/2} &= \epsilon_F^{5/2}\left[1-\frac{5\pi^2}{24}\left(\frac{T}{T_F}\right)^2\right] \\ \mu^{1/2} &= \epsilon_F^{1/2}\left[1-\frac{\pi^2}{24}\left(\frac{T}{T_F}\right)^2\right] \end{aligned} \tag{9.126}$$

Substituting in eqn.(9.127) (and after some simple math)

$$\mathcal{G} = -\frac{2}{3}\frac{\sqrt{2}\,Vm^{3/2}}{\pi^2\hbar^3} \times \frac{2}{5}\epsilon_F^{5/2}\left[1+\frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right] \tag{9.127}$$

This gives the equation of state

$$P = \frac{2}{5}\epsilon_F\,\rho\left[1+\frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2\right] \tag{9.128}$$

We now compute the specific heat of the system and see if the behaviour agrees with what we anticipated (linear variation with T). For this, we need to compute the mean energy of the system. At this point, it is instructive to deduce a relationship between the pressure, volume and energy of a system of weakly interacting particles in equilibrium. This exercise also sheds light on the quantum mechanical interpretation of pressure. Consider a system of particles (it could be an interacting system) confined to volume V. This system will have quantum microstates r with energy eigenvalues E_r which will depend on the volume of the system. For instance, for a weakly interacting system, the energy eigenstates (microstates) of the system in the occupation number representation are labelled as $r \equiv \{n_\alpha\}$ where n_α is the occupation number of the single particle state $|\alpha\rangle$. Since the energy of a single particle state is given by eqn.(9.25), the energy of microstate $r \equiv \{n_\alpha\}$ is of the form

$$E\{n_\alpha\} = \frac{\pi^2\hbar^2}{2mV^{2/3}}f\{n_\alpha\} \tag{9.129}$$

where $f\{n_\alpha\}$ depends only on the occupation numbers. In particular, the volume dependence of the microstate is through the factor $V^{-2/3}$. For a general system, the volume dependence could be more complicated. In thermal equilibrium, the system is flipping between different microstates. Say, we change the volume of the system by dV. Assuming this process is 'slow', for any given microstate, the occupation numbers will not change, but the energy of the microstate will change due to a change of volume. This change in energy of microstate r will be

$$dE_r = \frac{\partial E_r}{\partial V} dV \tag{9.130}$$

This must be equal to $-P_r dV$, where P_r is the pressure exerted by the system when in microstate r. Then,

$$P_r dV = -\frac{\partial E_r}{\partial V} dV \tag{9.131}$$

from which it follows that the

$$P_r = -\frac{\partial E_r}{\partial V} \tag{9.132}$$

The mean pressure of the system will be given by an average over microstates, averaged over the probability distribution

$$P = -\overline{\frac{\partial E_r}{\partial V}} \tag{9.133}$$

Let us apply this to a system of weakly interacting particles. It is clear from (9.129) that for this system,

$$\frac{\partial E_r}{\partial V} = -\frac{2}{3V} E_r \tag{9.134}$$

Therefore,

$$\begin{aligned} P &= -\overline{\frac{\partial E_r}{\partial V}} \\ &= \frac{2}{3V} \overline{E_r} \\ &= \frac{2}{3V} E \end{aligned} \tag{9.135}$$

where E is the mean energy of the system. this can be written as

$$PV = \frac{2}{3} E \tag{9.136}$$

We now go back to computing the mean energy and heat capacity of the Fermionic system. It follows from eqns.(9.128) and (9.136) that

$$\begin{aligned} E &= \frac{3}{2} PV \\ &= \frac{3}{5} N \epsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{T}{T_F} \right)^2 \right] \end{aligned} \tag{9.137}$$

from which the heat capacity can be computed

$$\begin{aligned} C_V &= \left(\frac{E}{T} \right)_V \\ &= \frac{\pi^2}{2} N k_B \left(\frac{T}{T_F} \right) \end{aligned} \tag{9.138}$$

The molar specific heat is

$$c_v = \frac{\pi^2}{2} R \left(\frac{T}{T_F} \right) \tag{9.139}$$

which, apart from overall numerical factor, is the same as (9.104).

9.2.4 White dwarf stars

An interesting system that can be viewed as a system of weakly interacting electrons is a white dwarf star. These are abnormally dim and small for their colour. This is because they have used up all the hydrogen and what is left is mostly helium. In the absence of pressure due to nuclear fusion, the star contracts. What holds the star against gravitational collapse? The temperature and mass of a white dwarf star are of the order of solar temperature and mass. The temperature $T \sim T_{\odot} \sim 10^7\ {}^\circ K$ corresponds to energyb scale $\epsilon_0 \sim 1000eV$, which is much larger than the binding energy of electrons in the helium atom. Therefore, the helium atoms are completely ionized. The star is much more dense compared to the Sun, with mass density $\rho_M \sim 10^7 \rho_{\odot}$ where $\rho_{\odot} \sim 1g/cm^3$ is the solar density. Assuming that the star is mostly helium, the number density of electrons is of the order $\rho \sim 10^{30} cm^{-3}$. Assuming the expression for Fermi energy given in eqn.(9.94), given this density, the Fermi energy is of the order $\epsilon_F \sim 20MeV$ which is about forty times the rest mass energy of an electron $mc^2 = 0.5MeV$. Therefore, the electrons are highly relativistic. The Fermi temperature is of the order $T_F \sim 10^{11}\ {}^\circ K$. Since the temperarture of the star is $T \sim T_{\odot} \sim 10^7\ {}^\circ K$, the system of electrons is at temperature $T << T_F$, just as in metals. Then, to leading order, we can assume that the system is effectively at $T = 0$. A white dwarf star can then be modelled as a core of helium nuclei (which provide mass to the star) surrounded by an electron gas.

We now analyze the stability of such a star. Given a star in equilibrium, if we take a section of the star, it is held in equilibrium due to a competition between the pressure and gravitational forces. Consider the following section of thickness dr, area A and mass dm. The pressure forces on the two faces and the gravitational force have been highlighted

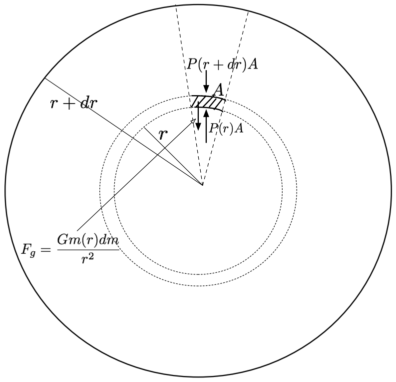


Figure 9.5: Stellar equilibrium.

In equilibrium

$$P(r+dr)A + \frac{Gm(r)dm}{r^2} - P(r)A = 0 \tag{9.140}$$

where $m(r)$ is the mass of the star upto radius r, given in terms of density $\rho(r)$ as

$$m(r) = 4\pi \int_0^r dr\ r^2 \rho(r) \tag{9.141}$$

The mass of the section is related to the density as $dm = \rho(r)Adr$. Substituting in the equilibrium equation gives

$$P(r+dr)A - P(r)A = -\frac{Gm(r)\rho(r)Adr}{r^2} \tag{9.142}$$

which gives a differential equation for the variation of pressure with radial distance from the centre of the star

$$\frac{dP}{dr} = -\frac{Gm(r)\rho(r)}{r^2} \tag{9.143}$$

From eqn.(9.141), we get

$$\frac{dm}{dr} = 4\pi r^2 \rho(r) \tag{9.144}$$

Given the equation of state $P = P(\rho)$,

$$\frac{dP}{dr} = \frac{dP}{d\rho}\,\frac{d\rho}{dr} \tag{9.145}$$

Then, eqn.(9.143) gives

$$\frac{d\rho}{dr} = -\frac{1}{dP/d\rho}\frac{Gm(r)\rho(r)}{r^2} \tag{9.146}$$

Equations (9.144) and (9.146) form a set of coupled first order differential equations in $\rho(r)$ and $m(r)$ and have a unique solution, given $\rho(0) = \rho_C$ (central density) and $m(0) = 0$ (mass contained within zero radial distance from the centre is zero). The solutions form a one-parameter family (with the central density being the parameter)

$$\begin{aligned} \rho &= \rho(r, \rho_C) \\ m &= m(r, \rho_C) \end{aligned} \tag{9.147}$$

The radius R and mass M of theb star are given by the conditions

$$\begin{aligned} \rho(R, \rho_C) &= 0 \\ m(R, \rho_C) &= M \end{aligned} \tag{9.148}$$

which can be solved to give the function $R = R(M)$. Then, given the mass of the star, its radius can be computed. A given equation of state can sustain equilibrium only up to a certain critical mass M_C. This critical mass is the solution to $R(M_C) = 0$. When the star exceeds this critical mass, it becomes unstable, collapses, till some new Physical process takes over, and changes the equation of state such that a new equilibrium is attained.

In case of a white dwarf star, the equation of state is just an expression for the electron degeneracy pressure as a function of the electron density at (effectively) $T = 0$. Since the electrons are relativistic, we need to recompute the relativistic single particle states. For this, we visualise the electron system to be confined to a torus of volume $V = L^3$. Given translational invariance, the energy eigenstates for an electron are just the momentum states with wavefunctions of the form

$$\psi_{\vec{p}}(\vec{r}) \sim e^{i\vec{p}\cdot\vec{r}/\hbar} \tag{9.149}$$

The energy eigenvalues for this state will be

$$\epsilon(\vec{p}) = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \tag{9.150}$$

Given the periodicity of the torus imposes the constraint

$$\psi_{\vec{p}}(\vec{r} + \vec{L}) = \psi_{\vec{p}}(\vec{r}) \tag{9.151}$$

which results in the condition

$$e^{i\vec{p}\cdot\vec{L}/\hbar} = 1 \tag{9.152}$$

This constrains the momentum eigenvalues to

$$\vec{p}_n = \frac{2\pi}{L}\hbar\,\vec{n} \tag{9.153}$$

where $\vec{n} = n_x\hat{i} + n_y\hat{j} + n_z\hat{k}$; $n_x, n_y, n_z = 0, \pm 1, \pm 2, \pm 3, ...$ These momentum states are the single particle states. The sum over single particle states can be reduced to integrals

$$\begin{aligned}
\sum_\alpha f(\epsilon_\alpha) &= 2\sum \Delta n_x \Delta n_y \Delta n_z f(\epsilon_\alpha) \\
&= 2\left(\frac{L}{2\pi\hbar}\right)^3 \sum \Delta p_x \Delta p_y \Delta p_z f(\epsilon) \\
&\longrightarrow 2\left(\frac{L}{2\pi\hbar}\right)^3 \int dp_x dp_y dp_z \; f(\epsilon) \qquad (9.154)
\end{aligned}$$

Since the summand depends only on the energy, we can reduce the momentum integrals to a spherical integral

$$\begin{aligned}
\sum_\alpha f(\epsilon_\alpha) &\longrightarrow 2\left(\frac{L}{2\pi\hbar}\right)^3 \times 4\pi \int_0^\infty dp \; p^2 f(\epsilon) \\
&= \frac{V}{\pi^2\hbar^3} \int_0^\infty dp \; p^2 \; f(\epsilon) \qquad (9.155)
\end{aligned}$$

Here, ϵ is given by eqn.(9.150). We assume that electron system to be at $T = 0$. Then, the occupation number distribution is given by eqn.(9.92). Then

$$\begin{aligned}
N &= \sum_\alpha \overline{n}_\alpha \\
&= \frac{V}{\pi^2\hbar^3} \int_0^{p_F} dp \; p^2
\end{aligned} \qquad (9.156)$$

where p_F is the magnitude of momentum of the highest energy single particle state for which the Fermi energy will be

$$\epsilon_F = \sqrt{p_F^2 c^2 + m^2 c^4} \qquad (9.157)$$

The above integral gives the 'Fermi momentum'

$$p_F = \hbar(3\pi^2\rho)^{1/3} \qquad (9.158)$$

The Landau potential at $T = 0$ is

$$\begin{aligned}
\mathcal{G} &= \sum_\alpha (\epsilon_\alpha - \epsilon_F) \\
&= \frac{V}{\pi^2\hbar^3} \int_0^{p_F} dp \; p^2 (\epsilon(p) - \epsilon_F) \\
&= \frac{V}{\pi^2\hbar^3} \int_0^{p_F} dp \; p^2 \left[\sqrt{p^2c^2 + m^2c^4} - \epsilon_F\right] \qquad (9.159)
\end{aligned}$$

Introducing dimensionless variable $x = p/mc$ gives

$$\mathcal{G} = \frac{Vm^4c^5}{\pi^2\hbar^3} \int_0^{x_F} dx \; x^2 \left[\sqrt{1 + x^2} - x_F\right] \qquad (9.160)$$

where $x_F = p_F/mc$. Since the electrons are highly relativistic, $p_F >> mc$. Therefore, $x_F >> 1$. Then

$$\begin{aligned}
\frac{\epsilon_F}{mc^2} &= \sqrt{1 + x_F^2} \\
&\simeq x_F + \frac{1}{2x_F} \qquad (9.161)
\end{aligned}$$

Then

$$\begin{aligned}\int_0^{x_F} dx\ x^2\sqrt{1+x^2} &= \int_0^{x_F} dx\ x^3\sqrt{1+(1/x^2)} \\ &\simeq \int_0^{x_F} dx\ x^3\left[1+\frac{1}{2x^2}\right] \\ &= \frac{x_F^4}{4}+\frac{x_F^2}{4} \qquad (9.162)\end{aligned}$$

This gives

$$\mathcal{G} = -\frac{Vm^4c^5}{12\pi^2\hbar^2}\left[x_F^4 - x_F^2\right] \qquad (9.163)$$

Since $\mathcal{G} = -PV$, we get the equation of state

$$P = \frac{m^4c^5}{12\pi^2\hbar^2}\left[x_F^4 - x_F^2\right] \qquad (9.164)$$

where

$$\begin{aligned}x_F &= \frac{p_F}{mc} \\ &= \frac{\hbar}{mc}(3\pi^2\rho)^{1/3} \qquad (9.165)\end{aligned}$$

Then, the equation of state is of the form

$$P = c_1\rho^{4/3} - c_2\rho^{2/3} \qquad (9.166)$$

where c_1 and c_2 are microscopic parameters. We can use this equation of state in the coupled differential equations for density and mass functions to determine the critical mass upto which a white dwarf star can sustain equilibrium. However, we shall do an approximate analysis to estimate this mass.

The first approximation we make is to assume a uniform density and pressure. First, imagine that the Fermion gas is enclosed in a spherical ragion of radius R, and we 'switch off' gravity. If we compress the volume by radius dR, the change in the kinetic energy of the system is

$$dE_K = -P4\pi R^2 dR \qquad (9.167)$$

In presence of gravity, we do not need to artificually restrict the gas to this volume (gravity does it). Then, this change should be equal to the change in the potential energy of the gas, given by (on dimensional grounds)

$$dE_g = \alpha\frac{GM^2}{R^2}dR \qquad (9.168)$$

where α is a dimensionless number. Since $dE_0 = -dE_g$, therefore

$$P = \frac{\alpha}{4\pi}\frac{GM^2}{R^4} \qquad (9.169)$$

This pressure should equal the Fermi pressure (9.164). The mass of the star is $M \sim 2Nm_p$ where N is the number of electrons and m_p is the proton mass.Furthur, the radius of the star is $R = (3V/4\pi)^{1/3}$ where V is the volume. Then, the number density of electrons is given by

$$\rho = \frac{3M}{8\pi m_p R^3} \qquad (9.170)$$

Then

$$\begin{aligned} x_F &= \frac{\hbar}{mc}\frac{1}{R}\left(\frac{9\pi}{8}\frac{M}{m_p}\right)^{1/3} \\ &= \frac{\overline{M}^{1/3}}{\overline{R}} \end{aligned} \tag{9.171}$$

where $\overline{M}$ is the mass of the star measured in units of proton mass

$$\overline{M} = \frac{9\pi}{8}\frac{M}{m_p} \tag{9.172}$$

and $\overline{R}$ is its radius measured in units of the compton wavelength of electrons, $\lambda_C = \hbar/(mc)$

$$\overline{R} = \frac{R}{\lambda_C} \tag{9.173}$$

Then, the Fermi pressure can be written as

$$P = k\left[\frac{\overline{M}^{4/3}}{\overline{R}^4} - \frac{\overline{M}^{2/3}}{\overline{R}^2}\right] \tag{9.174}$$

where

$$k = \frac{m^4c^5}{12\pi^2\hbar^3} \tag{9.175}$$

The pressure is also given by

$$\begin{aligned} P &= \frac{\alpha}{4\pi}\frac{GM^2}{R^4} \\ &= \frac{\alpha}{4\pi}G\left(\frac{8m_p}{9\pi}\right)^2\left(\frac{mc}{\hbar}\right)^4\frac{\overline{M}^2}{\overline{R}^4} \\ &= k'\,\frac{\overline{M}^2}{\overline{R}^4} \end{aligned} \tag{9.176}$$

Equating these gives

$$k\left[\frac{\overline{M}^{4/3}}{\overline{R}^4} - \frac{\overline{M}^{2/3}}{\overline{R}^2}\right] = k'\,\frac{\overline{M}^2}{\overline{R}^4} \tag{9.177}$$

which gives

$$\overline{M} = \overline{M}^{1/3}\left[1 - \left(\frac{\overline{M}}{M_0}\right)^{2/3}\right] \tag{9.178}$$

where

$$M_0 = \left(\frac{k}{k'}\right) \tag{9.179}$$

A calculation shows that $M_0 \sim M_\odot$ where $M_\odot$ is the solar mass. Equation (9.178) shows that the critical mass is $M_c = M_0 \sim M_\odot$. A more precise calculation gives $M_c \sim 1.4 M_\odot$. This is known as the *Chandrasekhar limit.*