

Lecture Notes #1

1. The problem

Starting with a large system (size of order Avogadro number $\sim 10^{23}$ particles), we are interested in the macroscopic behavior of this system. The standard method to solve such a problem will be to use a **dynamical approach**. In this approach the properties of the system are given in the form of a hamiltonian or a set of equations of motion, supplemented by the initial conditions for all the individual particles in the system. There are two major problems with the dynamical approach. The first is a technical problem that involves solving an Avogadro-number coupled differential equations in a reasonable time. Clearly this is impossible.

However, even if the technical problem of solving the equations of motion with the given initial conditions is solved, we still have a conceptual problem. How are we going to store the data (i.e. the initial conditions)? As an example, take a two-level system. Such a system can be e.g. a $\{0, 1\}$ bit realized by the two-states of the z-component of a spin- $\frac{1}{2}$ magnetic moment. However, no piece of apparatus can store the data on itself. Thus, to store the bit we need a unit that is larger than the bit. In the magnetic moment example we need a mechanism to change and keep the state of each individual moment. However, this unit that helps us store the bit interacts with our system and therefore its components have to be added to the whole system, thus adding equations of motion and more initial conditions that now have to be stored by new units. This line of argument will continue forever with no solution. We therefore conclude that the dynamical approach of making predictions based on initial data is only applicable to small and isolated systems.

2. Separation of degrees of freedom

There are cases for which dynamical methods do work. These are cases for which internal correlations in the macroscopic system can help us separate out the few degrees of freedom that we are interested in. For example, take a macroscopic rigid body in a gravitational field. the Hamiltonian of the body is:

$$\mathcal{H} = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m_i} + \sum_{i,j=1}^N V(\vec{q}_j - \vec{q}_i) - \vec{g} \cdot \sum_{i=1}^N m_i \vec{q}_i \quad (1)$$

Here \vec{p}_i is the momentum of the i^{th} particle, \vec{q}_i is its position, and \vec{g} denotes the gravitational field (taken to be uniform). Here we are interested in tracking the fall of this body in the gravitational field. Since the body is rigid, the relative positions among the particles do not change. Thus, using center of mass coordinate (\vec{Q}) and momentum (\vec{P}) we write:

$$\vec{Q} = \frac{1}{M} \sum_{i=1}^N m_i \vec{q}_i; \quad \vec{P} = M \dot{\vec{Q}}; \quad \vec{\pi}_i = m_i (\dot{\vec{q}}_i - \dot{\vec{Q}}) \equiv m_i \dot{\vec{\rho}}_i \quad (2)$$

where the total mass of the body is $M = \sum_{i=1}^N m_i$.

The hamiltonian now is:

$$\mathcal{H} = \frac{\vec{P}^2}{2M} - M\vec{g} \cdot \vec{Q} + \sum_{i=1}^N \frac{\vec{\pi}_i^2}{2m_i} + \sum_{i,j=1}^N V(\vec{q}_j - \vec{q}_i) \quad (3)$$

Calculating the equation of motion for the center of mass from the above hamiltonian:

$$\dot{\vec{P}} = -\frac{\partial \mathcal{H}}{\partial \vec{Q}} = M\vec{g} \quad (4)$$

and thus the equation of motion $\dot{\vec{P}} = M\ddot{\vec{Q}} = M\vec{g}$, which solves to give:

$$\vec{Q}(t) = \vec{Q}(0) + \frac{\vec{P}(0)}{M}t + \frac{1}{2}\vec{g}t^2 \quad (5)$$

we therefore note that the rest of the degrees of freedom of the body do not matter if we are only interested in the free fall of the body as a whole. the truth is that using common sense and observation we realized that we can separate out the center of mass motion. while clearly this is only an approximation (after all, we know that the atoms in the rigid body do move), it is a good one. While we realize that not all macroscopic problems can be treated this way, we look for a method that is doing a similar process of giving up the possibility of knowing information about many degrees of freedom without losing the essential physics of the system. Statistical mechanics is exactly this kind of theory.

3. States of a physical system

We saw that missing information is what is preventing us from making accurate predictions about a system and its components. the missing information is a function of the number of possibilities at which we can find our system. We call each of these possibilities an accessible state, or simply a **state** of the system. In general we will be talking about stationary quantum states of a physical system. Knowing (or counting) all the quantum states accessible to a system should help us calculate the missing information

A state of a system specifies the state of each individual component of a system. A simple example is a collection of N $\{0, 1\}$ bits. Clearly the number of states of such system is 2^N .

A state of a physical system that is made of $\sim 10^{23}$ particles will involve specifying the quantum state of each particle (i.e. energy level). The total energy of the system will be given by summing up all the energies of the individual particles.

Take for example the case of a particle of mass m in a box of size $L \times L \times L$. Imposing proper boundary conditions on the wave function of the particle yield the discrete energy levels for the particles in the system

$$\epsilon = \frac{\hbar^2}{2m} \left(\frac{\pi}{L} \right)^2 [n_x^2 + n_y^2 + n_z^2] \quad (6)$$

with the n_x, n_y, n_z being integers. Each (n_x, n_y, n_z) vector determines an energy level, and therefore a state of the single particle. For a system consisting of N particles, a set of N such vectors should be specified. Note however that in the above example the single particle state of $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ are all degenerate (give the same energy, ϵ). A correct counting of all possible states of the system should take into account such degeneracies.

4. Information

We saw above that in the process of looking at the dynamics of the rigid body we ignored a lot of information about the body. In that case it was a good assumption since we figured out that the body stays rigid with no noticeable relative motion of the atoms during the free fall. However, this is not the case for a gas, or for a system of N spins. To assess the missing information and what it tells us about macroscopic observables, let us first define it.

Assume a (classical) ball placed in one of n boxes. To give “all the information” needed to describe the system is to specify the box in which the ball is placed. Without this knowledge, there are n possible realizations to the ball in a box state. Similarly, assume an English word encoded in a binary sequence containing N bits. To know all the information about the sequence of bits, or, to know the word, we need to specify the state of all N bits (i.e. 0 or 1). Without this knowledge there are 2^N possible words that we can write with the N bits. These two examples show that the amount of knowledge needed to characterize the state of a system, and therefore, the amount of missing information if this is not given, has to be an increasing function of the number of possibilities for the system.

Let $I(n)$ denote the missing information associated with a system of n possible realizations i.e. microscopic states). It is natural to assume that lack of information only increases with

the number of states, thus:

$$I(n) > I(m) \quad \text{if } n > m \quad (7)$$

Also, for a system with a single realization there is no missing information and thus

$$I(1) = 0 \quad (8)$$

Finally, let us consider a composite system. which is separable. For example, take n boxes, each with a with very large number of compartments m . In that case the total number of possibilities is nm . Nevertheless, we can reach each possibility by first concentrating on an individual box and only then on the specific compartment in it. This separation lead us to require that the information will be additive:

$$I(nm) = I(n) + I(m) \quad (9)$$

This requirement can now be generalized to positive rational numbers of the form n/m

$$I\left(\frac{n}{m}\right) \equiv I(n) - I(m) \quad (10)$$

Using the above hypothesis we can generalize $I(n)$ to be defined also for any irrational number x since any x can be written as a limit of an infinite series of rational numbers. Thus for continuous variables x and y we require

$$I(x) > I(y) \quad \text{if } x > y; \quad I(xy) = I(x) + I(y); \quad I(1) = 0 \quad (11)$$

A solution to the above equation is:

$$I(x) = k \log(x) \quad (12)$$

where k is a constant to be determined later. Thus we conclude that for a system that can be found in n possible states, the missing information characterizing our knowledge of the state of the system is:

$$I(n) = k \log(n) \quad (13)$$

5. Consequences

The realization of how much missing information we have is at the heart of Statistical physics. Clearly the amount of missing information changes depending on our knowledge of the system that we are investigating. For example, consider a typed line in English. A line will have

approximately 80 characters while the typewriter has about 75 characters to choose from (including upper/lower case numbers, symbols etc.). With no knowledge of the English language the number of possible typed lines is 75^{80} , and the missing information is then $I = k\log(75^{80}) = 80k\log(75)$. This is clearly the maximum missing information possible. However, since we know that the language used is English and thus the typed line will consist of words and letters that “make sense”, this reduces the number of possibilities enormously and thus the missing information. In this example however it is difficult to implement the knowledge that we have about the English language to decrease the number of possible realizations of the typed line. There are other cases however in which we can eliminate possible states of the system by using a **probability distribution**.

6. Missing information with a given probability distribution

We return now to the problem of the missing information when n possibilities are given. Let us now impose on these n possibilities a probability distribution such that some possibilities are more likely to happen than others. Denote the probability of the i^{th} possibility (state) by P_i , clearly we require that

$$P_i \geq 0; \quad \text{and} \quad \sum_{i=1}^n P_i = 1 \quad (14)$$

Since the information given to us is in a fashion of a probability distribution, we know that it's actual use will be only if we try and use it many times, say N times, with $N \rightarrow \infty$. For a finite N however, each possibility, i , will appear NP_i times. Define the total number of possibilities that can be realized by the N replicas of the systems by Γ_N . It is easy to see that

$$\Gamma_N = \binom{N}{NP_1} \binom{N - NP_1}{NP_2} \binom{N - NP_1 - NP_2}{NP_3} \cdots = \frac{N!}{\prod_{i=1}^n (NP_i)!} \quad (15)$$

Define now the missing information for a system with n states that is replicated N times by $\mathcal{I}_N(n)$. The missing information is therefore

$$\mathcal{I}_N(n) = k\log\Gamma_N = k \left[\log N! - \sum_{i=1}^n \log(NP_i)! \right] \quad (16)$$

For $N \rightarrow \infty$ we can approximate

$$\log N! = N\log N - N + \mathcal{O}(\log N) \quad (17)$$

Thus

$$\mathcal{I}_N(n) = k \left[N \log N - N - \sum_{i=1}^n (N P_i \log N P_i - N P_i) \right] \quad (18)$$

which gives (after some algebra)

$$\mathcal{I}_N(n) = -Nk \sum_{i=1}^n P_i \log P_i + \mathcal{O}(\log N) \quad (19)$$

The missing information is therefore proportional to the number of times we can realize our system. To come back to the single system we started with we need to divide by N . Here we find that in the limit of $N \rightarrow \infty$ the expression for $I(n)$ depicts the fact that there is missing information in the fact that we were supplied with a probability density and not with actual data about individual possibilities. We get:

$$I(n) = \lim_{N \rightarrow \infty} \frac{\mathcal{I}_N(n)}{N} = -k \sum_{i=1}^n P_i \log P_i \quad (20)$$

Note that unlike no information at all for which $I = \lim_{n \rightarrow \infty} k \cdot \log(n) = +\infty$, when a distribution function is given

$$I = -k \sum_{i=1}^{\infty} P_i \log P_i < \infty \quad (21)$$

7. Probabilities and Ensemble averages

Given a system, a microscopic state of the system is given by specifying the states of the individual particles and the respective degeneracies. Knowing the state of the system at any given time allows us to calculate observables of the system. However, since this approach is impossible, we instead specify the probability distribution of the system to be in any given state. Physically we envision that we can replicate our system many times where each time we prepare the system in a different state. The rules that we use to replicate the system depend on the type of system and the relevant constraints. For example, in the case of the particle in a box a relevant constraint is the volume of the system. The frequency with which we find a replica of the system in this process will match exactly the probability distribution.

We call the collection of replicas of the system **ensemble**, and the average with respect to this collection of replicated systems **ensemble average**. In the future we will encounter many types of ensembles, depending on the physical situation at hand.

Finally we note that any physical system is dynamical and the microscopic states of the particles will change in time according to the relevant equations of motion with initial conditions. However, the assumption about probability distribution and ensemble average should not depend on the initial condition of the system. Thus, to make sense of our discussion we expect that the time evolution of the system will follow the ensemble as well, or that time average will be equivalent to ensemble average. We will elaborate more on this point later in the course.

Lecture Notes #2

1. The paramagnet in magnetic field

To continue with our discussion relating information (or missing information) to statistical physics, we will use a specific physical system consisting of particles, each having two possible states. The system that we have in mind is a collection of N magnetic moments $\vec{\mu}_i = g\mu_0\vec{S}_i$, each of spin $|\vec{S}_i| = 1/2$ (g is the gyromagnetic factor which we will take to be equal to 2, and μ_0 is a Bohr-magneton). We place this system in a magnetic field $\vec{H} = H\hat{z}$. In this simple example the hamiltonian of the system is the sum of the single-moment hamiltonians (\hat{h}_i) such that

$$\mathcal{H} = -\sum_{i=1}^N \vec{\mu}_i \cdot \vec{H} = -g\mu_0 \sum_{i=1}^N \vec{S}_i \cdot \vec{H} \equiv \sum_{i=1}^N \hat{h}_i \quad (1)$$

where the i^{th} moment satisfy the Schrodinger equation

$$\hat{h}_i \psi_i = \epsilon_i \psi_i \quad (2)$$

Note that in the above equations we used “hats” to denote operators. In the future we will omit the “hat” as it will be obvious when do we mean an operator.

where the hamiltonian of the single spin is:

$$\hat{h}_i = -\vec{\mu}_i \cdot \vec{H} \quad (3)$$

In our case of a spin 1/2 moment there are only two possible projections of the spin along the direction of the field. Define the variable $\sigma_i = 2S_i^z$, it is easy to see that $\sigma_i = \pm 1$ and that the energy states of a single moment are

$$\epsilon_i = -\mu_0 H \sigma_i \quad (4)$$

A system of N such magnetic moments will be found in a state that is given by a N -component vector specifying the σ -s of all the moments; $(\sigma_1, \sigma_2, \dots, \sigma_N) \equiv \{\sigma\}$. The energy of such state is then:

$$E(\sigma_1, \sigma_2, \dots, \sigma_N) = -\mu_0 H \sum_{i=1}^N \sigma_i \quad (5)$$

and the magnetization of the system in the direction of the magnetic field (i.e. the \hat{z} -direction), associated with that particular state is:

$$M(\sigma_1, \sigma_2, \dots, \sigma_N) = \sum_{i=1}^N \vec{\mu}_i \cdot \hat{z} = \mu_0 \sum_{i=1}^N \sigma_i \quad (6)$$

The total number of states that such a system can be found in is 2^N since each moment can be found in two states. Thus, in the absence of any input on the way the system is distributed among all these possible states the missing information is given by:

$$I(N) = k \log(2^N) = Nk \log 2 \quad (7)$$

This is therefore the maximum uncertainty (or missing information) that is associated with such a system of N spins.

As before, in a dynamical approach we derive the equations of motion from the hamiltonian of the system, Eqn. 1. The equations of motion for the i^{th} spin is given by the commutator of \vec{S}_i with the hamiltonian:

$$\dot{\vec{S}}_i = \frac{i}{\hbar} [\mathcal{H}, \vec{S}_i] \quad (8)$$

or, classically it is given by the spin precession equation of motion:

$$\dot{\vec{S}}_i = -\frac{g\mu_0}{\hbar} \vec{H} \times \vec{S}_i \quad (9)$$

To obtain the state of the system we need to solve the equations of motion for all the spins in the system using a set of given initial conditions. However, as discussed above this task is impossible. Moreover, we will be interested in macroscopic quantities such as the magnetization of the system and thus knowing the actual microscopic states of each of the spins may not be necessary.

2. Probabilities and averages

The solution to obtaining the macroscopic behavior of the system of spins is using statistics. In that approach we will obtain the state of the system by first postulating a probability distribution for the spins, i.e. a set of probabilities $\{P(\sigma_1, \sigma_2, \dots, \sigma_N)\}$ that satisfy all the constraints on our system.

Given the set of probabilities $\{P\}$, the problem of evaluating averages become a mere technicality. For any observable that depends on the set of spins, $\{A(\sigma_1, \sigma_2, \dots, \sigma_N)\}$, the average is:

$$\langle A \rangle = \sum_{\{\sigma\}} P(\sigma_1, \sigma_2, \dots, \sigma_N) A(\sigma_1, \sigma_2, \dots, \sigma_N) \quad (10)$$

In particular, the magnetization of the system in the direction of the magnetic field is:

$$\langle M \rangle = \left\langle \sum_{i=1}^N \vec{\mu}_i \cdot \hat{z} \right\rangle = \sum_{\{\sigma\}} \left(\mu_0 \sum_{i=1}^N \sigma_i \right) P(\sigma_1, \sigma_2, \dots, \sigma_N) = \mu_0 \sum_{\sigma_1, \sigma_2, \dots} (\sigma_1 + \sigma_2 + \dots + \sigma_N) P(\sigma_1, \sigma_2, \dots, \sigma_N) \quad (11)$$

where in the last equality we specified that we sum over all possible spin configurations. However, since all spins are equivalent and $\sigma_i = \pm 1$ for all spins S_i , the above sum can be contracted to one spin, say σ_1 :

$$\langle M \rangle = N\mu_0 \sum_{\sigma_1} \sigma_1 \sum'_{\{\sigma\}} P(\sigma_1, \sigma_2, \dots, \sigma_N) \equiv N\mu_0 \sum_{\sigma_1} \sigma_1 P(\sigma_1) \quad (12)$$

where the “prime” denotes summing over all spins except for σ_1 . However, since the label of the spin is arbitrary, the above result can also be written as:

$$\langle M \rangle = N\mu_0 \langle \sigma \rangle \quad (13)$$

where the average now is on the microscopic spin:

$$\langle \sigma_i \rangle = \sum_{\sigma_i = \pm 1} \sigma_i P(\sigma_i) \quad (14)$$

In the same manner we can write the probability that two spins, i and j will have spins σ_i and σ_j :

$$P(\sigma_i, \sigma_j) = \sum''_{\{\sigma\}} P(\sigma_1, \sigma_2, \dots, \sigma_N) \quad (15)$$

where the double prime denotes skipping i and j in the summation.

3. Isolated paramagnet

As we shifted from specifying the initial conditions of the individual particles and the solution of their equations of motion to a more statistical approach, we need to specify the relevant probability distribution or the type of ensemble. This will be derived based on the constraints on the system. The fixed volume in the case of a gas of particles in a box merely describes the constraint that no particles can be found outside the box.

An important set of constraints that we need to impose on the system is all the conservation laws. These should be applied to the individual particles and hence to all the states of the ensemble. For an isolated system, a system that cannot exchange energy with the environment, conservation of energy is the first conservation law that comes to mind. Clearly the time evolution of such system, and hence all the members of the ensemble describing such system, should have an energy equal to the initial energy of the system.

For a paramagnet, conservation of energy means that a given number of spins align with the field (up spins) while the rest of the spins are aligned opposite to the field (down spins) such that the sum

$$E = \sum_{i=1}^N \langle \hat{h}_i \rangle = \sum_{i=1}^N \epsilon_i \quad (16)$$

is constant.

Since all the spins are equivalent it does not matter which spins point up and which point down. Thus, a constant energy constraint is equivalent to specifying the number of up spins in the system. Since all the spins in the system are equivalent, it is reasonable to assume that the probabilities of finding any set of up spins out of the total number of spins are equal. This line of thinking is not unique to the paramagnet in a field but it is a rather general feature of a closed (isolated) systems.

For an isolated system the probabilities of all microscopic states are equal

or

An isolated system is equally likely to be in any of the quantum states accessible to it

Thus, if for a given energy the total number of microscopic states is $\Gamma(E)$, the probability for any microscopic state of the isolated paramagnet is:

$$P(\sigma_1, \sigma_2, \dots, \sigma_N) = \frac{1}{\Gamma(E)} \quad (17)$$

which defines a uniform distribution function.

An ensemble of systems that describes the possible realizations of an isolated system is called the **Microcanonical Ensemble**.

The calculation of $\Gamma(E)$ is simple since the energy of a state is determined only by the relative number of up and down spins. Denote the number of up spins by N_\uparrow and the number of down spins by N_\downarrow , $N = N_\uparrow + N_\downarrow$, and the energy of the system is:

$$E = -\mu_0 H (N_\uparrow - N_\downarrow) \quad (18)$$

Since the energy E is constant and the number of up and down spins adjust themselves to obtain this energy, it is convenient to write these numbers in terms of the energy. Solving the above two equations we obtain

$$N_{\uparrow} = \left(\frac{N}{2} - \frac{E}{2\mu_0 H} \right); \quad \text{and} \quad N_{\downarrow} = \left(\frac{N}{2} + \frac{E}{2\mu_0 H} \right) \quad (19)$$

Finding $\Gamma(E)$ amounts to calculating how many possibilities there are to choose N_{\uparrow} out of a total of N spins. This is easily found to be

$$\Gamma(E) = \binom{N}{N_{\uparrow}} = \frac{N!}{N_{\uparrow}!N_{\downarrow}!} = \frac{N!}{\left(\frac{N}{2} - \frac{E}{2\mu_0 H}\right)! \left(\frac{N}{2} + \frac{E}{2\mu_0 H}\right)!} \quad (20)$$

Using the number of up and down spins the probability for a state is

$$P(\{\sigma\}) = \frac{N_{\uparrow}!N_{\downarrow}!}{N!} \quad (21)$$

which can be used to calculate the probability for a state of any individual spin.

An easier way however is to realize that the fact that the energy determines the number of up spins means that it also determines the probability of any spin σ_i to be up or down. Thus we should have:

$$P(\uparrow) = \frac{N_{\uparrow}}{N}; \quad \text{and} \quad P(\downarrow) = \frac{N_{\downarrow}}{N} = \frac{N - N_{\uparrow}}{N} \quad (22)$$

so the energy of the system is simply

$$E = -\mu_0 H \sum_{i=1}^N \langle \sigma_i \rangle = -\mu_0 H \sum_{i=1}^N \sum_{\sigma_i = \pm 1} \sigma_i P(\sigma_i) = -N\mu_0 H [1 \cdot P(\uparrow) + (-1) \cdot P(\downarrow)] = -\mu_0 H (N_{\uparrow} - N_{\downarrow}) \quad (23)$$

as expected. Note also that $P(\uparrow)$ can also be determined by noting that if we fix a particular spin to be up, there are only $(N - 1)$ spins left to assign to $(N_{\uparrow} - 1)$ up spins. Thus:

$$P(\uparrow) = \frac{\binom{N-1}{N_{\uparrow}-1}}{\binom{N}{N_{\uparrow}}} = \frac{(N-1)!N_{\uparrow}!(N-N_{\uparrow})!}{N!(N_{\uparrow}-1)!(N-N_{\uparrow})!} = \frac{N_{\uparrow}}{N} \quad (24)$$

as expected. Note also that in terms of the energy we can write:

$$P(\uparrow) = \frac{1}{2} - \frac{E/N}{2\mu_0 H}; \quad P(\downarrow) = \frac{1}{2} + \frac{E/N}{2\mu_0 H} \quad (25)$$

Examination of this expressions shows that the minimum energy of a set of N spins is $-N\mu_0H$ for which $P(\sigma_i = \uparrow) = 1$ and the maximum energy of the system is $N\mu_0H$ for which $P(\sigma_i = \uparrow) = 0$. We also note that for $E = 0$, $P(\sigma_i = \uparrow) = P(\sigma_i = \downarrow) = 1/2$. Finally, the average moment per spin is:

$$\langle \mu \rangle = \sum_{i=1}^N \mu_i P(\mu_i) = \mu_0 P(\uparrow) + (-\mu_0) P(\downarrow) = \mu_0 \left(\frac{N_{\uparrow}}{N} - \frac{N_{\downarrow}}{N} \right) = -\frac{E}{HN} \quad (26)$$

note that this result gives for the total magnetization

$$M = N \langle \mu \rangle = -\frac{E}{H} \quad (27)$$

, or, as we remember it from E&M:

$$E = -MH \quad (28)$$

4. Entropy

The missing information for the system of N spins is given by the expression:

$$I(N) = -k \sum_{i=1}^N \sum_{\sigma_i = \pm 1} P(\sigma_i) \log P(\sigma_i) = -kN [P(\uparrow) \log P(\uparrow) + P(\downarrow) \log P(\downarrow)] \quad (29)$$

We call the amount of missing information in a physical system the entropy and denote it by S . We further identify the constant k with the *Boltzman factor* and denote it by $k_B = 1.38 \times 10^{-16} \text{erg/K}$. In our case $S = S(E)$ and we find:

$$S(E) = -k_B N \left(\frac{1}{2} - \frac{E/N}{2\mu_0 H} \right) \log \left(\frac{1}{2} - \frac{E/N}{2\mu_0 H} \right) - k_B N \left(\frac{1}{2} + \frac{E/N}{2\mu_0 H} \right) \log \left(\frac{1}{2} + \frac{E/N}{2\mu_0 H} \right) \quad (30)$$

Examining again the origin of the expression for the mission information I as the logarithm of the number of states, we can also write:

$$S = k_B \log \Gamma(E) \quad (31)$$

In our physical system the limit of maximum missing information is the limit of zero given energy (i.e. zero magnetic field). In that limit there is no tendency to align the spin in any particular direction. Some rearrangement of the terms in Eqn. 30, we find

$$\lim_{E \rightarrow 0} S = N k_B \log 2 \quad (32)$$

which is the expected result based on the fact that the total number of possibilities for a two-states system consisting of N particles is $\Gamma = 2^N$.

5. Analysis of $\Gamma(E)$

In the above analysis we postulate that all states with energy E of the system are equally probable. Mathematically it means that the relevant probability distribution is the Uniform distribution. In other words, if the number of accessible states is $\Gamma(E)$, the probability for each state is $[\Gamma(E)]^{-1}$. To test how justified is this assumption we have to show that $\Gamma(E)$ is sharply peaked around the given energy E .

In the language of information theory we have to show that indeed the system is well described by the given probability distribution. Clearly we saw that this is true in the limit of infinite number of states. However, for a finite number of particles the number of states will also be finite and thus $\Gamma(E)$ will have a width. The law of large numbers tells us that as $N \rightarrow \infty$, the system will have the exact energy E .

To prove the above statement we start from the expression:

$$\Gamma(E) = \Gamma(N_\uparrow) = \frac{N!}{N_\uparrow!(N - N_\uparrow)!} \quad (33)$$

It is more convenient however to use the logarithm of Γ which in the large N is approximated as:

$$\begin{aligned} \log \Gamma(N_\uparrow) &\approx N \log N - N - N_\uparrow \log N_\uparrow + N_\uparrow - (N - N_\uparrow) \log(N - N_\uparrow) + (N - N_\uparrow) = \\ &= N \log N - N_\uparrow \log N_\uparrow - (N - N_\uparrow) \log(N - N_\uparrow) \end{aligned} \quad (34)$$

Searching for a maximum of $\log \Gamma$ we find:

$$\frac{d \log \Gamma}{dE} = \frac{d \log \Gamma}{dN_\uparrow} \cdot \frac{dN_\uparrow}{dE} \quad (35)$$

using:

$$N_\uparrow = \frac{N}{2} - \frac{E}{2\mu_0 H}; \quad \text{and} \quad \frac{dN_\uparrow}{dE} = -\frac{1}{2\mu_0 H} \quad (36)$$

which gives:

$$\frac{d \log \Gamma}{dE} = -\frac{1}{2\mu_0 H} [-\log(N_\uparrow) - 1 + \log(N - N_\uparrow) + 1] = 0 \quad (37)$$

The above equation gives the solution for the maximum of $\log\Gamma$ at:

$$\tilde{N}_{\uparrow} = \frac{N}{2}; \quad \tilde{E} = 0 \quad (38)$$

as expected. To find the width of the peak we first calculate the second derivative:

$$\frac{d^2 \log\Gamma}{dE^2} = \frac{1}{(2\mu_0 H)^2} \left[-\frac{1}{N_{\uparrow}} - \frac{1}{N - N_{\uparrow}} \right] = -\frac{N}{N_{\uparrow}(N - N_{\uparrow})} \frac{1}{(2\mu_0)^2} \quad (39)$$

Indeed the second derivative is negative at $\tilde{N}_{\uparrow} = N/2$ and thus $\Gamma(E)$ is *max*. Expanding now around \tilde{E} we find:

$$\log\Gamma(E) = \log(\tilde{E}) + \frac{d\log\Gamma}{dE} \Big|_{E=\tilde{E}} (E - \tilde{E}) + \frac{1}{2} \frac{d^2 \log\Gamma}{dE^2} \Big|_{E=\tilde{E}} (E - \tilde{E})^2 + \dots \quad (40)$$

The first derivative vanishes at the maximum. Keeping only the first two terms we exponentiate the result to get:

$$\Gamma(E) = \Gamma(0) e^{-\frac{E^2}{2(\mu_0 H)^2 N}} \quad (41)$$

Which is a gaussian distribution with width $\sigma_E = \mu_0 H \sqrt{N}$. The above result is no surprise since in the large N limit, the law of large numbers tells us that all distributions, including a uniform one that describes the isolated system, can be replaced with a gaussian distribution. Physically what it means is that if the energy is allowed to vary, the entropy (or missing information) is maximum at the average energy of the system. This constitutes an intimate relation between the internal energy of the system and its entropy. For an isolated system with a given energy the entropy is maximum. We will use this relation further in the next lecture to define temperature.

Lecture Notes #3

1. Bringing two isolated systems to contact

In the previous lecture we discussed an isolated system and the way the states of this system are distributed. we saw that the missing information associated with such system depends on the constraint that we impose, i.e. the (conserved) energy of the system. For the particular case of the isolated paramagnet we saw that $E = 0$ corresponds to maximum missing information, or, maximum entropy, while $E = \pm N\mu_0 H$ correspond to the minimum or maximum possible energy of such system and for these extrema the entropy is zero, i.e. no missing information. Clearly a hidden requirement behind all the above results is that the system is in **equilibrium**, that is to say that the macroscopic parameters of the system do not vary in time.

Assume now that we take two isolated paramagnets and bring them in contact with each other to produce a larger system. Denote the two paramagnetic systems as:

System 1, denoted by \mathcal{S}_1 , with N_1 moments and energy E_1 , and

System 2, denoted by \mathcal{S}_2 , with N_2 moments and energy E_2 , and

For simplicity assume that both systems are placed in the same external magnetic field H . Initially, the two systems are completely isolated from each other, so each one is characterized by an energy E_i^0 ($i = 1, 2$). Denote the combined system by \mathcal{S} , the energy and number of moments of the combined system is:

$$E = E_1^0 + E_2^0 = E_1 + E_2; \quad N = N_1 + N_2 \quad (1)$$

Where:

$$E_i = -\vec{H} \cdot \sum_{i=1}^{N_i} \vec{\mu}_i \quad (2)$$

Assume now that the number of available states of system \mathcal{S}_1 is $\Gamma_1(E_1)$, while the number of available states of system \mathcal{S}_2 is $\Gamma_2(E_2) = \Gamma_2(E - E_1)$. At a first glance (see lecture Notes #1), the missing information of the combined system \mathcal{S} will be a function of the total number

of available states: $\Gamma(E) = \Gamma_1(E_1, N_1) \cdot \Gamma_2(E_2, N_2)$. However, this is true only if we assume that E_1 and E_2 remain unchanged. If we allow energy to flow from one system to the other, E_1 will change and force E_2 to change with it keeping $E = E_1 + E_2$ as constant. Thus, to obtain the true number of new states we need to factor all possible realizations of E_1 (we could choose any of the two energies, depending on whether we want to concentrate our attention to \mathcal{S}_1 or \mathcal{S}_2). Thus

$$\Gamma(E) = \sum_{E_1 \leq E} [\Gamma_1(E_1, N_1) \cdot \Gamma_2(E_2, N_2)] \equiv \sum_{E_1 \leq E} \Gamma(E, E_1) \quad (3)$$

The missing information of the combined system will be proportional to the logarithm of $\Gamma(E)$.

To proceed, we use the fact that when the two systems are brought to contact, energy flows from one to the other until the combined system reaches equilibrium. At this point, it is reasonable to assume that the energy of \mathcal{S}_1 will be mostly found around a new equilibrium value \tilde{E}_1 and similarly, \mathcal{S}_2 will be found with an equilibrium energy $\tilde{E}_2 = E - \tilde{E}_1$. For this argument to make sense, the number of states of the combined system should peak at that equilibrium value, and the peak should be very sharp. To test whether such value exists we need to find whether there is a term in the sum of Eqn. 3 that contributes most of the weight of the new distribution function. We therefore need to find out the maximum of $\log\Gamma(E, E_1)$, or solve:

$$d\log\Gamma(E, E_1) = \frac{\partial \log\Gamma_1}{\partial E_1} dE_1 + \frac{\partial \log\Gamma_2}{\partial E_2} dE_2 = 0 \quad (4)$$

or, since $E = E_1 + E_2$, $dE_1 = -dE_2$:

$$d\log\Gamma(E, E_1) = \left[\frac{\partial \log\Gamma_1}{\partial E_1} - \frac{\partial \log\Gamma_2}{\partial E_2} \right] dE_1 = 0 \quad (5)$$

The solution of the above maximization is that the condition for equilibrium is

$$\frac{\partial \log\Gamma_1}{\partial E_1} = \frac{\partial \log\Gamma_2}{\partial E_2} \quad (6)$$

which, using the definition of entropy for each individual isolated system $S_i(E) = k_B \log\Gamma_i(E_i)$ gives:

$$\left(\frac{\partial S_1}{\partial E_1} \right) = \left(\frac{\partial S_2}{\partial E_2} \right) \equiv \frac{1}{T} \quad (7)$$

Where we introduced a new quantity T , that we call temperature that is a measure of how energy was redistributed after combining the two systems. A simple dimensional analysis shows us that $[S][T] = [k_B][T] = [E]$. Thus, it is convenient to use the combination;

$$\beta = \frac{\partial \log \Gamma}{\partial E} = \frac{1}{k_B T} \quad (8)$$

such that β^{-1} has the dimensions of energy.

2. Hot, Cold, and Entropy change

We see above that temperature is a measure of how the flow of energy in or out of the system affects the distribution of available states for the system. We further call a system with large T - hot, and a system with smaller T - cold. Of course hot and cold are relative quantities and have meaning only when we compare two systems.

Let us now be specific on where the energy flows when we bring the two together. Assume that a positive amount of energy, ΔE flows from \mathcal{S}_1 to \mathcal{S}_2 when we bring the two systems together. Assume further that the temperature of \mathcal{S}_1 is larger than the temperature of \mathcal{S}_2 , i.e. $T_1 > T_2$. Calculating the entropy change of the combined system we find:

$$\Delta S = \left(\frac{\Delta S_1}{\Delta E_1} \right) (-\Delta E) + \left(\frac{\Delta S_2}{\Delta E_2} \right) \Delta E = k_B [-\beta_1 + \beta_2] \Delta E = \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \Delta E \quad (9)$$

For $\beta_2 > \beta_1$, or $T_1 > T_2$, the quantity in the parentheses is positive. Thus, the total entropy change is positive when the direction of the the energy flow is from the “hot” system to the “cold” system.

3. Distribution of states of a composite system

Let us now solve explicitly the condition for maximum, i.e. solve equation 5. However, we change our starting point and use the result of the previous lecture, writing the number of states as a gaussian distribution around the mean energy of the individual system. In that case:

$$\log \Gamma(E, E_1) = \log(\Gamma_1(0)\Gamma_2(0)) - \frac{E_1^2}{2(\mu_0 H)^2 N_1} - \frac{(E - E_1)^2}{2(\mu_0 H)^2 N_2} \quad (10)$$

and to find the maximum we require

$$\frac{\partial \log \Gamma}{\partial E_1} = -\frac{E_1}{(\mu_0 H)^2 N_1} - \frac{(E - E_1)}{(\mu_0 H)^2 N_2} = 0 \quad (11)$$

for which the solution is

$$\frac{\tilde{E}_1}{N_1} = \frac{E - \tilde{E}_1}{N - N_1} = \frac{E}{N} \quad (12)$$

Using the above result we obtain now for the maximum of Γ :

$$\Gamma_{max} = \Gamma_1(0)\Gamma_2(0)e^{-\frac{E^2}{2N(\mu_0 H)^2}} \quad (13)$$

We can now calculate the behavior of $\Gamma(E)$ near the max. Starting from Eqn. 10:

$$\Gamma = \Gamma_1(0)\Gamma_2(0)e^{-\frac{1}{2(\mu_0 H)^2} \left[\frac{E_1^2}{N_1} + \frac{(E-E_1)^2}{N_2} \right]} \quad (14)$$

and using the result for the max. (Eqn. 12) we get

$$\Gamma = \Gamma_{max} e^{-\frac{N}{2(\mu_0 H)^2 N_1(N-N_1)} (E_1 - \tilde{E}_1)^2} \quad (15)$$

This is a gaussian distribution with a width:

$$\langle (\Delta E_1)^2 \rangle = \frac{(\mu_0 H)^2 N_1(N-N_1)}{N} \quad (16)$$

and

$$\left| \frac{\sqrt{(\Delta E_1)^2}}{\langle E_1 \rangle} \right| = \frac{(\mu_0 H) \sqrt{N_1(N-N_1)/N}}{\mu_0 H |E| N_1/N} \quad (17)$$

Where we used $E = -\mu_0 H N$. We therefore see that

$$\left| \frac{\sqrt{(\Delta E_1)^2}}{\langle E_1 \rangle} \right| = \sqrt{\frac{N_1(N-N_1)}{N_1^2}} \frac{1}{N} \sim \frac{1}{\sqrt{N}} \quad (18)$$

We therefore see that in the limit $N \gg 1$, the distribution function, i.e. the number of states for the combined system with energy E while \mathcal{S}_1 has an energy E , becomes extremely sharp around $E_1 = \tilde{E}_1$. This therefore justifies our initial assumption that most of the contribution to the sum in equation ?? comes from very few terms for which $E_1 \sim \tilde{E}_1$. Thus, in that limit we can approximate the sum with an integral over all E_1

$$\Gamma(N, E) = \sum_{E_1 \leq E} \Gamma_1(E_1)\Gamma_2(E_2) \rightarrow \int_{-\infty}^{\infty} \Gamma(E, E_1) dE_1 \quad (19)$$

which gives

$$\Gamma(N, E) = \Gamma_{max} \int_{-\infty}^{\infty} e^{-\frac{N}{2(\mu_0 H)^2 N_1(N-N_1)} (E_1 - \tilde{E}_1)^2} \propto e^{-\frac{E^2}{2N(\mu_0 H)^2}} \quad (20)$$

This result is easily understood since, for the new system which is isolated and has energy E , the states distribution function will follow the gaussian behavior we derived in the last lecture.

4. Probabilities and temperature for the paramagnet in a field

Applying the above analysis to the paramagnet problem we find that:

$$\frac{\partial \log \Gamma}{\partial E} = -N \left[-\frac{1}{2\mu_0 H N} \log \left(\frac{1}{2} - \frac{E/N}{2\mu_0 H} \right) - \frac{1}{2\mu_0 H N} + \frac{1}{2\mu_0 H N} \log \left(\frac{1}{2} + \frac{E/N}{2\mu_0 H} \right) + \frac{1}{2\mu_0 H N} \right] \quad (21)$$

Thus,

$$\frac{\partial \log \Gamma}{\partial E} = \frac{1}{2\mu_0 H} \log \left[\frac{\frac{1}{2} - \frac{E/N}{2\mu_0 H}}{\frac{1}{2} + \frac{E/N}{2\mu_0 H}} \right] \equiv \beta \quad (22)$$

Remembering that $P(\uparrow) = \frac{1}{2} - \frac{E/N}{2\mu_0 H}$, and that $P(\downarrow) = \frac{1}{2} + \frac{E/N}{2\mu_0 H}$, we can rewrite the above result as

$$\beta = \frac{1}{2\mu_0 H} \log \left[\frac{P(\uparrow)}{P(\downarrow)} \right] \quad (23)$$

which gives the final result for the ratio of the up and down spins probabilities as

$$\frac{P(\uparrow)}{P(\downarrow)} = e^{2\mu_0 H \beta} = e^{-\beta(\epsilon_{\uparrow} - \epsilon_{\downarrow})} \quad (24)$$

where we used the fact that for an individual spin $\epsilon_{\uparrow} = -\mu_0 H$ and $\epsilon_{\downarrow} = +\mu_0 H$

We conclude that the ratio of the two probabilities is equal to the ratio of the exponential factors of the respective states $e^{-\beta\epsilon_{\uparrow}}$ and $e^{-\beta\epsilon_{\downarrow}}$.

5. The law of increased entropy

Let us examine equation 3 and compare it to the initial number of states of the two systems. Denoting the initial energies of the two systems by E_1^0 and E_2^0 for \mathcal{S}_1 and \mathcal{S}_2 respectively, the following inequality has to hold:

$$\Gamma(E) = \prod_{E_1 \leq E} [\Gamma_1(E_1, N_1) \cdot \Gamma_2(E_2, N_2)] \geq \Gamma_1(E_1^0, N_1) \cdot \Gamma_2(E_2^0, N_2) \quad (25)$$

This has to be true because the set of energies E_1^0 and E_2^0 are only one term in the product, and all terms are ≥ 1 . The special case for which the equality sign holds is when upon bringing them together, there is no flow of energy between the two systems and $\tilde{E}_1 = E_1^0$.

The above result can be rewritten in terms of the logarithm of the number of states, i.e. the entropy.

$$S(E) \geq S_1(E_1) + S_2(E_2) \quad (26)$$

This is the law of increased entropy. It basically states that the entropy of a system tends to increase if a constraint applied to the system is removed. In the case we analyzed above the constraint was the separation of the two subsystems.

Example:

While the above analysis is obvious, it may be easier to see it in a specific example. Consider two systems, each containing 3 spins (a total of 6 spins). Assume that the energy of each system with an applied magnetic field is $-\mu_0 H$. This energy is achieved with two spins pointing up and one pointing down. For each system there are 3 such possibilities and thus the total number of states available to the two systems with the constraint that the total energy is $E = E_1 + E_2 = -2\mu_0 H$ is $3 \times 3 = 9$.

Now let us combine the two systems together. All the previous states are still available. However, new states in which all the spins of one systems are up and only one of the other system, are now available. The total number of such states is 6, so altogether we have 15 states for the combined system. This is larger than the initial value of 9.

Lecture Notes #4

1. The Partition Function

We previously showed for a paramagnet in a magnetic field that by allowing energy to change, e.g. by making contact to another system, the entropy of the system changes such that the temperature of the two systems become equal. The temperature is then defined via

$$\left(\frac{\partial S}{\partial E}\right) = \frac{1}{T} \quad (1)$$

Defining $\beta = (k_B T)^{-1}$, we also found that the ratio between up and down pointing spins for such a system is:

$$\frac{P(\uparrow)}{P(\downarrow)} = e^{2\mu_0 H \beta} = e^{-\beta(\epsilon_{\uparrow} - \epsilon_{\downarrow})} \quad (2)$$

The two probabilities should also add up to 1:

$$P(\uparrow) + P(\downarrow) = 1 \quad (3)$$

which now can help us solve for the two probabilities:

$$P(\uparrow) = \frac{e^{\beta\mu_0 H}}{e^{\beta\mu_0 H} + e^{-\beta\mu_0 H}}; \quad P(\downarrow) = \frac{e^{-\beta\mu_0 H}}{e^{\beta\mu_0 H} + e^{-\beta\mu_0 H}} \quad (4)$$

We identify the denominator of both probabilities as a sum over the exponential factors of the single spin energies. We call this sum the single particle **partition function** and denote it by Z_1 :

$$Z_1 = \sum_{\epsilon_i = \mp \mu_0 H} e^{-\beta \epsilon_i}, \quad (5)$$

and the exponential factors $e^{-\beta \epsilon_i}$, **Boltzmann Factors**.

Using Z_1 of the two-states paramagnet, the probabilities for the different states of the system can be written as:

$$P(\mp \mu_0 H) = \frac{1}{Z_1} e^{\pm \beta \mu_0 H} \quad (6)$$

We can generalize the above observation to any system that its temperature is fixed. For such a system the probability for any state α which is characterized by an energy E_α is:

$$P(E_\alpha) = \frac{e^{-\beta E_\alpha}}{Z} \quad (7)$$

where Z is now the full system partition function given by:

$$Z = \sum_{\{\alpha\}} e^{-\beta E_\alpha} \quad (8)$$

For the case of non-interacting particles, such as our paramagnet in a field, we can calculate the partition function in terms of the single particle energy levels. For the paramagnet, $\sum_{\{\alpha\}} = \sum_{\sigma_1, \sigma_2, \dots, \sigma_i = \pm 1}$, and thus:

$$1 = \sum_{\sigma_1, \sigma_2, \dots, \sigma_i = \pm 1} P(\{\sigma_i\}) = \sum_{\sigma_1, \sigma_2, \dots, \sigma_i = \pm 1} \frac{1}{Z} e^{-\beta \sum_{i=1}^N (-\mu_0 H \sigma_i)} \quad (9)$$

However, since all the particles are non-interacting and equivalent, and the index of the particles is a “dummy” index over which we sum, the above equation is a product of N “single-particle” sums:

$$1 = \frac{1}{Z} \left[\sum_{\sigma_i = \pm 1} e^{\beta(-\mu_0 H \sigma_i)} \right]^N \quad (10)$$

We therefore conclude that for non-interacting particles:

$$Z = \left[\sum_{\epsilon_i} e^{-\beta \epsilon_i} \right]^N \equiv Z_1^N \quad (11)$$

Where we identified the single particle partition function as Z_1 , given also in equation 5.

2. The Canonical Ensemble

So far, the two systems that we brought together and observed their equilibration process, had no specific requirement. We saw that upon bringing the two systems together, energy will flow from one system to the other, changing the entropy of the combined system such that at equilibrium the temperatures of the two subsystems are constant.

assume now that while both systems are macroscopic in size, one system is much larger than the other. We call the larger of the two systems, **reservoir** (denoted by \mathcal{R}), and we connect our small system \mathcal{S} to it. The reservoir will be taken as much larger than our system by considering all extensive properties of the systems, such as the number of particles (or volume), the energy, the entropy, etc. Denoting these quantities with subscripts R or S , we require for examples that the energies of the two systems will satisfy: $E^R \gg E^S$. The total

system, $\mathcal{R} + \mathcal{S}$ is an isolated system with energy $E_0 = E^R + E^S$. Our aim now is to study the properties of the system \mathcal{S} as it is brought into contact with the reservoir. To do that we study $\mathcal{R} + \mathcal{S}$ with \mathcal{S} in any of its available states $\{\alpha\}$. If \mathcal{S} is forced to be in a particular state, γ , with energy E_γ^S , the total number of states of the combined system is:

$$\Gamma_{\mathcal{R}+\mathcal{S}} = \Gamma_{\mathcal{R}} \cdot 1 = \Gamma_{\mathcal{R}}(E_0 - E_\gamma^S) \quad (12)$$

where we used the fact that the number of states of the reservoir depends on the reservoir's energy $E^R = E_0 - E_\gamma^S$.

Clearly, the probability for the system \mathcal{S} to be in any of its available states will be proportional to

$$P(E_\gamma^S) = \frac{\Gamma_{\mathcal{R}}(E_0 - E_\gamma^S)}{\Gamma_{\mathcal{R}+\mathcal{S}}} \quad (13)$$

Thus, the ratio of the probability that the system is in a particular state γ to the probability that it is in a state τ is:

$$\frac{P(E_\gamma^S)}{P(E_\tau^S)} = \frac{\Gamma_{\mathcal{R}}(E_0 - E_\gamma^S)}{\Gamma_{\mathcal{R}}(E_0 - E_\tau^S)} \quad (14)$$

, Or, using the definition of entropy, $S = k_B \log \Gamma$,

$$\frac{P(E_\gamma^S)}{P(E_\tau^S)} = e^{\frac{1}{k_B} [S^{\mathcal{R}}(E_0 - E_\gamma^S) - S^{\mathcal{R}}(E_0 - E_\tau^S)]} \equiv e^{\frac{1}{k_B} [\Delta S^{\mathcal{R}}]} \quad (15)$$

Using now the fact that both, E_γ^S and E_τ^S are much smaller than E_0 , we can expand each one of the entropies:

$$S^{\mathcal{R}}(E_0 - E_\gamma^S) = S^{\mathcal{R}}(E_0) - E_\gamma^S \left(\frac{dS^{\mathcal{R}}}{dE} \right)_{E_0} + \frac{1}{2} (E_\gamma^S)^2 \left(\frac{d^2 S^{\mathcal{R}}}{dE^2} \right)_{E_0} + \dots = S^{\mathcal{R}}(E_0) - \beta E_\gamma^S + \dots \quad (16)$$

and

$$S^{\mathcal{R}}(E_0 - E_\tau^S) = S^{\mathcal{R}}(E_0) - E_\tau^S \left(\frac{dS^{\mathcal{R}}}{dE} \right)_{E_0} + \frac{1}{2} (E_\tau^S)^2 \left(\frac{d^2 S^{\mathcal{R}}}{dE^2} \right)_{E_0} + \dots = S^{\mathcal{R}}(E_0) - \beta E_\tau^S + \dots \quad (17)$$

Using the above results in equation 15 we get:

$$\frac{P(E_\gamma^S)}{P(E_\tau^S)} = \frac{e^{-\beta E_\gamma^S}}{e^{-\beta E_\tau^S}} \quad (18)$$

Thus, just like the case of the paramagnet in a field, we showed that for the general case of a system in contact with a heat reservoir the probability to find the system in any of its accessible states is

$$P(E_\alpha) = \frac{e^{-\beta E_\alpha}}{Z} \quad (19)$$

with the partition function given by:

$$Z(\beta) = \sum_{\{\alpha\}} e^{-\beta E_\alpha} = \sum_{\{\alpha\}} e^{-\frac{E_\alpha}{k_B T}} \quad (20)$$

Note that the system \mathcal{S} is not an isolated system. Thus, there is no conservation law for its energy. The quantity that is kept constant for this system is the temperature T which is regulated by the reservoir (sometimes called heat bath). We can however use Z to calculate the average energy of the system:

$$\langle E \rangle = \sum_{\{\alpha\}} E_\alpha P(E_\alpha) = \sum_{\{\alpha\}} E_\alpha \frac{e^{-\beta E_\alpha}}{Z} = \frac{1}{Z} \sum_{\{\alpha\}} E_\alpha e^{-\beta E_\alpha} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\left(\frac{\partial \log Z}{\partial \beta} \right)_{V,N} \quad (21)$$

where in the last equality we emphasize that the volume of the system and the number of particles are kept constant. To calculate the entropy of the system we remember that:

$$S = -k_B \sum_{\{\alpha\}} P(\alpha) \log P(\alpha) = -k_B \sum_{\{\alpha\}} P(\alpha) \log \frac{e^{-\beta E_\alpha}}{Z} = k_B \beta \sum_{\{\alpha\}} E_\alpha P(\alpha) + k_B \log Z \sum_{\{\alpha\}} P(\alpha) \quad (22)$$

Or:

$$S = -k_B \beta \left(\frac{\partial \log Z}{\partial \beta} \right)_{V,N} + k_B \log Z \quad (23)$$

In a similar manner we can calculate the heat capacity of the system:

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T} \right)_{V,N} = -k_B \beta^2 \frac{\partial}{\partial \beta} \langle E \rangle = k_B \beta^2 \left(\frac{\partial^2 \log Z}{\partial \beta^2} \right)_{V,N} \quad (24)$$

In the above discussion we introduced a new kind of distribution function over which our missing information, or entropy will be defined. The ensemble that describes this constant temperature distribution function is called the **canonical ensemble**.

3. Thermodynamics of the paramagnet in a field

Writing explicitly equation 11, the partition function of a system of N spins in a magnetic field H is:

$$Z = \left[\sum_{\epsilon_i} e^{-\beta \epsilon_i} \right]^N = \left[\sum_{\sigma_i = \pm 1} e^{-\beta(-\mu_0 H) \sigma_i} \right]^N = \left[e^{\beta \mu_0 H} + e^{-\beta \mu_0 H} \right]^N = 2^N [\cosh(\beta \mu_0 H)]^N \quad (25)$$

Thus

$$\langle E \rangle = -\mu_0 H N \tanh(\beta \mu_0 H) \quad (26)$$

Similarly, the magnetization of the system is:

$$M = \sum_{i=1}^N \langle \mu_i \rangle = \sum_{i=1}^N \sum_{\mu_i = \pm \mu_0} \mu_i P(\mu_i) = \frac{1}{Z} \sum_{i=1}^N \sum_{\mu_i = \pm \mu_0} \mu_i e^{-\beta \mu_i H} = \frac{1}{\beta} \frac{\partial \log Z}{\partial H} = N \mu_0 \tanh(\beta \mu_0 H) \quad (27)$$

and the entropy is

$$S = N k_B \log 2 + N k_B \log [\cosh(\beta \mu_0 H)] + k_B (\beta \mu_0 H) \tanh(\beta \mu_0 H) \quad (28)$$

Note that the entropy satisfies both expected limiting cases. For very high temperatures, as the system tends to become more random, $\beta \rightarrow 0$ and $S \rightarrow N k_B \log 2$.

For the low temperatures case, $T \rightarrow 0$, or $\beta \rightarrow \infty$, and

$$S = N k_B \log 2 + N k_B \log \left[\frac{e^{\beta \mu_0 H} + e^{-\beta \mu_0 H}}{2} \right] + N k_B (\beta \mu_0 H) \left[\frac{e^{\beta \mu_0 H} - e^{-\beta \mu_0 H}}{e^{\beta \mu_0 H} + e^{-\beta \mu_0 H}} \right] \quad (29)$$

which becomes

$$S \approx N k_B e^{-2\beta \mu_0 H} \quad (30)$$

which vanishes at zero temperatures in accordance with the laws of Thermodynamics.

Calculating the heat capacity we get:

$$C_V = N k_B (\beta \mu_0 H)^2 \frac{e^{2\beta \mu_0 H}}{(1 + e^{2\beta \mu_0 H})^2} \quad (31)$$

The two limits of the heat capacity are:

For $k_B T \ll \mu_0 H$:

In this case $\beta \mu_0 H \ll 1$ and

$$C_V \approx N k_B (\beta \mu_0 H)^2 = N k_B \left(\frac{\mu_0 H}{k_B T} \right)^2 \quad (32)$$

For $k_B T \gg \mu_0 H$:

In this case

$$C_V \approx N k_B (\beta \mu_0 H)^2 e^{-2\beta \mu_0 H} = N k_B \left(\frac{\mu_0 H}{k_B T} \right)^2 e^{-\frac{2\mu_0 H}{k_B T}} \quad (33)$$

Finally, the susceptibility of the system, i.e. the reaction of the system to the application of a magnetic field is given by

$$\chi = \left(\frac{\partial M}{\partial H} \right)_{H=0} \rightarrow \frac{\mu_0^2}{k_B T} \quad (34)$$

as $T \rightarrow 0$.

4. Helmholtz free energy

For the canonical ensemble introduced above we showed that the probability for a state $\{\alpha\}$ of the system is given by equation 7. The system, which is in contact with a heat reservoir at temperature T has an averaged energy E (from this point on we will drop the averaging symbols $\langle \rangle$ to denote averages of macroscopic quantities, unless introduced for the first time.) Taking the energy as a continuous variable, the probability to find the system at an energy between $E - \Delta E \leq E \leq E + \Delta E$ is:

$$P(E) \Delta E = \sum_{E_\alpha - \Delta E \leq E_\alpha \leq E_\alpha + \Delta E} P(E_\alpha) = \Gamma(E) \frac{e^{-\beta E}}{Z} \Delta E \quad (35)$$

where the second equality is the number of states within an energy ΔE around the energy of the system, E , or simply the degeneracy of that energy.

Using the fact that $S(E) = k_B \log \Gamma(E)$, equation 35 can be rewritten as:

$$P(E) = \frac{1}{Z} e^{-\beta E + \frac{1}{k_B} S} = \frac{1}{Z} e^{-\beta(E - TS)} \equiv \frac{1}{Z} e^{-\beta F} \quad (36)$$

Since E is the energy of the system, $P(E)$ is the most probable state for the system, i.e. it is maximum in that case that all variables except for the energy are fixed. However, for $P(E)$ to be maximum, the argument in the exponential has to be maximum, or, the new quantity that we defined:

$$F = E - TS \quad (37)$$

has to be minimum. This requirement is also self consistent with the definition of the temperature of the system, since the above requirement reads:

$$\delta(P(E)) = 0 = \delta\left(\frac{1}{Z}e^{-\beta(E-TS)}\right) = -\beta\left(\frac{1}{Z}e^{-\beta(E-TS)}\right) \cdot \delta[F(E)] \quad (38)$$

which requires:

$$\delta F = \delta E - T\delta S = 0 \quad (39)$$

or, since all quantities except for the energy were held constant:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,T,N} \quad (40)$$

We call F , the **Helmholtz free energy**, and from equation 22 we identify that:

$$F = -k_B T \log Z \quad (41)$$

5. Generalized force and Thermodynamic work

Our aim now is to express the thermodynamic work in terms of the partition function. First we note that the work performed by a system is related to the variation of external parameters. For example, in a gas the work is related to change in volume while in the case of a paramagnet the work is related to the change in magnetic field.

Assume that we have identified an external parameter that will, in most cases, depend on an external body, which affects the energies of the states of the system. For example, let us look again at the case of a particle in a box. We showed in lecture notes #1 that the energy levels of the particle in a box of size $L \times L \times L$ are:

$$\epsilon_{n_x, n_y, n_z} = \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 [n_x^2 + n_y^2 + n_z^2] = \frac{\hbar^2}{2m} \frac{\pi^2}{V^{2/3}} [n_x^2 + n_y^2 + n_z^2] \quad (42)$$

$(n_x, n_y, n_z = 0, \pm 1, \pm 2, \dots)$ which clearly depend on the volume of the system. Changing one length will immediately change the energies of the same quantum numbers.

Another example is the case of a paramagnet in a field. the energy levels of the single moment are:

$$\epsilon_\sigma = -\mu_0 H \sigma \quad (43)$$

$(\sigma = \pm 1)$ which depend on the magnetic field as an external; parameter.

Let us look at a general case for which the energy levels, E_α of the system depend on an external parameter X :

$$E_\alpha = E_\alpha(X) \quad (44)$$

The work done by changing the external parameter is related to the existence of a force:

$$F_\alpha = -\frac{\partial E_\alpha}{\partial X} \quad (45)$$

The force described in equation 45 is a generalized force. Also, the sign is the same as in Newtonian mechanics where the external parameter is the displacement (e.g.. changing the volume by displacing a piston). Suppose the system is in a state α . If the increase in X causes a decrease in energy, the system will accommodate this change by applying a force F_α on the external body, directed along X . In this case F_α will be positive and $\frac{\partial E_\alpha}{\partial X}$ is negative. For the force given by equation 45, the work done by the system in a state i on the external body, when the external coordinate changes from X to $X + dX$ is:

$$\delta W_\alpha = F_\alpha dX = -\frac{\partial E_\alpha}{\partial X} \cdot dX \quad (46)$$

the thermodynamic work performed by the macroscopic system will be:

$$\delta W = \sum_\alpha (\delta W_\alpha) P(E_\alpha) = \frac{1}{Z} \sum_\alpha (\delta W_\alpha) e^{-\beta E_\alpha} \quad (47)$$

Where in the last equality we used the probabilities of the canonical ensemble. Using now equation 46 we get

$$\delta W = \frac{1}{Z} \sum_\alpha (\delta W_\alpha) e^{-\beta E_\alpha} = \frac{1}{Z} \sum_\alpha \left(-\frac{\partial E_\alpha}{\partial X} \right) dX e^{-\beta E_\alpha} = \frac{1}{\beta} \frac{1}{Z} \frac{\partial}{\partial X} \sum_\alpha e^{-\beta E_\alpha} dX = \frac{1}{\beta} \frac{\partial \log Z}{\partial X} dX \quad (48)$$

Example: Paramagnet in a field:

In this case:

$$E_\alpha = \sum_{i=1}^N -\mu_i H = -\mu_0 H \sum_{i=1}^N \sigma_i \quad (49)$$

Thus

$$\delta W_\alpha = -\frac{\partial E_\alpha}{\partial H} dH = -\left(-\mu_0 \sum_{i=1}^N \sigma_i\right) dH \quad (50)$$

and using equation 47:

$$\delta W = -\frac{1}{Z} \sum_\alpha \frac{\partial E_\alpha}{\partial H} e^{-\beta E_\alpha} dH = \frac{1}{\beta} \frac{1}{Z} \frac{\partial Z}{\partial H} dH = \frac{1}{\beta} \frac{\partial \log Z}{\partial H} dH \quad (51)$$

Comparing this result to equation 27 we get:

$$\delta W = M dH \quad (52)$$

Example: Ideal gas in a box:

In this case:

$$E_\alpha = \sum_{i=1}^N \epsilon_i(n_x, n_y, n_z) \quad (53)$$

where ϵ_i is given in equation 42. Thus

$$\delta W_\alpha = -\frac{\partial E_\alpha}{\partial V} dV = -\left(-\frac{2}{3} \frac{1}{V} \sum_{i=1}^N \epsilon_i(n_x, n_y, n_z)\right) dV \quad (54)$$

and

$$\delta W = \frac{1}{Z} \sum_\alpha \left(\frac{2}{3} \frac{1}{V} \sum_{i=1}^N \epsilon_i(n_x, n_y, n_z)\right) e^{-\beta E_\alpha} dV = \frac{2}{3} \frac{1}{V} \sum_\alpha E_\alpha e^{-\beta E_\alpha} dV = \frac{2}{3} \frac{1}{V} E dV \quad (55)$$

where we identified the averaged energy of the system as the last equality. We use now results from the theory of ideal gas that while we are very familiar with, we will prove only later. This is that the equation of state of an ideal gas is:

$$PV = Nk_B T = \frac{2}{3}E \quad (56)$$

Thus we conclude that for an ideal gas, upon changing the volume of the gas

$$\delta W = PdV \quad (57)$$

Comparison now with equation 48 we find that

$$\delta W = \frac{1}{\beta} \frac{\partial \log Z}{\partial V} dV = PdV \quad (58)$$

and

$$P = \frac{1}{\beta} \frac{\partial \log Z}{\partial V} \quad (59)$$

6. Work and the laws of thermodynamics

When we previously discussed an **isolated system** we assumed that the energy of the system is fixed, and thus equilibrium required that the entropy of the system ($S = S(E, V) = k_B \log \Gamma(E)$ where Γ measures the number of available states to the system), is maximum. If we now allow the volume of the system to change, with no change of the number of accessible states, the entropy is constant. In that case we can look at the differential of the energy

$$dE(S, V) = \left(\frac{\partial E}{\partial V} \right)_S dV \quad (60)$$

At the same time, since the system is isolated and $S = \text{const.}$, conservation of energy requires

$$dE + \delta W = dE + PdV = 0 \quad (61)$$

Which implies

$$P = - \left(\frac{\partial E}{\partial V} \right)_S \quad (62)$$

Assume now that in general the entropy can change. However, let us look at a special case for which the entropy does not change. To achieve that we select dE and dV such that $dS = 0$. In this case:

$$dS(E, V) = 0 = \left(\frac{\partial S}{\partial E} \right)_V (\delta E)_S + \left(\frac{\partial S}{\partial V} \right)_E (\delta V)_S \quad (63)$$

where we emphasize that we chose dE and dV such that S does not change. The above equation gives:

$$\left(\frac{\delta E}{\delta V}\right)_S \left(\frac{\partial S}{\partial E}\right)_V + \left(\frac{\partial S}{\partial V}\right)_E = 0 \quad (64)$$

$$\left(\frac{\delta E}{\delta V}\right)_S = -\frac{\left(\frac{\partial S}{\partial V}\right)_E}{\left(\frac{\partial S}{\partial E}\right)_V} = -\frac{1}{T} \quad (65)$$

Which gives:

$$P = T \left(\frac{\partial S}{\partial V}\right)_E \quad (66)$$

In the general case, we calculate the differential of the entropy:

$$dS = \left(\frac{\partial S}{\partial E}\right)_V dE + \left(\frac{\partial S}{\partial V}\right)_E dV = \frac{1}{T}dE + \frac{P}{T}dV \quad (67)$$

or

$$TdS = dE + PdV = dE + \delta W \quad (68)$$

This modified conservation of energy is the first law of thermodynamics.

Classical mechanics tells us that conservation of energy requires that the change in energy of a system equals the work it does (equation 61). However, classical mechanics also assumes that we know everything about the system by solving its equations of motion with proper initial conditions. In the case that work done by the system not only changes its internal energy, but also changes the amount of missing information about the system, a new conservation of energy law exists, where the difference between the change in internal energy of the system and the work done by the system is the **heat** that represent the change in missing information of the system:

$$\delta Q \equiv TdS \quad (69)$$

The way we introduced δQ as the heat, also agrees with the second law of thermodynamics.

We finally note here that for a system that is described by the canonical ensemble, i.e. a system at a constant temperature T , and volume, V , equation 68 implies:

$$dE - TdS + \delta W = dF + \delta W = 0 \quad (70)$$

In words, the classical law of conservation of energy remains if we replace the internal energy of the system, E , by the free energy of the system F .

7. Properties of $F(T, V, N)$:

The Helmholtz free energy that we introduced above is minimum for a system at constant temperature and constant external parameters. It is therefore the fundamental function of the canonical ensemble and satisfies:

$$dF = d(E - TS) = dE - TdS - SdT = -SdT - \delta W \quad (71)$$

If we keep the temperature constant we get:

$$(dF)_T = -\delta W \quad (72)$$

and, if no work is done, the system is in equilibrium at a constant temperature if

$$dF = 0 \quad (73)$$

In general, since the variables that are kept constant are T and V , a small variation of F means that:

$$dF = \left(\frac{\partial F}{\partial T} \right)_{V,N} dT + \left(\frac{\partial F}{\partial V} \right)_{T,N} dV \quad (74)$$

At the same time:

$$dF = d(E - TS) = dE - TdS - SdT = (TdS - PdV) - TdS - SdT = -SdT - PdV \quad (75)$$

where we used the differential of dE from equation 60.

Comparing the two results we identify:

$$S = - \left(\frac{\partial F}{\partial T} \right)_V; \quad P = - \left(\frac{\partial F}{\partial V} \right)_T \quad (76)$$

Finally, since dF is an exact differential (see problem set #3), the cross second derivatives should be equal. We thus obtain the following “Maxwell Relation”:

$$\left(\frac{\partial S}{\partial V} \right)_T = \left(\frac{\partial P}{\partial T} \right)_V \quad (77)$$

8. Laws of Thermodynamics - first look

i) *Zeroth law:*

If two systems are in thermal equilibrium with a third system, they must be in thermal equilibrium with each other.

ii) *First law:*

Heat is a form of energy, and obey the conservation of energy:

$$dE + \delta W = \delta Q \quad (78)$$

[Here we distinguish quantities with an exact differential like the energy from quantities that do not have an exact differential like the heat Q and the work W].

iii) *Second law:*

If a closed system is in a configuration that is not the equilibrium one, the most probable consequence is that the entropy of the system will increase monotonically with time until the system reaches equilibrium with:

$$\delta Q = TdS \quad (79)$$

iv) *Third law:*

The entropy of a system approaches a constant value as the temperature approaches zero.

Lecture Notes #5

1. One-dimensional harmonic oscillator

As a first application of the formalism of the canonical ensemble we will discuss the one-dimensional harmonic oscillator. Assume a collection of N non-interacting one-dimensional harmonic oscillators. For example let us consider N identical masses connected to springs, with all spring constants, K are equal, and we measure the displacement of the oscillator x from its equilibrium position. We will next solve this problem both, quantum mechanically and classically.

In quantum mechanics the hamiltonian of the single oscillator is:

$$\hat{h} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \quad (1)$$

where $\omega = \sqrt{K/m}$ is the natural frequency of the oscillator.

Solving the Schrodinger equation for this hamiltonian we obtain the energy states of the oscillator that can be represented by a single quantum number $n = 0, 1, 2, \dots$:

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

Following the recipe for calculating thermodynamic properties in the canonical ensemble, we first find the partition function:

$$Z_1 = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n} = \sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})\hbar\omega} = e^{-\beta\frac{1}{2}\hbar\omega} \sum_{n=0}^{\infty} e^{-\beta n\hbar\omega} = \frac{e^{-\beta\frac{1}{2}\hbar\omega}}{1 - e^{-\beta\hbar\omega}} = \frac{1}{2\sinh\left(\frac{1}{2}\beta\hbar\omega\right)} \quad (3)$$

In the limit of high temperatures which represents the classical limit we can use $\sinh(x) \approx x$ and obtain the result:

$$Z_1 = (\beta\hbar\omega)^{-1} = \frac{k_B T}{\hbar\omega}, \quad \text{if } k_B T \gg \hbar\omega \quad (4)$$

The energy of the system of N oscillators is given by the energy of the single oscillator:

$$E = -\frac{\partial \log Z}{\partial \beta} = -\frac{\partial \log(Z_1)^N}{\partial \beta} = -N \frac{\partial \log Z_1}{\partial \beta} = N \left[\frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \right] \quad (5)$$

which in the high temperature limit could also be obtained directly from equation 4:

$$E \approx Nk_B T \quad \text{for} \quad k_B T \gg \hbar\omega \quad (6)$$

In the low temperature limit we obtain

$$E \approx N\frac{1}{2}\hbar\omega + N\hbar\omega e^{-\frac{\hbar\omega}{k_B T}} \rightarrow N\frac{1}{2}\hbar\omega \quad \text{as} \quad T \rightarrow 0 \quad (7)$$

Thus, at zero temperature the energy is just the sum of all “zero-point energies” of the individual oscillators.

The heat capacity can be calculated from the energy:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = - \frac{N\hbar\omega e^{\beta\hbar\omega} \left(-\frac{\hbar\omega}{k_B T^2} \right)}{(e^{\beta\hbar\omega} - 1)^2} = Nk_B (\beta\hbar\omega)^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} \quad (8)$$

the two limits of the heat capacity are:

For high temperatures (the classical limit), $k_B T \gg \hbar\omega$, and thus

$$C_V \approx Nk_B (\beta\hbar\omega)^2 \frac{1}{(1 + (\beta\hbar\omega) - 1)^2} = Nk_B \quad (9)$$

For low temperatures $k_B T \ll \hbar\omega$ and thus

$$C_V \approx Nk_B \left(\frac{\hbar\omega}{k_B T} \right)^2 e^{-\frac{\hbar\omega}{k_B T}} \quad (10)$$

which tends to zero exponentially as $T \rightarrow 0$.

Finally, an important quantity to calculate is $\langle n \rangle$, or, the average degree of excitation of the single oscillator. Using the partition function we get:

$$\langle n \rangle = \frac{1}{Z} \sum_{n=0}^{\infty} n e^{-\beta(n+\frac{1}{2})\hbar\omega} = \frac{1}{Z} \sum_{n=0}^{\infty} (n + \frac{1}{2}) e^{-\beta(n+\frac{1}{2})\hbar\omega} - \frac{1}{2} \frac{1}{Z} \sum_{n=0}^{\infty} e^{-\beta(n+\frac{1}{2})\hbar\omega} = -\frac{\partial \log Z}{\partial(\beta\hbar\omega)} - \frac{1}{2} \quad (11)$$

which gives:

$$\langle n \rangle = \frac{1}{e^{\beta\hbar\omega} - 1} \quad (12)$$

This is nothing but the “Bose-Einstein” distribution which we will discuss in details when we will study quantum gasses. an important feature to note from the result for $\langle n \rangle$ is that the average energy of a single oscillator can also be written as:

$$\langle \epsilon \rangle = \left(\langle n \rangle + \frac{1}{2} \right) \hbar \omega \quad (13)$$

a result we could expect from the averaged energy calculated before.

2. Einstein model of the Solid

In his attempt to understand the specific heat of simple solids, Einstein invented the following model for the vibrations of a solid. In this model the crystal is made up of N atoms placed on the sites of a crystal lattice, and are connected to their neighbors with springs. The springs represent the binding forces that hold the crystal together and we assume that they are all equal and that they all vibrate about an equilibrium position which is the lattice point. To make the model simple Einstein assumed further that all the masses are the same and that each mass represents three one-dimensional harmonic oscillators associated with the three lattice axis. The result is therefore a model of $3N$ identical harmonic oscillators, all of which have the same frequency, ω .

The solution to the Einstein model follows directly from the previous section. The energy of the solid is:

$$E = 3N \left[\frac{1}{2} \hbar \omega + \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \right] \quad (14)$$

which at high temperatures gives:

$$\langle E \rangle \simeq 3N k_B T, \quad (15)$$

which the result we would get from the equipartition theorem for $3N$ oscillators.

Define the Einstein temperature, $\theta_E = \hbar \omega / k_B$, the result for the specific heat is then:

$$c = 3R \left(\frac{\theta_E}{T} \right)^2 \frac{e^{\theta_E/T}}{[e^{\theta_E/T} - 1]^2} \quad (16)$$

Where we defined the specific heat as the heat capacity per mole, and R is the gas constant ($R = N_a k_B = 8.314 \text{ joule/mole-K}$, where $N_a = 6.02 \times 10^{23} \text{ molecules/mole}$ is the Avogadro number).

The high temperature limit of the Einstein specific heat gives:

$$c = 3R \quad (17)$$

which is the so-called ‘‘Dulong-Petit’’ law for the specific heat of the solid.

It is important to note that while the Einstein model fails to describe real solids at low temperatures, it does give partial description for the some vibrational modes of complex

solids. the failure at low temperatures is a direct consequence of the neglect of interactions among the masses through the springs (i.e. collective modes). This discrepancy is solved in the Debye theory of solids that we will discuss in a future lecture.

3. Statistical Mechanics of Classical Systems

Adopting the canonical ensemble to classical system involves going from discrete to continuous variables, and thus from discrete probabilities to probability density functions.

Given the particles position and momenta, the energy of a state of the system of N particles is:

$$E(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \sum_{i=1}^N \frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \quad (18)$$

If the particles do not interact it is possible to write the potential energy as a sum of single particle potential energies. In this case:

$$E(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \sum_{i=1}^N \left[\frac{p_i^2}{2m} + u_i(\vec{r}_i) \right] \equiv \sum_{i=1}^N \epsilon_i(\vec{p}_i, \vec{r}_i) \quad (19)$$

For a system in contact with a heat reservoir at a constant temperature T (i.e. a system in the canonical ensemble), the probability density of finding the system in a state with momenta and positions between $(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ and $(\vec{p}_1 + d\vec{p}_1, \vec{p}_2 + d\vec{p}_2, \dots, \vec{p}_N + d\vec{p}_N, \vec{r}_1 + d\vec{r}_1, \vec{r}_2 + d\vec{r}_2, \dots, \vec{r}_N + d\vec{r}_N)$ would be proportional to the Boltzman factor:

$$P(\vec{p}_1, \dots, \vec{p}_N, \vec{r}_1, \dots, \vec{r}_N) d^3 p_1 \dots d^3 p_N d^3 r_1 \dots d^3 r_N \propto e^{-\beta \sum_{i=1}^N \left[\frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \right]} d^3 p_1 \dots d^3 p_N d^3 r_1 \dots d^3 r_N \quad (20)$$

and for the case of non-interacting particles, the probability density that the single particle is found in a state with momentum and position between (\vec{p}, \vec{r}) and $(\vec{p} + d\vec{p}, \vec{r} + d\vec{r})$ is:

$$P(\vec{p}, \vec{r}) d^3 p d^3 r \propto e^{-\beta \left[\frac{p^2}{2m} + u(\vec{r}) \right]} d^3 p d^3 r \quad (21)$$

(note that for a d-dimensional space the infinitesimal volumes in phase space will be $d^d p$ and $d^d r$).

Following the recipe for discrete probabilities, the classical partition function is:

$$Z = C \int d^3 p_1 \dots \int d^3 p_N \int d^3 r_1 \dots \int d^3 r_N e^{-\beta \sum_{i=1}^N \left[\frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \right]} \quad (22)$$

Where we introduced the constant C to ensure that the partition function is dimensionless. In fact, it is easy to see that by defining a momentum scale p_0 , and a position scale r_0 , the constant C can be written as $C = (p_0 r_0)^{-3N}$.

For non-interacting particles the integrals in equation 22 is a product of N independent integrals, and thus the partition function of the N particles is $Z = [Z_1]^N$ with

$$Z_1 = C_1 \int d^3p \int d^3r e^{-\beta \left[\frac{p^2}{2m} + u(\vec{r}) \right]} \quad (23)$$

Where, using the above notations: $C_1 = (p_0 r_0)$.

The partition function defined above allows us to write the probability for a state of the sample, i.e. the probability that the sample will be found in a state with momenta and positions between $(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N, \vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ and $(\vec{p}_1 + d\vec{p}_1, \vec{p}_2 + d\vec{p}_2, \dots, \vec{p}_N + d\vec{p}_N, \vec{r}_1 + d\vec{r}_1, \vec{r}_2 + d\vec{r}_2, \dots, \vec{r}_N + d\vec{r}_N)$.

$$P(\vec{p}_1, \dots, \vec{p}_N, \vec{r}_1, \dots, \vec{r}_N) d^3p_1 \dots d^3p_N d^3r_1 \dots d^3r_N = \frac{C}{Z} e^{-\beta \sum_{i=1}^N \left[\frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \right]} d^3p_1 \dots d^3p_N d^3r_1 \dots d^3r_N \quad (24)$$

and for non-interacting particles:

$$P(\vec{p}, \vec{r}) d^3p d^3r = \frac{C_1}{Z_1} e^{-\beta \left[\frac{p^2}{2m} + u(\vec{r}) \right]} d^3p d^3r \quad (25)$$

The above probability allows us to calculate averages in a similar way that we did for discrete probabilities. For example, the average of $A(\vec{p}, \vec{r})$ for a non-interacting particle is:

$$\langle A \rangle = \frac{1}{Z_1} C_1 \int d^3p \int d^3r A(\vec{p}, \vec{r}) e^{-\beta \left[\frac{p^2}{2m} + u(\vec{r}) \right]} \quad (26)$$

Example: Gas in a gravitational field

As an example let us consider a gas of non-interacting identical particles in a gravitational field. This problem can represent the distribution of gas in our atmosphere. Assuming a column of gas with gravity acting along the $-z$ direction. Let us measure the potential energy due to gravity from the bottom of the column (i.e. the surface of the earth). The energy of a gas molecule is:

$$\epsilon(\vec{p}, \vec{r}) = \frac{p^2}{2m} + mgz \quad (27)$$

The probability to find a molecule with height between z and $z + dz$ is:

$$P(z)dz = \frac{\int d^3p \int dx dy e^{-\beta \left[\frac{p^2}{2m} + mgz \right]}}{\int d^3p \int d^3r e^{-\beta \left[\frac{p^2}{2m} + mgz \right]}} = \frac{e^{-\beta mgz}}{\int_0^\infty dz e^{-\beta mgz}} = \left(\frac{k_B T}{mg} \right) e^{-\beta mgz} \quad (28)$$

Example: One-dimensional harmonic oscillator

Let us come back now to the system of N harmonic oscillators. For a system of non-interacting oscillators the energy of a single oscillator is given by

$$\epsilon(p, x) = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \quad (29)$$

where again, $\omega = \sqrt{K/m}$ is the natural frequency of the oscillator.

Before we continue with the calculation, we want to introduce the concept of **phase space** which is the $6N$ -space representing $3N$ momentum, and $3N$ position coordinates of the N particles of the system. A state of the system is therefore represented by an infinitesimal hypercube of size $d^{3N}p d^{3N}r$ around the $(\vec{p}_1, \dots, \vec{p}_N, \vec{r}_1, \dots, \vec{r}_N)$ of the system.

For non-interacting systems it is enough to look at the 6-dimensional space of a single particle, and for the one dimensional harmonic oscillator phase space has two coordinates. A contour of constant energy in this $p - x$ space is an ellipse satisfying the canonical equation:

$$\frac{p^2}{(\sqrt{2m\epsilon})^2} + \frac{x^2}{(\sqrt{2\epsilon/m\omega^2})^2} = 1 \quad (30)$$

The classical partition function for the harmonic oscillator is

$$Z_1 = C_1 \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx e^{-\beta \left(\frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \right)} \quad (31)$$

The above integral is easily calculated as two gaussian integrals to give:

$$Z_1 = C_1 \sqrt{2\pi \frac{m}{\beta}} \cdot \sqrt{2\pi \frac{1}{\beta m \omega^2}} = C_1 \frac{2\pi}{\omega} k_B T \quad (32)$$

Comparing this result to the classical limit of the quantum calculation for the same Z_1 , we identify the constant C_1 and find that its origin is pure quantum mechanical:

$$C_1 = (p_0 x_0)^{-1} = h^{-1} \quad (33)$$

Which remind us of the uncertainty principle as it really represents the scale of an elementary cell size in phase space.

Following the recipes that we developed above, and using $Z = [Z_1]^N$, we find that:

$$F = -k_B T \log Z = -N k_B T \log \left(\frac{2\pi m k_B T}{h} \right) \quad (34)$$

The entropy of the system is:

$$S = - \left(\frac{\partial F}{\partial T} \right)_V = N k_B \log \left(\frac{2\pi m k_B T}{h} \right) + N k_B \quad (35)$$

The energy of the system can now be calculated from either $E = F + TS$, or from the derivative of the partition function to give us the result

$$E = N k_B T \quad (36)$$

independent of the properties of the oscillators and the same as we got in the high temperatures limit of the quantum mechanical calculation.

Lecture Notes #6

1. Classical ideal gas - classical approach

For a classical ideal gas on non-interacting molecules the energy of the single molecule is:

$$\epsilon(\vec{p}, \vec{r}) = \frac{p^2}{2m}, \quad \text{and} \quad 0 \leq x, y, z \leq L \quad (1)$$

The volume of the container is therefore $V = L \times L \times L = L^3$

The single-molecule partition function is

$$Z_1 = C \int_{all \vec{p}} d^3p \int_{0 \leq x, y, z \leq L} d^3r e^{-\beta \frac{p^2}{2m}} = CV \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z e^{-\beta \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)} \quad (2)$$

Where the integral over space yielded simply the volume of the container. The result is a product of three identical gaussian integrals. Thus we obtain:

$$Z_1 = CV \left(\int_{-\infty}^{\infty} dp_i e^{-\beta \frac{p_i^2}{2m}} \right)^3 = CV (2\pi m k_B T)^{3/2} \quad (3)$$

Using the above partition function, the average energy of a classical particle in an ideal gas is:

$$\langle \frac{p^2}{2m} \rangle = \frac{1}{Z_1} C \int_{all \vec{p}} d^3p \int_{0 \leq x, y, z \leq L} d^3r \frac{p^2}{2m} e^{-\beta \frac{p^2}{2m}} = -\frac{\partial \log Z_1}{\partial \beta} = -\frac{\partial}{\partial \beta} \log [CV (2\pi m k_B T)^{3/2}] \quad (4)$$

Thus

$$\langle \epsilon \rangle = \frac{3}{2} k_B T \quad (5)$$

and

$$E = \langle \sum_{i=1}^N \frac{p_i^2}{2m} \rangle = N \langle \frac{p^2}{2m} \rangle = \frac{3}{2} N k_B T \quad (6)$$

The pressure of the ideal gas is calculated from the partition function:

$$P = \frac{1}{\beta} \frac{\partial \log Z}{\partial V} = N k_B T \frac{\partial \log Z_1}{\partial V} = \frac{N k_B T}{V} \quad (7)$$

which gives the well known equation of state:

$$PV = Nk_B T \quad (8)$$

The entropy is given by:

$$S = - \left(\frac{\partial F}{\partial T} \right)_V = Nk_B \left\{ \log [CV(2\pi mk_B T)^{3/2}] + \frac{3}{2} \right\} \quad (9)$$

The heat capacity is given by

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \left(\frac{\partial(F + TS)}{\partial T} \right)_V = \left(\frac{\partial E}{\partial T} \right)_V + S + T \left(\frac{\partial S}{\partial T} \right)_V = T \left(\frac{\partial S}{\partial T} \right)_V \quad (10)$$

and

$$C_V = \frac{3}{2} Nk_B \quad (11)$$

It is convenient to consider the specific heat as an intrinsic property of the gas. For example, the specific heat of a single molecule is the heat capacity per molecule of the gas:

$$c_V = \frac{C_V}{N} = \frac{3}{2} k_B \quad (12)$$

We can also discuss the specific heat per mole gas. In that case the number of moles of gas is the number of molecules divided by the Avogadro number, $N_a = 6.02 \times 10^{23} \text{ molecules/mole}$.

$$C_V = \frac{3}{2} Nk_B = \frac{3}{2} \frac{N}{N_a} N_a k_B \equiv \frac{3}{2} \nu RT \quad (13)$$

Where ν is the number of moles in the gas and $R = N_a k_B = 8.3 \text{ Joule/mole}$ is the gas constant.

In that case we write for the molar specific heat:

$$c_V = \frac{3}{2} R \quad (14)$$

Coming back to equation 10, we realize that the heat capacity at constant pressure is simply:

$$C_P = T \left(\frac{\partial S}{\partial T} \right)_P \quad (15)$$

The entropy at constant pressure can be easily obtained from equation 9 and substituting the equation of state for the volume.

$$S = Nk_B \left\{ \log \left[C \frac{Nk_B T}{P} (2\pi m k_B T)^{3/2} \right] + \frac{3}{2} \right\} = Nk_B \left\{ \log \left[C \frac{Nk_B}{P} (2\pi m k_B)^{3/2} T^{5/2} \right] + \frac{3}{2} \right\} \quad (16)$$

and

$$C_P = T \left(\frac{\partial S}{\partial T} \right)_P = \frac{5}{2} Nk_B \quad (17)$$

The specific heat at constant pressure per molecule is therefore

$$c_P = \frac{C_P}{N} = \frac{5}{2} k_B = c_V + k_B \quad (18)$$

Finally, we apply the above formalism to the case of a gas that is a mixture of M different molecules, each of mass m_j where each component has N_j molecules. Since the particles do not interact, it is easy to see that the total partition function is a product of single particle partition functions, each raised to the power of its number of particles.

$$Z = \prod_{j=1}^M [Z_1^j]^{N_j} \quad (19)$$

The Free energy is now:

$$F = -k_B T \sum_{j=1}^M N_j \log Z_1^j \quad (20)$$

and

$$P = \frac{1}{\beta} \frac{\partial \log Z}{\partial V} = k_B T \sum_{j=1}^M N_j \frac{\partial \log Z_1^j}{\partial V} = k_B T \sum_{j=1}^M N_j \frac{1}{V} = \sum_{j=1}^M P_j \quad (21)$$

This is the so-called Dalton law that says that the pressure of a gas that is made up of a mixture of several gasses is the sum of all “partial pressures” of all the components.

2. Classical ideal gas - quantum approach

In the quantum mechanical approach we start with solving the Schrodinger equation of a particle in a box.

$$\frac{1}{2m} \left(\frac{\nabla}{i} \right)^2 \Psi = \epsilon \Psi \quad (22)$$

with the potential energy

$$U(x, y, z) = \begin{cases} 0 & 0 \leq x, y, z \leq L \\ \infty & \text{otherwise} \end{cases} \quad (23)$$

The potential energy is simply the box in which we place the gas, thus it imposes the boundary condition on the wave function to be zero on the walls. It is easy to see that the Schrodinger equation can be separated into three equations for each of the three coordinates. Thus:

$$\Psi = \varphi_x(x)\varphi_y(y)\varphi_z(z) \quad (24)$$

The x component for example we try the solution:

$$\varphi_x(x) = A\sin(k_x x) + B\cos(k_x x) \quad (25)$$

and using $\varphi_x(0) = \varphi_x(L) = 0$, we get the solutions $B = 0$, and

$$k_x L = n_x \pi; \quad n_x = 0, 1, 2, \dots \quad (26)$$

Note that in principle negative integers will work as well, however, they will not produce a new set of eigenstates [i.e. $\sin(k_x x) = -\sin(-k_x x)$]. Thus, only positive integers also imply only positive k_x, k_y and k_z .

A similar solution also exists for the y and z components. Using the above wave-functions in the Schrodinger equation we find:

$$\epsilon_{n_x, n_y, n_z} = \frac{\hbar^2}{2m} \left(\frac{\pi}{L} \right)^2 [n_x^2 + n_y^2 + n_z^2] \quad (27)$$

The single-particle partition function is:

$$Z_1 = \sum_{n_x} \sum_{n_y} \sum_{n_z} e^{-\beta \frac{\hbar^2}{2m} \left(\frac{\pi}{L} \right)^2 [n_x^2 + n_y^2 + n_z^2]} = \sum_{k_x} \sum_{k_y} \sum_{k_z} e^{-\beta \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2)} \quad (28)$$

Density of States:

The sums are difficult to calculate. However, we realize that since we are interested in the classical limit, i.e. high-temperature limit, the contributing Boltzmann factors are those with large n_i s. Thus, in that limit $\Delta n_i = 1 \ll n_i$ for $i = x, y, z$. We therefore approximate:

$$\Delta n_x \Delta n_y \Delta n_z = \left(\frac{L}{\pi} \right)^3 \Delta k_x \Delta k_y \Delta k_z \rightarrow \left(\frac{L}{\pi} \right)^3 d^3 k \quad (29)$$

We finally note that the Boltzmann factors depend quadratically on the wave-vector. Thus, we can change the above requirement of only positive k_i s to $-\infty < k_i < \infty$ by introducing a factor of $\frac{1}{2}$ for each component.

We therefore conclude that

$$\Delta n_x \Delta n_y \Delta n_z \rightarrow \left(\frac{L}{2\pi}\right)^3 d^3k = \frac{V}{8\pi^3} 4\pi k^2 dk = \frac{V}{2\pi^3} k^2 dk \equiv \rho k dk \quad (30)$$

Where $\rho(k)$ is the density of states. Note that since the energies that appear in the Boltzmann factors are all quadratic in k , it is more convenient to use spherical coordinates for which $dk_x dk_y dk_z = 4\pi k^2 dk$.

Equation 28 now reads:

$$Z_1 = \frac{V}{2\pi^2} \int_0^\infty k^2 e^{-\beta \frac{\hbar^2}{2m} k^2} dk = \frac{V}{h^3} (2\pi m k_B T)^{3/2} \quad (31)$$

Comparing now the above quantum calculation to the classical one, we identify the constant C_1 as:

$$C_1 = h^{-3} \quad (32)$$

Exactly the same as we found for the harmonic oscillator case.

3. The Gibbs paradox

Now that we saw that the classical ideal gas can be calculated from knowing the quantum states of the individual particles, we can examine the results we obtained. In particular we are interested in the behavior of the extensive quantities that describe the system such as the free energy:

$$F = -k_B T \log Z = -N k_B T \log \left[\frac{V}{h^3} (2\pi m k_B T)^{3/2} \right] \quad (33)$$

and the entropy

$$S = - \left(\frac{\partial F}{\partial T} \right)_V = -N k_B \log \left[\frac{V}{h^3} (2\pi m k_B T)^{3/2} \right] + \frac{3}{2} N k_B \quad (34)$$

Two problems can be found when we analyze these expressions. First we note that the above equations do not describe a true extensive quantity. This is because if we increase our system by some scale factor b , i.e. $V' = bV$, $N' = bN$, we expect these extensive quantities to be: $F' = bF$, and $S' = bS$. However, what we obtain is:

$$F' = F'(T, V', N') = F'(T, bV, bN) = bF(T, V, N) - bNk_B T \log b \quad (35)$$

and

$$S' = S'(T, V', N') = S'(T, bV, bN) = bS(T, V, N) - bNk_B \log b \quad (36)$$

The second problem with the above expressions for F and S arise when we consider the third law of thermodynamics. According to the third law the entropy should vanish as the temperature tends to zero. However, our result for the ideal gas predict that the entropy diverges as the temperature decreases towards zero. In principle this problem will not be solved until we will consider the gasses as true quantum gasses. We will therefore concentrate on solving the extensivity problem, or as it is also called, the “Gibbs paradox”.

The origin of the Gibbs paradox is in the fact that the gas is made of N molecules that are identical. Therefore, in bringing together two containers of the same gas at the same temperature T , there should not be any entropy associated with the mixing since we cannot tell the difference between molecules that came from one container or the other. Mathematically it means that for N particles, it is possible to permute their positions $N!$ times without changing the state of the system. To correct for this mistake we have to divide the partition function by this number, or:

$$Z = \frac{Z_1^N}{N!} \quad (37)$$

Note that the above modification for Z does not affect the calculation of thermal averages. this is because

$$\log Z = N \log Z_1 - \log(N!) \quad (38)$$

and any direct derivative of $\log Z$ with respect to β (to calculate E), or with respect to V (to calculate P) will ignore the new term.

With the above correction, the free energy is:

$$F = -Nk_B T \left[\log V + \log \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] + k_B T (N \log N - N) \quad (39)$$

where we used the Stirling approximation for $\log N!$.

Collecting all the terms, the free energy can be written as:

$$F = -Nk_B T \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{2\pi m k_B T}{h^2} \right) + 1 \right] \quad (40)$$

The entropy is now:

$$S = Nk_B \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{2\pi m k_B T}{h^2} \right) + \frac{5}{2} \right] \quad (41)$$

Both the free energy and entropy are proportional to the number of particles, i.e. the size of the system, while the dependence on the volume comes now through a dependence on the density which is an intensive quantity.

4. Entropy of mixing

The above result can now be generalized to a mixture of several gases each made of N_j molecules of mass m_j in the same volume V . The corrected partition function is:

$$Z = \prod_{j=1}^M \frac{[Z_1^j]^{N_j}}{N_j!} \quad (42)$$

The Free energy is now:

$$F = -k_B T \sum_{j=1}^M N_j \left[\log \left(\frac{V}{N_j} \right) + \frac{3}{2} \log \left(\frac{2\pi m_j k_B T}{h^2} \right) + 1 \right] \quad (43)$$

and the entropy:

$$S = k_B \sum_{j=1}^M N_j \left[\log \left(\frac{V}{N_j} \right) + \frac{3}{2} \log \left(\frac{2\pi m_j k_B T}{h^2} \right) + \frac{5}{2} \right] \quad (44)$$

Let us consider now two containers, each of volume V and equal number of particles N , but each has a different gas with different mass m_j . We now open a partition between the two gases and let them mix.

The entropy of the system before the mixing was:

$$S = k_B \sum_{j=1}^2 N \left[\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{2\pi m_j k_B T}{h^2} \right) + \frac{5}{2} \right] \quad (45)$$

while after the mixing it is

$$S' = k_B \sum_{j=1}^2 N \left[\log \left(\frac{2V}{N} \right) + \frac{3}{2} \log \left(\frac{2\pi m_j k_B T}{h^2} \right) + \frac{5}{2} \right] \quad (46)$$

The difference in entropy is now:

$$\Delta S = S' - S = 2Nk_B \log 2 \quad (47)$$

This is the entropy of mixing and it arises because the removal of the partition and the mixing of the two gases is not a reversible process. It is different than the case if the two gases were the same, as we discussed above.

5. The Chemical potential

When discussing the canonical ensemble we introduced the concept of temperature in a natural manner, through the requirement that the condition for equilibrium between two isolated systems, that are in thermal contact and can thus exchange energy, be the state for which

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2}; \quad \text{or} \quad \frac{1}{T_1} = \frac{1}{T_2} \quad (48)$$

We now want to ask a similar question concerning the number of particles. That is to say, what if we allow also for particles to exchange between the two systems. Assume again two isolated systems, \mathcal{S}_1 , and \mathcal{S}_2 . The two systems will be described by their energies E_1 and E_2 , by their respective volumes V_1 and V_2 , and by their number of particles N_1 and N_2 . The total entropy of the composite system will depend on all six variables. However, only three are going to be independent since:

$$\begin{aligned} E &= E_1 + E_2 \\ V &= V_1 + V_2 \\ N &= N_1 + N_2 \end{aligned} \quad (49)$$

From the requirement of maximum entropy we obtain the following conditions:

$$dS = \left[\left(\frac{\partial S_1}{\partial E_1} \right) dE_1 + \left(\frac{\partial S_2}{\partial E_2} \right) (-dE_1) \right] + \left[\left(\frac{\partial S_1}{\partial V_1} \right) dV_1 + \left(\frac{\partial S_2}{\partial V_2} \right) (-dV_1) \right] + \left[\left(\frac{\partial S_1}{\partial N_1} \right) dN_1 + \left(\frac{\partial S_2}{\partial N_2} \right) (-dN_1) \right] \quad (50)$$

and since $dS = 0$, each of the differentials in the square parentheses should equal zero separately. We get the following conditions for equilibrium:

$$\begin{aligned} \left(\frac{\partial S_1}{\partial E_1} \right) &= \left(\frac{\partial S_2}{\partial E_2} \right) & \text{or} & \quad \frac{1}{T_1} = \frac{1}{T_2} \\ \left(\frac{\partial S_1}{\partial V_1} \right) &= \left(\frac{\partial S_2}{\partial V_2} \right) & \text{or} & \quad \frac{P_1}{T_1} = \frac{P_2}{T_2} \\ \left(\frac{\partial S_1}{\partial N_1} \right) &= \left(\frac{\partial S_2}{\partial N_2} \right) & \text{or} & \quad \frac{-\mu_1}{T_1} = \frac{-\mu_2}{T_2} \end{aligned} \quad (51)$$

The first equation describes the equality of the temperature and the second describes the equality of the pressure between the two systems. The third equality involves a new quantity, the chemical potential, $\mu = -T \left(\frac{\partial S}{\partial N} \right)_{E,V}$

Note that in general, since $S = S(E, V, N)$:

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN \quad (52)$$

Let $V = \text{const.}$ for the process under discussion. In that case $dV = 0$. Next let us select dS , dE , and dN such that $dT = 0$ (i.e. $T = \text{const.}$). For this situation:

$$(\delta S)_T = \left(\frac{\partial S}{\partial E}\right)_N (\delta E)_T + \left(\frac{\partial S}{\partial N}\right)_N (\delta N)_T \quad (53)$$

and

$$\frac{(\delta S)_T}{(\delta N)_T} (\delta E)_T = \left(\frac{\partial S}{\partial E}\right)_N \frac{(\delta E)_T}{(\delta N)_T} + \left(\frac{\partial S}{\partial N}\right)_E \quad (54)$$

Which can also be written as:

$$\left(\frac{\partial S}{\partial N}\right)_{T,V} = \frac{1}{T} \left(\frac{\partial E}{\partial N}\right)_{T,V} + \left(\frac{\partial S}{\partial N}\right)_{E,V} \quad (55)$$

and by multiplying by T and moving the terms we get:

$$T \left(\frac{\partial S}{\partial N}\right)_{E,V} = T \left(\frac{\partial S}{\partial N}\right)_{T,V} - \left(\frac{\partial E}{\partial N}\right)_{T,V} = \left[\frac{\partial}{\partial T} (TS - E) \right]_{T,V} = - \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad (56)$$

Which lead us to the relation for the chemical potential:

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \quad (57)$$

The above relation shows that the chemical potential is an intensive quantity which has dimensions of energy.

6. The Chemical potential of an ideal gas

Starting from the free energy of an ideal gas, now corrected for the Gibbs paradox we calculate:

$$\mu = \frac{\partial}{\partial N} \left[-Nk_B T \left\{ \log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{2\pi m k_B T}{h^2} \right) + 1 \right\} \right]_{T,V} \quad (58)$$

Which gives:

$$\mu = k_B T \log \left[n \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right] \quad (59)$$

Where we used $n = N/V$ as the density of the gas. Note that upon exponentiating this expression we get:

$$n = \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} e^{\frac{\mu}{k_B T}} \quad (60)$$

We call the quantity $\lambda = e^{\frac{\mu}{k_B T}}$ the activity, and its inverse, $y = e^{-\frac{\mu}{k_B T}}$ the fugacity of the system.

Lecture Notes #7

1. A system at constant temperature, pressure and number of particles

While the Helmholtz free energy is an extensive quantity proportional to the size of the system, we found that:

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} \quad \text{but} \quad \mu \neq \frac{F}{N} \quad (1)$$

Inspection of this observation shows that the reason is that $F = F(T, V, N)$, thus depending on two extensive variables, V and N . To have μ as a free energy per particle we need to change variables, such except for the dependence on N , the rest of the “knobs” we have will all be intensive variables, this can easily done if we change the dependence on volume, to the dependence on its control variable, the pressure. Starting with

$$dF = \left(\frac{\partial F}{\partial T} \right)_{V,N} dT + \left(\frac{\partial F}{\partial V} \right)_{T,N} dV + \left(\frac{\partial F}{\partial N} \right)_{T,V} dN = -SdT - PdV + \mu dN \quad (2)$$

For a system at constant pressure ($P = \text{const.}$) we can write:

$$PdV = d(PV) - VdP \quad (3)$$

and thus

$$dF = -SdT - PdV + \mu dN = -SdT - [d(PV) - VdP] + \mu dN = -SdT + VdP + \mu dN - d(PV) \quad (4)$$

or:

$$d[F + PV] = -SdT + VdP + \mu dN \quad (5)$$

Define a new free energy, $G = G(T, P, N) = F + PV$ which we call the **Gibbs free energy**. From the dependence on its variables we write:

$$dG = \left(\frac{\partial G}{\partial T} \right)_{P,N} dT + \left(\frac{\partial G}{\partial P} \right)_{T,N} dP + \left(\frac{\partial G}{\partial N} \right)_{T,P} dN = -SdT + VdP + \mu dN \quad (6)$$

We therefore find:

$$-S = \left(\frac{\partial G}{\partial T} \right)_{P,N}; \quad V = \left(\frac{\partial G}{\partial P} \right)_{T,N}; \quad \mu = \left(\frac{\partial G}{\partial N} \right)_{T,P} \quad (7)$$

G was obtained by a Legendre transformation from the Helmholtz free energy. Thus, dG is an exact differential, and from the equality of the cross derivative we obtain another set of Maxwell relations:

$$-\left(\frac{\partial S}{\partial P} \right)_{T,N} = \left(\frac{\partial V}{\partial T} \right)_{P,N} \quad (8)$$

Coming back to the dependence of G on the number of particles, N , we note that while

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{T,P} \quad (9)$$

we also require that $G(T,P,N)$ will be an extensive quantity. However, since the only parameter that represents the size of the system upon which G depends on is N , we conclude that:

$$G(T, P, N) = N \cdot g(T, P) \quad (10)$$

Comparing the above equations we conclude that

$$\mu = g(T, P) = \frac{G(T, P, N)}{N} \quad (11)$$

For a single component system, the chemical potential is the Gibbs free energy per particle. Since the temperature T and the number of particles N are constant, the Gibbs free energy is still built upon the Canonical ensemble.

2. A system at constant temperature, volume and chemical potential

Following similar arguments as above, one can imagine leaving as the only extensive variable upon which the free energy depends on as the volume V . In that case we will want to have the chemical potential fixed, instead of the number of particles. Starting from equation 2, we note that for a fixed chemical potential we can write:

$$\mu dN = d(\mu N) - Nd\mu \quad (12)$$

Substituting in equation 2 we get:

$$dF = -SdT - PdV + [d(\mu N) - Nd\mu] = -SdT - PdV - Nd\mu + d(\mu N) \quad (13)$$

or:

$$d[F - \mu N] = -SdT - PdV - Nd\mu \quad (14)$$

Define now a new free energy $\Omega(T, V, \mu) = F - \mu N = E - TS - \mu N$, which we call the grand-canonical free energy (for reasons to be explained later). From the dependence on its variable we write:

$$d\Omega = \left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} dT + \left(\frac{\partial\Omega}{\partial V}\right)_{T,\mu} dV + \left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} d\mu = -SdT - PdV + Nd\mu \quad (15)$$

We therefore find:

$$S = -\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu}; \quad P = -\left(\frac{\partial\Omega}{\partial V}\right)_{T,\mu}; \quad N = -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} \quad (16)$$

3. The grand canonical ensemble

Trying to examine the new free energy that we introduced, we notice that we may encounter problems in calculating the partition function if we still adopt the usual Boltzmann factor approach. This is because since N is not fixed, we cannot determine a state of the system if we do not know how many particles there are. To overcome this problem, let us go back to the lecture notes where we introduced the canonical ensemble. We will use similar argument here as we introduce two systems which are in thermal contact (same temperature T) and in diffusive contact (same chemical potential μ).

Assume now that while both systems are macroscopic in size, one system is much larger than the other. We call the larger of the two systems, **reservoir** (denoted by \mathcal{R}), and we connect our small system \mathcal{S} to it. The reservoir will be taken as much larger than our system by considering all extensive properties of the systems, such as the number of particles (or volume), the energy, the entropy, etc. Denoting these quantities with subscripts R or S , we require for examples that the energies of the two systems will satisfy: $E^R \gg E^S$. The total system, $\mathcal{R} + \mathcal{S}$ is an isolated system with energy $E_0 = E^R + E^S$ and total number of particles $N_0 = N^R + N^S$. Our aim now is to study the properties of the system \mathcal{S} as it is brought into contact with the reservoir and is able to exchange energy and particles with the reservoir. To do that we study $\mathcal{R} + \mathcal{S}$ with \mathcal{S} in any of its available states $\{\alpha\}$. If \mathcal{S} is forced to be in a particular state, γ , with energy E_γ^S , and number of particles N_γ^S the total number of states of the combined system is:

$$\Gamma_{\mathcal{R}+\mathcal{S}} = \Gamma_{\mathcal{R}} \cdot 1 = \Gamma_{\mathcal{R}}(E_0 - E_\gamma^S, N_0 - N_\gamma^S) \quad (17)$$

where we used the fact that the number of states of the reservoir depends on the reservoir's energy $E^{\mathcal{R}} = E_0 - E_\gamma^{\mathcal{S}}$, and its number of particles $N^{\mathcal{R}} = N_0 - N_\gamma^{\mathcal{S}}$.

Clearly, the probability for the system \mathcal{S} to be in any of its available states will be proportional to

$$P(E_\gamma^{\mathcal{S}}, N_\gamma^{\mathcal{S}}) = \frac{\Gamma_{\mathcal{R}}(E_0 - E_\gamma^{\mathcal{S}}, N_0 - N_\gamma^{\mathcal{S}})}{\Gamma_{\mathcal{R}+\mathcal{S}}} \quad (18)$$

Thus, the ratio of the probability that the system is in a particular state γ to the probability that it is in a state τ is:

$$\frac{P(E_\gamma^{\mathcal{S}}, N_\gamma^{\mathcal{S}})}{P(E_\tau^{\mathcal{S}}, N_\tau^{\mathcal{S}})} = \frac{\Gamma_{\mathcal{R}}(E_0 - E_\gamma^{\mathcal{S}}, N_0 - N_\gamma^{\mathcal{S}})}{\Gamma_{\mathcal{R}}(E_0 - E_\tau^{\mathcal{S}}, N_0 - N_\tau^{\mathcal{S}})} \quad (19)$$

Using the fact that $S = k_B \log \Gamma(E, N)$, we can exponentiate the above ratio noting that

$$\frac{P(E_\gamma^{\mathcal{S}}, N_\gamma^{\mathcal{S}})}{P(E_\tau^{\mathcal{S}}, N_\tau^{\mathcal{S}})} = e^{\frac{1}{k_B} \Delta S} \quad (20)$$

Expanding now around both, E_0 and N_0 to first order, as we did in the case of the canonical ensemble we get:

$$\Delta S = -(E_\gamma^{\mathcal{S}} - E_\tau^{\mathcal{S}}) \left(\frac{\partial S}{\partial E} \right)_N - (N_\gamma^{\mathcal{S}} - N_\tau^{\mathcal{S}}) \left(\frac{\partial S}{\partial N} \right)_E \quad (21)$$

But, $\left(\frac{\partial S}{\partial E} \right)_N = \frac{1}{T}$, and $\left(\frac{\partial S}{\partial N} \right)_E = -\frac{\mu}{T}$. Thus:

$$\Delta S = \frac{(N_\gamma^{\mathcal{S}} - N_\tau^{\mathcal{S}})\mu}{T} - \frac{(E_\gamma^{\mathcal{S}} - E_\tau^{\mathcal{S}})}{T} \quad (22)$$

We therefore conclude that if a system is in thermal and diffusive contact with a reservoir at temperature T and chemical potential μ , the probability to find the system at an energy E_α and number of particles N_α is:

$$P(N_\alpha, E_\alpha) = \frac{1}{\mathcal{Z}} e^{\beta(N_\alpha \mu - E_\alpha)} \quad (23)$$

The exponential factor is called the **Gibbs factor**. The normalization factor, \mathcal{Z} is the partition function of this new ensemble that we call the **Grand canonical ensemble** and from the requirement that $\sum_{states-\alpha} P(N_\alpha, E_\alpha) = 1$ we get:

$$\mathcal{Z} = \sum_{N_\alpha} \sum_{E_\alpha} e^{\beta(N_\alpha \mu - E_\alpha)} = \sum_{N_\alpha} e^{\beta(N_\alpha \mu)} Z(T, V, N_\alpha) = \sum_{N_\alpha} e^{\beta(N_\alpha \mu)} e^{-\beta F(T, V, N_\alpha)} \quad (24)$$

or:

$$\mathcal{Z}(T, V, \mu) = \sum_{N_\alpha} e^{-\beta[F(T, V, N_\alpha) - \mu N_\alpha]} \quad (25)$$

In equation 23 we showed that the probability for a state $\{\alpha\}$ of the system is given by the respective Gibbs factor. The system, which is in contact with a heat reservoir at temperature T and a particle reservoir at constant chemical potential μ , has an averaged energy $E = \langle E \rangle$ and average number of particles $N = \langle N \rangle$. Taking the energy and the number of particles as continuous variables, the probability to find the system at an energy between $E - \Delta E \leq E \leq E + \Delta E$, and number of particles between $N - \Delta N \leq N \leq N + \Delta N$ is

$$P(N, E) \Delta E \Delta N = \sum_{N_\alpha - \Delta N \leq E_\alpha \leq N_\alpha + \Delta N} \sum_{E_\alpha - \Delta E \leq E_\alpha \leq E_\alpha + \Delta E} P(N_\alpha, E_\alpha) = \Gamma(E) \frac{e^{-\beta(E - \mu N)}}{\mathcal{Z}} \Delta E \Delta N \quad (26)$$

where the second equality is the number of states within an energy ΔE and number of particles ΔN around the average energy of the system, E , and the average number of particles of the system N . This is simply the degeneracy of that energy and number of particles.

Using the fact that $S(E) = k_B \log \Gamma(N, E)$, equation 26 can be rewritten as:

$$P(N, E) = \frac{1}{\mathcal{Z}} e^{-\beta(E - \mu N) + \frac{1}{k_B} S} = \frac{1}{\mathcal{Z}} e^{-\beta(E - TS - \mu N)} \equiv \frac{1}{\mathcal{Z}} e^{-\beta \Omega(T, V, \mu)} \quad (27)$$

Since E and N are the energy and number of particles of the system, $P(N, E)$ is the most probable state for the system, i.e. it is maximum in that case that all variables except for the energy are fixed. However, for $P(N, E)$ to be maximum, the argument in the exponential has to be maximum, or, the new quantity that we defined:

$$\Omega = F(T, V, N) - \mu N \quad (28)$$

has to be minimum. This requirement is also self consistent with the definition of the temperature of the system, since the above requirement reads:

$$\delta(P(N, E)) = 0 = \delta \left(\frac{1}{\mathcal{Z}} e^{-\beta(E - TS - \mu N)} \right) = -\beta \left(\frac{1}{\mathcal{Z}} e^{-\beta(E - TS - \mu N)} \right) \cdot \delta[\Omega(N, E)] \quad (29)$$

which requires:

$$\delta \Omega = \delta E - T \delta S - \mu \delta N = 0 \quad (30)$$

or, since all quantities except for the energy and number of particles were held constant:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right); \quad \text{and} \quad \frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right) \quad (31)$$

We call Ω , the **Grand canonical free energy**, and:

$$\Omega = -k_B T \log \mathcal{Z} \quad (32)$$

4. Fluctuations

So far we discussed the averaged thermodynamic quantities, mostly in the canonical ensemble. For example we calculated the average energy or magnetization. In several cases we verified indeed that the averaged quantity that we calculate is indeed peaked very sharply with very small variance. In this section we want to explore further the behavior of the fluctuations around the average.

Magnetization fluctuations

In the microcanonical ensemble the magnetization is determined by the number of moments pointing up vs. how many moments point down. the system is isolated and therefore this number is given. Thus

$$\langle (M - \langle M \rangle)^2 \rangle = \langle M^2 \rangle - \langle M \rangle^2 = 0 \quad (33)$$

. However, in the canonical ensemble this is different since the energy is not fixed and thus the number of spins pointing in any directions can fluctuate. In the canonical ensemble the magnetization was calculated as:

$$\langle M \rangle = N \langle \mu_i \rangle = N \frac{1}{Z_1} \sum_{\mu_i = \pm \mu_0} \mu_i e^{-\beta \mu_i H} = \frac{N}{\beta} \frac{\partial \log Z_1}{\partial H} = \frac{1}{\beta} \frac{\partial \log Z}{\partial H} \quad (34)$$

To calculate $\langle M^2 \rangle$ we note that since the spins do not interact it is enough to calculate $\langle \mu_i^2 \rangle$, or:

$$\langle M^2 \rangle = N \frac{1}{Z_1} \sum_{\mu_i = \pm \mu_0} (\mu_i)^2 e^{-\beta \mu_i H} = \frac{N}{Z_1} \frac{1}{\beta^2} \frac{\partial^2 Z_1}{\partial H^2} = \frac{N}{\beta^2} \left[\frac{\partial^2 \log Z_1}{\partial H^2} + \frac{1}{Z_1^2} \left(\frac{\partial Z_1}{\partial H} \right)^2 \right] \quad (35)$$

Therefore:

$$\langle M^2 \rangle - \langle M \rangle^2 = N \left[\langle \mu^2 \rangle - \langle \mu \rangle^2 \right] = \frac{N}{\beta^2} \left[\frac{\partial^2 \log Z_1}{\partial H^2} + \frac{1}{Z_1^2} \left(\frac{\partial Z_1}{\partial H} \right)^2 - \frac{1}{Z_1^2} \left(\frac{\partial Z_1}{\partial H} \right)^2 \right] = \frac{1}{\beta} \frac{\partial \langle M \rangle}{\partial H} \quad (36)$$

Define the magnetic susceptibility: $\chi = \frac{1}{N} \frac{\partial \langle M \rangle}{\partial H}$, we find the final result that connects the fluctuations in the macroscopic magnetization to a response function to the magnetic field, i.e. the susceptibility:

$$\langle M^2 \rangle - \langle M \rangle^2 = k_B T N \chi \quad (37)$$

Note now that $\langle M \rangle = N \langle \mu \rangle$, and thus:

$$\frac{\sqrt{\langle M^2 \rangle - \langle M \rangle^2}}{\langle M \rangle} = \frac{\sqrt{N k_B T \chi}}{N \langle \mu \rangle} = \sqrt{\frac{k_B T \chi}{\langle \mu \rangle^2}} \frac{1}{\sqrt{N}} \quad (38)$$

In the limit $N \rightarrow \infty$ the fluctuations vanish and the result of the canonical ensemble coincides with the result of the microcanonical ensemble. We conclude that the two ensembles are equivalent up to corrections of order $1/\sqrt{N}$.

Energy Fluctuations

Let us now look at energy fluctuations in the canonical ensemble clearly for the paramagnet these will be similar to the magnetization fluctuations because $E = -MH$, and the field is constant.

We already found that

$$\langle E \rangle = - \frac{\partial \log Z}{\partial \beta} \quad (39)$$

For the average of the square we calculate:

$$\langle E^2 \rangle = \frac{1}{Z} \sum_{\alpha} E_{\alpha}^2 e^{-\beta E_{\alpha}} = \frac{1}{Z} \frac{\partial^2}{\partial \beta^2} \sum_{\alpha} e^{-\beta E_{\alpha}} = \frac{\partial^2 \log Z}{\partial \beta^2} + \left(\frac{\partial \log Z}{\partial \beta} \right)^2 = - \frac{\partial \langle E \rangle}{\partial \beta} + \langle E \rangle^2 \quad (40)$$

We finally obtain:

$$\langle E^2 \rangle - \langle E \rangle^2 = - \frac{\partial \langle E \rangle}{\partial \beta} = k_B T^2 \left(\frac{\partial E}{\partial T} \right) = k_B T^2 C_V = N k_B T^2 c_V \quad (41)$$

Where in the last equality we used the specific heat as the heat capacity per particle. following the same procedure as before we find that the relative fluctuations are:

$$\frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} = \frac{\sqrt{Nk_B T^2 c_V}}{N\langle \mu \rangle} = \sqrt{\frac{Nk_B T^2 c_V}{N^2 \langle \epsilon \rangle^2}} = \sqrt{\frac{k_B T^2 c_V}{\langle \epsilon \rangle^2}} \frac{1}{\sqrt{N}} \quad (42)$$

Where we used the fact that the particles do not interact and we can use the average single particle energy ϵ such that $\langle E \rangle = N\langle \epsilon \rangle$.

Again we notice that the fluctuations will vanish in the limit of $N \rightarrow \infty$ (the thermodynamic limit), and in that limit the energy in the microcanonical ensemble, will equal to the energy of the canonical ensemble.

Fluctuations in the Number of particles

As an example to the fluctuations in the size of the system we calculate the number fluctuations in the grand-canonical ensemble. We need to calculate: $\langle N^2 \rangle - \langle N \rangle^2$.

We already showed earlier that:

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial \log \mathcal{Z}}{\partial \mu} \quad (43)$$

In a similar way we calculate;

$$\langle N^2 \rangle = \frac{1}{\mathcal{Z}} \sum_{\alpha} N_{\alpha}^2 e^{\beta(N_{\alpha}\mu - E_{\alpha})} = \frac{1}{\beta^2} \frac{1}{\mathcal{Z}} \frac{\partial^2 \mathcal{Z}}{\partial \mu^2} = \frac{1}{\beta^2} \left[\frac{\partial^2 \log \mathcal{Z}}{\partial \mu^2} + \left(\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \mu} \right)^2 \right] \quad (44)$$

Combining the above results we obtain;

$$\langle N^2 \rangle = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu} + \langle N \rangle^2 \quad (45)$$

We next change quantities and use the quantities that are fixed. In this case it is the volume of the system, V . We further define the number-density of our system as $\langle n \rangle = \langle N \rangle / V$. In that case:

$$\frac{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}{\langle N \rangle} = \frac{\sqrt{k_B T V \frac{\partial \langle n \rangle}{\partial \mu}}}{\langle N \rangle} = \sqrt{\frac{V k_B T \frac{\partial \langle n \rangle}{\partial \mu}}{\langle N \rangle^2}} = \sqrt{k_B T \frac{1}{\langle n \rangle} \frac{\partial \langle n \rangle}{\partial \mu} \frac{1}{\sqrt{\langle N \rangle}}} \quad (46)$$

The response function now is the “density of states”, $\frac{\partial \langle n \rangle}{\partial \mu}$ of the system at the chemical potential. Again we notice that in the limit of infinite number of particles the relative fluctuations of the system vanish. We will discuss further these kind of fluctuations when we study quantum ideal gases. However, this result can be expressed in another way that reveals some of the physics associated with the notion of the chemical potential.

Using maxwell relations one can easily derive from equation 16:

$$\frac{\partial N}{\partial V} = \frac{\partial P}{\partial \mu} \quad (47)$$

This result can also be derived from equation 7 where we get the Maxwell relation:

$$\frac{\partial V}{\partial N} = \frac{\partial \mu}{\partial P} \quad (48)$$

Using the above result, equation 45 gives:

$$\langle N^2 \rangle - \langle N \rangle^2 = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu} = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial V} \frac{\partial V}{\partial \mu} = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial V} \frac{\partial V}{\partial P} \frac{\partial P}{\partial \mu} = \frac{1}{\beta} \frac{\partial V}{\partial P} \left(\frac{\partial N}{\partial V} \right)^2 \quad (49)$$

However, $\left(\frac{\partial N}{\partial V} \right)^2 = n^2$, and the compressibility of the gas is given by:

$$\kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_T \quad (50)$$

We therefore get:

$$\frac{\sqrt{\langle N^2 \rangle - \langle N \rangle^2}}{\langle N \rangle} = \frac{\sqrt{k_B T V |\kappa| n^2}}{\langle N \rangle} = \sqrt{\frac{V k_B T |\kappa| n^2}{\langle N \rangle^2}} = \sqrt{k_B T |\kappa| n} \frac{1}{\sqrt{\langle N \rangle}} \quad (51)$$

Which indicates that the compressibility and the density of states are intimately connected, as both are describing how hard or easy it is to change the size of the system using an external field (pressure or chemical potential).

Lecture Notes #8

1. Quantum statistical physics

Similar to our previous considerations also in quantum mechanics it is impossible to do a quantum mechanical treatment of a macroscopic system with $N \geq 10^{20}$ particles. It is impossible to define the initial conditions since we cannot define both coordinates and momenta. again we will attempt to use the laws of thermodynamics instead of equations of motion.

1.1 Reminder

- 1) $|m\rangle$ - is a vector in Hilbert space.
- 2) The scalar product $\langle m|n\rangle = \int \varphi_m^* \varphi_n d^d r$ (in d-dimensions).
- 3) $\langle m|n\rangle = \langle n|m\rangle^*$
- 4) If $|m\rangle$ are eigenvectors (eigenfunctions) of the operator \hat{a} then: $\hat{a}|m\rangle = \alpha_m|m\rangle$.
- 5) $\hat{a}^\dagger = \hat{a}$ is a hermitian operator for which the eigenvalues are real.
- 6) Using the complete set $\{|m\rangle\}$, the unit operator \hat{I} such that: $\langle n|\hat{I}|n\rangle = \langle n|n\rangle = 1$ is: $\hat{I} = \sum_m |m\rangle\langle m|$.

1.2 The density operator

To start our discussion, let us assume that our system, instead of being in an exact state, it is distributed among many states with a probability P_m to be in the pure state $|m\rangle$. $\{|m\rangle\}$ is a complete set of normalized states such that $\langle m|m\rangle = 1$.

Clearly for all m $P_m \geq 0$ and $\sum_m P_m = 1$.

Define the statistical operator:

$$\hat{\rho} = \sum_m |m\rangle P_m \langle m| \quad (1)$$

To show that $\hat{\rho}$ is indeed a statistical operator, let ρ be an eigenvalue of $\hat{\rho}$ such that:

$$\hat{\rho}|\rho\rangle = \rho|\rho\rangle \quad (2)$$

Therefore:

$$\hat{\rho}|\rho\rangle = \sum_m |m\rangle P_m \langle m| = \rho|\rho\rangle \quad (3)$$

Therefore, for some particular state $|\rho'\rangle$ we calculate the expectation value:

$$\langle \rho' | \hat{\rho} | \rho' \rangle = \rho' \langle \rho' | \rho' \rangle = \sum_m \langle \rho' | m \rangle P_m \langle m | \rho' \rangle = \sum_m P_m |\langle \rho' | m \rangle|^2 \geq 0 \quad (4)$$

Thus we conclude that $\rho' \geq 0$ for all $|\rho'\rangle$, i.e. all the eigenvalues of $\hat{\rho}$ are positive.

We next calculate the trace of $\hat{\rho}$. to do that we choose some basis of states $\{|a\rangle\}$

$$Tr \{\hat{\rho}\} = \sum_a \langle a | \hat{\rho} | a \rangle = \sum_a \sum_m \langle a | m \rangle P_m \langle m | a \rangle = \sum_m P_m \sum_a \langle a | m \rangle \langle m | a \rangle \quad (5)$$

However, $\langle a | m \rangle \langle m | a \rangle = |\langle a | m \rangle|^2 = \langle m | a \rangle \langle a | m \rangle$ and thus:

$$Tr \{\hat{\rho}\} = \sum_m P_m \sum_a \langle m | a \rangle \langle a | m \rangle = \sum_m P_m \langle m | \hat{I} | m \rangle = \sum_m P_m = 1 \quad (6)$$

We therefore confirmed that $Tr \{\hat{\rho}\} = \sum_m P_m = 1$.

If the system is in a pure state $|m\rangle$, the expectation value of a linear hermitian operator \hat{A} is: $\langle A \rangle = \langle m | \hat{A} | m \rangle$. If the system is distributed among a set of pure states, $\{|m\rangle\}$ with corresponding probabilities $\{P_m\}$, then:

$$\langle A \rangle = \sum_m P_m \langle m | \hat{A} | m \rangle = \sum_m \sum_a P_m \langle m | a \rangle \langle a | \hat{A} | m \rangle \quad (7)$$

Where we inserted the unit operator in the sum. Since $\langle F \rangle$ is real, we also find that

$$\langle A \rangle = \sum_m \sum_a P_m \langle a | \hat{A} | m \rangle \langle m | a \rangle = Tr \{\hat{A} \hat{\rho}\} \quad (8)$$

In a similar way we could write:

$$\langle A \rangle = \sum_m \sum_a P_m \langle m | \hat{A} | a \rangle \langle a | m \rangle = \sum_m \sum_a P_m \langle a | m \rangle \langle m | \hat{A} | a \rangle = Tr \{\hat{\rho} \hat{A}\} \quad (9)$$

We therefore found that $\langle A \rangle = Tr \{\hat{A} \hat{\rho}\} = Tr \{\hat{\rho} \hat{A}\}$.

1.3 Time dependence

If the system is in a state $|m\rangle$, this state will develop in time according to the Schrodinger equation:

$$i\hbar \frac{\partial}{\partial t} |m\rangle = \hat{\mathcal{H}} |m\rangle \quad (10)$$

Where $\hat{\mathcal{H}}$ is the hamiltonian of the system. Thus:

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = i\hbar \frac{\partial}{\partial t} \sum_m |m\rangle P_m \langle m| = i\hbar \sum_m \left[\left(\frac{\partial}{\partial t} |m\rangle \right) P_m \langle m| + |m\rangle \frac{\partial P_m}{\partial t} \langle m| + |m\rangle P_m \left(\frac{\partial}{\partial t} \langle m| \right) \right] \quad (11)$$

Since the P_m s are a property of the system that does not change in time, the second term is zero. Using equation 10 we can rewrite the above result as:

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = \sum_m \left[\hat{\mathcal{H}} |m\rangle P_m \langle m| - |m\rangle P_m \langle m| \hat{\mathcal{H}} \right] = [\hat{\mathcal{H}}, \hat{\rho}] \quad (12)$$

If the system is stationary (i.e. in equilibrium) so that $\hat{\rho}$ does not depend explicitly on time, then

$$[\hat{\mathcal{H}}, \hat{\rho}] = 0 \quad (13)$$

In that case, for any two states, m and n , we can calculate the expectation value of $\hat{\rho}$ in the following way.

$$\langle n | [\hat{\mathcal{H}}, \hat{\rho}] | m \rangle = \langle n | \hat{\mathcal{H}} \hat{\rho} | m \rangle - \langle n | \hat{\rho} \hat{\mathcal{H}} | m \rangle = (E_n - E_m) \langle n | \hat{\rho} | m \rangle \quad (14)$$

Since $E_n \neq E_m$ if there is no degeneracy. We obtain $\langle n | \hat{\rho} | m \rangle = 0$. Thus, if the energy spectrum is nondegenerate then $\hat{\rho}$ is diagonal.

1.4 The adiabatic theorem

Let us assume that the hamiltonian of the system depends on an external parameter a (e.g. magnetic field, electric field, etc.). In this case we write the (time independent) Schrodinger equation as:

$$\hat{\mathcal{H}}(a) |n, a\rangle = E_n(a) |n, a\rangle \quad (15)$$

Therefore:

$$\frac{\partial}{\partial a} [\hat{\mathcal{H}}(a)|n, a\rangle] = \frac{\partial \hat{\mathcal{H}}}{\partial a} |n, a\rangle + \hat{\mathcal{H}}(a) \frac{\partial}{\partial a} |n, a\rangle = \frac{\partial E_n}{\partial a} |n, a\rangle + E_n \frac{\partial}{\partial a} |n, a\rangle \quad (16)$$

Sandwiching the above equation with $\langle m, a|$, after some algebra we find:

$$\langle m, a| \frac{\partial \hat{\mathcal{H}}}{\partial a} |n, a\rangle = \frac{\partial E_n}{\partial a} \delta_{m,n} + (E_n - E_m) \langle m, a| \frac{\partial}{\partial a} |n, a\rangle \quad (17)$$

We distinguish now the following two cases. For $n = m$ we find:

$$\langle n, a| \frac{\partial \hat{\mathcal{H}}}{\partial a} |n, a\rangle = \frac{\partial E_n}{\partial a} \quad (18)$$

If the system is in a pure state $|n\rangle$, even if $a \rightarrow a + \delta a$, we find:

$$\delta E_n = \frac{\partial E_n}{\partial a} \delta a = \langle n, a| \frac{\partial \hat{\mathcal{H}}}{\partial a} \delta a |n, a\rangle = \langle n, a| \delta \hat{\mathcal{H}} |n, a\rangle \quad (19)$$

For $n \neq m$, but with no degeneracies we find:

$$\langle m, a| \frac{\partial}{\partial a} |n, a\rangle = \frac{\langle m, a| \frac{\partial \hat{\mathcal{H}}}{\partial a} |n, a\rangle}{E_n - E_m} \quad (20)$$

The case of degenerate energy levels can be taken into account either by another field that lifts the degeneracies or using a density of levels near each energy. Since degeneracies can be treated, we will assume no degeneracies in the future considerations, and will use the result of equation 20 as always true.

We therefore showed that if the system is in a pure state $|n\rangle$, and an external parameter applied to the system is varying from a to $a + \delta a$, the energy of the system changes by $\delta E(a)$ given by equation 65, and the wavefunction of the system changes from $|n, a\rangle$ to $|n, a + \delta a\rangle$. The question is whether the system will stay in the state $|n\rangle$, or will change to a different state, say $|m\rangle$. The probability of changing state is then given by equation 20.

Starting from a general state of the system:

$$\Psi = \sum_m c_m(t) |m, a\rangle \quad (21)$$

Schrodinger equation reads:

$$i\hbar \frac{\partial}{\partial t} \sum_m c_m(t) |m, a\rangle = \hat{\mathcal{H}} \sum_m c_m |m, a\rangle \quad (22)$$

or

$$i\hbar \sum_m \dot{c}_m(t) |m, a\rangle + i\hbar \sum_m (t) \dot{a} \frac{\partial}{\partial a} |m, a\rangle = \sum_m c_m E_m |m, a\rangle \quad (23)$$

Sandwiching the above equation with $\langle n, a|$, we get for $n = m$:

$$\dot{c}_m(t) = -\frac{i}{\hbar} c_m E_m \quad (24)$$

while for $n \neq m$

$$\dot{c}_m(t) + \dot{a} \sum_{n \neq m} c_n \langle m, a | \frac{\partial}{\partial a} |n, a\rangle = -\frac{i}{\hbar} c_m E_m \quad (25)$$

or, using equation 20 and a “prime” to denote summation except for $n = m$,

$$\dot{c}_m(t) + \frac{i}{\hbar} c_m E_m + \dot{a} \sum_n' c_n \frac{\langle m, a | \frac{\partial \hat{\mathcal{H}}}{\partial a} |n, a\rangle}{E_n - E_m} = 0 \quad (26)$$

Assume now that at $t = 0$ the state of the system was $|n_0, a\rangle$. Assume further that the change in a is small and takes very long time, i.e. $\dot{a} \rightarrow 0$. Equation 26 becomes:

$$\frac{dc_{n_0}}{dt} + \frac{i}{\hbar} c_{n_0}(t) E_{n_0}(a) = 0 \quad (27)$$

or

$$c_{n_0}(t) = e^{-\frac{i}{\hbar} E_{n_0}(a)t} \quad (28)$$

and

$$c_m t \approx 0 \quad \text{for} \quad m \neq n_0 \quad (29)$$

This is called the adiabatic theorem. In simple words it states that if the system is in a certain state at $t = 0$, then upon a very slow and small variation of external parameters the system will tend to stay in this state. If the system is distributed among many states with probabilities P_m , the system will stay distributed among these probabilities without shifting towards one state or another.

1.5 Adiabatic processes and Entropy

The expectation value of the energy is given by:

$$\langle E \rangle = \sum_m P_m E_m; \quad \text{where :} \quad E_m = \langle m | \hat{\mathcal{H}} | m \rangle \quad (30)$$

this is true providing the distribution is defined over the states which are eigenstates of the hamiltonian $\hat{\mathcal{H}}$. If an external parameter is changed such that the energy changes we get:

$$d\langle E \rangle = \sum_m E_m \delta P_m + \sum_m P_m \delta E_m \quad (31)$$

If this change is *slow*, the adiabatic theorem says: $\delta P_m = 0$, and

$$d\langle E \rangle = \sum_m P_m dE_m(a) \quad (32)$$

let us call the force conjugated to the external parameter a , A . In that case:

$$\langle A \rangle = - \sum_m P_m \frac{\partial E_m}{\partial a} \quad (33)$$

The work is:

$$\delta W = \langle A \rangle da = - \sum_m P_m \frac{\partial E_m}{\partial a} \delta a = - \sum_m P_m \delta E_m = -d\langle E \rangle \quad (34)$$

In the language of thermodynamics this is a reversible adiabatic process of the expectation value of A which is a function of a only. In other words, $\int \langle A(a) \rangle da$ depends only on the end points of the process. with the choice of equal initial and final condition it is possible to reverse the state of the system.

According to thermodynamics the entropy does not change in a reversible process, but the process is characterized by the distribution $\{P_n\}$. Thus, the entropy has to be a function of the P_n 's. $S = S(P_1, P_2, P_3, \dots)$. S is therefore a new quantity, non-mechanical. In other words, we cannot determine S by knowing the mechanical properties of the system. However, there are a few simple conditions that S has to fulfill:

1) S has to be symmetric with respect to the P_n s, i.e. it should not depend on the order we distribute the system among them.

2) S should have a maximum so we can describe an isolated system in equilibrium.

3) S should be differentiable and satisfy the differential equation:

$$dS(P_1, P_2, P_3, \dots) + dS(Q_1, Q_2, Q_3, \dots) = dS(P_1 Q_1, P_2 Q_2, P_3 Q_3, \dots) \quad (35)$$

This is because the change in entropy of two systems with no interactions should be described by a distribution that is the joint distribution.

All the above requirements *do not* determine S . For example, the family of solutions:

$$-(\sum_n P_n^\lambda \log P_n) / (\sum_n P_n^\lambda) + \text{Const.} \quad (36)$$

or:

$$-\log(\sum_n P_n^\lambda) + \text{Const.} \quad (37)$$

satisfy the above requirements for $\lambda \geq 1$. To find what is the correct form of the entropy we need to compare to a known model such as ideal gas, or simply ask that whatever we choose should reduce to what we already derived before in the classical limit. In that case we have to choose:

$$S = -k_B \sum_n P_n \log P_n + \text{Const.} \quad (38)$$

Using the definition of the statistical operator $\hat{\rho}$, the entropy can also be written as:

$$S = -k_B \cdot \text{Tr} \{ \hat{\rho} \log \hat{\rho} \} + \text{Const.} \quad (39)$$

S is therefore defined on the spectrum of $\hat{\rho}$. There are other ways that we can justify the above equation as described the entropy of the system. However, we will simply take the above result which is very intuitive to us by now and use it. the discussion below will all rely on the assumption that nature tends to maximize the entropy, subject to the constraint that the ensemble average of the hamiltonian has certain prescribed values.

Another way to look at the problem is to look at an ensemble of N members and ask what is the density matrix for a completely random ensemble. This is easy to guess as:

$$\hat{\rho} = \frac{1}{N} M \quad (40)$$

where M is a $N \times N$ diagonal matrix with $M_{ii} = 1$ for all i .

Clearly $\text{Tr} \{ \hat{\rho} \} = 1$, while $\rho_{ii} = \frac{1}{N}$. Can we construct a quantity, on the spectrum of $\hat{\rho}$ that will be maximum in the case of maximum randomness and at the same time will have the properties described above. the answer is equation 39, and the constant can be chosen as zero.

1.6 The canonical ensemble

Let us assume that our system reached equilibrium such that $\frac{\partial \hat{\rho}}{\partial t} = 0$. $\hat{\rho}$ and $\hat{\mathcal{H}}$ can be diagonalized simultaneously. Choose a complete system of energy eigenvalues $|m\rangle$. such that:

$$\hat{\rho} = \sum_m |m\rangle P_m \langle m| \quad (41)$$

and

$$\hat{\mathcal{H}}|m\rangle = E_m|m\rangle \quad (42)$$

The energy of the system is calculated using $\hat{\rho}$.

$$E = \langle \hat{\mathcal{H}} \rangle = Tr \{ \hat{\rho} \hat{\mathcal{H}} \} \quad (43)$$

With the normalization condition:

$$\sum_m \langle m | \hat{\rho} | m \rangle \equiv \sum_m \rho_{mm} = \sum_m P_m = 1 \quad (44)$$

The requirement for maximum of S , with the above constraints, can be written using two Lagrange multipliers, β and γ :

$$\delta \left[- \sum_m \rho_{mm} \log \rho_{mm} - \beta \sum_m \rho_{mm} E_m - \gamma \sum_m \rho_{mm} \right] = 0 \quad (45)$$

or:

$$\sum_m \delta \rho_{mm} [(\log \rho_{mm} + 1) + \beta E_m + \gamma] = 0 \quad (46)$$

The above has to be true for any variation in any matrix element and thus the solution is:

$$\rho_{mm} \equiv P_m = e^{-\beta E_m - \gamma - 1} \quad (47)$$

The constraint γ can be eliminated using the normalization condition. Thus:

$$P_m = \rho_{mm} = \frac{e^{-\beta E_m}}{\sum_m e^{-\beta E_m}} \quad (48)$$

The above result is the state probability in the canonical ensemble. We recognize the partition function Z as the denominator in the expression for the probabilities with $\beta = \frac{1}{k_B T}$, and thus:

$$Z = Tr \{ e^{-\beta \hat{\mathcal{H}}} \} \quad (49)$$

In this ensemble, the statistical operator $\hat{\rho}$ can now be written as:

$$\hat{\rho} = \frac{e^{-\beta\hat{\mathcal{H}}}}{Z} \quad (50)$$

The ensemble average of any observable A is then:

$$\langle A \rangle = \frac{Tr \left\{ \hat{A} e^{-\beta\hat{\mathcal{H}}} \right\}}{Z} \quad (51)$$

1.7 The grand canonical ensemble

Equilibrium of a system which is in thermal contact with a heat reservoir at temperature T , and in contact with a particles reservoir is that for which the thermodynamic free energy, Ω is minimum. This is an open system and the number of particles N can vary. this property expressed in “quantum-mechanics language” means that

$$[\hat{N}, \hat{\mathcal{H}}] = 0 \quad (52)$$

The number of particles in each state can now be found:

$$\hat{N}|m\rangle = n_m|m\rangle \quad (53)$$

The number of particles in the system will be calculated using $\hat{\rho}$:

$$N = \langle \hat{N} \rangle = Tr \left\{ \hat{\rho} \hat{N} \right\} = \sum_m \rho_{mm} n_m \quad (54)$$

Adding now the constraint of the number of particles, we need to maximize:

$$\delta \left[- \sum_m \rho_{mm} \log \rho_{mm} - \beta \sum_m \rho_{mm} E_m + \alpha \sum_m \rho_{mm} n_m - \gamma \sum_m \rho_{mm} \right] = 0 \quad (55)$$

or:

$$\sum_m \delta \rho_{mm} [(\log \rho_{mm} + 1) + \beta E_m - \alpha n_m + \gamma] = 0 \quad (56)$$

The above has to be true for any variation in any matrix element and thus the solution is:

$$\rho_{mm} \equiv P_m = e^{-\beta E_m + \alpha n_m - \gamma - 1} \quad (57)$$

Again, the constraint γ can be eliminated using the normalization condition. Thus:

$$P_m = \rho_{mm} = \frac{e^{-\beta(E_m - \mu n_m)}}{\sum_m e^{-\beta(E_m - \mu n_m)}} \quad (58)$$

The above result is the state probability in the grand canonical ensemble. We recognize the grand canonical partition function \mathcal{Z} as the denominator in the expression for the probabilities with $\beta = \frac{1}{k_B T}$, and $\alpha = \beta\mu$, thus:

$$\mathcal{Z} = Tr \left\{ e^{-\beta(\hat{\mathcal{H}} - \mu\hat{N})} \right\} \quad (59)$$

In this ensemble, the statistical operator $\hat{\rho}$ can now be written as:

$$\hat{\rho} = \frac{e^{-\beta(\hat{\mathcal{H}} - \mu\hat{N})}}{\mathcal{Z}} \quad (60)$$

The ensemble average of any observable A is then:

$$\langle A \rangle = \frac{Tr \left\{ \hat{A} e^{-\beta(\hat{\mathcal{H}} - \mu\hat{N})} \right\}}{\mathcal{Z}} \quad (61)$$

2. Many particles system

We shift our attention now to the many-particle nature of the quantum mechanical system. While in classical mechanics one could imagine labeling the particles in order to track them, this becomes impossible for a quantum system. In particular we note that the indistinguishability of the particles makes it impossible to follow them individually in the region where they both can be found simultaneously (i.e. their wave-function overlap).

We therefore make the basic assumption that any complete set of dynamical variables which describes the behavior of a single particle, can also be employed for N particles of the same kind. This is a very strong assumption since it implies that the composite constitutes a complete set for the many-particle system. Clearly, a most general such state will be a linear combination of the single particle states. We therefore describe a state of the system in terms of the occupancy of the single particle states.

$$\Psi = |n_1, n_2, \dots, n_i, \dots\rangle \quad (62)$$

To this we also add the vacuum state:

$$\Psi^0 \equiv |0\rangle = |0, 0, 0, \dots, 0, \dots\rangle \quad (63)$$

In this language, a single particle state in which the particle exists in the state i is:

$$\Psi_i^1 = |0, 0, 0, \dots, n_i = 1, \dots\rangle \quad (64)$$

Finally we note that in almost all of our discussion we assume that if the system is made of interacting particles we can still expand it in terms of the states of the non-interacting particles.

While we will not review here in details the theory of many particle system, we would like to sketch here some of the ingredients of this theory. Starting from the vacuum state, we can create particles in any given state or annihilate them using creation and annihilation operators:

$$a_i^\dagger \Psi^0 = a_i^\dagger |0, 0, 0, \dots, 0, \dots\rangle \propto |0, 0, 0, \dots, n_i = 1, \dots\rangle \quad (65)$$

Similarly

$$a_i \Psi^1 = a_i |0, 0, 0, \dots, n_i = 1, \dots\rangle \propto |0, 0, 0, \dots, 0, \dots\rangle \quad (66)$$

This now allow us to occupy all states with a given number of particles until we exhaust all the

$$N = \sum_i n_i \quad (67)$$

particles.

the above relations we also supplement with the fact that if we annihilate the vacuum we get zero, and that if we annihilate any many particle state that has no occupation in a particular state we also get zero:

$$a_i \Psi^0 = 0 \quad \text{and} \quad a_j \Psi_i^1 = 0 \quad (68)$$

While the phase of the wave-function is arbitrary, such as in equation 65, the successive creation of two particles introduces a relative phase that may not be arbitrary. In general, for any basis state Ψ we can write:

$$a_i^\dagger a_j^\dagger \Psi = \lambda a_j^\dagger a_i^\dagger \Psi \quad (69)$$

In fact, it can be shown that the only two possibilities for λ are: $\lambda = \pm 1$. The plus sign will hold for Bose-Einstein particles (“bosons”) and will introduce the commutator:

$$a_i^\dagger a_j^\dagger - a_j^\dagger a_i^\dagger \equiv [a_i^\dagger, a_j^\dagger] = 0 \quad (70)$$

It will also be supplemented by the other coomutation relations:

$$[a_i, a_j] = 0 \quad \text{and} \quad [a_i, a_j^\dagger] = 1 \cdot \delta_{ij} \quad (71)$$

The minus sign will hold for Fermi-Dirac particles (“fermions”) and will introduce the anti-commutator:

$$a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger \equiv \{a_i^\dagger, a_j^\dagger\} = 0 \quad (72)$$

It will also be supplemented by the other anti-coomutation relations:

$$\{a_i, a_j\} = 0 \quad \text{and} \quad \{a_i, a_j^\dagger\} = 1 \cdot \delta_{ij} \quad (73)$$

It is easy to see now that the operator $\hat{N}_i \equiv a_i^\dagger a_i$ counts the number of particles in the state $-i$, such that:

$$\langle n_1, n_2, n_3, \dots, n_i, \dots | \hat{N}_i | n_1, n_2, n_3, \dots, n_i, \dots \rangle = \langle n_1, n_2, n_3, \dots, n_i, \dots | a_i^\dagger a_i | n_1, n_2, n_3, \dots, n_i, \dots \rangle = n_i \quad (74)$$

We therefore conclude that:

$$a_i | n_1, n_2, \dots, n_i + 1, \dots \rangle = e^{i\alpha} \sqrt{n_i} | n_1, n_2, \dots, n_i, \dots \rangle \quad (75)$$

and that

$$a_i^\dagger | n_1, n_2, \dots, n_i - 1, \dots \rangle = e^{-i\alpha} \sqrt{n_i} | n_1, n_2, \dots, n_i, \dots \rangle \quad (76)$$

The number of particles operator is now given in terms of the single particle occupation number operators:

$$\hat{N} = \sum_i \hat{N}_i \quad (77)$$

The hamiltonian is now given in terms of the single particle occupation number operators:

$$\hat{\mathcal{H}} = \sum_i \hat{N}_i \epsilon_i \quad (78)$$

3. Quantum statistical mechanics with many particles

To introduce statistical mechanics at this stage is rather straightforward, but may not be easy to solve. If we are describing a system at constant temperature T , the appropriate ensemble is the canonical ensemble and thus to calculate the partition function we use equation 49. Thus

$$Z = \text{Tr} \left\{ e^{-\beta \hat{\mathcal{H}}} \right\} = \text{Tr} \left\{ e^{-\beta \sum_i \hat{N}_i \epsilon_i} \right\} = \sum_{\{n_i\}} e^{-\beta \sum_i n_i \epsilon_i} \quad (79)$$

Where we emphasized that we sum over all single particle states. Since we need to keep track on the occupation numbers to satisfy the condition: $\sum_i n_i = N$, it is clear that

$$Z \neq \sum_{n_1} e^{-\beta n_1 \epsilon_1} \cdot \sum_{n_2} e^{-\beta n_2 \epsilon_2} \dots \quad (80)$$

In addition, if we started from the states of the total energy of the system and summed over all states of the full system:

$$Z = \sum_{all\{k\}} \frac{n_1! n_2! n_3! \dots n_k! \dots}{N!} e^{-\beta E(\{k\})} \quad (81)$$

we would end up with the same problem.

To solve the problem of calculation equation 79 with the constraint $N = \sum_i n_i$, we will use the fact that in the thermodynamic limit all ensembles are equivalent. Thus we will change the ensemble from the canonical ensemble that we were using so far to the grand canonical ensemble. In this ensemble the variable that is kept fixed is the chemical potential, not the number of particles. We will see later that this approach is justified by the fact that the number of particles will determine the fixed chemical potential via the equation:

$$N = \frac{1}{\beta} \frac{\partial \log \mathcal{Z}}{\partial \mu} \quad (82)$$

Where the partition function is given by:

$$\mathcal{Z} = \sum_{n_1} \sum_{n_2} \dots \sum_{n_k} \dots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} \quad (83)$$

Thus, μ here plays the role of a Lagrange multiplier which ensures the constraint on the number of particles. The grand partition function *can* be factorized and we find:

$$\mathcal{Z} = \left[e^{-\beta \sum_1 n_1 (\epsilon_1 - \mu)} \right] \left[e^{-\beta \sum_2 n_2 (\epsilon_2 - \mu)} \right] \dots = \prod_k \left(\sum_{n_k} e^{-\beta (\epsilon_k - \mu) n_k} \right) \quad (84)$$

Our objective now is to perform the sum with the constraint put by the different types of particles.

4. Different types of statistics

Equation 84 is our starting point to explore the different statistics. Based on the previous analysis of the many particles wave-functions, we showed that there are only two types of statistics possible. One for which the commutation relation of the creation and annihilation operators hold, and the other for which the anti-commutation relations hold. Quantum mechanics tells us that if the commutation relation holds, we can occupy in any quantum state any number of particles. This is the case for Bose-Einstein statistics. On the other hand, if the anti-commutation relations hold, the particles are fermions and thus no more than one particle can occupy any quantum state. this is the Fermi-Dirac statistics.

Derivation of Bose-Einstein partition function:

In this case $n_k = 0, 1, 2, 3, \dots$ for each quantum state with energy ϵ_k . Thus:

$$\sum_{n_k=0}^{\infty} e^{-\beta(\epsilon_k - \mu)n_k} = \frac{1}{1 - e^{-\beta(\epsilon_k - \mu)}} \quad (85)$$

and

$$\mathcal{Z} = \prod_k \left(\frac{1}{1 - e^{-\beta(\epsilon_k - \mu)}} \right) \quad (86)$$

The grand canonical free energy is:

$$\Omega = -k_B T \log \mathcal{Z} = k_B T \sum_k \log \left(1 - e^{-\beta(\epsilon_k - \mu)} \right) \quad (87)$$

and the self consistent equation for the chemical potential is derived while calculating the average number of particles in the system:

$$\langle N \rangle = N = \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} = \sum_k \langle n_k \rangle = \sum_k \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1} \quad (88)$$

The occupation number of the k^{th} state is:

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1} \quad (89)$$

Derivation of Fermi-Dirac partition function:

In this case $n_k = 0$ or 1 , for each quantum state with energy ϵ_k . Thus:

$$\sum_{n_k=0}^1 e^{-\beta(\epsilon_k-\mu)n_k} = 1 + e^{-\beta(\epsilon_k-\mu)} \quad (90)$$

The grand canonical free energy is:

$$\Omega = -k_B T \log \mathcal{Z} = -k_B T \sum_k \log \left(1 + e^{-\beta(\epsilon_k-\mu)} \right) \quad (91)$$

and the self consistent equation for the chemical potential is derived while calculating the average number of particles in the system:

$$\langle N \rangle = N = \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} = \sum_k \langle n_k \rangle = \sum_k \frac{1}{e^{\beta(\epsilon_k-\mu)} + 1} \quad (92)$$

The occupation number of the k^{th} state is:

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k-\mu)} + 1} \quad (93)$$

The limit of Maxwell-Boltzmann statistics:

We already saw how to achieve this limit using the canonical ensemble partition function. There, the classical limit is achieved by requiring the occupation numbers to be very small. Thus we require $\langle n_k \rangle \ll 1$ for all k , which, independent of the statistics, leads to:

$$e^{\beta(\epsilon_k-\mu)} \gg 1; \quad \text{and} \quad n_k \simeq e^{-\beta(\epsilon_k-\mu)} \quad (94)$$

Thus:

$$N = \sum_k n_k \simeq \sum_k e^{-\beta(\epsilon_k-\mu)} = e^{\beta\mu} \sum_k e^{-\beta\epsilon_k} = e^{\beta\mu} Z_1 \quad (95)$$

where we identified the sum of the Boltzmann factors as the single particle partition function of the ideal gas.

The chemical potential is therefore:

$$\mu = -k_B T \log \left(\frac{Z_1}{N} \right) = k_B T \log \left[\frac{N}{V} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right] \quad (96)$$

Which is the correct expression for the chemical potential of the ideal gas **with** the correction to the Gibbs paradox already included.

Note that the grand canonical free energy in this limit is:

$$\Omega = -k_B T \log \mathcal{Z} = \pm k_B T \sum_k \log \left(1 \mp e^{-\beta(\epsilon_k - \mu)} \right) \simeq -k_B T \sum_k e^{-\beta(\epsilon_k - \mu)} = -N k_B T \quad (97)$$

Where in the last equality we used the expression for N in the limit of $n_k \ll 1$.

We therefore recover a very interesting relation between the grand canonical free energy and the energy of the system (ideal gas) which is:

$$\Omega = -\frac{2}{3} E \quad (98)$$

Finally, the Helmholtz free energy is given by:

$$F = \Omega + \mu N = -N k_B T \left\{ \log \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] + 1 \right\} \quad (99)$$

Which is the familiar expression for the free energy corrected for the Gibbs paradox.

Lecture Notes #9

1. Statistical physics of electromagnetic radiation**1.1 General considerations**

We consider a cavity made of matter. Inside the empty cavity there exists electromagnetic radiation, since the material of the walls contain charges that move at finite temperature and therefore radiate. In general radiation in the cavity is emitted and absorbed in the walls until equilibrium is reached. At that time the radiation in the cavity is characterized by an equilibrium distribution function, $\rho(\omega)d\omega$, which represents the amount of energy of radiation per unit volume with frequency between ω and $\omega + d\omega$, at temperature T .

An important property of $\rho(\omega)$, which makes it possible to calculate it, is the fact that at equilibrium it is independent of the structure of the cavity or of the materials the cavity is made of.

The above property can be divided into two statements, both of which can be proven using the second law of thermodynamics.

a) The distribution of radiation in the cavity is independent of position.

This can easily be proven in the following way. Assume two closed cavities, A and B . Bringing the two cavities together and isolating this composite system will bring the two to a constant temperature T . If we now open a hole in the wall that connects the two cavities, radiation will go through the hole in both directions. Assume now that due to the non-uniformity of the radiation, more radiation crosses the hole from A to B , than in the opposite direction. In this case the radiation will no longer be in thermal equilibrium with the walls on both sides of the hole. Equilibrium will be reached, but with a new temperature that is different than the original temperature, all this without any work done. In other words, heat passes between the two parts of the system which is at the same temperature, without performing any work, which is in contradiction to the second law of thermodynamics.

b) The distribution of the radiation in the cavity is independent of the direction of the wave vector or polarization of the radiation.

This can be proved in a similar way. In a cavity we place two identical prisms that are coated with perfectly reflecting material except for one surface. We assume that in the first place all components of this cavity are at a constant temperature T and the full system is isolated. The non-reflecting surface can absorb radiation at one particular frequency, ω_0 , and reflect all other frequencies. We place the two prisms perpendicular to each other, such that wave vectors \vec{q}_1 hit one surface, while wave vectors \vec{q}_2 hit the other surface. We further assume that $\omega_0 = c|\vec{q}_1| = c|\vec{q}_2|$. If indeed $\rho(\vec{q}_1, T) > \rho(\vec{q}_2, T)$, while the frequency is the same, the second body will absorb more radiation and thus its temperature will increase, again, since no work has been done, it is in contradiction to the second law of thermodynamics.

1.2 Photons and the electromagnetic field

Unlike the case of particles in a box, or even magnetic moments, classically radiation is represented by wave solutions to the Maxwell equations with no “real particles” to do statistical mechanics. Indeed, photons, which are the quanta of radiation are a quantum-mechanical consequence of the particle-wave duality. Photons appear as a consequence of quantization of the electromagnetic field.

To discuss any property of the electromagnetic field we first recall Maxwell equations in vacuum:

$$\nabla \cdot \vec{\mathcal{E}} = 0 \tag{1}$$

$$\nabla \cdot \vec{H} = 0 \tag{2}$$

$$\nabla \times \vec{\mathcal{E}} + \frac{1}{c} \frac{\partial \vec{H}}{\partial t} = 0 \tag{3}$$

$$\nabla \times \vec{H} - \frac{1}{c} \frac{\partial \vec{\mathcal{E}}}{\partial t} = 0 \tag{4}$$

Where $\vec{\mathcal{E}}$ is the electric field and \vec{H} is the magnetic field. The two fields are most conveniently expressed in terms of the potentials \vec{A} and Φ such that:

$$\vec{\mathcal{E}} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \Phi; \quad \text{and} \quad \vec{H} = \vec{\nabla} \times \vec{A} \tag{5}$$

Starting from the classical expression for the electromagnetic energy:

$$E = \frac{1}{8\pi} \int [\mathcal{E}^2 + H^2] \quad (6)$$

Define now a canonical momentum $\vec{P} = -\frac{1}{4\pi c}\vec{\mathcal{E}}$ which is conjugate to the vector potential \vec{A} , the hamiltonian of the system is then;

$$\mathcal{H} = \int \left[2\pi c^2 P^2 + \frac{1}{8\pi} (\vec{\nabla} \times \vec{A})^2 \right] d^3r \quad (7)$$

In order for us to go to a quantum mechanical description, we need to change the hamiltonian to an operator. In that case the canonical field variables also become operators and will satisfy the commutation relation:

$$[A_i(\vec{r}, t), A_j(\vec{r}', t)] = [P_i(\vec{r}, t), P_j(\vec{r}', t)] = 0; \quad \text{and} \quad [A_i(\vec{r}, t), P_j(\vec{r}', t)] = i\hbar \delta_{ij} \delta^3(\vec{r} - \vec{r}') \quad (8)$$

In a standard way we can calculate the equation of motion for a typical component of \vec{A} :

$$i\hbar \dot{A}_i(\vec{r}, t) = [A_i(\vec{r}, t), \mathcal{H}] = 4\pi c^2 i\hbar P_i(\vec{r}, t) \quad (9)$$

It is convenient now to describe the potentials in terms of a complete set of plane waves that are taken to be vector functions of \vec{r} and are polarized perpendicular to the propagation vector \vec{k} so that the conditions $\nabla \cdot \vec{A} = \nabla \cdot \vec{P} = 0$. We use the following plane waves:

$$\varphi_{\vec{k}, \lambda} = \frac{1}{\sqrt{V}} \vec{\epsilon}_{\vec{k}, \lambda} e^{i\vec{k} \cdot \vec{r}} \quad (10)$$

Where $\vec{\epsilon}_{\vec{k}, \lambda}$ is a unit vector such that $\vec{k} \cdot \vec{\epsilon}_{\vec{k}, \lambda} = 0$. V is the volume of the system and $\lambda = \pm 1$ are the two polarization states.

It is easy to see now that

$$\int \varphi_{\vec{k}, \lambda}^* \varphi_{\vec{k}', \lambda'} d^3r = \delta_{\vec{k}\vec{k}'} \delta_{\lambda, \lambda'} \quad (11)$$

The expansion of the two fields will include amplitudes $q_{\vec{k}, \lambda}$ and $p_{\vec{k}, \lambda}$ that by imposing the commutation relations (8) will become operators. Thus:

$$\vec{A}(\vec{r}, t) = \sum'_{\vec{k}, \lambda} \left[q_{\vec{k}, \lambda} \varphi_{\vec{k}, \lambda} + q_{\vec{k}, \lambda}^\dagger \varphi_{\vec{k}, \lambda}^* \right] \quad \text{and} \quad \vec{P}(\vec{r}, t) = \sum'_{\vec{k}, \lambda} \left[p_{\vec{k}, \lambda} \varphi_{\vec{k}, \lambda} + p_{\vec{k}, \lambda}^\dagger \varphi_{\vec{k}, \lambda}^* \right] \quad (12)$$

Where the prime denotes summation only over positive \vec{k} in order not to duplicate the negative ones with the complex conjugate. It is easy to see now that

$$[q_{\vec{k}, \lambda}, p_{\vec{k}, \lambda}^\dagger] = [q_{\vec{k}, \lambda}^\dagger, p_{\vec{k}, \lambda}] = i\hbar \delta_{\vec{k}\vec{k}'} \delta_{\lambda, \lambda'} \quad (13)$$

In order for equation 8 to be satisfied.

Substituting now the expressions for \vec{A} and \vec{P} into the field hamiltonian we get:

$$\hat{\mathcal{H}} = \sum'_{\vec{k},\lambda} \left(4\pi c^2 p_{\vec{k},\lambda} p_{\vec{k},\lambda}^\dagger + \frac{k^2}{4\pi} q_{\vec{k},\lambda} q_{\vec{k},\lambda}^\dagger \right) \quad (14)$$

We note that both operators, $q_{\vec{k},\lambda}$ and $p_{\vec{k},\lambda}$ are time dependent. In fact, to find their equation of motion we can perform their commutation relation with the hamiltonian.

$$i\hbar \dot{q}_{\vec{k},\lambda} = [q_{\vec{k},\lambda}, \hat{\mathcal{H}}] = 4\pi i\hbar c^2 p_{\vec{k},\lambda}; \quad \text{and} \quad i\hbar \dot{p}_{\vec{k},\lambda} = [p_{\vec{k},\lambda}, \hat{\mathcal{H}}] = \frac{i\hbar k^2}{4\pi} q_{\vec{k},\lambda} \quad (15)$$

The two equations can be combined to give the equation for $q_{\vec{k},\lambda}$:

$$\ddot{q}_{\vec{k},\lambda} = -k^2 c^2 q_{\vec{k},\lambda} = -\omega_k^2 q_{\vec{k},\lambda} \quad (16)$$

which is the equation of motion for a harmonic oscillator with frequency $\omega_k = ck$.

Thus, in a similar way we treated a one-dimensional harmonic oscillator, also here we will define creation and annihilation operators that are time independent such that:

$$q_{\vec{k},\lambda} = \sqrt{\frac{4\pi\hbar c}{k}} \left(a_{\vec{k},\lambda} e^{-i\omega_k t} + a_{-\vec{k},\lambda}^\dagger e^{i\omega_k t} \right) \quad (17)$$

and

$$p_{\vec{k},\lambda} = i\sqrt{\frac{\hbar k}{4\pi c}} \left(-a_{\vec{k},\lambda} e^{-i\omega_k t} + a_{-\vec{k},\lambda}^\dagger e^{i\omega_k t} \right) \quad (18)$$

where again we remember to use only half of the \vec{k} s. The above equations can be inverted to recover the familiar relation for the harmonic oscillator creation and annihilation operators:

$$a_{\vec{k},\lambda} = \frac{1}{2} \left(\sqrt{\frac{k}{4\pi\hbar c}} q_{\vec{k},\lambda} + i\sqrt{\frac{4\pi c}{\hbar k}} p_{\vec{k},\lambda} \right) e^{i\omega_k t}; \quad \text{and} \quad a_{-\vec{k},\lambda} = \frac{1}{2} \left(\sqrt{\frac{k}{4\pi\hbar c}} q_{\vec{k},\lambda} - i\sqrt{\frac{4\pi c}{\hbar k}} p_{\vec{k},\lambda} \right) e^{-i\omega_k t} \quad (19)$$

where again we are consistent in using only positive \vec{k} values.

Using the new operators, the hamiltonian can now rewritten as:

$$\hat{\mathcal{H}} = \sum'_{\vec{k},\lambda} \hbar\omega_k \left(a_{\vec{k},\lambda} a_{\vec{k},\lambda}^\dagger + a_{-\vec{k},\lambda} a_{-\vec{k},\lambda}^\dagger \right) = \sum'_{\vec{k},\lambda} \hbar\omega_k \left(a_{\vec{k},\lambda}^\dagger a_{\vec{k},\lambda} + a_{-\vec{k},\lambda}^\dagger a_{-\vec{k},\lambda} + 1 \right) \quad (20)$$

Where in the second equality we used the commutation relation to change the order of the operators. It is easy to see that we can have just one sum without the “prime” if we sum over all \vec{k} . In that case:

$$\hat{\mathcal{H}} = \sum_{\vec{k}, \lambda} \hbar \omega_k \left(a_{\vec{k}, \lambda}^\dagger a_{\vec{k}, \lambda} + \frac{1}{2} \right) \equiv \sum_{\vec{k}, \lambda} \hbar \omega_k \left(\hat{N}_{\vec{k}, \lambda} + \frac{1}{2} \right) \quad (21)$$

The interpretation of this final result is quite straightforward. Using plane-wave description we identified the eigen-states of the electromagnetic field as a collection of excitation, each described by a harmonic oscillator of unique frequency ω_k and a propagation wave vector \vec{k} . The hamiltonian therefore describe the number of each type of such oscillators. We identify these oscillators as **photons** and thus the hamiltonian counts the number of photons of each frequency and multiply it by its energy to obtain the total energy of the system.

1.3 Finding $\rho(\omega)$:

The result of the previous section is that the hamiltonian of the electromagnetic field is a hamiltonian of harmonic oscillators representing photons. It is also obvious that the number of photons is not fixed, and will be determined by the conditions we put the system. For each frequency ω_k there are n_k photons with that frequency. The total energy of the photons gas is calculated from the expectation value of the hamiltonian:

$$E = \langle \hat{\mathcal{H}} \rangle = \sum_{\vec{k}, \lambda} \hbar \omega_k \left(\langle \hat{N}_{\vec{k}, \lambda} \rangle + \frac{1}{2} \right) = \sum_{\vec{k}, \lambda} \hbar \omega_k \left(n_{\vec{k}, \lambda} + \frac{1}{2} \right) \quad (22)$$

Neglecting the zero-point energy of the harmonic oscillators $\sum_{\vec{k}, \lambda} \frac{1}{2} \hbar \omega_k$, we assume that there are photons of all frequencies in the cavity with total energy:

$$E = \sum_{\vec{k}, \lambda} \hbar \omega_k n_{\vec{k}, \lambda} \quad (23)$$

The fact that the number of photons is not fixed in the cavity implies that the chemical potential for photons is zero (i, other words, since there is no explicit dependence of the free energy on N , $\mu = 0$). The thermodynamic problem is therefore one of a gas of bosons with $\mu = 0$ for which the partition function is:

$$\mathcal{Z}(T, V) = Z(T, V) = \prod_{\vec{k}, \lambda} \left(\frac{1}{1 - e^{-\beta \hbar \omega_k}} \right) \quad (24)$$

where we noted that for this case the grand canonical partition function equals to the canonical one.

The average occupation number of the $\{\vec{k}, \lambda\}$ -state is:

$$\langle n_{\vec{k}, \lambda} \rangle = \frac{1}{e^{\beta \hbar \omega_k} - 1} \quad (25)$$

The energy of the photon gas in the cavity is therefore:

$$E = \sum_{\vec{k}, \lambda} \hbar \omega_k n_{\vec{k}, \lambda} = \sum_{\vec{k}, \lambda} \frac{\hbar \omega_k}{e^{\beta \hbar \omega_k} - 1} \quad (26)$$

The summation is complicated and thus we want to integrate over all states. Thus, we first need to calculate the density of states.

Density of states:

While the dispersion relation, i.e. the relation between the energy of the photon and its wave vector is a continuous function of k , the actual wave vectors, or, the actual photons that will fit in the cavity will be determined by imposing boundary conditions. For simplicity we assume a cubic cavity of side L . Imposing e.g. periodic boundary conditions on any component of the generalize coordinates A_i or P_i , the boundary conditions imply:

$$e^{ik_i L} = e^{ik_i \cdot 0} = 1; \quad \text{or} \quad k_i L = 2\pi n_i \quad (27)$$

Where $i = x, y, z$ and $n_i = 0, 1, 2, \dots$ is an integer associated with the i^{th} component. The sum in (26) is over all the possible \vec{k} s and λ s. The sum over $\lambda = \pm 1$ is easy because it just multiply the result by a factor of 2 for the two polarizations. The sum over the possible \vec{k} s leads to a sum over the n_i s, which result in the following density of states:

$$\Delta n_x \Delta n_y \Delta n_z = \frac{L^3}{(2\pi)^3} \Delta k_x \Delta k_y \Delta k_z \rightarrow \frac{V}{8\pi^3} 4\pi k^2 dk \quad (28)$$

Here we use the fact that the volume of the system is $V = L^3$, and in the last equality we used spherical coordinates. The density of states is:

$$g(\omega) d\omega = 2 \frac{V}{8\pi^3} 4\pi k^2 dk = \frac{V}{\pi^2 c^3} \omega^2 d\omega \quad (29)$$

The factor of two comes from the two polarization states.

Energy density:

Continuing now with the equation for the energy we find:

$$E = \sum_{\vec{k}, \lambda} \frac{\hbar \omega_k}{e^{\beta \hbar \omega_k} - 1} \rightarrow \frac{V}{\pi^2 c^3} \int_0^\infty \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} \omega^2 d\omega \quad (30)$$

The definition of $\rho(\omega)$ is the energy density per unit volume. Comparing this to equation 30 we find:

$$\rho(\omega)d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1} d\omega \quad (31)$$

Which is also called the Planck distribution function.

1.4 Properties of a black body

To integrate equation 30 we first change variable to $x = \beta\hbar\omega$. We get:

$$E = \frac{V}{\pi^2 c^3} \beta^{-4} \hbar^{-3} \int_0^\infty \frac{x^3}{e^x - 1} dx = V \frac{(k_B T)^4}{\pi^2 (\hbar c)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx \quad (32)$$

After evaluating the integral, we can calculate the energy density in the black body $u = E/V$:

$$u = \frac{\hbar \pi^2}{15 c^3} \left(\frac{k_B T}{\hbar} \right)^4 \equiv \frac{4\sigma}{c} T^4; \quad \sigma = 5.67 \times 10^{-8} \text{Joule} \cdot \text{sec}^{-1} \text{m}^{-2} \text{K}^{-4} \quad (33)$$

Which is the so-called Stephan-Boltzman formula for the energy density a black body radiates. The constant σ is the Stephan-Boltzman coefficient.

We next investigate the formula for $\rho(\omega)$. We note that this density of radiation is very low at low frequencies and at very high frequencies where it decays exponentially. In fact, the distribution has a maximum that can be found by calculating $d\rho(\omega)/d\omega = 0$. the result is the so-called Wien's law that:

$$\frac{\nu_{max}}{T} = 5.88 \times 10^{10} \text{ sec}^{-1} \text{K}^{-1} \quad (34)$$

where $\nu = \omega/2\pi$.

Emissivity

In the discussion of the homogeneity of the radiation in a cavity we made a construction of two cavities with a small aperture connecting them. We were looking at the amount of radiation that was escaping from the aperture going from one cavity to the other. To find the relation between this quantity, which is called the emissivity of the cavity to the energy density, $\rho(\omega)$, we make the following arguments.

The amount of radiation energy with frequency between ω and $\omega + d\omega$ in a volume element dV is: $\rho(\omega)d\omega dV$. The radiation propagates from this volume element in all directions with velocity c . The part of the radiation which is directed towards the surface element in the aperture is: $(\Delta A \cos\theta)/4\pi r^2$ which is the ratio of the surface element to the area of the whole sphere multiplied by the tilt of the area element ($\cos\theta$).

the energy of radiation that will cross the unit area during time Δt is therefore:

$$\Delta E = \int_0^{c\Delta t} dr \int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi r^2 \sin\theta \rho(\omega) \Delta\omega \frac{\Delta A \cos\theta}{4\pi r^2} = \frac{1}{2} \Delta A \Delta\omega \rho(\omega) \int_0^{c\Delta t} dr \int_0^{\pi/2} \sin\theta \cos\theta d\theta \quad (35)$$

here we used the infinitesimal volume element in spherical coordinates: $dV = r^2 dr \sin\theta d\theta d\phi$.

Integrating over the azimuthal angle and over the radius we get:

$$\Delta E = \frac{c}{4} \rho(\omega) \Delta\omega \Delta t \Delta A \quad (36)$$

Thus, the emissivity of the cavity (the black body) is given by the energy emitted from a unit area of the aperture per unit time:

$$I(\omega)d\omega = \frac{c}{4} \rho(\omega)d\omega \quad (37)$$

2. Statistical mechanics of lattice vibrations

2.1 General considerations

Previously we discussed the Einstein model for solids in which we modeled the bound atoms in the solid as masses and springs, but considered only the independent vibration of each mass irrespective of the other masses. While the Einstein model gave the correct answer at high temperatures, i.e. the classical limit, it resulted with a specific heat that is much lower than the one observed experimentally. We understood this discrepancy by the fact that at low temperatures only excitations with low energies are possible and these will involve collective modes of the masses and springs. Thus we need to go beyond the Einstein model to find the correct low temperatures behavior of a solid.

The Debye model was constructed to correctly account for the possible long wavelength modes that can be excited at low temperatures. To develop this model we consider a three dimensional cubic solid of N atoms and a lattice constant a . We expand the potential between every two adjacent atoms around the equilibrium position (distance - a) up to the harmonic

term. The harmonic term gives the “spring constant K of the connection between adjacent atoms. Due to the cubic symmetry we can decouple the problem into three one-dimensional problems of masses and springs. thus consider one of these directions (say, x), we denote by $u_j(t)$ the displacement of the j^{th} atom, the equation of motion of this atom that also takes into account the coupling to the $j - 1$ and $j + 1$ masses in the chosen direction is:

$$m \frac{d^2 u_j}{dt^2} = -K(u_j - u_{j-1}) + K(u_{j+1} - u_j) = K(u_{j+1} + u_{j-1} - 2u_j) \quad (38)$$

Since we are looking for solutions that oscillatory in time, we try:

$$u_j(t) = \tilde{u}_j e^{-i\omega t} \quad (39)$$

which leads to the difference equation:

$$-m\omega^2 \tilde{u}_j = K(\tilde{u}_{j+1} + \tilde{u}_{j-1} - 2\tilde{u}_j) \quad (40)$$

We want to go now to the continuum limit. In this case we note that in going from site j to $j + 1$, the distance increases by one lattice constant. Thus, $x_j = ja$. Thus we try a solution of the type:

$$\tilde{u}_{j+1} = \tilde{u}_j e^{\pm ika}; \quad \tilde{u}_{j-1} = \tilde{u}_j e^{\mp ika} \quad (41)$$

Substitution into equation 40 gives:

$$-m\omega^2 \tilde{u}_j = K(e^{ika} + e^{-ika} - 2)\tilde{u}_j \quad (42)$$

Which, after some algebra gives the dispersion relation:

$$\omega^2 = \frac{4K}{m} \sin^2 \left(\frac{ka}{2} \right) \quad (43)$$

We note that in the high frequency limit, i.e. in the limit where the wavelength is of order of lattice constant, $ka/2 \sim \pi/2$, and the frequency is independent of the wavelength, just like we assumed for the einstein model. However, in the long wavelengths limit, $ka \ll 1$ and we can use the approximation $\sin(\frac{ka}{2}) \sim \frac{ka}{2}$. In this limit we find

$$\omega \sim \sqrt{\frac{K}{m}} ka = \sqrt{\frac{K/a}{m/a^3}} k = \sqrt{\frac{B}{\rho}} k \equiv v_s k \quad (44)$$

Here we identified the Bulk modulus B and the density ρ of the solid. v_s is the sound velocity in that solid.

The result we obtained showed that if the frequency is not too high, the dispersion relation of the lattice vibrations is similar to that of electromagnetic radiation, except that the speed

of light is replaced by the speed of sound in the solid. Moreover, in a similar way to the treatment of the electromagnetic field we can treat here the strain field of the deformations in the solid and subsequently quantize the resulting excitations in the same way we did for the electromagnetic field. the resulting quanta of lattice vibrations are called **phonons** and they resemble photons in all aspects except for the fact that there is no reason to forbid longitudinal waves. Thus in the case of lattice vibrations we expect 3 polarizations, two transverse associated with shear modes and one longitudinal associated with compressional mode.

Another important distinction between the quanta of lattice vibrations and photons is the fact that the number of lattice vibrations is bound from above by the total number of degrees of freedom for individual atom. i.e., by the fact that we cannot have more modes than $3N$ in three dimensions.

2.2 The Debye model

In the Debye approximation we use the dispersion relation $\omega = v_s k$ for all the modes. Thus it will be a good approximation for the low temperatures limit of lattice vibrations. In analogy with photons, phonons are massless bosons and thus their energy is given by:

$$E = \sum_{\vec{k}, \lambda} \hbar \omega_k n_{\vec{k}, \lambda} \quad (45)$$

while now, $\lambda = 1, 2, 3$ for the three polarizations.

The occupation number of the $\{\vec{k}, \lambda\}$ state is given by:

$$\langle n_{\vec{k}, \lambda} \rangle = \frac{1}{e^{\beta \hbar \omega_k} - 1} \quad (46)$$

where again we used the fact that since the number of excited phonons is not fixed, the chemical potential is zero.

The density of states for phonons will be similar to that of photons, except for the use of the speed of sound and the different number of polarizations. If we make the approximation that the speed of sound of the three polarizations is the same, the density of states is:

$$g(\omega) d\omega = 3 \frac{V}{8\pi^3} 4\pi k^2 dk = \frac{3V}{2\pi^2 v_s^3} \omega^2 d\omega \quad (47)$$

The difference with the gas of photons is that the number of phonons has to be equal to the number of degrees of freedom in the system. Since there are N atoms in a three-dimensional space the number of degrees of freedom is $3N$. Thus we must satisfy:

$$3N = \int_0^{\omega_{max}} g(\omega) d\omega \quad (48)$$

where the maximum frequency is of order v_s/a as expected from equation 43. It is common to call this cutoff frequency the Debye frequency, i.e. $\omega_{max} \equiv \omega_D$. Using the above expression for $g(\omega)$ we get:

$$3N = \frac{3V}{2\pi^2} \frac{1}{v_s^3} \frac{\omega_D^3}{3} \quad (49)$$

which gives the self-consistent expression for ω_D :

$$\omega_D = (6\pi^2 n)^{1/3} v_s \quad (50)$$

and the corresponding Debye wave vector:

$$k_D = \omega_D/v_s = (6\pi^2 n)^{1/3} \quad (51)$$

where $n = N/V$ is the number density of atoms in the solid.

The expression for the energy becomes;

$$E = \frac{3V}{2\pi^2} \frac{\hbar}{v_s^3} \int_0^{\omega_D} \frac{\omega^3}{e^{\beta\hbar\omega} - 1} d\omega \quad (52)$$

Again we note that the distribution function inside the integrand has the form of the planck distribution function. Changing variables to $x = \beta\hbar\omega$ we get:

$$E = \frac{3V}{2\pi^2} \frac{(k_B T)^4}{\hbar^3 v_s^3} \int_0^{x_D} \frac{x^3}{e^x - 1} dx \quad (53)$$

with $x_D = \beta\hbar\omega_D \equiv \frac{\theta_D}{T}$.

Using now the expression for ω_D , we get:

$$E = 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \int_0^{\frac{\theta_D}{T}} \frac{x^3}{e^x - 1} dx \quad (54)$$

High temperatures limit

In this limit $k_B T \gg \hbar\omega_D$. We can approximate the integrand as $e^x - 1 \simeq x$ and thus:

$$E \approx 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \frac{1}{3} \left(\frac{\theta_D}{T} \right)^3 = 3Nk_B T \quad (55)$$

This result is in accordance with the equipartition theorem and with the einstein model. There are $3N$ one-dimensional harmonic oscillators in the system, each contributes $2 \times \frac{1}{2} k_B T$ to the energy for the two quadratic degrees of freedom.

the result for the specific heat in this high-temperatures limit is:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = 3Nk_B; \quad \text{and} \quad c_V = 3k_B \quad (56)$$

which is indeed the expected classical limit for $3N$ harmonic oscillators. This limit is also called the *Dulong – Petit* limit.

Low temperatures limit

In that limit we can set $x_D \rightarrow \infty$. the value of the integral is then $\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}$ and the energy is:

$$E \approx 9Nk_B T \left(\frac{T}{\theta_D} \right)^3 \frac{\pi^4}{15} = \frac{3\pi^4}{5} Nk_B T \left(\frac{T}{\theta_D} \right)^3 \propto T^4 \quad (57)$$

The proportionality to the fourth power of temperature is similar to the one we found for a gas of photons.

The heat capacity of the system at low temperatures is:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\theta_D} \right)^3 \quad (58)$$

Thus in the case of phonons, at low temperatures the specific heat is proportional to T^3 , in contrast to the result of the Einstein model that predicted exponentially vanishing specific heat at low temperatures.

Lecture Notes #10

1. Statistical physics of gas of fermions

1.1 General considerations

As was discussed before, Pauli exclusion principle imposes that at most, only one fermion can occupy an allowed state. Thus, $n_k = 0$ or 1 , for each quantum state with energy ϵ_k and we obtained:

$$\Omega = -k_B T \log \mathcal{Z} = -k_B T \sum_k \log \left(1 + e^{-\beta(\epsilon_k - \mu)} \right) \quad (1)$$

The occupation number of the k^{th} state is:

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1} \quad (2)$$

and the self consistent equation for the chemical potential is derived while calculating the average number of particles in the system:

$$N = \sum_k \langle n_k \rangle = \sum_k \frac{1}{e^{\beta(\epsilon_k - \mu)} + 1} \quad (3)$$

The energy of the system is:

$$E = \sum_k \langle n_k \rangle \epsilon_k = \sum_k \frac{\epsilon_k}{e^{\beta(\epsilon_k - \mu)} + 1} \quad (4)$$

To proceed we need to discuss the nature of the single particle states, so we can either perform the summation, or change it into an integral. Let us first discuss a gas of fermions for which the single particle hamiltonian is:

$$\hat{h} = \frac{\hat{p}^2}{2m} = -\frac{1}{2m} \left(\frac{\hbar}{i} \vec{\nabla} \right)^2 \quad (5)$$

The resulting eigenstates are momentum states with $\vec{p} = \hbar \vec{k}$ and the single particle energies (the dispersion relation) are:

$$\epsilon_k = \frac{\hbar^2 k^2}{2m} \quad (6)$$

The density of states is therefore:

$$g(\epsilon)d\epsilon = g_0 \frac{V}{8\pi^3} 4\pi k^2 dk = g_0 \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon^{1/2} d\epsilon \quad (7)$$

where g_0 is the spin degeneracy which we will take as equal to 2 which applies to a gas of electrons. In the continuum limit, the occupation number becomes a continuous function of the single particle energy which we denote by $f(\epsilon)$, also called the Fermi function.

$$f(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} + 1} \quad (8)$$

the number of particles and energy can now be expressed in terms of the density of states and the fermi function:

$$N = \int_0^\infty g(\epsilon) f(\epsilon) d\epsilon \quad \text{and} \quad E = \int_0^\infty \epsilon g(\epsilon) f(\epsilon) d\epsilon \quad (9)$$

1.2 Properties of the Fermi function

In examining equation 8 we note that the limit of $T \rightarrow 0$ depends on how large is the energy with respect to the chemical potential. In fact, for $\epsilon > \mu(T=0)$ the denominator diverges and thus the limit of zero temperatures is $f = 0$. On the other and, if $\epsilon < \mu(T=0)$, the exponential will vanish at zero temperature and $f = 1$. We therefore conclude that at $T = 0$ $f(\epsilon)$ is a step function

$$f(\epsilon)_{T=0} = \begin{cases} 1 & \epsilon < \epsilon_F \\ 0 & \epsilon > \epsilon_F \end{cases} \quad (10)$$

This result is very easy to interpret. Since we are dealing with N Fermi particles at zero temperature, the particles are in their ground state. we therefore start to fill the single particle levels from the lowest level, one particle per energy state (or in fact two, if we cannot distinguish the two spin states). The energy level filled by the last particle (the N^{th} particle is the Fermi energy, ϵ_F . This interpretation also makes sense from the point of view of the chemical potential being the energy needed to add a particle to the system.

At finite temperature $f(\epsilon, T)$ is modified. Starting from a very low temperature we see that not much should happen for $k_B T \ll \epsilon_F$. Since the only appreciable contribution from the exponential is if it is of order unity, or the argument is very small, the contribution to f will come only from energies satisfying $|\epsilon - \mu(T)|/k_B T \leq 1$. Or since $\mu(0) = \epsilon_F$, the contribution will come from a region of $\Delta\epsilon \sim k_B T$ around the Fermi energy.

The best way to see the behavior of $f(\epsilon)$ as a function of temperature is to look at its derivative $\partial f/\partial T$. For $T = 0$ the derivative is a delta-function. For a finite temperature we calculate:

$$-\frac{\partial f}{\partial \epsilon} = \frac{\beta e^{\beta(\epsilon-\mu)}}{[e^{\beta(\epsilon-\mu)} + 1]^2} \quad (11)$$

Which again peaks only within a $k_B T$ around $\mu(0)$, and in the limit of $T \rightarrow 0$ becomes a delta function. we therefore conclude that

$$\lim_{T \rightarrow 0} \left[-\frac{\partial f}{\partial \epsilon} \right] = \delta(\epsilon - \mu(0)) = \delta(\epsilon - \epsilon_F) \quad (12)$$

1.3 Calculation of the integrals

It is easy to see that both integrals in 32 involve integrals of the type $\int \epsilon^\alpha f(\epsilon) d\epsilon$. Thus, we can develop a simple method to calculate these integrals, based on the properties of the Fermi function discussed above. We first note that

$$\int_0^\infty f(\epsilon) \epsilon^\alpha d\epsilon = \left[f(\epsilon) \frac{\epsilon^{\alpha+1}}{\alpha+1} \right]_0^\infty - \frac{1}{\alpha+1} \int_0^\infty \left(\frac{df}{d\epsilon} \right) \epsilon^{\alpha+1} d\epsilon \quad (13)$$

The first term vanishes at zero due to the $\epsilon^{\alpha+1}$ term and at infinity due to $f(\epsilon)$. The second term has in it a derivative of almost a step function where the step is at $\epsilon = \mu(T)$. It therefore suggests that we may want to expand the function $\epsilon^{\alpha+1}$ around μ . This gives:

$$\epsilon^{\alpha+1} = \sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{d^m}{d\epsilon^m} \epsilon^{\alpha+1} \right)_{\epsilon=\mu} (\epsilon - \mu)^m \quad (14)$$

such that:

$$\int_0^\infty f(\epsilon) \epsilon^\alpha d\epsilon = -\frac{1}{\alpha+1} \int_0^\infty \left(\frac{df}{d\epsilon} \right) \epsilon^{\alpha+1} d\epsilon = -\frac{1}{\alpha+1} \sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{d^m}{d\epsilon^m} \epsilon^{\alpha+1} \right)_\mu \int_0^\infty \left(\frac{df}{d\epsilon} \right) (\epsilon - \mu)^m d\epsilon \quad (15)$$

Let us now change variables to $x = \beta(\epsilon - \mu)$. The above result reads:

$$\int_0^\infty f(\epsilon) \epsilon^\alpha d\epsilon = \frac{1}{\alpha+1} \sum_{m=0}^{\infty} \frac{(k_B T)^m}{m!} \left(\frac{d^m}{d\epsilon^m} \epsilon^{\alpha+1} \right)_\mu I_m(x) \quad (16)$$

Where

$$I_m(x) = \int_{-\beta\mu}^{\infty} \frac{e^x}{(e^x + 1)^2} x^m dx \quad (17)$$

Looking in tables we find that $I_0 = 1$, and $I_2 = \pi^2/3$.

As argued above, we do not expect μ to change much from ϵ_F if the temperature is low. Thus, in that limit we can replace the lower bound with $-\infty$.

$$\underline{T = 0}$$

We already showed that at $T = 0$ the Fermi function is a step function and acquires the unity value below the Fermi energy. In this case the calculation of both integrals for the number and energy are very simple and we get:

$$N = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\epsilon_F} \epsilon^{1/2} d\epsilon = \frac{V}{3\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon_F^{3/2} \equiv V \frac{k_F^3}{3\pi^2} \quad (18)$$

Where we identified the Fermi wave vector k_F , which through the density of particles is:

$$\epsilon_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3} = \frac{\hbar^2 k_F^2}{2m} \quad (19)$$

The energy at $T = 0$ IS NOT $N\epsilon_F$, but rather integrates over all energies including low energies. We therefore get:

$$\frac{E(T=0)}{V} = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\epsilon_F} \epsilon^{3/2} d\epsilon = \frac{3}{5} n \epsilon_F \quad (20)$$

For an electron gas in a metal the density of electrons is $n \sim 10^{23} e/cm^3$. Using an electron mass of $\sim 10^{-27} gr$, we get $\epsilon_F \sim 10^5 K$, much higher than a typical melting temperature of metals. The respective fermi wave vector is $k_F \sim 10^8 cm^{-1}$.

1.4 The chemical potential

In this case $\alpha = 1/2$, and, since we are interested in the low temperatures limit it is enough to keep the first temperature correction. Using equation 16, and dividing by the volume, V , we get:

$$n = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \left[\frac{2}{3} \mu^{3/2} I_0 + \frac{1}{4} (k_B T)^2 \mu^{-1/2} I_2 \right] \quad (21)$$

To solve for $\mu(T, n)$, we need to invert the equation. Because the $T = 0$ result gives $\mu = \epsilon_F$, the easiest way to invert the equation is to write: $\mu = \epsilon_F + \delta\mu$, and expand in the small parameter $\delta\mu/\epsilon_F$. The result for μ to first order in the temperature correction is:

$$\mu(T, n) = \epsilon_F - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 \epsilon_F \quad (22)$$

As we can see, the correction to the chemical potential is of order $\left(\frac{k_B T}{\epsilon_F} \right)^2$, which for electrons in metal, at room temperature can be a correction which is $\sim 10^{-5}$ times the leading term. Also, we see that the chemical potential is reduced from the Fermi energy. This is possible because as the temperature increases, electrons from near the Fermi level are excited to higher energies, hence making it available for particles to enter the Fermi sea even with energies slightly lower than ϵ_F .

1.5 The energy and specific heat

To calculate the energy we need to substitute $\alpha = 3/2$ in equation 16. Again, since we are interested in the low temperatures limit it is enough to keep only the first temperature correction. We find:

$$E = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \left[\frac{2}{5} \mu^{5/2} I_0 + \frac{3}{4} (k_B T)^2 \mu^{1/2} I_2 \right] \quad (23)$$

We can use now the expression for $\mu(T)$ to find for the energy:

$$E = \frac{3}{5} N \epsilon_F + \frac{\pi^2}{4} N \left(\frac{k_B T}{\epsilon_F} \right)^2 \epsilon_F \quad (24)$$

The heat capacity is:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V = \frac{\pi^2}{4} N \left(\frac{k_B T}{\epsilon_F} \right) k_B \quad (25)$$

The specific heat, i.e. the heat capacity per particle is then:

$$c_V = \frac{\pi^2}{4} \left(\frac{k_B T}{\epsilon_F} \right) k_B \ll k_B \quad (26)$$

We therefore see that the specific heat of the fermions at low temperatures is very small, much smaller than the specific heat expected from a classical gas of particles for which $c_V \sim k_B$. The reason for that is that because of the Fermi statistics we cannot utilize all the particles in the system to carry heat. Only particles near the Fermi energy will be able to be excited. using this logic, it is easy to obtain the above results using a simple argument. Assume a system of N fermions. To find the difference in energy due to temperature we note that only a fraction of $\frac{k_B T}{\epsilon_F}$ of the particles will be excited, each with average energy $k_B T$. Thus, the average increase in energy of the system is:

$$\Delta E \approx \left(N \frac{k_B T}{\epsilon_F} \right) k_B T \quad (27)$$

This is exactly the result of equation 24 leading to:

$$c_V \approx k_B \frac{k_B T}{\epsilon_F} \quad (28)$$

1.6 Low temperatures specific heat of solids

From the Debye calculation that gave us a phonons contribution to the specific heat that varies as T^3 and from the above calculation as applied to the electron gas in a metal we find that the general expression for the low-temperatures specific heat (per particle) of a solid will be:

$$c_V = \gamma T + \beta T^3 \quad (29)$$

Thus, plotting the low temperatures specific heat in the form c_V/T vs. T^2 gives β as the slope and γ as the intercept of the straight line.

1.7 Effect of magnetic field

The Fermi-Dirac statistics that was applied to a gas of fermions such as electrons in metal can also be used to explain the discrepancy between the rather large predicted paramagnetic susceptibility that one would calculate for a gas of electrons and the experimentally observed small one.

Assume an ideal Fermi-dirac gas of spin $\frac{1}{2}$ particles, such as electrons in a metal. In the presence of a magnetic field the single particle hamiltonian is:

$$\hat{h} = \frac{\hat{p}^2}{2m} - \mu_0 \cdot \vec{H} \quad (30)$$

Where μ_0 is the magnetic moment of the particle. Taking the magnetic field to be in the z -direction will give us the two possible spin components in the form of a splitting of the zero field energy to give us:

$$\epsilon_{k,\sigma} = \frac{\hbar^2 k^2}{2m} - \sigma \mu_0 H \quad (31)$$

with $\sigma = \pm 1$.

Adding the magnetic field will change now the density of states as it lifts the 2-fold degeneracy due to the spin (the $g_0 = 2$ in $g(\epsilon)$). Thus, instead of the single equation for the number of particles, N , we have now two equations for the two σ components:

$$N_+ = \int_{-\mu_0 H}^{\infty} g_+(\epsilon) f(\epsilon) d\epsilon \quad N_- = \int_{\mu_0 H}^{\infty} g_-(\epsilon) f(\epsilon) d\epsilon \quad (32)$$

It is easy to see from equation 31 that $g_+(\epsilon) = \frac{1}{2}g(\epsilon + \mu_0 H)$ and that $g_-(\epsilon) = \frac{1}{2}g(\epsilon - \mu_0 H)$, where the factor of $\frac{1}{2}$ is introduced to compensate for $g_0 = 2$ that appears in equation 7.

Since we are interested in calculating the susceptibility, we apply a very small field. In fact, the susceptibility is defined in the limit of $H \rightarrow 0$ and thus we can expand the density of states such that:

$$g_{\pm} = \frac{1}{2}g(\epsilon \pm \mu_0 H) = \frac{1}{2}g(\epsilon) \pm \frac{1}{2}\mu_0 H \left(\frac{dg}{d\epsilon} \right)_{\epsilon} + \frac{1}{2} \frac{1}{2} (\mu_0 H)^2 \left(\frac{d^2 g}{d\epsilon^2} \right)_{\epsilon} + \dots \quad (33)$$

Clearly the first derivative term is of order $\frac{\mu_0 H}{\epsilon} \approx \frac{\mu_0 H}{\epsilon_F}$, and the second term is smaller by the same factor.

To find the susceptibility we need first to calculate the magnetization:

$$M = \mu_0(N_+ - N_-) \quad (34)$$

Where

$$N_+ = \frac{1}{2} \int_{-\mu_0 H}^{\infty} g(\epsilon + \mu_0 H) f(\epsilon) d\epsilon = \frac{1}{2} \int_0^{\infty} g(\epsilon + \mu_0 H) f(\epsilon) d\epsilon + \frac{1}{2} \int_{-\mu_0 H}^0 g(\epsilon + \mu_0 H) f(\epsilon) d\epsilon \quad (35)$$

and

$$N_- = \frac{1}{2} \int_{\mu_0 H}^{\infty} g(\epsilon - \mu_0 H) f(\epsilon) d\epsilon = \frac{1}{2} \int_0^{\infty} g(\epsilon - \mu_0 H) f(\epsilon) d\epsilon + \frac{1}{2} \int_{\mu_0 H}^0 g(\epsilon - \mu_0 H) f(\epsilon) d\epsilon \quad (36)$$

It is therefore clear that in taking the difference between the above equations for N_+ and N_- , the first terms will cancel to first order in the correction $\mu_0 H$ since this term appears with opposite sign in the density of states and thus will appear with opposite sign in the first term of the expansion. the second term will be of order $\left(\frac{\mu_0 H}{\epsilon_F} \right)^2$ and thus negligible. Thus we find:

$$M = \mu_0(N_+ - N_-) \approx \frac{\mu_0^2 H}{2} \int_{-\mu_0 H}^0 \left(\frac{dg}{d\epsilon} \right)_{\epsilon} f(\epsilon) d\epsilon - \frac{\mu_0^2 H}{2} \int_{\mu_0 H}^0 \left(\frac{dg}{d\epsilon} \right)_{\epsilon} f(\epsilon) d\epsilon \quad (37)$$

Integration by parts give:

$$M \approx \frac{\mu_0^2 H}{2} \int_{-\mu_0 H}^0 \left(\frac{df}{d\epsilon} \right)_\epsilon g(\epsilon) d\epsilon - \frac{\mu_0^2 H}{2} \int_{\mu_0 H}^0 \left(\frac{df}{d\epsilon} \right)_\epsilon g(\epsilon) d\epsilon = \mu_0^2 H g(\epsilon_F) \quad (38)$$

Where in the last equality we used the fact that the derivative of f for $H \rightarrow 0$ becomes a delta function.

Substituting now $g(\epsilon_F)$ we obtain the expression for Pauli paramagnetism:

$$M \approx \frac{3N\mu_0^2}{2k_B T_F} H \equiv \chi_P H \quad (39)$$

Where we denote $k_B T_F = \epsilon_F$.

Note that the above result is different from the result we obtained for “classical” moments in a field. It is reduced by a factor $\frac{T}{T_F}$.

Lecture Notes #11

1. Statistical physics of gas of bosons

1.1 General considerations

A gas of bosons does not have any restrictions on the number of particles that can occupy any single quantum state. Thus, $n_k = 0, 1, 2, \dots$ for each quantum state with energy ϵ_k and we obtained:

$$\Omega = -k_B T \log \mathcal{Z} = k_B T \sum_k \log \left(1 - e^{-\beta(\epsilon_k - \mu)} \right) \quad (1)$$

The occupation number of the k^{th} state is:

$$\langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1} \quad (2)$$

and the self consistent equation for the chemical potential is derived while calculating the average number of particles in the system:

$$N = \sum_k \langle n_k \rangle = \sum_k \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1} \quad (3)$$

The energy of the system is:

$$E = \sum_k \langle n_k \rangle \epsilon_k = \sum_k \frac{\epsilon_k}{e^{\beta(\epsilon_k - \mu)} - 1} \quad (4)$$

As before, the energy eigenstates are momentum states with $\vec{p} = \hbar \vec{k}$ and the single particle energies (the dispersion relation) are:

$$\epsilon_k = \frac{\hbar^2 k^2}{2m} \quad (5)$$

The density of states is therefore:

$$g(\epsilon) d\epsilon = g_0 \frac{V}{8\pi^3} 4\pi k^2 dk = g_0 \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon^{1/2} d\epsilon \quad (6)$$

where g_0 is the degeneracy which we will take as equal to 1.

1.2 Occupation of the ground state

Let us look first at the ground state energy, $\epsilon_0 = 0$. The occupation number for this energy is:

$$n_0 = \frac{1}{e^{\beta(\epsilon_0 - \mu)} + 1} = \frac{1}{e^{-\beta\mu} - 1} \quad (7)$$

Since occupation numbers cannot be negative, we require $n_0 > 0$. This implies that $\mu < 0$. Also, since the number of particles is independent of temperature, $\frac{\partial N}{\partial T} = 0$. Therefore:

$$\frac{\partial N}{\partial T} = 0 = \sum_k \frac{[-k_B\beta^2(\epsilon_k - \mu) - \beta\frac{\partial\mu}{\partial T}]e^{\beta(\epsilon_k - \mu)}}{(e^{\beta(\epsilon_k - \mu)} - 1)^2} \quad (8)$$

Since we already proved that $\mu < 0$, the expression above implies that $\frac{\partial\mu}{\partial T} < 0$. We therefore conclude that the chemical potential of an ideal Bose gas is negative, and it monotonically increases as the temperature decreases towards zero. Clearly the case of $\mu = 0$ is pathological because the ground state level occupation number diverges. This may suggest that as the temperature is lowered, μ increases and the ground state level start to become macroscopically occupied.

We can calculate now at what temperature $\mu = 0$. We call this temperature T_0 , or $\beta_0 = (k_B T_0)^{-1}$. Here we need to solve:

$$\sum_k \frac{1}{e^{\beta_0\epsilon_k} - 1} \rightarrow \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{\epsilon^{1/2} d\epsilon}{e^{\beta_0\epsilon} - 1} = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \beta_0^{-3/2} \int_0^\infty \frac{x^{1/2} dx}{e^x - 1} \quad (9)$$

Evaluating the dimensionless integral: $\int_0^\infty x^{1/2}(e^x - 1)^{-1} dx = \Gamma(3/2)\zeta(3/2)$, the solution for T_0 is:

$$T_0 = 0.084 \frac{\hbar^2}{k_B m} n^{2/3} \quad (10)$$

Where we use the particles number density $n = N/V$. We therefore find that at a finite temperature the chemical potential vanishes which seems to indicate a macroscopic occupation of the ground state level.

1.3 Bose-Einstein condensation

Let us assume that indeed the ground state is macroscopically occupied. Let us assume that the occupation of this state is: $\langle n_0 \rangle = N/a$ where $a > 1$ is a number of order unity. Solving equation 7 for the chemical potential in the ground state level we get:

$$\mu = \frac{1}{\beta} \log \left(1 + \frac{a}{N} \right) \approx -k_B T \frac{a}{N} \quad (11)$$

which is of course in agreement with the $\mu < 0$ that we found.

The free energy associated with the ground state occupation, Ω_0 is (to order $(a/N \ll 1)$):

$$\Omega_0 = k_B T \log \left(1 - e^{\beta\mu} \right) \approx k_B T \log \left(1 - e^{-\frac{a}{N}} \right) \approx k_B T \log \left(\frac{a}{N} \right) = \mathcal{O}(\log N) \quad (12)$$

This contribution is only of order $\log N$ and thus small compared to the rest of the terms that contribute order N to the sum.

This situation is not satisfied for the calculation of N itself. This is because the occupation of the ground state level is $\langle n_0 \rangle \equiv N_0 = N/a = \mathcal{O}(N)$. However, this macroscopic contribution, which is present in the sum form for N , is not represented in the integral form of N since there is no divergence of the integral for $\epsilon = 0$:

$$N = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} d\epsilon \quad (13)$$

We therefore conclude that we “lost the contribution at the lower limit of the integration. This was a direct consequence of going from a sum to an integral with density of states that multiplied the macroscopically occupied $\epsilon = 0$ state by zero. Starting back from the sum in equation 3 we write:

$$N = \sum_k \langle n_k \rangle = \frac{1}{e^{\beta(\epsilon_0-\mu)} - 1} + \sum_{k>0} \frac{1}{e^{\beta(\epsilon_k-\mu)} - 1} \quad (14)$$

Where $\epsilon_0 = \epsilon_{k=0} = 0$. Going now to an integral form, the sum with $k > 0$ will still give the same integral with the same density of states, and thus, the correct integral equation for N has to include the ground-state level contribution to give:

$$N = N_0 + \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} d\epsilon \quad (15)$$

and

$$\Omega = \Omega_0 + k_B T \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \epsilon^{1/2} \log \left(1 - e^{-\beta(\epsilon-\mu)} \right) d\epsilon \quad (16)$$

Where N_0 and Ω_0 are as defined above. In particular, the calculation of N at any temperature $T < T_0$ yields:

$$N = N_0 + \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (k_B T)^{3/2} \int_0^\infty \frac{x^{1/2}}{e^x - 1} dx = N_0 + N \left(\frac{T}{T_0} \right)^{3/2} \quad (17)$$

Thus, the number of particles that occupy the ground state $\epsilon = 0$ is:

$$N_0 = N \left[1 - \left(\frac{T}{T_0} \right)^{3/2} \right] \quad (18)$$

The above discussion points to a phase transition in which particles condense into the ground state in macroscopic numbers. This condensation is not like vapor condensation because it occurs in momentum space. In this transition N_0 plays the role of an order parameter that measures the fractional amount of condensation. This is the so-called **Bose-Einstein condensation**.

Since the energy of the ground state is $E_0 = N_0 \epsilon_0 = 0$, the energy at any given temperature $T < T_0$ is just the energy of the remaining uncondensed particles. This is easily calculated as:

$$E = \sum_{k>0} \frac{\epsilon_k}{e^{\beta \epsilon_k} - 1} \rightarrow \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{\epsilon^{3/2}}{e^{\beta \epsilon} - 1} d\epsilon = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} (k_B T)^{5/2} \int_0^\infty \frac{x^{3/2}}{e^x - 1} dx \quad (19)$$

Evaluating the integral: $\int_0^\infty x^{3/2} (e^x - 1)^{-1} dx = \Gamma(5/2) \zeta(5/2)$, we find:

$$E = 0.77 \left(\frac{T}{T_0} \right)^{3/2} N k_B T = 0.77 \left(\frac{T}{T_0} \right)^{5/2} N k_B T_0 \quad (20)$$

which gives a heat capacity:

$$C_V = \frac{5}{2} \frac{E}{T} \propto T^{3/2} \quad (21)$$

We therefore conclude that the specific heat is continuous through the transition, and is vanishing at zero temperature.