Statistical foundations of Thermodynamics

States of a system and importance of degeneracy :

A classical mechanical system composed of discrete particles is completely described by positions and momenta of all the particles constituting the system. A set of values assigned to the position - momentum pairs defines the state of the system. For a quantum mechanical system, apart from the above, other quantum mechanical attributes like spin may additionally be required to be specified.

Each state of a system has a definite energy. The energy of a given state is unique i.e. a given state of the system cannot have more than one energies. However, more than one "quantum states" can have the same energy. Such states which are distinct in configuration but have the same energy are called "degenerate".

We used the term quantum state in the previous paragraph. As mentioned before, the state of the system is specified by position-momentum pairs of all the particles. Such states are called "microstates". Additionally, the quantum mechanical attributes define analogous "quantum states". The state of the system described by conventional macroscopically observed properties of the system (e.g. T, p etc) is called the "macrostate". Degeneracy plays an important role in observing the macroscopic properties. We establish this as follows.

Consider a system containing 'N' particles such that the particles can have either '$+m$' or '$-m$' as the value of a property (say magnetic moment). The energy of the particle depends upon the total magnetic moment of all particles combined. Each particle can be in one of the two states. Hence, we refer the system as a binary-state system.

For a system with only one particle, the number of possible states of the system can be represented as:

state 1 : $\uparrow$

state 2 : $\downarrow$

A two-particle system will have the following microstates.

state 1 : $\uparrow_1 \uparrow_2$

state 2 : $\uparrow_1 \downarrow_2$

state 3 : $\downarrow_1 \uparrow_2$

state 4 : $\downarrow_1 \downarrow_4$

The subscripts refer to the particles. Since every particle can be in one of the two states, it can be easily established that the total number of microstates of an N-particle system equals 2^N.

For a 3-particle system, the total of $2^3 = 8$ microsystems can be generated following the multiplication rule as given below:

$$(\uparrow_1 + \downarrow_1)(\uparrow_2 + \downarrow_2)(\uparrow_3 + \downarrow_3)$$

$$= \uparrow_1 \uparrow_2 \uparrow_3$$

$$+ \uparrow_1 \uparrow_2 \downarrow_3 + \uparrow_1 \downarrow_2 \uparrow_3 + \downarrow_1 \uparrow_2 \uparrow_3$$

$$+ \uparrow_1 \downarrow_2 \downarrow_3 + \downarrow_1 \uparrow_2 \downarrow_3 + \downarrow_1 \downarrow_2 \uparrow_3$$

$$+ \downarrow_1 \downarrow_2 \downarrow_3$$

It should be noted that the states are not being added. This is merely an algorithm to identify all possible microstates. On generalization, for an N-particle binary-state system, the microstates can be identified as

$$(\uparrow_1 + \downarrow_1)(\uparrow_2 + \downarrow_2) \ldots (\uparrow_{N-1} + \downarrow_{N-1})(\uparrow_N + \downarrow_N)$$

Now consider a system in which it is possible to count the number of particles but it is not possible to distinguish the particles. If the two microstates under consideration are $\uparrow_1 \downarrow_2$ and $\downarrow_1 \uparrow_2$ and the particles are identicle then the two microstates become identicle. Hence, the two states $\uparrow_1 \downarrow_2$ and $\downarrow_1 \uparrow_2$ give rise to the same energy and, therefore, are degenerate. The following table summarizes the states of a two-, three- and four-particle system.

No. of particles	No. of states	No. of degenerate states
2	4	1
3	8	2
4	16	3

For any system with a larger number of particles, a similar analysis can be carried out.

The energy of the system depends upon the sum of energy contributions from individual particle's magnetic moments. Consider an arbitrary state of the system in which 'p' number of particles have +m magnetic moment and 'N-p' number of particles have -m magnetic moment.

Total energy of the system =

$$\alpha m p + \alpha(-m)(N-p)$$

where α is a constant. The total number of possible energies of the system can be found out as follows:

- All particles with $+m$: $\alpha m N$
- $(N-1)$ particles with $+m$ and one particle with $-m$:

$$\alpha(N-1)m + \alpha(-m) = \alpha m(N-2)$$

- $(N-2)$ particles with $+m$ and two particles with $-m$:

$$\alpha(N-2)m + 2\alpha(-m) = \alpha m(N-4)$$

$\vdots$

- All particles with $-m$ $= -\alpha m N$

From the above analysis, the followings can be concluded.

Total number of possible values of energies $= N+1$

Total number of possible states of the system $= 2^N$

From the above analysis, it can be seen that $N+1$ values of energies must be distributed among 2^N states. For $N > 1$, 2^N is always greater than $N+1$. This means that many different states will possess the same energy. For a 10-particle system,

Total number of possible values of energies $= 10 + 1 = 11$

Total number of possible states of the system $= 2^{10} = 1024$.

Therefore, 11 different values of energies must be distributed among 1024 states (imagine now for Avogadro number of particles!). obviously, a large number of states will have the same energy.

The above analysis only tells that more than one state can have the same energy. It is required to know the distribution of the states among the available energies i.e. it is desired to know how many states possess a particular energy. For this, we first define "excess property". Consider a system with even number of particles. The "spin-up" and "spin-down" states can be counted as:

$$N_\uparrow = \frac{1}{2}N + s$$

$$N_\downarrow = \frac{1}{2}N - s$$

$$N_\uparrow - N_\downarrow = 2s$$

This difference 2s is called the excess property (spin excess in this case). It was previously established that the total number of possible states of the system consisting of indistinguishable particles can be found out using the generating function $(\uparrow + \downarrow)^N$. This can be expanded using binomial theorem. For a general x and y,

$$(x+y)^N = \sum_{t=0}^{N} \frac{N!}{(N-t)!\,t!} x^{N-t} y^{t} \quad -(1)$$

When $t = \frac{1}{2}N - s$

$$(x+y)^N = \sum_s \frac{N!}{(\frac{1}{2}N+s)!.(\frac{1}{2}N-s)!} x^{\frac{1}{2}N+s} y^{\frac{1}{2}N-s} \quad -(2)$$

$$\Rightarrow (\uparrow+\downarrow)^N = \sum_s \frac{N!}{(\frac{1}{2}N+s)!(\frac{1}{2}N-s)!} \uparrow^{\frac{1}{2}N+s} \downarrow^{\frac{1}{2}N-s}$$

$$-(3)$$

The coefficient of the term $\uparrow^{\frac{1}{2}N+s} \downarrow^{\frac{1}{2}N-s}$ gives the number of states having spin-up $N_\uparrow = \frac{1}{2}N + s$ and the number of states having spin down $N_\downarrow = \frac{1}{2}N - s$ resulting in the spin excess of 2s. The net energy of such a system will be $2sm\alpha$.

The coefficients of the binomial expansion are regarded as the "degeneracy function".

$$g(N,s) = \frac{N!}{(\frac{1}{2}N+s)!\,(\frac{1}{2}N-s)!} = \frac{N!}{N_\uparrow!\,N_\downarrow!} \quad -(4)$$

$$(\uparrow + \downarrow)^N = \sum_{s=-\frac{1}{2}N}^{\frac{1}{2}N} g(N,s)\, \uparrow^{\frac{1}{2}N+s} \downarrow^{\frac{1}{2}N-s} \quad -(5)$$

The value of the degeneracy function will give the number of states with the same spin excess. Therefore, a distribution of states can be obtained from g(N, s). The distribution has a nature shown in the following figure.

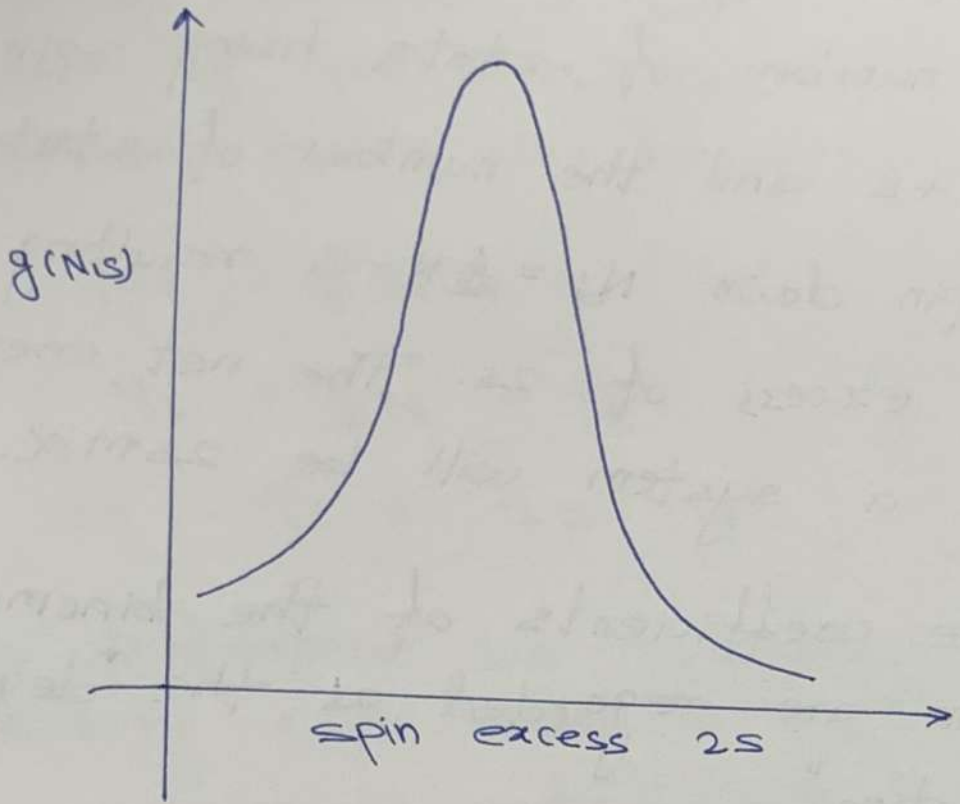


Figure : Variation of the degeneracy function with spin excess.

The distribution shown in the above figure can provide vital information on the property of the system observed at a macroscopic level. If the distribution is very sharp i.e. the width of the peak is very narrow then a very large number of states has the same property and that value of the property is observed macroscopically.

For any material, the number of particles present is of the order of Avogadro number. In order to analyze such large numbers, it is convenient to deal with their logarithms.

$$\log\{g(N,s)\} = \log N! - \log(\tfrac{1}{2}N+s)! - \log(\tfrac{1}{2}N-s)! \quad - (6)$$

Alternatively,

$$\log\{g(N,s)\} = \log N! - \log N_\uparrow! - \log N_\downarrow! \quad - (7)$$

The term $\log N!$ can be evaluated using stirling approximation.

$$N! \approx (2\pi N)^{1/2} N^N \exp\{-N + 1/(12N) + \cdots\} \quad - (8)$$

For large N,

$$\log N! \approx \frac{1}{2}\log 2\pi + (N+\tfrac{1}{2})\log N - N \quad -(9)$$

Similarly,

$$\log N_\uparrow! = \frac{1}{2}\log 2\pi + (N_\uparrow + \tfrac{1}{2})\log N_\uparrow - N_\uparrow \quad -(10)$$

$$\log N_\downarrow! = \frac{1}{2}\log 2\pi + (N_\downarrow + \tfrac{1}{2})\log N_\downarrow - N_\downarrow \quad -(11)$$

But $N_\uparrow = \frac{1}{2}N + s$ and $N_\downarrow = \frac{1}{2}N - s$. Hence, Eq^n (9) can be written as

$$\log N! = \frac{1}{2}\log(2\pi/N) + \left(N_\uparrow + \frac{1}{2} + N_\downarrow + \frac{1}{2}\right)\log N - (N_\uparrow + N_\downarrow) \qquad -(12)$$

Using Eq^n (10), (11) and (12). Eq^n (7) can be written as

$$\log\{g(N,s)\} \approx \frac{1}{2}\log(1/2\pi N) - \left(N_\uparrow + \frac{1}{2}\right)\log(N_\uparrow/N) - (N_\downarrow + 1/2)\log(N_\downarrow/N) \qquad -(13)$$

Now $\log(N_\uparrow/N) = \log\frac{1}{2}\left(1 + \frac{2s}{N}\right)$

$$= -\log 2 + \log\left(1 + \frac{2s}{N}\right)$$

$$= -\log 2 + \left(\frac{2s}{N}\right) - \left(\frac{2s^2}{N^2}\right) + \cdots$$

Similarly,

$$\log(N_\downarrow/N) = -\log 2 - \left(\frac{2s}{N}\right) - \left(\frac{2s^2}{N^2}\right) \cdots$$

$$\Rightarrow \log\{g(N,s)\} = \frac{1}{2}\log(2/\pi N) + N\log 2 - \frac{2s^2}{N} \qquad -(14)$$

$$\Rightarrow g\{(N,s)\} = g(N,0)\exp(-2s^2/N) \qquad -(15)$$

where $g(N,0) = \left(\frac{2}{\pi N}\right)^{1/2} 2^N$

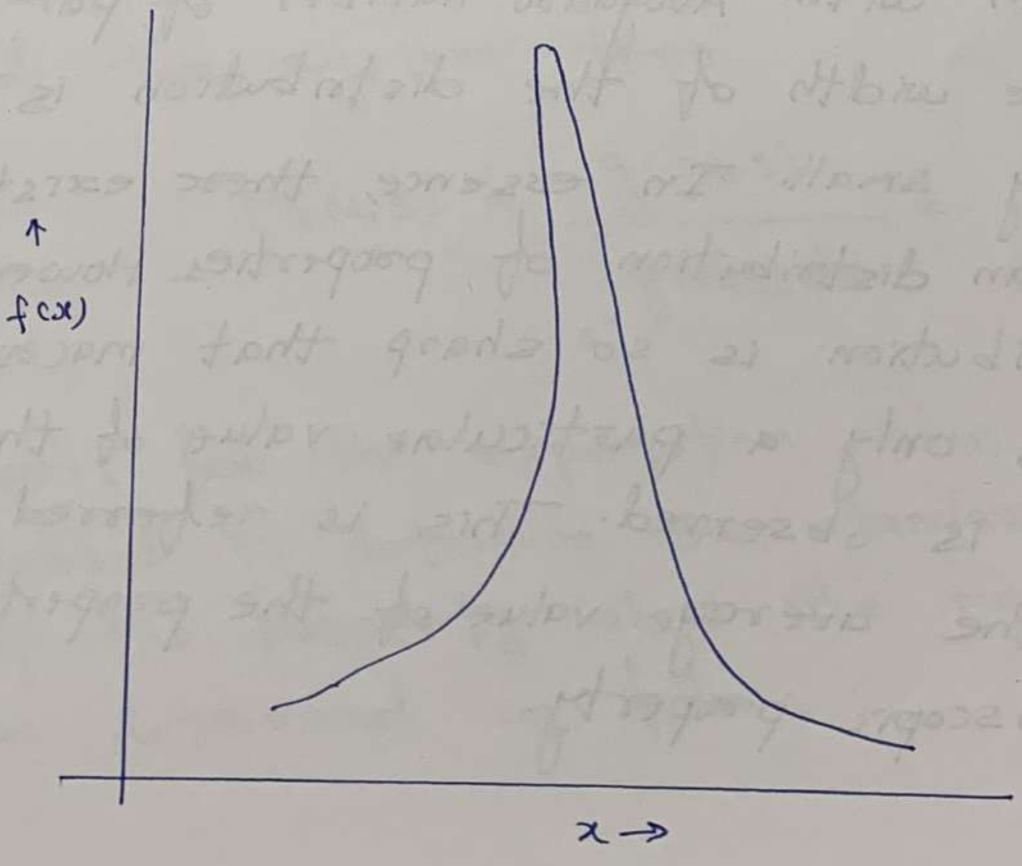
The nature of the curve shown by Eq^n (15) is bell-shaped and the distribution is called Gaussian distribution. The $g(N,0)$ value is the peak of the distribution giving the population of zero spin excess. The general expression for a Gaussian distribution is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\} \qquad -(16)$$

where μ = mean

σ = standard deviation.

The general nature of the curve is as shown below:



From the comparison of Eq^{n}s (15) and (16), it can be seen that for the present case, $\mu = 0$ and $\sigma = \sqrt{N}/2$. For a typical system, N is of the order of Avogadro number. Consider $N = 10^{24}$. In such a case, the standard deviation σ will be of the order of 10^{12}. The total width of the distribution is signified by all the spin excesses those can be sampled.

$\Rightarrow$ Width of the distribution $= N + 1$

$\approx 10^{24}$

Thickness of the distribution $\approx 10^{12}$

Therefore, the fractional width of the distribution is merely 10^{-12}. Therefore for a system with Avogadro number of particles the width of the distribution is extremely small. In essence, there exists a Gaussian distribution of properties. However, the distribution is so sharp that macroscopically, only a particular value of the property is observed. This is referred to as the average value of the property or macroscopic property.

If the probability distribution function for a certain quantity 'f' is given as $P(s)$ then the average value of the function is given as

$$\langle f \rangle = \sum_{s} f(s)\, P(s) \qquad -(17)$$

If f is continuous then

$$\langle f \rangle = \int_{s} f(s)\, P(s)\, ds \qquad -(18)$$

$\langle f \rangle$ is called the "expectation value" of the property f. The degeneracy function can be used to signify the probability distribution of the states. When there is an equal probability of accessing any stat of the system then

$$P(N,s) = \frac{g(N,s)}{\int_{s} g(N,s)\, ds} = \frac{g(N,s)}{2^{N}} \qquad -(19)$$

Apart from the expectation value, it is also desired to determine the standard deviation of the property about the mean. This can be determined as follows:

$$\langle s^2 \rangle = \frac{\left(\frac{2}{\pi N}\right)^{1/2} 2^N \int s^2 \exp(-2s^2/N)\, ds}{2^N}$$

$$\Rightarrow \langle s^2 \rangle = \left(\frac{2}{\pi N}\right)^{1/2} \left(\frac{N}{2}\right)^{\frac{3}{2}} \int_{-\infty}^{\infty} x^2 \exp(-x^2)\, dx$$

$$\Rightarrow \langle s^2 \rangle = \left(\frac{2}{\pi N}\right)^{1/2} \left(\frac{N}{2}\right)^{3/2} \left(\frac{\pi}{4}\right)^{1/2}$$

$$\Rightarrow \langle s^2 \rangle = \frac{N}{4}$$

$$\Rightarrow \langle (2s)^2 \rangle = N \qquad -(20)$$

The quantity $\langle (2s)^2 \rangle$ signifies the "mean square spin excess". The RMS spin excess is, therefore, equal to $\sqrt{N}$. If the fractional fluctuation 'f' is defined as root mean square deviation per particle then

$$f = \frac{\langle (2s)^2 \rangle^{1/2}}{N} = \frac{1}{\sqrt{N}} \qquad -(21)$$

It can be seen from the above equation that if the system has a large number of particles then the fluctuations will be very small. For $N \approx 10^{24}$, $f \approx 10^{-12}$ which is a very small quantity signifying a very sharp distribution.

To summarize, different particles in a system of a large collection of particles possess different microscopic properties. There is an extremely large difference between the total number of microscopic states of the system and the total number of available values of a given property. As a result, a large number of states possess the same property. This results into observation of a particular value of a property populated by a large number of states. This value of the property is observed macroscopically and is often referred to as the average property of the system.

The distribution of states among different values of a given property follows a Gaussian distribution. The sharpness of this Gaussian distribution gives rise to the macroscopically observed average property. The sharpness of the peak increases with an increase in the number of the particles in the system. Therefore, a macroscopic system, in principle, is allowed to show fluctuation in the properties, large fluctuations are not observed due to the sharpness of the peak.

The state of equilibrium

Consider a system consisting of two subsystems A and B such that they are in perfect thermal contact i.e. transfer of energy (and not particles) is allowed between the two sub-systems. The complete system is isolated with a total energy of U and the energies of the subsystems as U_A and U_B. Since the subsystems are in thermal contact, a re-distribution of energy is expected to take place. It is desired to know what governs this re-distribution and when will the re-distribution stop.

Due to the condition of isolation, the imposed constraint on the system is the total energy U of the system is a constant. However, there are a large number of ways in which the energy re-distribution can take place between the sub-systems. The re-distribution will take place such that the total number of accessible quantum states on every change increases. The equilibrium state therefore, is a state which has the maximum number of accessible quantum states. We illustrate this using a binary-state system described previously.

Consider two binary-state spin systems described previously kept in a perfect thermal contact. The two systems may have different spin excesses initially resulting in different energies. When the two systems are kept in thermal contact then the energies of the individual systems may change subject to the condition that the total energy of the system remains unchanged. If the spin excesses of the individual and the total system are denoted as S_A, S_B and S, respectively then

$$U_A(S_A) + U_B(S_B) = U(S) \quad -(1)$$

Another constraint imposed due to isolation is on the number of particles.

$$N_A + N_B = N \quad -(2)$$

N_A and N_B are also individually constants. Since total energy U is a function of s only, the third constraint is

$$s_A + s_B = s \quad -(3)$$

However, on thermal contact, a re-distribution of spin excesses is allowed subject to all of the above constraints.

Consider at any instance of time, the spin excesses of A and B as s_A and s_B, respectively. The accessible states of the two systems individually with these spin excesses will be given by the degeneracy functions $g_A(N_A, s_A)$ and $g_B(N_B, s_B)$, respectively. For the combined system, the total number of accessible states becomes the multiplication of the individual states. This is because for one state of A, 'p' number of states of B will give rise to 'p' number of states of the combined system and so on. Therefore, the degeneracy function for the combined system can be written as

$$g(N,S) = \sum_{S_A=-\frac{1}{2}N_A}^{\frac{1}{2}N_A} g_A(N_A,S_A)\, g_B(N_B,S-S_A) \qquad -(4)$$

Different values of S_A will give different configurations of the system (i.e. states). The above summation is subject to the constraints of $S = \text{constant}$, $N_A = \text{constant}$ and $S_B = \text{constant}$. There may exist a value of S_A such that $g(N,S)$ attains a maxima. It is desirable to check if this situation exists at all. In essence, we want to check if there exists a particular value of total spin excess which will make the total number of accessible quantum states the maximum. Let this value be $\hat{S}_A$. The configuration at which the total degeneracy function becomes maximum is called the most probable configuration.

We have established

$$g_i(N_i, s_i) = g_i(N_i, 0)\exp(-2s_i^2/N_i) \qquad -(5)$$

$$\Rightarrow g_A(N_A, S_A)\, g_B(N_B, S_B)$$

$$= g_A(N_A,0)\, g_B(N_B,0)\exp\left\{-\left(\frac{2S_A^2}{N_A^2} + \frac{2S_B^2}{N_B^2}\right)\right\}$$

$$-(6)$$

Eq^n (6) gives the total number of accessible quantum states of the combined system when the total spin excess is $2S$ and the spin excesses of the individual systems is $2S_A$ and $2S_B$. It is now required to find the maxima of the above function with respect to S_A. We use the logarithmic scale with the first & the second derivatives.

$$\log\left\{g_A(N_A,S_A)\, g_B(N_B,S-S_A)\right\}$$

$$= \log(g_A g_B)_0 - \frac{2S_A^2}{N_A} - \frac{2(S-S_A)^2}{N_B}$$

$$\Rightarrow \frac{\partial}{\partial S_A}\left\{g_A(N_A,S_A)\, g_B(N_B,S-S_A)\right\}$$

$$= -\frac{4S_A}{N_A} + \frac{4(S-S_A)}{N_B} \qquad -(7)$$

$$\Rightarrow \frac{\partial^2}{\partial S_A^2}\left\{g_A(N_A,S_A)\, g_B(N_B,S-S_A)\right\}$$

$$= -4\left(\frac{1}{N_A} + \frac{1}{N_B}\right) \qquad -(8)$$

Since $N_A > 0$ and $N_B > 0$ always, the second derivative is negative indicating a maxima. The condition for maxima from Eq^n (7) can now be written as:

$$\frac{\hat{S}_A}{N_A} = \frac{\hat{S}_B}{N_B} = \frac{\hat{S}}{N} \qquad -(9)$$

Accordingly, the most probable configuration follows

$$\left[g_A(N_A, S_A)\, g_B(N_B, S_B)\right]_{MAX} = g_A(N_A, 0)\, g_B(N_B, 0) \exp\left(\frac{-2S^2}{N}\right) \quad -(10)$$

Eq^n (10) gives the most probable configuration of the combined system. The re-distribution will stop when this configuration has been attained.

Eq^n (10) is a Gaussian distribution. Therefore, it is required to know about the width of the distribution to have an idea about the spread and the probability of finding the system outside the most probable configuration due to fluctuations. Consider a small fluctuation δ introduced to the system when it is in the most probable configuration such that

$$S_A = \hat{S}_A + \delta$$

$$S_B = \hat{S}_B - \delta$$

$$\Rightarrow \quad S_A^2 = \hat{S}_A^2 + 2 S_A \delta + \delta^2$$

$$\text{and} \quad S_B^2 = \hat{S}_B^2 - 2 S_B \delta + \delta^2$$

$$\Rightarrow g_A(N_A, S_A)\, g_B(N_B, S_B)$$

$$= (g_A g_B)_{MAX} \exp\left(\frac{-4\hat{S}_A \delta}{N_A} - \frac{2\delta^2}{N_A} + \frac{4\hat{S}_2 \delta}{N_B} - \frac{2\delta^2}{N_B}\right)$$

But at maxima, $\hat{S}_A / N_A = \hat{S}_B / N_B$. Therefore,

$$g_A(N_A, \hat{S}_A + \delta)\, g_B(N_B, \hat{S}_B - \delta)$$

$$= (g_A g_B)_{MAX} \exp\left(- \frac{2\delta^2}{N_A} - \frac{2\delta^2}{N_B} \right) \quad -(11)$$

Eq^n (11) gives the combined degeneracy function when a fluctuation δ takes place in the system so as to take the system away from equilibrium. Consider a system with $N_A = N_B = 10^{22}$. If a fluctuation of $\delta = 10^{12}$ is introduced to the system then the fraction fluctuation will be $\delta / N_A = 10^{-10}$. From Eq^n (11), it can be seen that such a small fractional fluctuation of 10^{-10} will reduce the generating function by e^{-400}. This means that the population of the system in state of fluctuation beyond the most probable configuration is e^{-400} times i.e. 1.92×10^{-174} times that of the most probable configuration. Therefore, the system at equilibrium does not deviate from equilibrium once equilibrium has been attained.

Thermal equilibrium and the concept of temperature:

Using the concept of combined degeneracy function, we can establish the concept of temperature. Consider two systems with energies U_A and U_B interacting within a constraint of total energy $U = U_A + U_B$. The combined degeneracy function for the system can be written as

$$g(N,U) = \sum_{U_A} g_A(N_A,U_A)\, g_B(N_B,U_B) \qquad -(12)$$

The above summation is over all possible values of $U_A \leq U$. Consider an infinitesimal change taking place in the system at equilibrium.

$$dg(N,U) = \left[\left(\frac{\partial g_A}{\partial U_A}\right)_{N_A} dU_A\right] g_B + \left[\left(\frac{\partial g_B}{\partial U_B}\right)_{N_B} dU_B\right] g_A = 0$$

In light of the constraint $dU_A + dU_B = 0$

$$\frac{1}{g_A}\left(\frac{\partial g_A}{\partial U_A}\right)_{N_A} = \frac{1}{g_B}\left(\frac{\partial g_B}{\partial U_B}\right)$$

$$\Rightarrow \left(\frac{\partial \ln g_A}{\partial U_A}\right)_{N_A} = \left(\frac{\partial \ln g_B}{\partial U_B}\right)_{N_B} \qquad -(13)$$

Eq^n (13) gives the condition for equilibrium of two systems in thermal contact. It will be proven shortly later that the term $\ln(g_i)$ is an indicator of the entropy of the system.

$$S_A \propto \ln(g_A)$$

$$S_A = k_B \ln(g_A) \qquad - (14)$$

where k_B = Boltzmann constant $= 1.38 \times 10^{-23}\ JK^{-1}$.

Eq^n (13) can, therefore, be written as

$$\frac{1}{k_B}\left(\frac{\partial S_A}{\partial U_A}\right)_{N_A} = \frac{1}{k_B}\left(\frac{\partial S_B}{\partial U_B}\right)_{N_B} \qquad - (15)$$

The quantity $(\partial S_i/\partial U_i)_N^{-1}$ can be regarded as the temperature. Remember the combined first and second law of thermodynamics

$$dU = TdS - pdV$$

$$\Rightarrow T = (\partial U/\partial S)_V$$

Hence, we regain our classical definition of thermal equilibrium as $T_A = T_B$. Therefore, it can be seen that the obvious condition of equality of temperatures arises naturally following the concepts of degeneracy and most probable configuration. We will deal with other types of equilibrium later in the discussion.

Directionality and entropy change for spontaneous processes:

It can be shown statistically that when two bodies are brought in thermal contact then the entropy other combined insulated system always increases. For an isolated system with subsystems A and B, $U_A + U_B = U$ = constant. The combined multiplicity function of the system in thermal contact is

$$g(N, U) = \sum_{U_A} g_A(N_A, U_A)\, g_B(N_B, U-U_A) \quad -(16)$$

None of the terms in the above equation is negative. Further, every term is greater than one. Therefore, the resultant degeneracy function $g(N, U)$ is always greater than the initial individual multiplicity functions. Therefore, it can be concluded that the total number of quantum states always increase on thermal contact. Since the number of quantum states or population of the states as obtained by degeneracy function is an indicator of entropy of the system, entropy can be concluded to always increase.

The fundamental postulate of statistical thermodynamics :

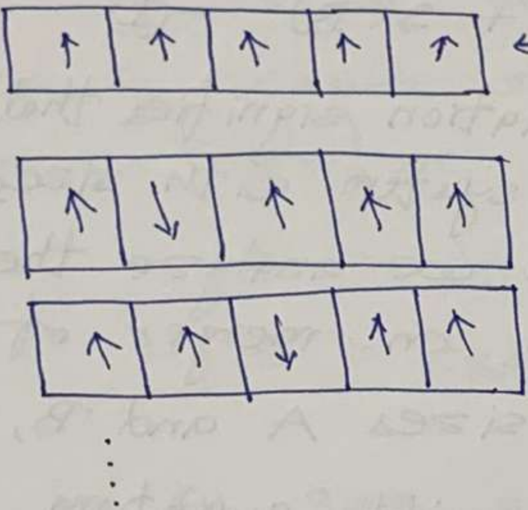
There are certain properties associated with a system which can be described without any reference to the detailed molecular-level construction of the system. For example, the volume of a system conveys exactly the same information about the system irrespective of a knowledge of molecular-level construction of the system i.e. whether the system is a continuum or consists of a collection of a large number of discrete particles. However, not all the properties of a system can be identified in this manner. It was realized by Boltzmann that the entropy of a system is related to the number of quantum states of the system. Therefore the concept of entropy can be understood only on the basis of a molecular level description of the matter. The theory of statistical mechanics, which considers matter to be composed of discrete particles, is developed on the basis of the following fundamental postulate.

"An isolated system at equilibrium is equally likely to be in any of the accessible quantum states".

The above postulate is also called the principle of equal apriori probability. It is important to note that in the fundamental postulate, only accessible quantum states are included. Accessibility of quantum states depends upon the constraints imposed on the system. The postulate is proposed without proof and the theory is developed on the basis of it. Since no experimental observation has been observed to contradict the theory developed on the basis of the postulate, the postulate is considered to be true.

In classical thermodynamics, the entropy of a system signifies disorder or randomness of the system. An increase in randomness of the system is manifested as an increase in the entropy of the system. We now quantify randomness and relate it to the concept of entropy.

Consider the previous case of a binary-state system. If all the particles are forced to be only spin-up (or equivalently spin-down) then the system loses all the randomness. As the particles are allowed to have different spins, the disorder or randomness of the system increases. We call this as an increase in entropy of the system.



← All spin-up, no disorder, no entropy

← Some spin-down, some disorder, increase in entropy

The spin-up or spin-down of the particles signifies the microstates or quantum states of the system and a collective information of the spin states gives the macrostate of the system. With an increase in different spins, the number of states also increases. From the scheme shown above, an increase in the number of states seems to be related in some fashion to the entropy of the system.

If the total number of accessible quantum states of the system is 'Ω' then the relationship between Ω and the entropy 's' of the system can be developed as follows:

Entropy is an extensive quantity. Therefore, the entropies of the system at sizes A and B of the system would be S(A) and S(B), respectively. Further,

$$S(A+B) = S(A) + S(B) \quad -(1)$$

A + B in the above equation signifies that now we have a larger system with sizes A and B merged. Now we analyze the accessible quantum states on merger of A and B. At system sizes A and B, the total number of accessible quantum states are $\Omega(A)$ and $\Omega(B)$, respectively. When A and B are merged, the resultant quantum states donot add up but get multiplied.

$$\Omega(A+B) = \Omega(A)\,\Omega(B) \quad -(2)$$

It was previously deduced from arguments that the entropy of the system is a function of the number of accessible quantum states.

$$S = f(\Omega)$$

$$\Rightarrow S(A) = f[\Omega(A)]$$

$$S(B) = f[\Omega(B)]$$

$$\Rightarrow S(A) + S(B) = f[\Omega(A)] + f[\Omega(B)]$$

$$\Rightarrow S(A+B) = f[\Omega(A)] + f[\Omega(B)]$$

$$\Rightarrow f[\Omega(A+B)] = f[\Omega(A)] + f[\Omega(B)]$$

$$\Rightarrow f[\Omega(A)\,\Omega(B)] = f[\Omega(A)] + f[\Omega(B)] \qquad -(3)$$

Eq^n (3) is of the form $f(xy) = f(x) + f(y)$. This is possible only if f is a logarithmic function. Therefore,

$$S \propto \ln\Omega$$

$$\Rightarrow S = k_B \ln\Omega \qquad -(4)$$

Eq^n (4) gives a quantitative relationship between entropy and accessible quantum states of the system. Therefore, one can in principle, "count" the total number of quantum states of a system and calculate the entropy of the system. At 0 K for a perfectly crystalline system, there is only one quantum state and therefore $\ln\Omega = 0$. Hence, entropy becomes zero, which is also the statement of the third law of thermodynamics

Partition function :

Consider an isolated state at equilibrium. The system has a large number of accessible quantum states. At any time t, suppose that the system is found in "ith" quantum state. According to the fundamental postulate, all the accessible quantum states are equally likely. Therefore the probability of the system to be in ith quantum state is

$$P_i = \frac{1}{\Omega} \quad - (5)$$

From Eqn (4)

$$S = k_B \ln \Omega$$

$$\Rightarrow \Omega = e^{S/k_B}$$

$$\Rightarrow P_i = e^{-S/k_B} \quad - (6)$$

Now consider a system A in thermal contact with another system B both together constituting an isolated system of total energy E_T. If at any time t, the energy of A is E_A then the energy of B will be $E_T - E_A$. Let 'i' and 'j' be two accessible quantum states of A with energies E_{Ai} and E_{Aj}. The quantum states accessible to B will become restricted to those having energies $E_T - E_{Ai}$ and $E_T - E_{Aj}$, respectively.

$P(E_{Ai})$ = probability of find the system A in a quantum state i with energy E_{Ai}.

$$P(E_{Ai}) \propto \Omega_A(E_{Ai})$$

Equivalently,

$$P(E_{Ai}) \propto \Omega_B(E_T - E_{Ai})$$

Similarly, for the quantum state j with energy E_{Aj},

$$P(E_{Aj}) \propto \Omega_B(E_T - E_{Aj})$$

$$\Rightarrow \frac{P(E_{Ai})}{P(E_{Aj})} = \frac{\Omega_B(E_T - E_{Ai})}{\Omega_B(E_T - E_{Aj})}$$

$$\Rightarrow \frac{P(E_{Ai})}{P(E_{Aj})} = \exp\left(\Delta S_B / k_B\right) \quad - (7)$$

where ΔS_B is the difference in entropy of the system B in two quantum states

$$\Delta S_B = S_B(E_T - E_{Ai}) - S_B(E_T - E_{Aj})$$

A Taylor series expansion of $S_B(E_T - E_{Ai})$ about E_T gives

$$S_B(E_T - E_{Ai}) = S_B(E_T) - E_{Ai}\left(\frac{\partial S_B}{\partial E}\right)_{V,N} + \cdots$$

$$\Rightarrow S_B(E_T - E_{Ai}) = S_B(E_T) - \frac{E_{Ai}}{T}$$

$$\Rightarrow \Delta S_B = -\left(\frac{E_{Ai} - E_{Aj}}{T}\right)$$

Substituting the above equation in Eq^n ⑦

$$\frac{P(E_{Ai})}{P(E_{Aj})} = \exp\left\{-\left(\frac{E_{Ai} - E_{Aj}}{k_B T}\right)\right\}$$

$$\Rightarrow P(E_{Ai}) \propto \exp(-E_{Ai}/k_B T)$$

The proportionality in the above equation can be converted to equality using normalization condition. The sum of probabilities of finding the system over all possible quantum states is unity. Therefore, if the proportionality constant is denoted by Z then

$$P(E_{Ai}) = Z \exp(-E_{Ai}/k_B T)$$

$$\Rightarrow \sum_i P(E_{Ai}) = \sum_i Z \exp(-E_{Ai}/k_B T)$$

$$\Rightarrow Z = \frac{1}{\sum_i \exp(-E_{Ai}/k_B T)}$$

$$\Rightarrow P(E_{Ai}) = \frac{\exp(-E_{Ai}/k_B T)}{\sum_i \exp(-E_{Ai}/k_B T)} \qquad -⑧$$

The sum in the previous equation is over all possible quantum states. The quantity $\exp(-E_A/k_B T)$ is called the "Boltzmann factor". The proportionality constant Z is called the "partition function". It is an extremely important quantity as it contains the information about all accessible quantum states of the system, condensed in a single expression. It can be used for the evaluation of different thermodynamic quantities, as will be shown later.

The probability distribution developed above can be used for developing the expectation values of the properties of the system. For any general property B, the expectation value can be obtained as

$$\langle B \rangle = \sum_i B_i \left[\frac{\exp(-E_{Bi}/k_B T)}{\sum_i B_i \exp(-E_{Bi}/k_B T)} \right] \quad -(9)$$

where the summation, as mentioned before, is over all possible quantum states of the system. Now we explore how to calculate various thermodynamic properties using partition function.

Thermodynamic properties from partition function

Average energy:

$$\langle E \rangle = \sum_i E_i P_i$$

$$= \sum_i E_i \left(\frac{e^{-E_i/k_BT}}{\sum_i e^{-E_i/k_BT}} \right)$$

Let $\frac{1}{k_BT} = \beta$

$$\Rightarrow \langle E \rangle = \frac{\sum_i E_i e^{-\beta E_i}}{Z}$$

$$= \frac{-\sum_i \frac{\partial}{\partial \beta} \left(e^{-\beta E_i}\right)_{N,V}}{Z}$$

$$= \left[\frac{-\partial}{\partial \beta} (Z) \right] \Big/ Z \Big|_{N,V}$$

$$\Rightarrow \langle E \rangle = \left(-\frac{\partial \ln Z}{\partial \beta} \right)_{N,V} \quad -(10)$$

But $\beta = \frac{1}{k_BT}$ or $T = \frac{1}{k_B\beta}$

$$\Rightarrow \left(\frac{\partial T}{\partial \beta} \right)_{N,V} = -\frac{1}{k_B \beta^2} = -k_BT^2 \quad -(11)$$

From Eq (10)

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial T}\,\frac{\partial T}{\partial \beta}$$

$$= \left(-\frac{\partial \ln Z}{\partial T}\right)\cdot\left(-k_B T^2\right)$$

$$\Rightarrow \langle E \rangle = k_B T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} \quad -(12)$$

Specific heat at constant volume :

$$C_V = \left(\frac{\partial \langle E \rangle}{\partial T}\right)_{N,V}$$

$$= \frac{\partial}{\partial T}\left\{k_B T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}\right\}_{N,V}$$

$$\Rightarrow C_V = 2k_B T\left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} + k_B T^2 \left(\frac{\partial^2 \ln Z}{\partial T^2}\right)_{N,V} \quad -(13)$$

Entropy :

$$S = \int_0^T \left(\frac{C_V}{T}\right) dT$$

$$\Rightarrow S = \int_0^T \left\{2k_B T\left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} + k_B T^2\left(\frac{\partial^2 \ln Z}{\partial T^2}\right)_{N,V}\right\} dT \quad -(14)$$

We evaluate the two terms on RHS individually.

$$\int_0^T 2k_B \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} dT = 2k_B \ln Z \Big|_0^T \qquad -(15)$$

$$\int_0^T k_B T \left(\frac{\partial^2 \ln Z}{\partial T}\right)_{N,V} dT$$

$$= k_B T \frac{\partial}{\partial T}(\ln Z)_{N,V}\Big|_0^T - \int_0^T k_B \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} dT$$

$$= k_B T \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}\Big|_0^T - k_B \ln Z \Big|_0^T \qquad -(16)$$

From Eqs (14), (15) and (16),

$$S = k_B \ln Z \Big|_0^T + k_B T \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}\Big|_0^T$$

$$\Rightarrow \quad S = k_B \ln Z + k_B T \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} \qquad -(17)$$

It is to be noted that Eq^n (17) gives the value of entropy S rather than the expectation value $\langle S \rangle$. This is because S_i. the entropy of i^{th} microstate is always zero. $\langle S \rangle$ is not defined.

Helmholtz free energy :

From the previous derivations,

$$S = k_B \ln Z + k_B T \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}$$

$$\langle U \rangle = k_B T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}$$

$$\langle A \rangle = \langle U \rangle - TS$$

$$\Rightarrow \langle A \rangle = k_B T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} - k_B T \ln Z$$

$$- k_B T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}$$

$$\Rightarrow \langle A \rangle = - k_B T \ln Z \qquad - (18)$$

Pressure :

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

$$= - \frac{\partial}{\partial V}\left(- k_B T \ln Z\right)_{T,N}$$

$$\Rightarrow \langle P \rangle = k_B T \left(\frac{\partial \ln Z}{\partial V}\right)_{T,N} \qquad - (19)$$

From all of the above derivations, it can be seen that partition function $Z(T,V,N)$ can be used as a tool to determine the macroscopic thermodynamic properties.

Diffusive contact and Gibbs sum :

The system considered till now are the ones in which only a rearrangement of energy was allowed between the subsystems. Boltzmann factor and partition function can be used only when the system is in a thermal contact with a reservoir i.e. when exchange of energy is allowed keeping the number of atoms in each individual compartment to be individually constant. A diffusive contact is defined as the one which allows exchanges of both energy and particles between the system and the reservoir. Therefore unlike thermal contacts in which the systems are characterized by (N, V, T), the systems under diffusive contacts are the characterized by (μ, V, T), μ being the chemical potential.

To analyze such systems, let us consider a composite body consisting of a small body S in diffusive contact with a reservoir. Now the reservoir not only acts as the reservoir for energy but also for the exchange of particles.

The composite body as a whole is perfectly insulated with total eneagy of E_0 and the total number of particles N_0. The conditions for equilibrium under diffusive contacts is

$$T_S = T_R$$

$$\mu_S = \mu_R$$

(* $\mu_i^S = \mu_i^R$ for a multicomponent system)

Further,

$$E_R + E_S = E_0$$

$$N_R + N_S = N_0$$

It is required to determine the probability of observing the system to have N particles with a quantum state of eneagy E_s.

$P(N, E_s)$ = probability that the system has N particles and it is in a quantum state with eneagy E_s.

$$P(N, E_s] \propto g_s(N, E_s)$$

The degeneracy function of the reservoir corresponding to the given degeneracy of the system will be $g_R(N_0 - N, E_0 - E_s)$

$$P(N, E_s) \propto g_R(N_0 - N, E_0 - E_s)$$

$$\Rightarrow \frac{P(N_1, E_1)}{P(N_2, E_2)} = \frac{g_R(N_0 - N_1, E_0 - E_1)}{g_R(N_0 - N_2, E_0 - E_2)} \quad -(20)$$

But entropy of the system is defined as

$$S(N, E) = k_B \ln g(N, E)$$

$$\Rightarrow S(N, E) = k_B \ln g_R(N_0 - N, E_0 - E)$$

$$\Rightarrow \frac{P(N_1, E_1)}{P(N_2, E_2)} = \frac{\exp\{S(N_0 - N_1, E_0 - E_1)/k_B\}}{\exp\{S(N_0 - N_2, E_0 - E_2)/k_B\}}$$

$$\Rightarrow \frac{P(N_1, E_1)}{P(N_2, E_2)} = \exp(\Delta S / k_B) \quad -(21)$$

where $\Delta S = S(N_0 - N_1, E_0 - E_1) - S(N_0 - N_2, E_0 - E_2)$

Using Taylor series expansion about N_0, E_0

$$S(N_0 - N, E_0 - E) = S(N_0, E_0) - N\left(\frac{\partial S}{\partial N_0}\right)_{E_0} - E\left(\frac{\partial S}{\partial E_0}\right)_{N_0} \cdots$$

$$\Rightarrow \Delta S = -(N_1 - N_2)\left(\frac{\partial S}{\partial N_0}\right)_{E_0} - (E_1 - E_2)\left(\frac{\partial S}{\partial E_0}\right)_{N_0} + \cdots$$

But $\left(\frac{\partial S}{\partial E_0}\right)_{N_0} = \frac{1}{T}$ and

$$\left(\frac{\partial S}{\partial N_0}\right)_{U_0} = -\frac{\mu}{T}$$

$$\Rightarrow \quad \Delta S = \frac{(N_1 - N_2)\mu}{T} - \frac{(E_1 - E_2)}{T} \qquad -(22)$$

Substitution of Eq (22) in Eq^n (21) gives

$$\frac{P(N_1, E_1)}{P(N_2, E_2)} = \frac{\exp\left\{\frac{N_1\mu - E_1}{k_B T}\right\}}{\exp\left\{\frac{N_2\mu - E_2}{k_B T}\right\}}$$

$$\Rightarrow \quad P(N, E) \propto \exp\left(\frac{N\mu - E}{k_B T}\right)$$

$$\Rightarrow \quad P(N, E) = \left(\frac{1}{\mathcal{Z}}\right)\exp\left(\frac{N\mu - E}{k_B T}\right) \qquad -(23)$$

$\mathcal{Z}$ is the proportionality constant, called Gibbs sum. It is also called grand sum or grand partition function or grand canonical partition function. It can be obtained by normalization condition

$$\left(1/\mathcal{Z}\right) = \frac{1}{\sum_{N=0}^{\infty}\sum_{S(N)} \exp\left\{\frac{(N\mu - E)}{k_B T}\right\}} \qquad -(24)$$

The sum is over all possible energies of the system and over all such possibilities with all possible number of particles in the system.

Average number of particles :

$$\langle N\rangle = \sum N \frac{\exp\{(\mu N-E)/k_BT\}}{\sum\sum\{\exp(\mu N-E)/k_BT\}}$$

$$\mathcal{Z} = \sum\sum \exp\left\{(\mu N-E)/k_BT\right\}$$

$$\Rightarrow \left(\frac{\partial \mathcal{Z}}{\partial \mu}\right)_{T,V} = \frac{1}{k_BT}\sum\sum N\exp\left\{\frac{(\mu N-E)}{k_BT}\right\}$$

$$\Rightarrow \langle N\rangle = \frac{k_BT}{\mathcal{Z}}\left(\frac{\partial \mathcal{Z}}{\partial \mu}\right)_{T,V}$$

$$\Rightarrow \langle N\rangle = k_BT\left(\frac{\partial \ln \mathcal{Z}}{\partial \mu}\right)_{T,V} \qquad -(25)$$

Average energy in grand canonical ensemble:

$$\langle E\rangle = \frac{\sum U_i \exp\{\beta(\mu N-E_i)\}}{\mathcal{Z}}$$

But $\langle \mu N - E\}\rangle = \mu\langle N\rangle - \langle E\rangle$

$$= \frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}}{\partial \beta}$$

$$\Rightarrow \langle E\rangle = \langle N\rangle\mu - \frac{\partial \ln \mathcal{Z}}{\partial \beta}$$

$$\Rightarrow \langle E\rangle = \mu k_BT\left(\frac{\partial \ln \mathcal{Z}}{\partial \mu}\right)_{T,V} - k_BT^2\left(\frac{\partial \ln \mathcal{Z}}{\partial T}\right)_{\mu,V} \quad -(26)$$

Fluctuations in thermodynamic variables :

The relations previously derived give only the magnitudes of the ensemble averages. However, the spread about the mean needs to be determined. This can be obtained from the variance defined as

$$\sigma_E^2 = \langle (E - \langle E \rangle)^2 \rangle \qquad - (27)$$

standard deviation is the square root of the variance, denoted as σ_E.

$$\langle (E - \langle E \rangle)^2 \rangle = \langle E^2 - 2E\langle E \rangle + \langle E \rangle^2 \rangle$$
$$= \langle E^2 \rangle - \langle E \rangle^2$$

But from the definition,

$$\langle E^2 \rangle = \sum E_i^2 P_i$$
$$\langle E \rangle^2 = \left(\sum E_i P_i\right)^2$$

In canonical ensemble, From Eq^n (12),

$$\langle E \rangle^2 = \frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)^2_{N,V}$$

For any integer ν,

$$\langle E^\nu \rangle = \frac{(-1)^\nu}{Z}\left(\frac{\partial^\nu Z}{\partial \beta^\nu}\right)_{N,V}$$

$$\Rightarrow \quad \langle E^2 \rangle = \frac{1}{Z}\left(\frac{\partial^2 Z}{\partial \beta^2}\right)_{N,V}$$

$$\Rightarrow \quad \sigma_E^2 = \frac{1}{Z}\left(\frac{\partial^2 Z}{\partial \beta^2}\right)_{N,V} - \frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)^2_{N,V} \qquad - (28)$$

Now let us consider the individual terms.

$$\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{N,V} = \frac{\partial}{\partial \beta}\left(\frac{1}{Z}\frac{\partial Z}{\partial \beta}\right)$$

$$= -\frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)^2 + \frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2}$$

Therefore,

$$\sigma_E^2 = \left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} \qquad - (29)$$

Further,

$$\left(\frac{\partial^2 \ln Z}{\partial \beta^2}\right)_{V,N} = \frac{\partial}{\partial \beta}\left(\left(\frac{\partial \ln Z}{\partial \beta}\right)_{V,N}\right)_{V,N}$$

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

$$\Rightarrow \quad \sigma_E^2 = k_B T^2 \frac{\partial \langle E \rangle}{\partial T} \qquad - (30)$$

Fluctuations in other thermodynamic quantities can be found in an analogous manner.

Partition function for many-body systems:

Let a system be composed of two bodies, each of which can have two possible values of energies. For the combined system, the total number of accessible quantum states will be

state 1 : $E_1 = E_1^1 + E_1^2$

state 2 : $E_2 = E_2^1 + E_2^2$

state 3 : $E_3 = E_1^1 + E_2^2$

state 4 : $E_4 = E_1^2 + E_2^1$

The subscripts in the above notation are for the energies while the superscripts are for the particles. The partition function is the summation over all accessible quantum states. Therefore,

$$Z = \sum_{i=1}^{4} e^{-\beta E_i}$$

$$= e^{-\beta(E_1^1 + E_1^2)} + e^{-\beta(E_2^1 + E_2^2)} + e^{-\beta(E_1^1 + E_2^2)} + e^{-\beta(E_1^2 + E_2^1)}$$

$$= e^{-\beta E_1^1} e^{-\beta E_1^2} + e^{-\beta E_2^1} e^{-\beta E_2^2} + e^{-\beta E_1^1} e^{-\beta E_2^2} + e^{-\beta E_1^2} e^{-\beta E_2^1}$$

$$\Rightarrow Z = e^{-\beta E_1^1}\left(e^{-\beta E_1^2} + e^{-\beta E_2^2}\right)$$

$$+ e^{-\beta E_2^1}\left(e^{-\beta E_1^2} + e^{-\beta E_2^2}\right)$$

$$= \left(e^{-\beta E_1^1} + e^{-\beta E_2^1}\right)\left(e^{-\beta E_1^2} + e^{-\beta E_2^2}\right)$$

$$= Z(\text{particle } 1) \times Z(\text{particle } 2)$$

On generalization to N particles

$$Z = \prod_{i=1}^{N} Z_i \qquad -(51)$$

Hence, for distinguishable particles which are non-interacting, the partition function is the product of single-particle partition functions.